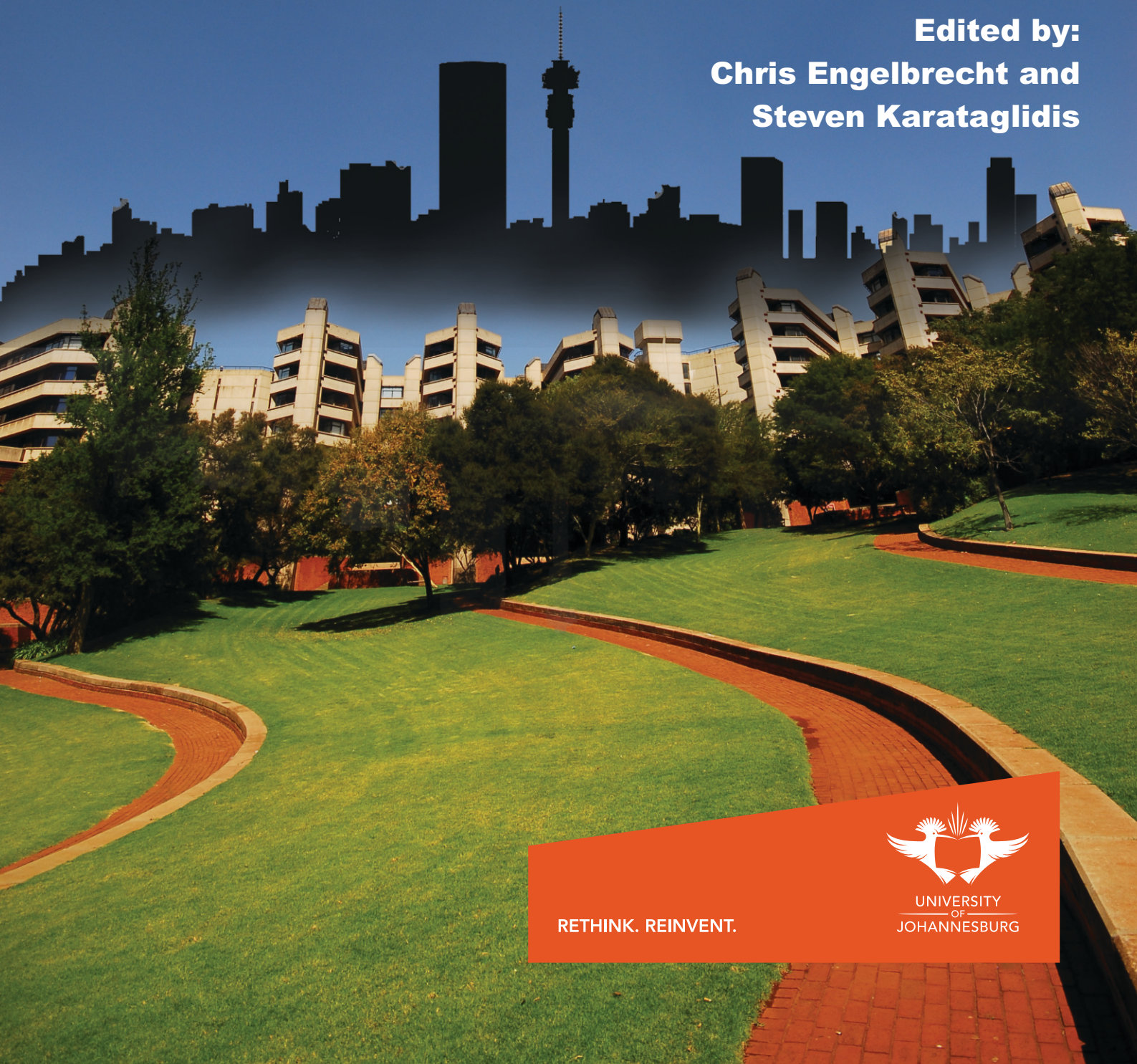




Proceedings of SAIP2014: the 59th annual conference of the South African Institute of Physics

University of Johannesburg, 7 - 11 July 2014

**Edited by:
Chris Engelbrecht and
Steven Karataglidis**



RETHINK. REINVENT.



PROCEEDINGS EDITORIAL TEAM

Editors: Chris Engelbrecht and Steven Karataglidis (University of Johannesburg)

Advisor: Ilsa Basson (UNISA)

Advisor and online administration: Roelf Botha (HartRAO and SAIP)

Online administration and PDF Compilation: Juan Grey (SAIP)

REVIEW PANEL

Roelf Botha (HartRAO and SAIP)

Emanuela Carleschi (University of Johannesburg)

Simon Connell (University of Johannesburg)

Bryan Doyle (University of Johannesburg)

Chris Engelbrecht (University of Johannesburg)

Steven Karataglidis (University of Johannesburg)

Paulus Masiteng (University of Johannesburg)

Azwinndini Muronga (University of Johannesburg)

Pap Nair (University of Johannesburg)

Sam Ramaila (University of Johannesburg)

Krish Reddy (University of Johannesburg)

Buyi Sondezi (University of Johannesburg)

Andre Strydom (University of Johannesburg)

Hartmut Winkler (University of Johannesburg)

PUBLISHER: The South African Institute of Physics (SAIP), <http://www.saip.org.za>

SAIP COPYRIGHT NOTICE: Copyright © 2015 by the South African Institute of Physics (SAIP)

The Proceedings of SAIP2014, the 59th annual conference of the South African Institute of Physics (SAIP) will only be available electronically on compact disk (CD) and on the SAIP website www.saip.org.za.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires specific permission and/ or a fee. Request required permissions from the SAIP Office, Tel. +27 (0)12 841 2627, Fax +27 (0)86 648 8474, or E-mail secretary@saip.org.za.

ISBN: 978-0-620-65391-6

SAIP2014

Proceedings of SAIP2014,
the 59th annual conference of the
South African Institute of Physics

Hosted by the University of Johannesburg

7 – 11 July 2014
University of Johannesburg
Johannesburg
South Africa

Edited by
Chris Engelbrecht and Steven Karataglidis

Table of Contents

<i>Message from the Conference Chair</i>	<i>xii</i>
<i>Message from the Review Committee</i>	<i>xiv</i>
<i>Conference Chairs and Committees</i>	<i>xv</i>
<i>List of Reviewers</i>	<i>xvi</i>

Full Research Papers

Division A – Division for Physics of Condensed Matter and Materials

Synthesis and characterization of the semiconducting intermetallic compound FeGa ₃	2
<i>M A M Ahmed, E Carleschi, G R Hearne and B P Doyle</i>	
Superconductivity in LaRh ₂ Sn ₂	7
<i>D Britz and A M Strydom</i>	
Effect of Mo content on the structural and physical properties of Cr _{100-x} Mo _x alloys	13
<i>A U Chavan, A R E Prinsloo, C J Sheppard and B Muchono</i>	
Magnetic and thermodynamic properties of Ce ₂₃ Ru ₇ Mg ₄ compound	19
<i>J C Debnath, A M Strydom, O Niehaus and R Pöttgen</i>	
Can Shapiro step subharmonics be "charged"?	24
<i>H Azemtsa Donfack, Y M Shukrinov and I R Rahmonov</i>	
Magnetic and Electronic studies on Cr _{100-x} Ir _x alloy single crystals	29
<i>P R Fernando, A R E Prinsloo and C J Sheppard</i>	
Combustion Synthesis of Dy ³⁺ -doped YVO ₄ phosphor	34
<i>K E Foka, B F Dejene and H C Swart</i>	
Pressure effects on the magnetic-electronic behavior of the local moment ferromagnet CeCuSi	41
<i>G R Hearne, G Diguët, A M Strydom, B Sondezi-Mhlungu, K Kamenev, F Baudelet and L Nataf</i>	
Probing the magnetic order in (Cr ₈₄ Re ₁₆) _{100-y} V _y alloys using neutron diffraction	47
<i>B S Jacobs, A R E Prinsloo, C J Sheppard, A M Venter and H E Maynard-Casely</i>	
Electrical properties of epitaxial Cr _{100-x} Co _x (100) alloy thin films on MgO(100)	53
<i>M B Kadam, C J Sheppard, A R E Prinsloo and E E Fullerton</i>	
Irradiation-induced improvement in crystalline quality of epitaxially grown InGaN thin films: A preliminary study	59
<i>M Madhuku, G Husnain, I Ahmad and H Saleem</i>	
The effect of silver (Ag) dopant on the structural and optical properties of sol gel prepared CdO nanoparticles	66
<i>M W Maswanganye, R V Makgobela, K E Rammutla, T E Mosuang and B W Mwakikunga</i>	

Electronic structures of oxygen adsorption on {110} nickel-rich pentlandite ($\text{Fe}_4\text{Ni}_5\text{S}_8$) mineral surface	72
<i>P P Mkhonto, H R Chauke and P E Ngoepe</i>	
Influence of zinc acetate concentration in the preparation of ZnO nanoparticles via chemical bath deposition	79
<i>F V Molefe, L F Koao, B F Dejene and H C Swart</i>	
Magnetic susceptibility studies of the $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloy system	85
<i>B Muchono, C J Sheppard, A R E Prinsloo and H L Alberts</i>	
Residual stress in polycrystalline thin Cr films deposited on fused silica substrates	91
<i>Z P Mudau, A R E Prinsloo, C J Sheppard, A M Venter, T P Ntsoane and E E Fullerton</i>	
Influence of solvent casting and weight ratios on the morphology and optical properties of inorganic-organic hybrid structures	97
<i>T F G Muller, A Ramashia, D E Motaung, F R Cummings, G F Malgas and C J Arendse</i>	
Structural characterisation of aluminium and yttrium co-doped tin oxide	103
<i>J N Ntimane, K E Rammutla and T E Mosuang</i>	
Structural and optical characterisation of double-doped TiO_2 nanoparticles	107
<i>O O Nubi, K E Rammutla and T E Mosuang</i>	
Synthesis and magnetic-electronic characterization of mixed-valence $\text{LuFe}_2\text{O}_{4.8}$: effect of stoichiometry δ	114
<i>M A Peck, W N Sibanda, G Diguët, C Martin, E Carleschi and G R Hearne</i>	
A study of crack formation and its effect on internal surface area using Micro-Focus X-ray Computerised Tomography and Fractal Geometry	120
<i>T Seakamela, G Nothnagel, F C de Beer and B P Doyle</i>	
Charge ordering dynamics under pressure in LuFe_2O_4	126
<i>W N Sibanda, E Carleschi, G Diguët, C Martin and G R Hearne</i>	
Pressure induced charge order collapse in Fe_2OBO_3	132
<i>W N Sibanda, E Carleschi, G Diguët, V Pischedda, J P Attfield and G R Hearne</i>	
Effects of catalyst:Ba molar ratio on the structure and luminescence properties of BaCO_3 :1% Eu^{3+} , 2% Dy^{3+} phosphors synthesized using sol-gel process	138
<i>A S Tebele, S V Motlounj and F B Dejene</i>	
Synthesis and characterisation of ZnO-CNFs hybrid nanostructures for hydrogen storage applications	144
<i>C T Thethwayo, S E Mavundla, T G Nyawo, P N Mbuyisa and O M Ndwandwe</i>	

Characterization of cerium doped yttrium gadolinium aluminate garnet (Y-Gd) ₃ Al ₅ O ₁₂ :Ce ³⁺ phosphor thin films fabricated by pulsed laser deposition	150
<i>A H Wako, F B Dejene and H C Swart</i>	
Division B - Nuclear, Particle and Radiation Physics	
Single top quark production in association with the Higgs: a feasibility study	157
<i>C Antel</i>	
Representation of the few-group homogenized cross sections of a MOX fuel assembly	163
<i>S Chifamba, D Botes, P M Bokov and A Muronga</i>	
Towards a crystal undulator	169
<i>S H Connell, J Härtwig, A Masvaure, D Mavunda and T N Tran Thi</i>	
The development of a General Purpose Processing Unit for the upgraded electronics of the ATLAS Tile Calorimeter	175
<i>M A Cox and B Mellado</i>	
Affordable and power efficient computing for high energy physics: CPU and FFT benchmarks of ARM processors	180
<i>M A Cox, R Reed and B Mellado</i>	
Cluster model analysis of exotic decay in actinide nuclei	185
<i>E J du Toit</i>	
Monte Carlo simulation of secondary gamma production during proton therapy for dose verification purposes	192
<i>J Jeyasugiththan and S Peterson</i>	
Radiation hardness of plastic scintillators for the Tile Calorimeter of the ATLAS detector	199
<i>H Jivan, R Erasmus, B Mellado, G Peters, K Sekonya and E Sideras-Haddad</i>	
Coupled-channel studies of nucleon scattering from oxygen isotopes	206
<i>S Karataglidis, J P Svenne, K Amos, L Canton, P R Fraser and D van der Knijff</i>	
Firmware development for the upgrade of the Tile Calorimeter of the ATLAS detector	211
<i>C O Kureba and B Mellado</i>	
Search for the Standard Model Higgs boson and its coupling to fermions with the ATLAS detector	217
<i>C A Lee</i>	
Characterization of damage in in-situ radiated plastic scintillators at the Tile calorimeter of the ATLAS	223
<i>L Maphanga, R Erasmus, H Jivan, B Mellado, G Peters and E Sideras-Haddad</i>	
Measurement of Higgs production in association with high p_T jets with the ATLAS detector	229
<i>B Mellado</i>	
The status of the LHeC project and its impact on Higgs Physics	235
<i>B Mellado</i>	

Validation of the performance of Geant4 in the simulation of neutron induced reactions relevant to reactor studies	241
<i>R Mudau, A Muronga and S Connell</i>	
Electron paramagnetic resonance analysis of plastic scintillators for the Tile Calorimeter of the ATLAS detector	248
<i>C Pelwan</i>	
An ATCA framework for the upgraded ATLAS read out electronics at the LHC	252
<i>R Reed</i>	
The ATLAS Tile Calorimeter hybrid demonstrator	257
<i>M Spoor and B Mellado Garcia</i>	
The Higgs as a portal to the hidden sector via an analysis of $H \rightarrow Z_d Z_d \rightarrow 4l$ using the ATLAS detector	263
<i>D A Unwuchola</i>	
Searching for BSM physics in Higgs to WW coupling with e^+e^- collisions	269
<i>S von Buddenbrock, A S Cornell and B Mellado</i>	
Memory performance of ARM processors and its relevance to High Energy Physics	275
<i>T Wrigley, G Harmsen and B Mellado</i>	
Octupole Correlation and Collective Coupling in the rare earth nucleus ^{154}Dy	281
<i>G L Zimba, S P Bvumbi, L P Masiteng, P Jones, S N T Majola, T S Dinoko, J F Sharpey-Schafer, E Lawrie, J J Lawrie, D Roux, P Papka, J E Easton, D Negi, O Shirinda and S Khumalo</i>	
Division C - Photonics	
Propagation characteristics control by variation of PCF structural parameters	288
<i>J F Jena, P Baricholo, T S Dlodlo and P K Buah-Bassuah</i>	
Division D1 - Astrophysics	
Radio astronomy in Africa: the case of Ghana	296
<i>B D Asabere, M J Gaylard, C Horellou, H Winkler and T Jarrett</i>	
Imaging & study of VLBI reference frame sources in the southern hemisphere	302
<i>S Basu, A De Witt, J Quick, A Bertarini and L Leeuw</i>	
Weighing dark matter in brightest cluster galaxies	308
<i>H Branken, S Loubser and K Sheth</i>	
The effect of an offset-dipole magnetic field on the Vela pulsar's γ -ray light curves	311
<i>M Breed, C Venter, A K Harding and T J Johnson</i>	

Spectral studies of Flat Spectrum Radio Quasars: constraining models at GeV energies	317
<i>R J G Britto, S Razzaque and B Lott</i>	
DFT, LS or GLS? A ‘road test’ of three period-finding techniques applied to stellar photometry	324
<i>C A Engelbrecht and F A M Frescura</i>	
A different approach to the perturbation of Astrophysical Fluids	330
<i>F A M Frescura and C A Engelbrecht</i>	
Comparing the results of two CCD image reduction packages	336
<i>R Kgoadi and C A Engelbrecht</i>	
Multi-wavelength classification of unidentified AGN in the Fermi 2LAC catalogue	341
<i>L Klindt, P J Meintjes and B van Soelen</i>	
Neutrino events at IceCube and the Fermi bubbles	347
<i>C Lunardini, S Razzaque and L Yang</i>	
Tidal effects on pulsation modes in close binaries	354
<i>M Predieri, F A M Frescura and C A Engelbrecht</i>	
Comparison of photometric and spectroscopic parameters of eclipsing contact binary stars	359
<i>P Skelton</i>	
Characterization of variable stars using the ASAS and SuperWASP databases	365
<i>D P Smits</i>	
The excitation of pulsation modes in rapidly rotating main sequence B-stars	371
<i>P van Heerden, C Engelbrecht and P Martinez</i>	
Spectral modelling of a H.E.S.S.-detected pulsar wind nebula	377
<i>C van Rensburg, P Krüger and C Venter</i>	
Search for Very High Energy candidate sources using South African observatories	383
<i>B van Soelen, L Klindt, I P van der Westhuizen, P J Meintjes, M Böttcher, L Hanlon, M Topinka and P Väisänen</i>	
Examining the behaviour of PKS 1424-418 during flaring	387
<i>P van Zyl, M Gaylard and S Colafrancesco</i>	
Modelling the average spectrum expected from a population of gamma-ray globular clusters	394
<i>C Venter and A Kopp</i>	
Spectral comparison between AGN at $z = 0.1, 0.2$ and 0.3	400
<i>H Winkler and J Tsuen</i>	
AUTOCAL: A software application for the calibration of stellar magnitudes	404
<i>D J Wium and B van Soelen</i>	

Division D2 - Space Science

- Influence of lightning on total electron content in the ionosphere using
WWLLN lightning data and GPS data 411
M M Amin, M Inggs and P J Cilliers

Division E - Physics Education

- Relationship between grade 10 learners' views about nature of science
and contextual factors 418
V M Baloyi, H I Nordhoff, W E Meyer, E Gaigher and M H W Braun
- Learners' conceptual resources for learning kinematics graphs 424
G Djan, M Lemmer and O N Morabe
- Student understanding of DC circuits: fine grained issues 430
I John and S Allie
- A CHAT perspective on the tensions and dynamics in the professional
development of Physical Sciences teachers in a mentoring relationship 436
S Ramaila and U Ramnarain
- First year university physics students' perceptions of the teaching-
learning environment: In search of a coherent pedagogic learning
orientation 442
S Ramaila and U Ramnarain
- Lesson planning perceptions and experiences of South African
Physical Sciences teachers in a new curriculum 449
S Ramaila and U Ramnarain
- Students' perceptions of the study process 455
S Ramaila, L Reddy, P Nair and S Bvumbi
- A critical assessment of first year entering university science students'
conceptual understanding 461
L Reddy, P Nair and S Ramaila
- A triggering strategy for improved pass rate in software-managed
evaluations of Physics practicals for the Engineering Programmes at
the University of Johannesburg 466
L Reddy, P Nair, S Ramaila and J Oelofse

Division F - Applied Physics

- A single DFB laser with multilevel directly modulated signal for high
speed optical fibre communication system 473
T V Chabata, E K Rotich Kipnoo, R R G Gamatham, A W R Leitch and T B Gibbon
- Overview of the Mineral-PET run-of-mine Diamond bearing rock
sorter 479
M Cook, M Tchonang, E Chinaka, M Bhamjee, F Bornman and S H Connell
- Interpretation of spectral electroluminescence images of photovoltaic
modules using modelling 485
J L Crozier, E E van Dyk and F J Vorster

Reproducing observed solar radiation characteristics in tropical regions using stochastic theoretical models	491
<i>M C Cyulinyana and H Winkler</i>	
Characterization of plume and thrust for the corona ionization space propulsion system	497
<i>P Ferrer</i>	
Modelling flow phenomena in time dependent store release from transonic aircraft	503
<i>D A MacLucas and I M A Gledhill</i>	
Polarization alignment system for quantum key distribution	509
<i>M Mariola, A Mirza and F Petruccione</i>	
Simulation of the Egyptian 2nd Testing Research Reactor (ETRR-2) experimental benchmark in aid of verification and validation of the OSCAR-4 system	515
<i>M Mashau, B Erasmus, R H Prinsloo and S H Connell</i>	
Analysis of a thermal conductivity measurement technique formulated as an inverse heat conduction problem	522
<i>V Ramnath</i>	
A quantum circuit modeling toolkit for high performance computing	528
<i>M Senekane, B Zulu and F Petruccione</i>	
The Alan Cousins Telescope - A robotic multi-purpose telescope for Sutherland	535
<i>P van Heerden, P Martinez and C Engelbrecht</i>	
Using a VCSEL to accurately measure the chromatic dispersion in single mode fibre by the phase shift technique	539
<i>S Wassin, E K Rotich Kipnoo, R R G Gamatham, A W R Leitch and T B Gibbon</i>	
Calculating solar irradiance to determine optimal angles of solar cells for De Aar	546
<i>G Webber and H Winkler</i>	
Division G - Theoretical and Computational Physics	
Evolution of quark masses and flavour mixings in the 2UED	552
<i>A Abdalgabar and A S Cornell</i>	
Structural and electronic properties of iron doped technetium sulphide	558
<i>M Abdulsalam and D P Joubert</i>	
An open quantum systems approach to Avian Magnetoreception	564
<i>B Adams, I Sinayskiy and F Petruccione</i>	
Computer assisted ‘proof ’ of the global existence of periodic orbits in the Rössler system	571
<i>A E Botha and W Dednam</i>	
Homogeneous open quantum walks on a line	578
<i>R C Caballar, I Sinayskiy and F Petruccione</i>	
Properties of the quark-gluon plasma observed at RHIC and LHC	584
<i>W A Horowitz</i>	

Finite-size key security of Phoenix-Barnett-Chefles 2000 quantum-key-distribution protocol	590
<i>M Mafu, K Garapo and F Petruccione</i>	
Quark-gluon plasma physics from string theory	596
<i>R Morad and W A Horowitz</i>	
First principle study of structural, thermal and electronic properties of the chalcopyrites AlAgX_2 ($\text{X}=\text{S}, \text{Se}, \text{Te}$)	602
<i>G M D Nguimdo and D P Joubert</i>	
Quantum Neural Networks - Our brain as a quantum computer?	608
<i>M Schuld, I Sinayskiy and F Petruccione</i>	
The stochastic Schrödinger equation approach to open quantum systems	614
<i>I Semina and F Petruccione</i>	
Molecular dynamics studies of some carbon nanotubes chiral structures	620
<i>M Shai, T Mosuang and E Rammutla</i>	
Parallel benchmarks for ARM processors in the high energy context	626
<i>J W Smith and A Hamilton</i>	
Methodology for the digital simulation of open quantum systems	631
<i>R B Sweke, I Sinayskiy and F Petruccione</i>	
Entanglement of two distant nitrogen-vacancy-center ensembles under the action of squeezed microwave field	637
<i>N Teper, L N Mbenza, I Sinayskiy and F Petruccione</i>	
Vacuum energies and frequency dependent interactions	643
<i>H Weigel</i>	

Message from the Conference Chair

The University of Johannesburg successfully hosted the annual conference of the South African Institute of Physics from 7 to 11 July 2014.

Winter schools in Astroparticle Physics and Magnetism preceded the actual SAIP conference. The former school was held in response to the major developments related to this field, such as the recent discoveries in particle physics and the ongoing astronomical infrastructure growth expected in southern Africa, the latter being a research niche at the University of Johannesburg. On that same Monday a photovoltaics workshop, the annual National Laser Centre rental pool meeting and a meeting of the AMS facility were also scheduled in parallel to the SAIP conference.

The SAIP 2014 annual conference was officially opened by the Honourable Minister of Science and Technology, Naledi Pandor, on the evening of Monday 7 July. The function was also addressed by our SAIP President, Igle Gledhill, and members of the University of Johannesburg senior management. The evening was concluded by musical entertainment and a buffet dinner.

The conference proper started on Tuesday 8 July, and included mesmerising hour-long plenary lectures from our invited dignitaries from Brazil, Germany, Japan, Sweden, the UK, the USA and South Africa on some of the currently most topical subjects in physics. Between these the conference participants got down to their business of presenting over 400 scientific talks and posters, discussing their results with their peers and in many cases submitting a full conference paper for these proceedings. On the final day of the conference 20 Science teachers joined the sessions, this being an initiative to grow the enthusiasm in physics within the ranks of our educators.

The management of the University of Johannesburg is sincerely thanked for making a large section of its Auckland Park (Kingsway) campus available for the conference at no rental cost. This included the massive and very comfortable Auditorium, which proved to be a most appropriate venue for the opening and plenary sessions. The large and well-lit foyer proved to be a very popular dining, snacking and mixing area (partly due to the exhibitor stalls there, the quality food and the freely available coffee), and future organisers would be well advised to incorporate a similar facility in their planning. The individual sessions for the seven divisions took place in spacious lecture halls in the adjacent building, which included the poster displays in the large common area. The only major negative was that the country was hit by a substantial cold snap in the week of the conference, and we apologise for the inadequate heating arrangements during the event.

A magnificent banquet at Helderfontein Estate in Kyalami concluded the conference. Prizes were here awarded to the top student presenters in each of the SAIP divisions. The highlight of the evening was the presentation of the De Beers Gold Medal to Manfred Hellberg for his outstanding lifetime service to the South African physics community. The evening concluded with the ceremonial handing over of the 'SAIP mace' to the organisers of the 2015 conference, Rhodes and Nelson Mandela Metropolitan Universities. I wish them very well for this, and know that it too will be an exceptional event.

In conclusion, I thank everyone that helped in making this conference and these proceedings a success, especially to the members of the Local Organising Committee, but also to the other helpers, the UJ support staff, the paper peer reviewers, and of course the conference participants and contributors to the proceedings.

A handwritten signature in black ink, appearing to be 'H. Winkler', with a long horizontal flourish extending to the right.

Hartmut Winkler

Chair: SAIP 2014 annual conference Local Organising Committee

Message from the Review Committee

The Editors of the SAIP2014 Proceedings received 159 manuscripts for consideration by the deadline of 6 July 2014. A total of 147 of these manuscripts met the relevant criteria and were submitted to a full, double-blind, peer-review process involving 124 individual reviewers (some reviewed more than one manuscript). Of these reviewers, 118 were from the South African physics and astronomy community and 6 were scientists from other countries. From the 124 reviewers, 122 hold a PhD in physics or a closely related field. Once two independent content reviews had been received for a particular manuscript through this process, a team of three referees (including the editors of the Proceedings) were involved in processing the reviewed manuscript through corrections, acceptance or rejection from the Proceedings. At the conclusion of the review process, 108 manuscripts were accepted into the final Proceedings. This represents an acceptance rate of 68% of received manuscripts.

The style of these proceedings is that of the (British) Institute of Physics Conference Series, similar to the three previous SAIP Proceedings. Authors were requested to ensure that the clearly defined layout and length specifications were adhered to in their submitted pdf documents. At the start of the reviewing process, an initial layout review was conducted on each manuscript. Manuscripts that deviated considerably from the specified layout specifications, while still broadly appropriate in their composition, were referred back to the authors for layout and length corrections before going out to Content Review.

The response from the relatively small community of South African physicists who possess PhD's in taking on the hard work of professionally reviewing the manuscripts has been very gratifying. To get more than 115 out of a few hundred suitably qualified people to participate in the process reflects admirably on their generosity and commitment to the common good. Most of the reviews received were done with great care and diligence and to the highest standards. The editors wish to voice their deeply-felt thanks to the participating reviewers for their pro bono work. We also wish to thank Prof I Basson (UNISA) for her valuable and helpful inputs during the review process. The meticulous reviewing process described above has ensured that these Proceedings contain thoroughly peer-reviewed manuscripts of a high professional standard, which report on novel work that has not been published elsewhere.

Finally, the editors wish to thank all of the authors for submitting their work to this rigorous process. It is our hope that the final product offered here constitutes a due outcome of their hard work.

CA Engelbrecht (on behalf of the SAIP2014 Review Committee)

Conference Chairs and Committees

SAIP2014 Conference Chair

Hartmut Winkler, Department of Physics, University of Johannesburg

SAIP2014 Scientific Chair

Steven Karataglidis, Department of Physics, University of Johannesburg

SAIP Division Chairs

Division A: Japie Engelbrecht (Nelson Mandela Metropolitan University)

Division B: Simon Mullins (iThemba LABS)

Division C: Erich Rohwer (University of Stellenbosch)

Division D1: Ilani Loubser (North West University)

Division D2: Sunil Maharaj (SANSa)

Division E: Sam Ramaila (University of Johannesburg)

Division F: Freddie Vorster (Nelson Mandela Metropolitan University)

Division G: Frederik Scholtz (National Institute for Theoretical Physics)

Proceedings Editorial Team

Editors: Chris Engelbrecht and Steven Karataglidis (University of Johannesburg)

Advisor: Ilsa Basson (UNISA)

Advisor and online administration: Roelf Botha (HartRAO and SAIP)

Online administration and PDF Compilation: Juan Grey (SAIP)

Review Panel

Roelf Botha (HartRAO and SAIP)

Emanuela Carleschi (University of Johannesburg)

Simon Connell (University of Johannesburg)

Bryan Doyle (University of Johannesburg)

Chris Engelbrecht (University of Johannesburg)

Steven Karataglidis (University of Johannesburg)

Paulus Masiteng (University of Johannesburg)

Azwinndini Muronga (University of Johannesburg)

Pap Nair (University of Johannesburg)

Sam Ramaila (University of Johannesburg)

Krish Reddy (University of Johannesburg)

Buyi Sondezi (University of Johannesburg)

Andre Strydom (University of Johannesburg)

Hartmut Winkler (University of Johannesburg)

List of Reviewers

Dr. ALBERS, Claudia	University of the Witwatersrand
Prof. ALLIE, Saalih	University of Cape Town
Dr. ASANTE, Joseph	Tshwane University of Technology
Mr. BALLESTRERO, Sergio	University of Johannesburg
Dr. BIETENHOLZ, Michael	York University (Canada)
Dr. BOTHA, André	University of South Africa
Dr. BOTTCHER, Markus	North-West University
Prof. BRAUN, Moritz	University of South Africa
Prof. BUFFLER, Andy	University of Cape Town
Prof. BURGER, Renier	North-West University
Prof. CARIGNAN, Claude	University of Cape Town
Dr. CARLESCHI, Emanuela	University of Johannesburg
Prof. CHEN, Andrew	University of the Witwatersrand
Prof. CHETTY, Nithaya	University of Pretoria
Prof. CLEYMANS, Jean	University of Cape Town
Prof. CONNELL, Simon	University of Johannesburg
Prof. CRESS, Catherine	University of the Western Cape
Dr. DEKADJEVI, David	University of Johannesburg
Dr. DOYLE, Bryan	University of Johannesburg
Dr. DU PLESSIS, Anton	Stellenbosch University
Dr. DUDLEY, Angela	CSIR National Laser Centre
Dr. ENGELBRECHT, Christian	University of Johannesburg
Prof. ENGELBRECHT, Japie	Nelson Mandela Metropolitan University
Dr. FALTENBACHER, Andreas	University of the Witwatersrand
Dr. FERRER, Phil	University of the Witwatersrand
Prof. FORBES, Andrew	Council for Scientific and Industrial Research
Dr. FÖRTSCH, Siegfried	iThemba Labs
Dr. FRANKLYN, Chris	South African Nuclear Energy Corporation (Necsa)
Dr. FRESCURA, Fabio	University of the Witwatersrand
Dr. GIBBON, Timothy	Nelson Mandela Metropolitan University
Dr. GILBANK, David	South African Astronomical Observatory (SAAO)
Dr. GLEDHILL, Irvy (Igle)	Council for Scientific and Industrial Research
Dr. GOVENDER, Indresan	University of Cape Town
Prof. GRAYSON, Diane	Council on Higher Education
Dr. HAMILTON, Andrew	University of Cape Town
Prof. HEARNE, Giovanni	University of Johannesburg
Prof. HOLLENBERG, Lloyd	University of Melbourne (Australia)
Dr. HOROWITZ, William	University of Cape Town
Dr. JAFARI, Hossein	University of South Africa
Dr. JONES, Pete	iThemba Labs
Prof. JOUBERT, Daniel	University of Witwatersrand
Dr. KALOSAKAS, George	University of Patras (Greece)
Prof. KARATAGLIDIS, Steven	University of Johannesburg
Prof. KEARTLAND, Jonathan	University of Witwatersrand
Dr. KEMP, Garreth	University of Johannesburg
Prof. KILKENNY, David	University of the Western Cape
Prof. KOEN, Chris	University of the Western Cape
Prof. KONRAD, Thomas	University of KwaZulu-Natal
Prof. KRAAN-KORTEWEG, Renee C.	University of Cape Town
Prof. KRIEK, Jeanne	University of South Africa
Dr. KROON, Ted	University of the Free State
Dr. KUMAR, Vijay	University of the Free State

Dr. LEMMER, Miriam	North-West University
Dr. LODYA, Lonzeche	Sasol Technology
Dr. LOUBSER, Ilani	North-West University
Dr. MADHUKU, Morgan	iThemba Labs
Dr. MAGNUS, Lindsay	South African Square Kilometre Array (SKA)
Dr. MANYALA, Ncholu	University of Pretoria
Dr. MASITENG, Paulus	University of Johannesburg
Dr. MATHE, Bhekumusa	University of the Witwatersrand
Dr. MATTHEWS, Alan	University of KwaZulu-Natal
Dr. MCBRIDE, Vanessa	University of Cape Town
Prof. MEDVED, Allan Joseph	Rhodes University
Prof. MEINTJES, Pieter	University of the Free State
Prof. MELLADO, Bruce	University of the Witwatersrand
Dr. MOEKETSI, Daniel Mojalefa	CSIR Center for High Performance Computing
Dr. MOMA, John	Mintek
Dr. MOYO, Thomas	University of KwaZulu-Natal
Prof. MURONGA, Azwinndini	University of Johannesburg
Dr. NAICKER, Vishnu	North-West University
Prof. NAIDOO, Deena	University of the Witwatersrand
Dr. NAIR, Padmanabhan	University of Johannesburg
Prof. NDWANDWE, Muzi	University of Zululand
Dr. NEETHLING, Pieter	Stellenbosch University
Dr. NEL, Jacqueline	University of Pretoria
Prof. NEMRAOUI, Ouassini	University of Zululand and iThemba Labs
Prof. NEWMAN, Richard	Stellenbosch University
Dr. NOLTING, Volkmar	Vaal University of Technology
Dr. NTSHANGASE, Sifiso Senzo	University of Zululand
Dr. O'CONNELL, Jacques	Nelson Mandela Metropolitan University
Dr. ORCE, Nico	University of the Western Cape
Dr. PAPKA, Paul	Stellenbosch University
Prof. PERESSI, Maria	University of Trieste (Italy)
Prof. PRINSLOO, Aletta	University of Johannesburg
Dr. PRINSLOO, Rian	South African Nuclear Energy Corporation (Necsa)
Prof. QI, Guoyuan	University of South Africa
Prof. RAKITIANSKI, Sergei	University of Pretoria
Dr. RAMAILA, Sam	University of Johannesburg
Prof. RAZZAQUE, Soebur	University of Johannesburg
Dr. REDDY, Leelakrishna	University of Johannesburg
Mr. REITSMA, Frederik	International Atomic Energy Agency (IAEA) (Austria)
Prof. SCHOLTZ, Frederik	National Institute for Theoretical Physics (NITheP)
Dr. SEFAKO, Ramotholo	South African Astronomical Observatory (SAAO)
Prof. SHARPEY-SCHAFER, John F.	University of the Western Cape
Dr. SHEPPARD, Charles	University of Johannesburg
Dr. SINAYSKIY, Ilya	University of KwaZulu-Natal and NITheP
Prof. SMITS, Derck	University of South Africa
Dr. SONDEZI, Buyi	University of Johannesburg
Dr. STEYN, Deoin	iThemba Labs
Prof. STRYDOM, Andre	University of Johannesburg
Prof. SWART, Hendrik	University of the Free State
Prof. TAME, Mark	University of KwaZulu-Natal
Dr. TAYLOR, Dale	University of Cape Town
Prof. TAYLOR, Geoffrey	University of Melbourne (Australia)
Prof. TCHOULA TCHOKONTE, Moise Bertin	University of the Western Cape
Prof. THERON, Chris	University of Pretoria
Prof. TRIAMBAK, Smarajit	University of the Western Cape

Prof. TWALA, Bhakisipho	University of Johannesburg
Dr. UYS, Hermann	CSIR National Laser Centre
Prof. VAN DER LINGEN, Elma	University of Pretoria
Prof. VAN DER WALT, Johan	North-West University
Prof. VAN DYK, Ernest	Nelson Mandela Metropolitan University
Mr. VAN SOELEN, Brian	University of the Free State
Prof. VENTER, Andre	Nelson Mandela Metropolitan University
Dr. VENTER, Christo	North-West University
Dr. VORSTER, Frederik	Nelson Mandela Metropolitan University
Dr. WAMWANGI, Daniel	University of the Witwatersrand
Prof. WEIGEL, Herbert	Stellenbosch University
Prof. WEIGERT, Heribert	University of Cape Town
Prof. WINKLER, Hartmut	University of Johannesburg
Prof. WOUTD, Patrick	University of Cape Town
Prof. WYNGAARDT, Shaun	Stellenbosch University
Dr. YACOOB, Sahal	University of KwaZulu-Natal
Dr. ZAMONSKY RIMOLDI, Oscar Mario	South African Nuclear Energy Corporation (Necsa)

Division A – Division for Physics of Condensed Matter and Materials

Synthesis and characterization of the semiconducting intermetallic compound FeGa_3

M A M Ahmed, E Carleschi, G R Hearne and B P Doyle

Department of Physics, University of Johannesburg PO Box 524, Auckland Park 2006,
Johannesburg, South Africa

E-mail: mustafa@aims.ac.za

Abstract. FeGa_3 polycrystalline samples were synthesised using the gallium self flux method. The stoichiometry was confirmed via energy dispersive spectroscopy and the tetragonal crystal structure was refined from X-ray diffraction data. Mössbauer effect spectroscopy was performed down to 120 K at ambient pressure, as well as at 3 K and 4 GPa in a diamond anvil pressure cell, and shows an absence of magnetic ordering at these temperatures and pressures. The isomer shift and quadrupole splitting are consistent with a 3+ charge state for Fe with a 5/2 spin state.

1. Introduction

Materials with complex band structure have very interesting electronic, magnetic and transport properties [1]. Recent phenomena such as the Kondo insulating state, heavy fermion behaviour and unconventional superconductivity have been reported for those materials [2]. Such phenomena are usually seen in the rare-earth compounds where the hybridization between different electrons in the atoms forms narrow electronic bands or pseudo-gaps at the Fermi level [3].

Ruthenium and iron compounds (such as RuAl_2 , RuGa_3 , FeSi , FeSb_2 , Fe_2Al_2 and FeGa_3) are semiconductor compounds in which the energy gap, of the order of 0.1 eV, is formed by the hybridization between the d states of the transition metal with the p states of the group 13 or 14 elements [2]. These compounds are of interest due to their unusual transport and magnetic behaviour [4]. Amongst the above compounds, FeSi and FeSb_2 show strong correlated electron properties similar to rare-earth based Kondo semiconductors. The distinguishing feature between a Kondo semiconductor and a band-gap semiconductor is that, in the Kondo semiconductor, the gap disappears upon heating already at temperatures and energies lower than the energy gap. This feature has been observed in both FeSi and FeSb_2 in the temperature dependence of the optical conductivity [5].

FeGa_3 is one of such intermetallic compounds which are composed of metallic elements, and yet their combination is a semiconductor. The band gap in FeGa_3 arises in particular from the hybridization between the Fe 3*d* and Ga 4*p* states [6]. FeGa_3 crystallizes in the tetragonal lattice with space group $P4_2/mnm$ (No. 136) [7]. In the tetragonal structure of FeGa_3 , Fe atoms occurs as dimers along the (110) and ($\bar{1}\bar{1}0$) directions, as it shown in figure 2 (b). The band gap of this compound has been measured experimentally using photoemission spectroscopy and predicted using density functional theory calculations within the local density approximation, where it was found to be in the range 0.3 - 0.5 eV [6, 8]. This value is in good agreement with the observed

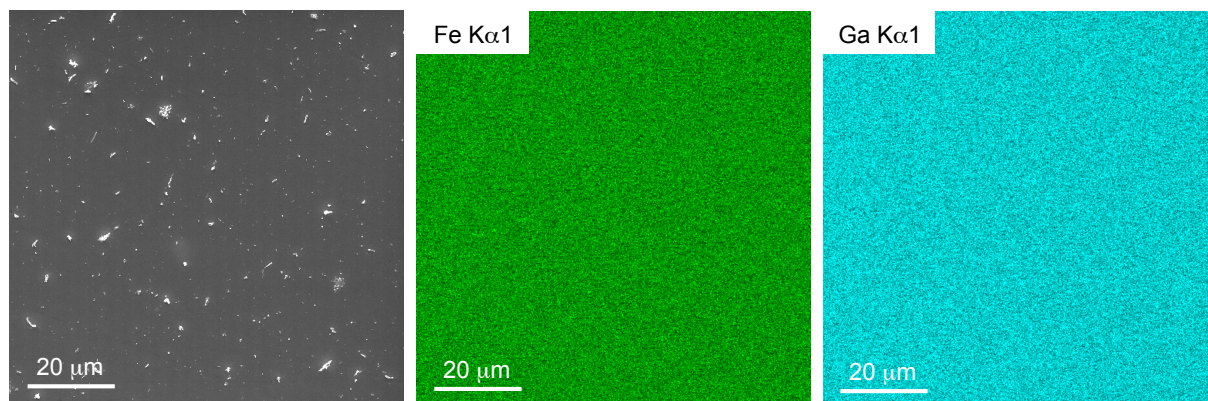


Figure 1. Representative SEM image and EDS compositional maps of a polished FeGa_3 sample. A uniform distribution of the elements Fe and Ga is observed.

gap of 0.25 - 0.47 eV previously reported for FeGa_3 [5, 7, 9, 10]. Lue *et al.* observed metallic conductivity above 10 K in their electrical resistivity measurements [8]. The resistivity exhibits a minimum at 160 K, below which it has negative temperature coefficient (semiconducting) and above 160 K there is a typical metallic-like trend (positive temperature coefficient) [6].

The magnetism of FeGa_3 is subject to debate and has not yet been directly observed. Some measurements such as conductivity, magnetic susceptibility, Mössbauer spectroscopy and specific heat have not shown strong magnetic correlations. This justifies the absence of magnetism at room temperature [8, 11]. Recently, weakly coupled local moments have been suggested in Co-doped FeGa_3 by Bittar *et al.* [12]. In particular they postulate that the doping creates spin-1/2 local moments and drives the compound from semiconducting to the metallic state [12]. Work done by Umeo *et al.* [4] pointed out that electron-type doping in this compound using Ge introduces a crossover to a correlated metallic state at $y \approx 0.006$ in $\text{FeGa}_{3-y}\text{Ge}_y$.

Here, we report on the synthesis and the characterisation of FeGa_3 polycrystalline samples. Mössbauer effect spectroscopy measurements have been performed down to 120 K at ambient pressure, as well as at 3 K and at a pressure of 4 GPa. All the spectra show the characteristic line shape of a quadrupole split doublet consistent with an Fe^{3+} charged state and a 5/2 spin state. No evidence for a magnetic state has been detected, contrary to theoretical and experimental work which propose spin polarons [11] and Fe dimers [13], respectively.

2. Experimental Details

FeGa_3 polycrystalline samples were prepared via the self flux method described in Ref. [8]. High purity Ga (from Alfa Aesar, 99.99999%) and Fe (from Sigma Aldrich, 99.9%) were mixed according to the molar ratio $\text{Fe}_{15}\text{Ga}_{85}$. The mixture was pressed into a pellet and loaded into an alumina crucible which was vacuum sealed ($p \sim 10^{-5}$ Pa) inside a quartz ampoule under argon gas. The sample was heated to 1000 °C at a rate of 500 °C/hr, held at this temperature for 40 hrs, cooled down to 850 °C for 12 hrs, cooled down again to 700 °C for 75 hrs and finally allowed to cool to room temperature. Excess gallium was removed by heating the pellet and washing with dilute HCl and deionised water. A photograph of the FeGa_3 sample (showing large facets) is shown in figure 2 (c).

The crystalline structure was checked by means of x-ray diffraction (XRD) measured with monochromatic $\text{Cu K}\alpha$ radiation on a powdered specimen using a Philips Analytical X-Pert PRO system. The XRD pattern was refined using the Jana 2006 program [14]. The stoichiometry and the compositional homogeneity were checked with energy dispersive spectroscopy (EDS) using a Tescan Vega3 electron microscope equipped with an Oxford EDS spectrometer. From the EDS multi-spot analysis it was consistently found to be at the nominal value, within the detection

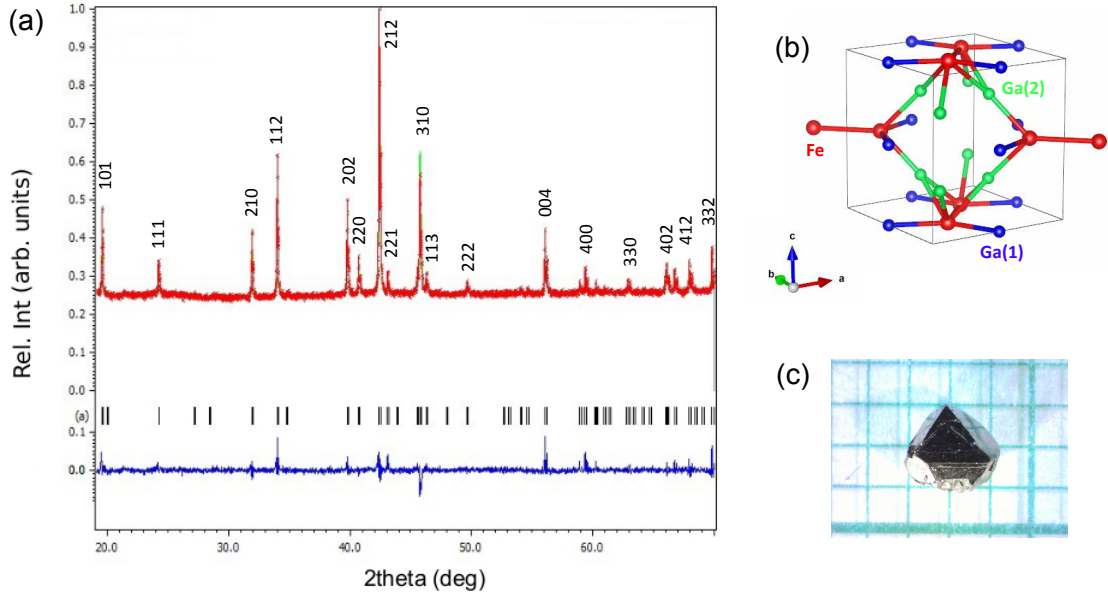


Figure 2. (a) Experimental indexed XRD pattern for FeGa₃ (red) along with the Rietveld refinement (green). The peak positions are represented by vertical bars (black), while the residual is in blue. (b) Schematics of the tetragonal unit cell of FeGa₃, where Fe atoms are represented by bigger red circles, and Ga(1) and Ga(2) atoms are represented by smaller blue and green circles respectively. The Fe-Fe dimers are shown connected via red bars. (c) One block in the photograph represents a scale of 1 mm.

Table 1. Structural parameters obtained from the Rietveld refinement of the crystal structure based on the x-ray powder diffraction pattern measured at room temperature.

Lattice parameters	values	atom	Wyckoff site	x	y	z
a(Å)	6.2636	Fe	4f	0.3426(11)	0.3444(11)	0
b(Å)	6.2636	Ga(1)	4c	0	0.5	0
c(Å)	6.5550	Ga(2)	8j	0.1565(5)	0.1565(5)	0.2633(7)

limits of the technique (FeGa_{2.97±0.03}). EDS elemental maps confirmed the homogeneity of the samples, as reported in figure 1.

The sample was crushed into powder to a thickness of ~ 11 mg/cm² to perform Mössbauer spectroscopy measurements. A constant acceleration spectrometer with a conventional ⁵⁷Co(Rh) source was used for Mössbauer measurements at ambient pressure. Measurements to low temperature were performed in a top-loading cryostat, with both source and sample kept at the same temperature. The Mössbauer spectrum at 3 K and $p = 4$ GPa was measured at the beamline ID18 at the European synchrotron radiation facility [15] using the synchrotron Mössbauer source (SMS). The sample was loaded into a Merrill-Basset diamond anvil cell (DAC). Sample was pressurized in 200 μ m cavity of a Re gasket pre-indented to a thickness of 20 μ m. Pressure was determined from the ruby fluorescence of tiny balls inserted in the cavity [16]. All spectra were calibrated with respect to α -Fe.

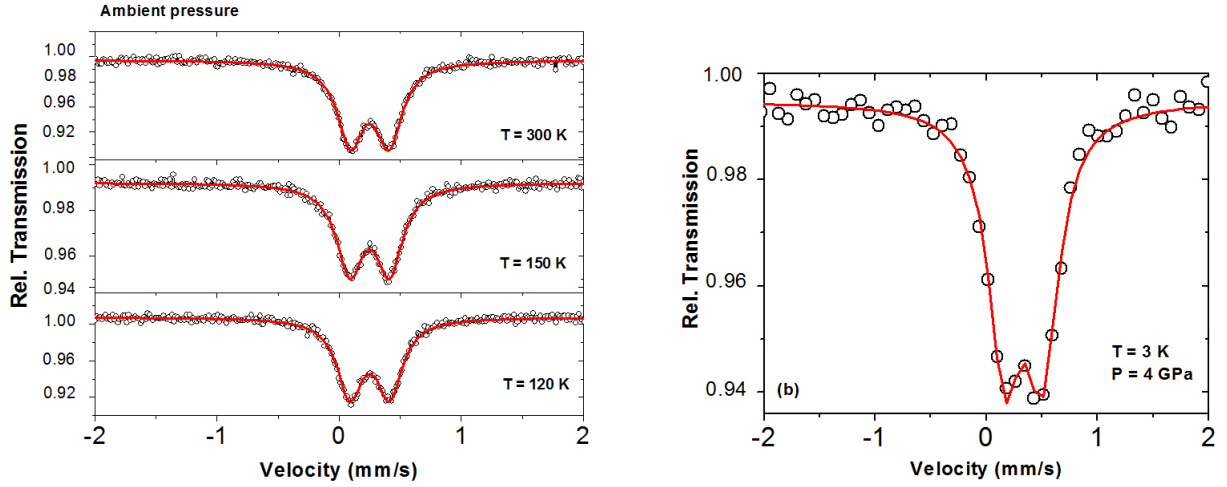


Figure 3. (a) Mössbauer spectra of the powdered sample recorded at 300 K, 150 K and 120 K at ambient pressure. (b) Mössbauer spectrum measured at ID18 at 3 K and 4 GPa. The solid red lines indicate the fit to the experimental data (open black circles).

3. Results and discussion

Figure 2 (a) shows the XRD pattern for FeGa₃. The Rietveld refinement of the diffraction pattern confirmed the tetragonal space group $P4_2/mnm$ and $Z = 4$. The lattice parameters and atomic positions obtained from the refinement are presented in table 1. These are in good agreement with values reported by Häussermann [6] and Lue [2].

The Mössbauer spectra of FeGa₃ recorded at 300 K, 150 K and 120 K are shown in figure 3. They have been fitted with a single quadrupole doublet, and the parameter values derived from fits to these spectra are presented in table 2. Our results at 300 K are similar to what has been previously reported by Tsujii *et al.* for FeGa₃ single crystals at room temperature [8]. The data presented in table 2 show that: (i) no deviation occurs in the isomer shift (δ) and the quadrupole splitting (ΔE_Q) at these temperatures; (ii) the value of the isomer shift (δ) and the quadrupole splitting ΔE_Q indicate that iron (III) is present in the 5/2 spin state.

A particular feature of the Mössbauer spectra is the quadrupole split doublet. The similar profile at all measured temperature indicates good phase purity and there is no indication of Fe-based amorphous/ disordered phases which would be difficult to detect by XRD. The sample under pressure shows broadened peaks, which are due to the SMS source. There is no indication of magnetic ordering down to temperatures of 120 K at ambient pressure and down to 3 K at 4 GPa (these would manifest as resonance dips in the wings of the spectrum).

Table 2. Parameters derived from fitting the Mössbauer spectra of FeGa₃.

	ΔE_Q (mm/s)	δ/Fe	Γ (mm/s)
T = 300 K, amb. p	0.32	0.28	0.26
T = 150 K, amb. p	0.33	0.38	0.26
T = 120 K, amb. p	0.33	0.38	0.25
T = 3 K, p = 4 GPa	0.33	0.33	0.35

4. Conclusion

In conclusion, we have synthesised and characterized FeGa₃. Single phase, homogeneous, polycrystalline FeGa₃ was obtained. This forms the basis for future studies of the electrical-transport behaviour under pressure, where a semiconductor-metal transition (correlation gap closure) is predicted to occur [1]. Importantly, Mössbauer-effect spectroscopy shows the absence of the magnetic ordering down to 120 K at ambient pressure and down to 3 K at 4 GPa. This sets the energy scale for magnetic coupling, which is either absence or is likely below ~ 3 K at pressure up to ~ 4 GPa. It is of interest to consider this in conjunction with recent electronic structure calculations which predict an Fe magnetic moment of $\sim 0.6\mu_B$ and antiferromagnetic coupling between the atoms in the Fe₂ pairs of the structure [13].

References

- [1] Osorio-Guillen J, Larrauri-Pizarro Y and Dalpian G 2012 *Phys. Rev. B* **86** 235202
- [2] Lue C S, Lai W and Kuo Y K 2005 *J. Alloys Compd.* **392** 72
- [3] Haldolaarachchige N, Karki A, Phelan W A, Xiong Y, Jin R, Chan J Y, Stadler S and Young D 2011 *J. Appl. Phys.* **109** 103712
- [4] Umeo K, Hadano Y, Narazu S, Onimaru T, Avila M A and Takabatake T 2012 *Phys. Rev. B* **86** 144421
- [5] Hadano Y, Narazu S, A Avila M, Onimaru T and Takabatake T 2009 *J. Phys. Soc. Jpn* **78** 013702
- [6] Häussermann U, Boström M, Viklund P, Rapp Ö and Björnängen T 2002 *J. Solid State Chem.* **165** 94
- [7] Imai Y and Watanabe A 2006 *Intermetallics* **14** 722
- [8] Tsujii N, Yamaoka H, Matsunami M, Eguchi R, Ishida Y, Senba Y, Ohashi H, Shin S, Furubayashi T, Abe H *et al.* 2008 *J. Phys. Soc. Jpn* **77** 024705
- [9] Amagai Y, Yamamoto A, Iida T and Takanashi Y 2004 *J. Appl. Phys.* **96** 5644
- [10] Arita M, Shimada K, Utsumi Y, Morimoto O, Sato H, Namatame H, Taniguchi M, Hadano Y and Takabatake T 2011 *Phys. Rev. B* **83** 245116
- [11] Storchak V G, Brewer J H, Lichti R L, Hu R and Petrovic C 2012 *J. Phys. Condens. Matter* **24** 185601
- [12] Bittar E, Capan C, Seyfarth G, Pagliuso P and Fisk Z 2010 *J. Phys. Conf. Ser.* **200** 012014
- [13] Yin Z P and Pickett W E 2010 *Phys. Rev. B* **82** 155202
- [14] Petricek V, Dusek M and Palatinus L 2014 *Z. Kristallogr.* **223** 345
- [15] Potapkin V, Chumakov A I, Smirnov G V, Celse J P, Rüffer R, McCammon C and Dubrovinsky L 2012 *J. Synch. Rad.* **19** 559
- [16] Forman R A, Piermarini G J, Barnett J D and Block S 1972 *Science* **176** 284–285

Superconductivity in LaRh_2Sn_2

D Britz and A M Strydom

Highly correlated matter research group, Physics Department, University of Johannesburg,
PO Box 524, Auckland Park 2006, South Africa

E-mail: amstrydom@uj.ac.za

Abstract. LaRh_2Sn_2 crystallizes in the centrosymmetric primitive tetragonal crystal structure $P4/nmm$ (number 129) commonly referred to as the CaBe_2Ge_2 -type which was confirmed on our synthesized samples by x-ray diffraction. We report the existence of this superconducting transition in exploratory work and in this study we proceed with a study into the physical properties of this superconducting ground state. Electrical resistance and heat capacity measurements reveal a sharp and well defined superconducting transition at $T_{\text{sc}} = 0.699(7)$ K. The low-temperature heat capacity measurements show LaRh_2Sn_2 to be a weakly coupled ($\lambda_{\text{el-ph}} \approx 0.5$) bulk BCS superconductor that has an s-wave singlet ground state with an isotropic energy gap such that $2\Delta/k_{\text{B}}T_{\text{sc}} = 3.38(6)$. From the field dependence of the electrical resistance the upper critical field was estimated to be 0.127(3) T. The calculated penetration depth and coherence length showed that LaRh_2Sn_2 is a type-II superconductor, which was confirmed with low temperature magnetization measurements.

1. Introduction

The family of compounds RT_2X_2 (R = rare-earth element, T = d-block element and X = p-block element) are split into two separate groups according to their crystal structure. One is the body-centred-tetragonal ThCr_2Si_2 -type (space group $I4/mmm$) and the other is the primitive tetragonal CaBe_2Ge_2 -type (space group $P4/nmm$). Most of the experimental and theoretical work has been done on the compounds that fall into the ThCr_2Si_2 -type structures, for this is where the heavy-fermion superconductors are located [1] and references therein. La-based superconductors are not uncommon in the ThCr_2Si_2 -type structure formed with Si and Ge, for instance: LaPd_2Ge_2 , LaPt_2Ge_2 [2], LaNi_2Ge_2 [3], LaRh_2Si_2 [4] but of the LaT_2Sn_2 compounds which crystallize in the primitive tetragonal structure, where T = Rh is the only one that has been found to be superconducting [5] to date. This work is therefore devoted to studying the physical properties of this particular superconductor.

2. Experimental

Polycrystalline samples were prepared with stoichiometric quantities of the starting materials: La (99.99 %), Rh (99.99 %) and Sn (99.9999 %) with the purities quoted in wt. %, on a water cooled Cu hearth in an arc-furnace manufactured by *Edmund Bühler*. In-situ ultra-high purity Ar gas (> 99.999 %) was used during melting procedure. The samples were melted and flipped 4 times in order to promote sample homogeneity. The samples were then wrapped in Ta foil (99.95 %) and sealed in evacuated quartz tubes. Heat treatment was conducted at 800 °C for 5 days and subsequently quenched in water. Powder X-ray diffraction measurements were

performed at room temperature using a Cu radiation source. The resistivity $\rho(T)$ and specific heat $C_P(T)$ measurements were performed on a commercial Physical Property Measurement System manufactured by *Quantum Design* (San Diego, USA) equipped with a ^3He and dilution refrigerator inserts for measurements down to 50 mK. The resistance was measured using the Electronic Transport Option (ETO) preamplifier setup with an ac-current (9.1 Hz with a current amplitude of 0.3 to 0.1 mA depending on the temperature with lower current values used at lower temperatures in order to reduce Joule heating at these temperatures) while the heat capacity was measured using the adiabatic time relaxation method. The magnetization was measured using a commercial Magnetic Property Measurement System equipped with a ^3He cryostat developed by *iQuantum* (Tsukuba, Japan) manufactured by *Quantum Design* (San Diego, USA).

3. Results

Powder x-ray diffraction of the polycrystalline LaRh_2Sn_2 was analyzed with a Rietveld refinement and crystallize in the CaBe_2Ge_2 -type structure (space group $P4/\text{nm}$) [6] with $a = 4.5139(1) \text{ \AA}$ and $c = 10.4824(5) \text{ \AA}$, shown in figure 1. Within this structure the La atoms occupy the single Ca site (2c), the Rh atoms occupy two distinct Be sites (2c and 2a) and the Sn atoms occupy two distinct Ge-sites (2c and 2b).

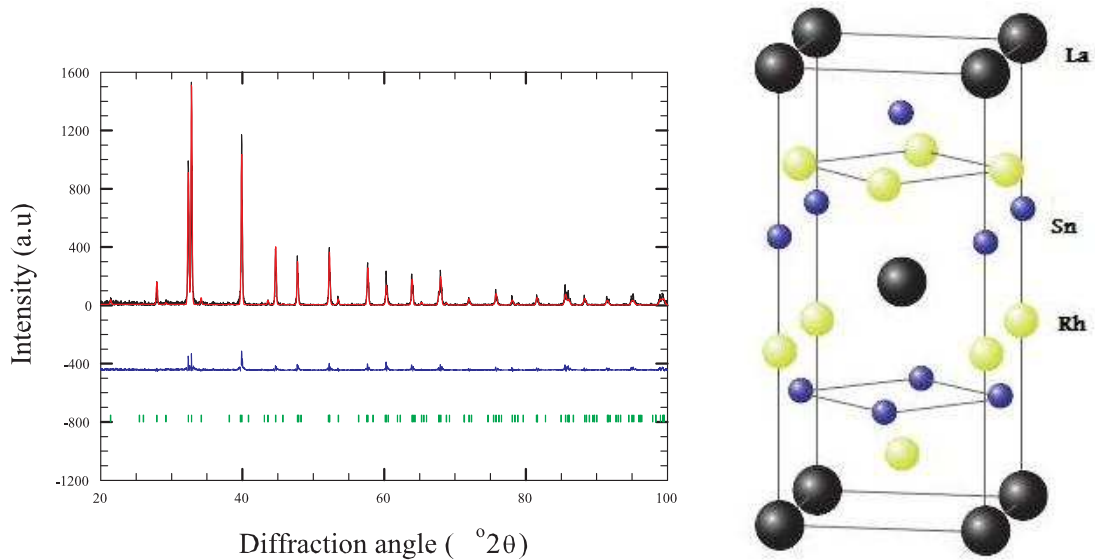


Figure 1. Rietveld refinement was performed using the General Structure Analysis System (G.S.A.S) [7] software on x-ray powder diffraction data of LaRh_2Sn_2 : the black line representing the experimental spectrum, the red line represents the least-squares fit of the refinement, the blue line is the difference between the experimental data and the fit while the green bars represent the expected peak positions for the $P4/\text{nm}$ structure. Ball and stick representation of the $P4/\text{nm}$ structure is shown on the right of the diffractogram.

In the results that follow we follow the convention that the applied field strength $\mu_0\mathbf{H} =$ magnetic induction \mathbf{B} . Figure 2 shows the a.c. electrical resistivity that were carried out on a polycrystalline sample of LaRh_2Sn_2 in the stated applied magnetic fields, with $\mathbf{B} \perp \mathbf{j}$ where \mathbf{j} is the current density. The superconducting transition is clearly visible as the sharp decrease in $\rho(T)$ around 0.7 K. The transition temperature, T_{sc} , was taken at the point where the resistance reached half its residual value in the metallic state. Under the application of an external magnetic field, the onset of superconductivity occurred at lower temperatures as expected due to

the pair breaking nature of the magnetic field on the Cooper pairs within the superconducting ground state. The broadening in the width of the transition as the magnetic field is increased is probably due to the appearance of the vortex state within a type-II superconductor.

Inset (a) of figure 2 shows the field dependence of T_{sc} and this was fitted by [8]:

$$H_c(T) = H_c(0) \left(1 - \left(\frac{T}{T_{sc}} \right)^2 \right); \quad (1)$$

where $H_c(0)$ is the critical field at zero temperature. The values obtained from the fit were $H_c(0) = 0.127(3)$ T and $T_{sc} = 0.699(7)$ K.

Inset (b) of figure 2 shows the $\mathbf{B} = 0.11$ T resistivity data where the superconducting transition is no longer visible. The temperature dependence of $\rho(T)$ becomes proportional to T^2 , which is shown by the dashed red line which serves as a guide to the eye. This temperature dependence in the electrical resistivity is indicative of electron-electron scatterings within a Fermi-liquid state. The application of magnetic field suppresses the superconducting ground state and a Fermi-liquid state emerges.

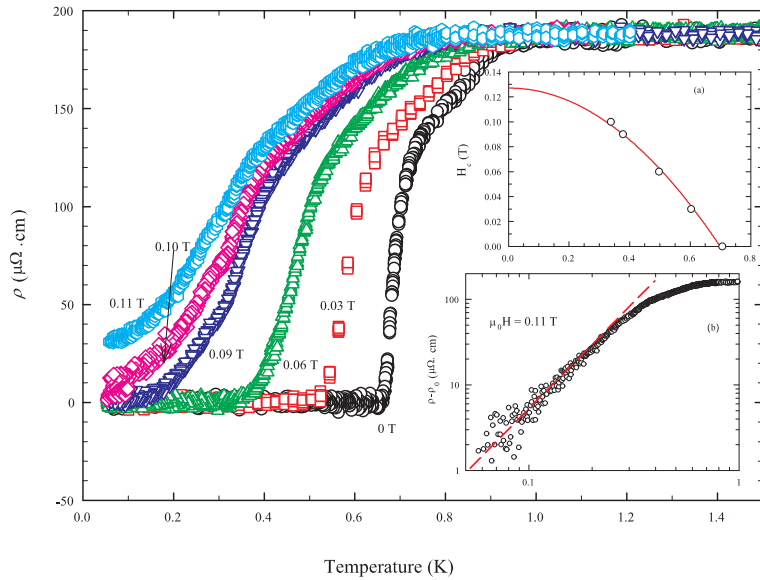


Figure 2. Resistivity of LaRh_2Sn_2 as a function of temperature in various applied magnetic fields. Inset (a): The critical field plotted as a function of temperature with a fit discussed in the text. Inset (b): Log-log plot of the $B = 0.11$ T resistivity data with the residual resistivity subtracted, revealing the T^2 dependence shown by the dashed red line which is discussed in the text.

The specific heat measured in applied magnetic fields in the range of 0 to 0.7 T is shown in figure 3. The Debye fit, $C_p(T)/T = \gamma_N + \beta T^2$, is shown in inset (a) of figure 3 and returned the following values: $\gamma_N = 11.796(9)$ mJ.mol $^{-1}$ K $^{-2}$ and $\beta = 0.816(2)$ mJ.mol $^{-1}$ K $^{-4}$ which corresponds to a Debye temperature $\theta_D = 228.4(5)$ K. Using the theory outlined by McMillan [9] the electron-phonon coupling is given by:

$$\lambda_{\text{el-ph}} = \frac{1.04 + \mu^* \ln(\theta_D/1.45T_{sc})}{(1 - 0.62\mu^*) \ln(\theta_D/1.45T_{sc}) - 1.04} \quad (2)$$

where μ^* represents the screened Coulomb part, which is generally taken to lie within the range of 0.1 to 0.15, and if we use the value for $\mu^* = 0.13$ [10] then $\lambda_{\text{el-ph}} \approx 0.5$ which implies that LaRh_2Sn_2 is a weakly coupled superconductor.

The linear field dependence of the non-superconducting electronic density of states γ_S , which was taken as the value of $C_p(T)_B/T$ at 0.36 K is shown in inset (b) of figure 3. The linear field dependence is expected for s-wave pairing revealing that LaRh_2Sn_2 is a fully gapped isotropic superconductor [10].

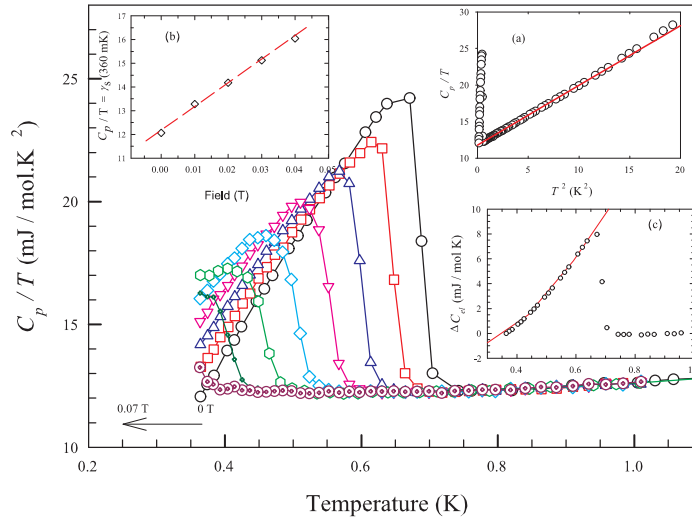


Figure 3. Specific heat of LaRh_2Sn_2 as a function of temperature in various applied magnetic fields (0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07 T) with the solid lines being guides to the eye. Inset (a): Debye fit to the paramagnetic region above the superconducting transition represented by the solid red line. Inset (b): The field dependence of the non-superconducting density of states within the superconducting state with a dashed red line serving as a guide to the eye for the linear-in- \mathbf{B} dependence. Inset (c): Temperature dependence of the electronic part of the specific heat, $\Delta C_{\text{el}}(T)$, for LaRh_2Sn_2 with the solid red line being a fit that assumed an isotropic s-wave BCS superconducting gap.

Inset (c) of figure 3 shows the electronic part of the specific heat which is obtained from subtracting the Debye spectrum, which was calculated from the specific heat data measured in zero applied field from the measured zero applied field specific heat. The solid red line is a fit of the expected BCS temperature dependence of the electronic specific heat in the superconducting region, namely [8]:

$$\Delta C_{\text{el}}(T) \approx \sqrt{T} e^{-\Delta/k_B T} \quad (3)$$

with $2\Delta/k_B T_{\text{SC}} = 3.38(6)$ which is in reasonable agreement with the expected weak coupling BCS value of 3.52. The critical field at 0 K is related to the Ginzburg-Landau coherence length (ξ_{GL}) by [11]:

$$\xi_{GL}(0) = \left(\frac{\Phi_0}{2\pi H_c(0)} \right)^{1/2}; \quad (4)$$

where $\Phi_0 = h/2e = 2.07 \times 10^{-15} \text{ T.m}^2$ is the flux quantum and by substituting in these respective values one finds $\xi_{GL}(0) = 50.93 \text{ nm}$. Together with the parameters obtained from the Debye fit to the specific heat data, inset (a) figure 3, one can find the thermodynamic critical field ($H_c^*(0)$) by using the following relation [11]:

$$H_c^*(0) = 4.23T_{SC}\sqrt{\gamma_N}; \quad (5)$$

with the paramagnetic normal state Sommerfeld coefficient $\gamma_N = 11.796(9) \text{ mJ.mol}^{-1}\text{K}^{-2}$ in the case of LaRh_2Sn_2 . Substituting in these values yields $H_c^*(0) = 10.155 \times 10^{-3} \text{ T}$. The Ginzburg-Landau parameter (κ_{GL}) at zero temperature is used to identify whether a superconductor is unstable toward vortex formation if $\kappa_{GL}(0) > 1/\sqrt{2}$ then the vortex state will be stable. The Ginzburg-Landau parameter is given by the following relation [11]:

$$\kappa_{GL}(0) = \frac{H_c(0)}{\sqrt{2}H_c^*} \quad (6)$$

and using the values which were calculated above the value for $\kappa_{GL}(0) = 8.84$. This value for $\kappa_{GL}(0)$ is above $1/\sqrt{2}$ therefore LaRh_2Sn_2 is classified as a type-II superconductor. These two relations allows one to find the Ginzburg-Landau penetration depth $\lambda_{GL}(0) = \xi_{GL}(0)\kappa_{GL}(0)$ which gives $\lambda_{GL}(0) = 450.22 \text{ nm}$.

The magnetization below the superconducting transition temperature is shown in figure 4. The low field region has a linear field dependence, which is a result of the perfect diamagnetism that exists in the Meissner state. This linear field dependence terminated around 25 Oe. This field where the Meissner state breaks down is known as the lower critical field H_{c1} in type-II superconductors and signals the nucleation and growth of the vortices in the vortex state. The slow increase of the magnetization with increasing applied magnetic field is due to the increased density of vortices within the mixed state of a type-II superconductor, as more flux quanta are allowed to penetrate the superconducting state. The hysteresis observed in opposed field runs (positive and negative field polarities), indicated by the arrows, is caused by vortex pinning at the surface of the superconductor. This behaviour confirms that the vortex state exists and that LaRh_2Sn_2 is a type-II superconductor for there are no vortices present in type-I superconductors.

The value of the gradient of the linear-in-field region (Meissner state) of the $M(T, B)$ data allows us to estimate the volume fraction that is superconducting. The gradient value corresponding to 100 % superconducting volume fraction is $-1/4\pi$ (perfect diamagnetism). The calculated gradient revealed that around 84 % of the sample is superconducting at 0.48 K, within error regarding the approximations made with the geometrical factor due to the irregular shape of the sample. This, along with the specific heat measurements, confirmed bulk superconductivity within LaRh_2Sn_2 .

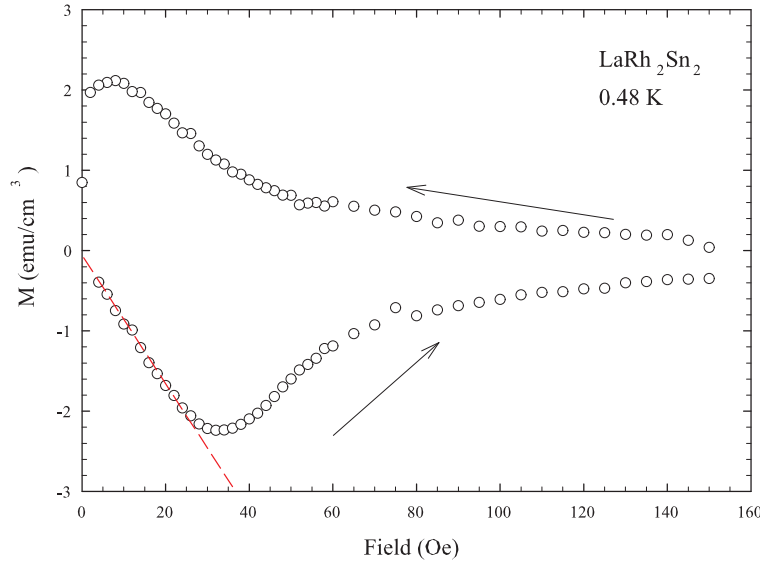


Figure 4. Magnetization as a function of applied magnetic field with the arrows indicating the direction of the field sweep. The dashed red line is a straight line to highlight the linear-in- B behaviour at low B .

4. Conclusion

We have examined the superconducting properties of LaRh_2Sn_2 and found it to be a weakly coupled BCS superconductor that is isotropically gapped at the Fermi surface with a transition temperature $T_{\text{SC}} = 0.699(7)$ K. Above 25 Oe at 0.48 K the vortex state is seen in magnetization measurements which indicated that LaRh_2Sn_2 is a type-II superconductor. This behaviour is in line with the Ginzburg-Landau parameter being larger than $1/\sqrt{2}$ which was calculated from specific heat capacity data. The upper critical field is estimated to be 0.127(3) T from electrical resistance measurements. This BCS behaviour is similar to the other non-4f RT_2X_2 ($\text{R} = \text{La}$, Y and $\text{X} = \text{Si}$, Ge) superconductors which form in the body-centered tetragonal structure.

5. References

- [1] Stockert *et al* O 2012 *J. Phys. Soc. Jpn.* **81** 011001
- [2] Hull G W, Wernick J H, Geballe T H, Waszczak J V and Bernardini J E 1981 *Phys. Rev. B* **24**(11) 6715–6718
URL <http://link.aps.org/doi/10.1103/PhysRevB.24.6715>
- [3] Wernick J, Hull G, Geballe T, Bernardini J and Waszczak J 1982 *Materials letters* **1** 71
- [4] Palstra T T M, Lu G, Mcnovsky A A, Nieuwenhuys G J, Kes P H and Mydosh J A 1986 *Phys. Rev. B* **34** 4566
- [5] Strydom A and Britz D 2012 *J. Phys. Soc. Jpn.* **81** SB018
- [6] Francois M, Venturini G, Mareche J, Malaman B and Roques B 1985 *J. Less-Common Met.* **113** 231
- [7] Larson A and Von Dreele R 1994 General structure analysis system (gsas) los alamos national laboratory report laur
- [8] Schrieffer J 1999 *Theory of Superconductivity* Advanced Book Program Series (UK: Advanced Book Program, Perseus Books)
- [9] McMillan W L 1968 *Phys. Rev.* **167**(2) 331–344
- [10] Anand V K, Hillier A D, Adroja D T, Strydom A M, Michor H, McEwen K A and Rainford B D 2011 *Phys. Rev. B* **83**(6) 064522 URL <http://link.aps.org/doi/10.1103/PhysRevB.83.064522>
- [11] Orlando T P, McNiff E J, Foner S and Beasley M R 1979 *Phys. Rev. B* **19**(9) 4545–4561 URL <http://link.aps.org/doi/10.1103/PhysRevB.19.4545>

Effect of Mo content on the structural and physical properties of $\text{Cr}_{100-x}\text{Mo}_x$ alloys

A U Chavan, A R E Prinsloo, C J Sheppard and B Muchono

Department of Physics, University of Johannesburg, PO Box 524, Auckland Park, 2006

Corresponding author e-mail: alettap@uj.ac.za

Abstract. Alloying Cr with Mo, which is isoelectric with it, shows an unexpected decrease in the Néel temperature (T_N) with an increase in Mo concentration. This is attributed to a delocalization of the 3-d bands in Cr through the introduction of 4-d electrons of Mo. In the present investigation the effect of Mo concentration on the structural, magnetic and electrical properties of Cr is systematically studied. A series of $\text{Cr}_{100-x}\text{Mo}_x$ alloys, with $x = 0, 3, 7, 15$ and 25 , was prepared and the actual concentrations established using electron microprobe analyses. XRD studies confirm the bcc structure of these alloys as in pure Cr and indicate an increase in lattice constant with an increase in Mo concentration. The crystallite sizes calculated from these results for the $\text{Cr}_{100-x}\text{Mo}_x$ alloys ranges between 15 and 30 nm. The physical properties of these alloys were investigated through magnetic susceptibility (χ), Seebeck coefficient (S), electrical resistivity (ρ) and Hall coefficient (R_H) as function of temperature (T) measurements. T_N values obtained from these measurements are comparable.

1. Introduction

Chromium is a fascinating metal to study because of its antiferromagnetic spin-density-wave (SDW) behaviour below 311 K [1]. The SDW results from the ‘nesting’ of the electron and hole sheets of the Fermi surface on cooling through the Néel temperature (T_N) [2, 3]. This nesting effect is sensitive to changes in the electron-to-atom (e/a) ratio and is influenced by the diluent elements used to dope Cr. In pure Cr, the periodicity of the SDW is incommensurate (I) with that of the lattice [1].

Chromium has six valence electrons and therefore its e/a ratio is 6. When it is alloyed with an element having e/a greater than 6, T_N increases with concentration and ordering changes to commensurate (C) [1]. This happens because the addition of electrons to the d-band of Cr enlarges the electron Fermi surface and brings to match with hole Fermi surface. This results in a larger nesting area, T_N then increases and the ordering becomes commensurate. In opposite case i.e. for e/a less than 6, T_N decreases and the magnetic state remains incommensurate, as in pure Cr. In this case the electron Fermi surface shrinks. There is a mismatch between the electron and hole Fermi surfaces causing T_N to decrease [1, 2, 4]. Few reports are available on Cr alloys in which the diluent elements have the same e/a values as in pure Cr [1]. Considering this, investigations into the physical properties of Cr alloys with isoelectronic elements are important, as it rules out the effect of e/a on the Fermi surface. [5].

The dependence of T_N on the electron concentration and nature of the solute atoms were described by Fedders and Martins [6], and Koehler *et al.* [7] with the help of the isotropic model in which two spherical pieces of Fermi surface with a radius k_c in different bands are connected by a wave vector q . They showed that T_N for such a model is given by an equation [6], $T_N = T_0 \exp(-1/\lambda)$, where T_0 is a function which depends on the band structure and $\lambda = \gamma^2 V(0) k_c^2 / 2\pi^2 v$, where γ is a mean overlap

matrix element for electrons in the same band, $V(0)$ is the average screened Coulomb potential and v is the arithmetic mean of the Fermi velocities in the two bands. The addition of Mo or W slightly modifies the Fermi surface, but since the d-wave functions in these elements are less localized than those in pure Cr, the value of γ is reduced and T_N falls with diluent concentration [1, 6, 7].

$\text{Cr}_{100-x}\text{Mo}_x$ alloys with Mo concentrations below 25 at.% were studied previously by several researchers [5, 10, 11, 12]. Mamiya *et al.* [5] reported extensive measurements on the electrical resistivity and low temperature specific heat for the $\text{Cr}_{100-x}\text{Mo}_x$ alloys in the range $0 \leq x \leq 36.6$ and determined the Néel temperatures from the positions of the anomalies. Mitchell *et al.* [8] compared the effect of Mo and Fe addition to Cr on T_N by measuring the electrical resistivity of Cr-Mo and Cr-Fe alloys, while the thermoelectric power studies for the $\text{Cr}_{100-x}\text{Mo}_x$ alloys with $x = 0, 5, 10$ and $x \geq 25$ were reported by Schröder and Tomaschke [9]. It was found that $\text{Cr}_{100-x}\text{Mo}_x$ alloys with $x \geq 25$ are paramagnetic. The effect of the type and concentration of solute atoms on transport properties of antiferromagnetic Cr alloys have been studied by Trego and Mackintosh [10] through electrical resistivity and thermoelectric power measurements. This study reported measurements for the $\text{Cr}_{100-x}\text{Mo}_x$ alloys with Mo concentration up to approximately 21 at.%. The magnetic susceptibility and Hall coefficient measurements were reported by Bender *et al.* [11] and Shabel *et al.* [12], respectively. However, the Hall coefficient measurements reported were only for the paramagnetic Cr-Mo alloys [12]. Mo was doped to form ternary alloys in order to study two interesting binary Cr alloys SDW systems [1, 13, 14]. The interesting behaviours in Cr-Al-Mo [13] and Cr-Si-Mo [14] ternary alloy systems might be related to the Mo in these and the need to understand the fundamental behaviour of Cr-Mo alloys better therefore arise. This study focus on measurements done using newly sophisticated instruments and extend existing knowledge to include more physical and structural properties of this $\text{Cr}_{100-x}\text{Mo}_x$ alloys, with $x = 0, 3, 7, 15, 25$.

2. Experimental

The $\text{Cr}_{100-x}\text{Mo}_x$ alloys, with $x = 0, 3, 7, 15$ and 25 , were prepared by arc melting in a purified low pressure argon atmosphere. The precursors used for the preparation of these binary alloys were Cr and Mo with purities 99.999% and 99.99%, respectively. Cr and Mo were weighed in required proportion and melted together into a button form in an arc melting furnace with argon atmosphere. After etching the samples with hydrochloric acid they were each sealed in a quartz ampoule filled with pure argon gas at low pressure and annealed at 1273 K for seven days and then quenched in ice bath. The prepared alloy samples were cut and polished using spark erosion techniques in preparation for their characterization.

Structural studies of these samples were carried out by XRD (Phillips PAN analytical X-pert Pro X-ray diffractometer) in the 2θ geometry, in the range $10-90^\circ$ using a $\text{Cu-K}_{\alpha 1}$ radiation ($\lambda = 1.54056\text{\AA}$). The XRD patterns were compared with standard Joint Council of Powder Diffraction Database (JCPDD) files of Cr (04-008-5987) and Mo (42-1120). In order to determine actual composition of the samples, microprobe analyses were carried out using a JEOL electron microprobe analyzer (EPMA) model JXA-8900. A Quantum Design (QD) Physical Property Measurement System (PPMS) incorporating appropriate measuring options was used for the electrical resistivity (ρ) and Seebeck coefficient (S) measurements in the temperature range 2 K to 350 K. Magnetic susceptibility measurements in the temperature range 2 K to 350 K were performed using a QD Magnetic Properties Measurement System (MPMS).

3. Results and Discussion

XRD patterns for the $\text{Cr}_{100-x}\text{Mo}_x$ alloy system with $x = 0, 7$ and 25 are shown in figure 1(a). These patterns confirm the phase synthesis of alloys with expected body centred cubic (bcc) crystal structure. The lattice parameters were computed from the XRD patterns for all the alloys. It is observed that lattice parameter increases with Mo content as shown in figure 1(b). The crystallite sizes were calculated by using Debye Scherrer's formula [15]:

$$D = \frac{0.9 \lambda}{\beta \cos \theta}, \quad (1)$$

where D is crystallite size, λ is the wavelength of the source, β is the full width at half maximum (FWHM) of the peak and θ corresponds to the Bragg angle. The calculated crystallite sizes for these $\text{Cr}_{100-x}\text{Mo}_x$ alloys ranges between 15 and 30 nm.

Electron microprobe analyses done on the $\text{Cr}_{100-x}\text{Mo}_x$ alloys show that all the samples have a uniform composition and the results are tabulated in table 1. Differences in the nominal and actual concentrations are attributed to weight loss during crushing and to evaporation of Cr during melting [5]. The bright and dark spots are observed in the backscattered SEM images shown in figure 2 (a) and (b), and the corresponding Cr and Mo concentrations for bright and dark spots are tabulated in table 1. Also the black spots are observed in SEM images were analysed, these are attributed to the presence of Cr oxide in some regions of the alloys. The area covered by the black spots indicates that the percentage of Cr oxide is approximately 1% to 5% and this should not affect the antiferromagnetic properties of the present alloy system. As far as can be ascertained this is the first report on the correlation between microprobe and SEM analysis for Cr-Mo alloys.

Figure 3 shows the electrical resistivity measurements as a function of temperature, $\rho(T)$, for the series of $\text{Cr}_{100-x}\text{Mo}_x$ alloys with $x \leq 15$. These curves show well defined anomalies in the form of clear minimum followed by hump in the $\rho(T)$ curves. The anomaly, which occurs in the vicinity of T_N is monotonically depressed with increasing Mo concentration as discussed by Mamiya *et al.* [5]. From previous measurements it is known that $\text{Cr}_{100-x}\text{Mo}_x$ alloys with $x \geq 25$ are paramagnetic [5].

T_N values were obtained from $d\rho/dT$ versus T curves by taking the temperature associated with the minimum in these curves [1, 14]. The inset of figure 3 shows $d\rho/dT$ versus T for the $\text{Cr}_{85}\text{Mo}_{15}$ alloy as an example. The increase in the electrical resistivity below T_N for Cr-Mo alloys is attributed to an opening up of the SDW energy gap over part of the Fermi surface [1, 5, 8, 16]. Mamiya *et al.* [5] and Trego *et al.* [10] published some resistivity studies on different $\text{Cr}_{100-x}\text{Mo}_x$ alloys, but no resistivity studies were previously reported for the particular alloys investigated in the present study.

The thermoelectric power is an essential tool for determining T_N in Cr and its alloys because it is more sensitive as compared to the electrical resistivity, to changes in electronic band structure when the SDW state is entered on cooling through T_N [1]. Previously Schröder *et al.* [9] reported thermoelectric measurements for the Cr-Mo alloys in the temperature range 100 K to 600 K while the present study gives the thermoelectric measurements for Cr-Mo alloys in the temperature range 2 K to 400 K. Figure 4 depicts the $S(T)$ curves of pure Cr and $\text{Cr}_{100-x}\text{Mo}_x$ alloys with $x = 3, 7, 15$ and 25. The experimental $S(T)$ curves show a prominent dome, just below T_N , also ascribed to the influence of the SDW energy gap [1, 10, 17]. From figure 4, it is seen that with an increase in Mo concentration in the

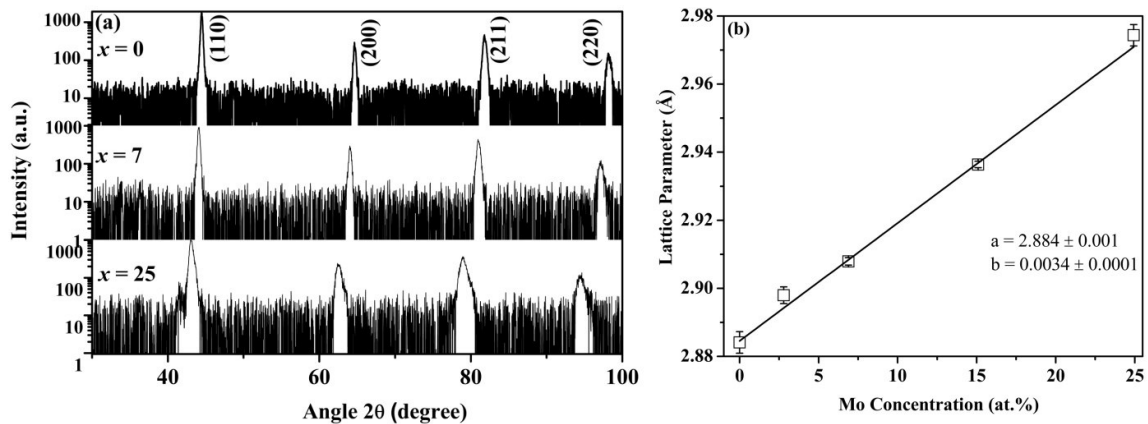
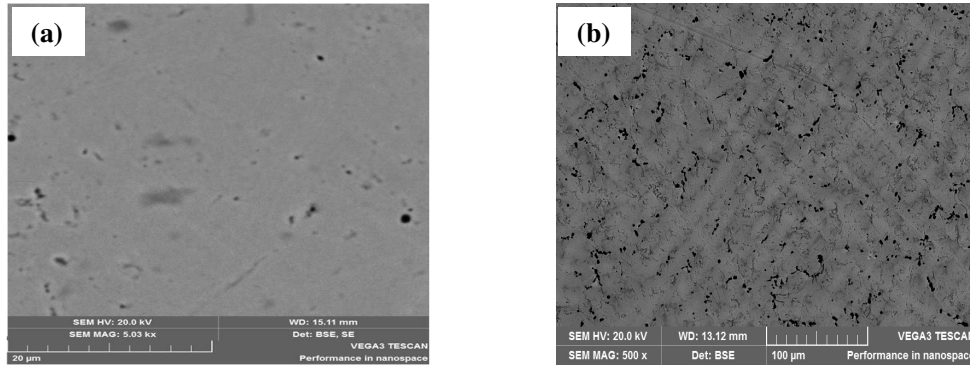
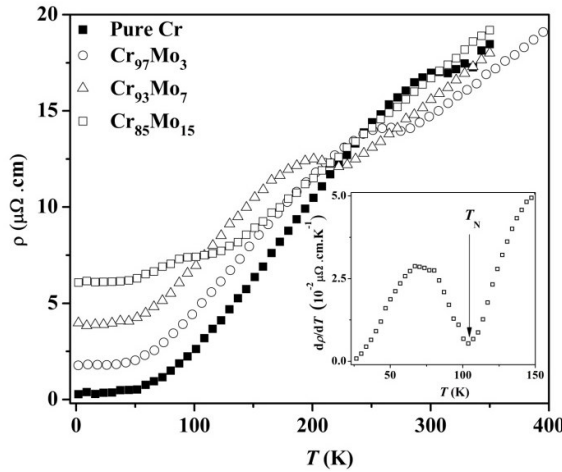
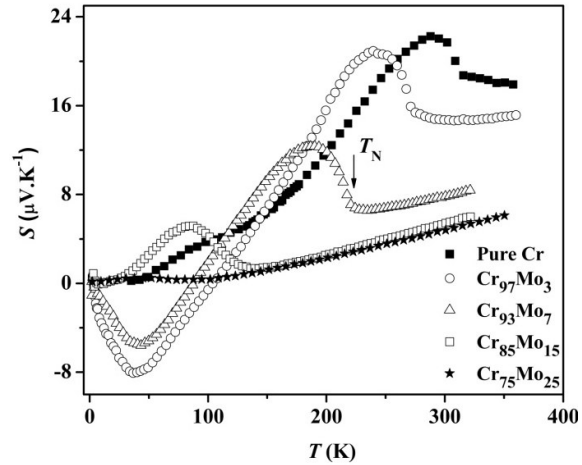


Figure 1. (a) The XRD patterns for the $\text{Cr}_{100-x}\text{Mo}_x$ alloys, with $x = 0, 7$ and 25 (b) The variation in the lattice parameter with Mo content.

Table 1. Elemental concentrations obtained from electron microprobe analyses for Cr-Mo alloys.

Sr. No.	Name of the sample	Nominal concentration (at.%)		Actual concentration (Microprobe) (at.%)		Bright Region		Dark Region	
		Cr	Mo	Cr	Mo	Cr	Mo	Cr	Mo
1	Cr ₉₇ Mo ₃	97	3	97.0 ± 0.2	2.8 ± 0.2	97.27 ± 0.04	2.72 ± 0.03	97.30 ± 0.07	2.70 ± 0.07
2	Cr ₉₃ Mo ₇	93	7	93.1 ± 0.6	6.9 ± 0.6	93.0 ± 0.2	7.05 ± 0.15	98.8 ± 0.8	1.2 ± 0.8
3	Cr ₈₅ Mo ₁₅	85	15	84.9 ± 0.6	15.1 ± 0.6	83.3 ± 0.4	16.4 ± 0.4	86.6 ± 0.6	13.4 ± 0.6
4	Cr ₇₅ Mo ₂₅	75	25	74 ± 1	26 ± 1	71.2 ± 0.3	28.8 ± 0.3	79 ± 1	21 ± 1


Figure 2. Backscattered SEM images of Cr_{100-x}Mo_x alloys with (a) $x = 3$ and (b) $x = 25$.

Figure 3. Electrical resistivity (ρ) as a function of temperature (T) for the Cr_{100-x}Mo_x alloys. Inset shows $d\rho/dT$ versus T for Cr₈₅Mo₁₅ alloy.

Figure 4. The Seebeck coefficient (S) as a function of temperature (T) for Cr_{100-x}Mo_x alloys.

Cr-Mo alloys, the dome is depressed and almost vanishes for sample with 25 at.% Mo. The T_N values were determined from the temperature at the minimum point on the dS/dT versus T curve [17]. For the Cr_{100-x}Mo_x alloys with $x = 3$ and 7, the valleys are observed below T_N , this can be attributed to phonon or magnon drag terms [10, 18].

The dependence of T_N as obtained from $\rho(T)$ and $S(T)$ measurements on Mo concentration is shown in figure 5. T_N decreases linearly with an increase in Mo concentration and disappears for an alloy with

25 at.% Mo. These results agree very well with results obtained in ref. [5]. In addition, the T_N values determined from resistivity and thermoelectric measurements are comparable with each other.

Magnetic susceptibility (χ) measurements as a function of temperature for pure Cr and the $\text{Cr}_{97}\text{Mo}_3$ alloy are shown in figure 6(a) and (b), respectively. The arrows in the $\chi(T)$ curves show the positions of T_N obtained from $\rho(T)$ and $S(T)$ measurements. Pure Cr (figure 6(a)) shows a sharp change in slope of the $\chi(T)$ curve on heating through T_N . $\chi(T)$ for pure Cr is distinctly depressed at $T < T_N$, and is attributed to the effects of the SDW energy gap. A discontinuous sharp step down is observed in the $\chi(T)$ curve of pure Cr at $T_{sf} \approx 123$ K where a spin-flip transition occurs. At T_{sf} the transverse (T) ISDW phase transforms on cooling to a longitudinal (L) ISDW phase [1, 17]. Figure 6(b) shows the increase in $\chi(T)$ on cooling from 350 K to T_N giving a maximum value for χ at T_N , followed by a sharp (or sudden, choose) decrease on further cooling below T_N . This results in an anomalous sharp peak near T_N . Sousa *et al.* [19] observed a similar behaviour in the $\text{Cr}_{100-x}\text{Al}_x$ alloy system with $x = 2.23$ and 2.83. They attributed the peaks in χ at T_N to a local moments resulting in Curie-Weiss paramagnetism above T_N [1, 19]. At $T \approx 65$ K there appears to be a spin-flip transition, not previously reported.

Figure 7 shows the $R_H(T)$ curves for pure Cr and the $\text{Cr}_{97}\text{Mo}_3$ alloy, respectively. The $R_H(T)$ behaviour for pure Cr well matches that reported in [17, 20, 21]. The prominent upturn in $R_H(T)$ in Cr alloys at $T < T_N$ is attributed to the formation of the SDW energy gap, while the downturn below about 150 K probably reveals the spin-flip transition at $T_{sf} = 123$ K. In $R_H(T)$ curve for the $\text{Cr}_{97}\text{Mo}_3$ alloy, it is observed that on cooling through T_N , there is a rise in R_H values. At $T \approx 200$ K the R_H values decrease

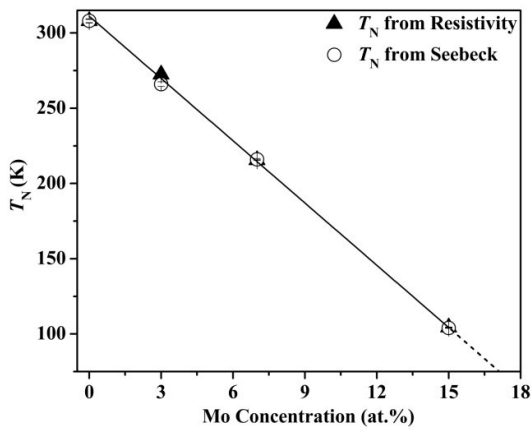


Figure 5. Variation of Néel temperatures (T_N) with respect to Mo content in $\text{Cr}_{100-x}\text{Mo}_x$ alloys.

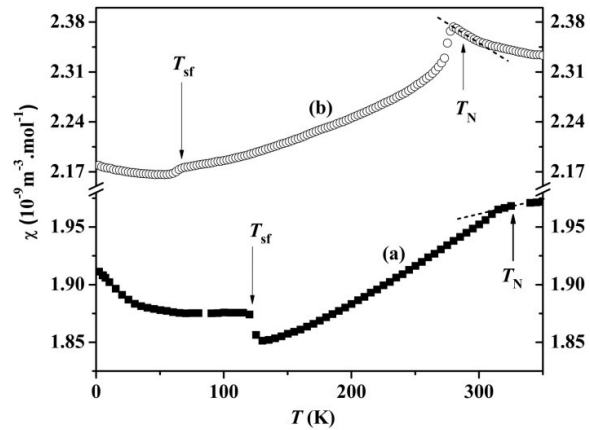


Figure 6. The magnetic susceptibility (χ) as function of temperature (T) for the $\text{Cr}_{100-x}\text{Mo}_x$ with (a) $x = 0$ and (b) $x = 3$.

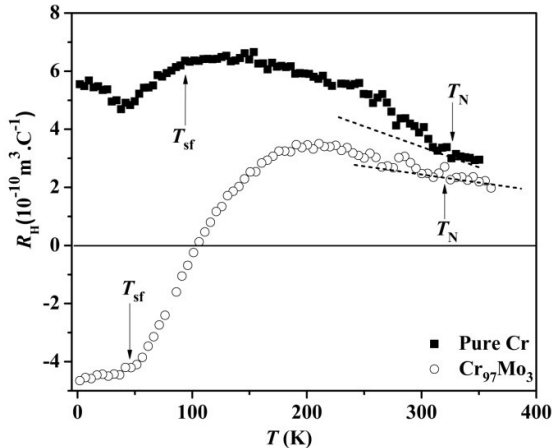


Figure 7. The Hall coefficient (R_H) as a function of temperature (T) for the pure Cr and $\text{Cr}_{97}\text{Mo}_3$ alloy. Arrows indicate the Néel temperatures (T_N) and the spin-flip transition temperature T_{sf} .

monotonically and become negative at $T \leq 100$ K. This behavior can be attributed to a change in the type of charge carriers below 100 K. The positive R_H indicates that the main charge carriers are holes and the hole mobility in the sample is greater than electron mobility [12]. R_H values level off at $T \approx 60$ K, corresponding to the temperature in $\chi(T)$ curve for the same sample where a step is observed. This can possibly be attributed to a spin-flip transition, not previously reported in Cr-Mo alloys. Shabel *et al.* [12] reported the Hall effect measurements for Cr-Mo alloys in the concentration range $x \geq 25$. No Hall measurements were reported previously for the Cr-Mo alloys with $x < 25$.

4. Conclusions

The present investigation reports on the $\text{Cr}_{100-x}\text{Mo}_x$ alloys with $x = 0, 3, 7, 15$ and 25 . The XRD results reveal the expected pure bcc structure of these alloys. The lattice parameters calculated from XRD patterns increase with Mo concentration. The crystallite sizes for the Cr-Mo alloys ranges from 15 to 30 nm. Prominent anomalies were observed in the electrical resistivity and Seebeck coefficient measurements as a function of temperature for these alloys. T_N values for each alloy, determined from these measurements agree very well with each other and decrease with an increase in Mo content. $\chi(T)$ measurements give nearly same value for T_N and also indicate a spin-flip transitions in pure Cr and $\text{Cr}_{97}\text{Mo}_3$ alloy. Hall measurements in $\text{Cr}_{97}\text{Mo}_3$ alloys reported here for first time. It shows positive as well as negative R_H values, indicating a change in number as well as type of charge carriers.

Acknowledgements

Financial support from the South African NRF (Grant numbers 80928 and 80626) and the Faculty of Science from the University of Johannesburg are acknowledged.

References

- [1] Fawcett E, Alberts H L, Galkin V Y, Noakes D R and Yakhmi J V 1994 *Rev. Mod. Phys.* **66** 25
- [2] Lomer W M 1962 *Proc. Phys. Soc.* **80** 489
- [3] Ström-Olsen J O and Wilford D F 1980 *J. Phys. F: Metal Phys.* **10** 1467
- [4] Jayaraman A 1998 *Curr. Sci.* **75** (11) 1200
- [5] Mamiya T and Masuda Y 1976 *J. Phys. Soc. Jpn.* **40** (2) 390
- [6] Fedders P A and Martin P C 1966 *Phys. Rev.* **143** (1) 245
- [7] Koehler W C, Moon R M, Trego A L and Mackintosh A R 1966 *Phys. Rev.* **151**(2) 405
- [8] Mitchell M A and Goff J F 1972 *Phys. Rev. B* **5** (3) 1163
- [9] Schröder K and Tomaschke H 1968 *Phys. Kondens. Materie* **7** 318
- [10] Trego A L and Mackintosh 1968 *Phys. Rev.* **166** (2) 495
- [11] Bender D and Muller J 1970 *Phys. Kondens. Materie* **10** 342
- [12] Shabel B and Schröder K 1967 *J. Phys. Chem. Solids* **28** 2169
- [13] Muchono B, Prinsloo A R E, Sheppard C J, Alberts H L and Strydom A M 2014 *J. Magn. Magn. Mater.* **354** 222
- [14] Sheppard C J, Prinsloo A R E, Alberts H L, Strydom A M *J. Appl. Phys.* 2011 **109** 07E104-1
- [15] Cullity B D 1956 *Elements of X-ray diffraction Addison-Wesley Publ. Comp., Inc. U. S. America* 97
- [16] Mitchell M A and Goff J F 1975 *Phys. Rev. B* **12** (5) 1858
- [17] Sheppard C J, Prinsloo A R E, Alberts H L, Muchono B and Strydom A M 2014 *J. Alloys Compd.* **595** 164
- [18] Reddy L, Prinsloo A, Sheppard C and Strydom A 2013 *J. Korean Phys. Soc.* **63** (3) 756
- [19] Sousa J B, Amado M M, Pinto R P, Pinheiro M F, Braga M E, Moreira J M, Hedman L E, Åström H U, Khlaif L, Walker P, Garton G and Hukin D 1980 *J. Phys. F: Metal Phys.* **10** 2535
- [20] Furuya Y 1976 *J. Phys. Soc. Jpn.* **40** 490
- [21] Furuya Y, Misui T 1976 *J. Phys. Soc. Jpn.* **41** 1938

Magnetic and thermodynamic properties of $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ compound

J C Debnath¹, A M Strydom¹, O Niehaus² and R Pöttgen²

¹Highly Correlated Matter Research Group, Department of Physics, University of Johannesburg, South Africa.

²Institut für Anorganische und Analytische Chemie, Universität Münster, Germany.

Author e-mail address: jdebnath@uj.ac.za

Abstract. The magnetic susceptibility and heat capacity of the novel compound $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ have been studied above room temperature to low temperature range and in the applied magnetic field up to 7 T. This compound crystallizes with the hexagonal non-centrosymmetric $\text{Pr}_{23}\text{Ir}_7\text{Mg}_4$ -type structure, with space group $P6_3mc$. The structure is built up from complex three dimensional networks of edge and corner-sharing RE_6Ru trigonal prisms. The magnetic susceptibility and specific heat both exhibit a distinct anomaly at about ~ 2 K which most probably suggests a paramagnetic to antiferromagnetic phase transition. The magnetic susceptibility revealed a magnetic moment $\mu_{\text{eff}} = 2.24 \mu_B/\text{Ce}$ which is close to the value for cerium in pure Ce metal ($\mu_{\text{eff}} = 2.54 \mu_B$), indicating a presence of well localized magnetic moments carried by the stable Ce^{3+} ions but with significant deviation from full moment. The magnitude of the electronic specific heat coefficient $\gamma = 127 \text{ mJ/Ce-mol K}^2$ suggests correlated electron behavior in this compound.

1. Introduction

Complex metal alloys have been shown to offer new possibilities in developing high efficiency thermoelectric material [1]. Among the correlated electron class of magnetic systems, the enhanced thermoelectric power characteristic of Kondo metals offers a distinct advantage in gaining thermoelectric efficiency. In recent years, the rare earth (RE)–transition metal (T)–magnesium systems has been studied a few with respect to phase analyses, crystal structures, and chemical bonding as well as magnetic and mechanical properties. The rare earth (RE) metal rich $\text{RE}_x\text{T}_y\text{Mg}_z$ (T = late transition metal) compounds show a strong tendency for a formation of transition metal centered RE_6 trigonal prisms with different condensation patterns and strong RE-T bonding [2].

These $\text{RE}_x\text{T}_y\text{Mg}_z$ intermetallics have technical importance for precipitation coarsening in modern light weight alloy systems [3]. The $\text{RE}_x\text{T}_y\text{Mg}_z$ compounds with the so far highest rare earth content are the hexagonal ones with $\text{Pr}_{23}\text{Ir}_7\text{Mg}_4$ -type structure, space group $P6_3mc$ [4]. They have similar structural features, i.e. a complex network of condensed RE_6T prisms and Mg_4 tetrahedra which fill voids.

The crystal chemical details have been reported in the literature [5 – 8]. The electronic structure, chemical bonding, and physical properties of $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ compound have also been reported [9-10]. Unfortunately physical properties at very low temperature (< 2 K) are not available for this compound. Keeping this in mind here we present the physical properties of $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ at very low temperature to room temperature. Here we present exploratory results of a study on the novel compound $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ which has 68 atoms per unit cell and therefore qualifies as a complex metal alloy.

2. Experimental

A polycrystalline $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ sample was synthesized from small cerium ingots (smart elements, > 99.9 %), ruthenium powder (Heraeus, *ca.* 200 mesh, > 99.9 %), and a magnesium rod (Johnson Matthey, \varnothing 16 mm, > 99.95 %). Pieces of the cerium ingot were first arc-melted to a small button under purified argon. The cerium button, the ruthenium powder, and pieces of the magnesium rod were then weighed in the 23 : 7 : 4 atomic ratios and arc-welded in a tantalum tube under an argon pressure of about 800 mbar. The ampoule was inductively heated (Hüttinger Elektronik, Freiburg, Type TIG 5/300) first for 2 minutes at about 1300 K, followed by 2 hours at *ca.* 920 K. Finally the tube was quenched to room temperature. The brittle $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ sample could easily be separated from the tube. No reaction with the tantalum container was observed. $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ crystallizes with the hexagonal $\text{Pr}_{23}\text{Ir}_7\text{Mg}_4$ type structure, space group $\text{P6}_3\text{mc}$, $a = 993.5(3)$, $c = 2243.9(8)$ pm.

3. Results and discussion

The temperature dependence of inverse magnetic susceptibility of $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ between 1.76 and 400 K measured in an applied field of 0.05 T is presented in figure 1. A fit of the inverse magnetic susceptibility in the region 125-400 K according to a modified Curie-Weiss expression $\chi^{-1}(T) = [\chi_0 + C/(T - \theta_p)]^{-1}$ revealed a temperature-independent contribution $\chi_0 = 0.0004168$ emu/Ce-mole, an experimental effective magnetic moment $\mu_{\text{eff}} = 2.24 \mu_B/\text{Ce}$, which is lower than the free ion value of Ce^{+3} $\mu_{\text{eff}} = 2.54 \mu_B$ and can be attributed to not all cerium atoms being in a trivalent state as the basic building unit of the $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ structure is ruthenium centered trigonal prisms formed by cerium atoms which are capped by three further cerium atoms on the rectangular faces. The monomeric building units (trigonal prisms) are condensed via common edges and corners to three-dimensional networks which leave cavities for Mg_4 tetrahedra. In the structure of $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ compound, Ce1, Ce2, Ce7, and Ce8 can be considered as trivalent, while Ce3, Ce4, Ce5, Ce6, and Ce9 tend towards intermediate cerium valence. Thus, 30 of the 46 cerium atoms per unit cell are in an intermediate valence state [9-10]. Extrapolation of the $\chi^{-1}(T)$ vs. T data lead to a Weiss constant of $\theta_p = -32.5$ K.

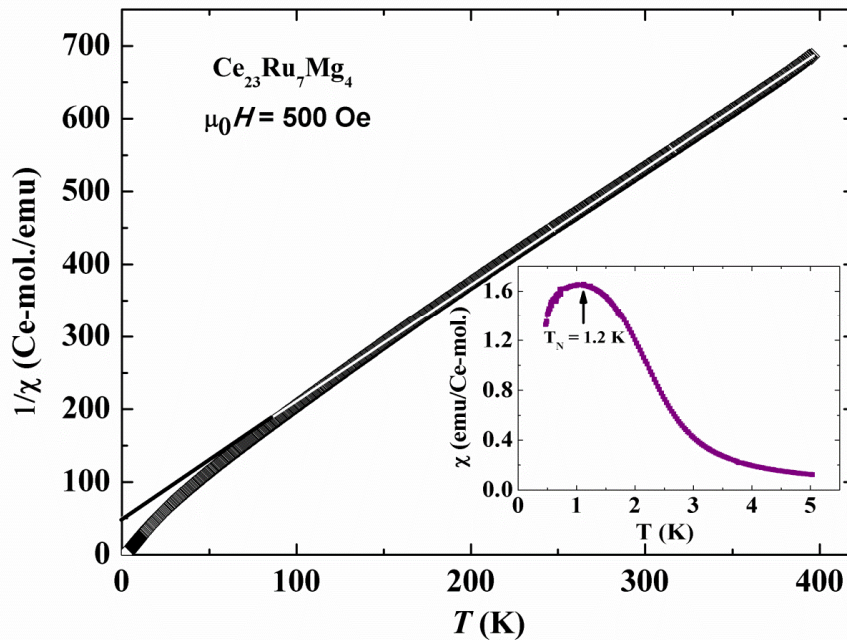


Figure 1. Temperature dependence of the inverse magnetic susceptibility of $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ measured in 500 Oe. The line represents the modified Curie-Weiss fit of $\chi^{-1}(T)$. Inset shows the low temperature data of $\chi(T)$, arrow points to the anomaly corresponding to the magnetic phase transition.

The observed negative sign of θ_p can be understood to arise from the development of antiferromagnetic-type correlations between the Ce moments. The attained μ_{eff} value is higher than the obtained value by Linsinger et al. [9]. According to Linsinger et al. μ_{eff} value is $2.01 \mu_B/\text{Ce}$ and the $\theta_p = -13.7 \text{ K}$ at applied field 1 T. Below 125 K, the inverse magnetic susceptibility $\chi^{-1}(T)$ considerably deviates from Curie-Weiss behavior, which could be attributed to crystal field splitting of the $J = 5/2$ ground state of Ce^{+3} and/or the onset of short-range magnetic interactions. With decreasing temperature the magnetic susceptibility exhibits a distinct anomaly due to magnetic ordering which we provisionally attribute to antiferromagnetic type ordering at $T_N = 1.2 \text{ K}$. The inset in figure 1 shows the low temperature dependence of $\chi(T)$ with better visible anomaly and the arrow indicating the magnetic phase transition.

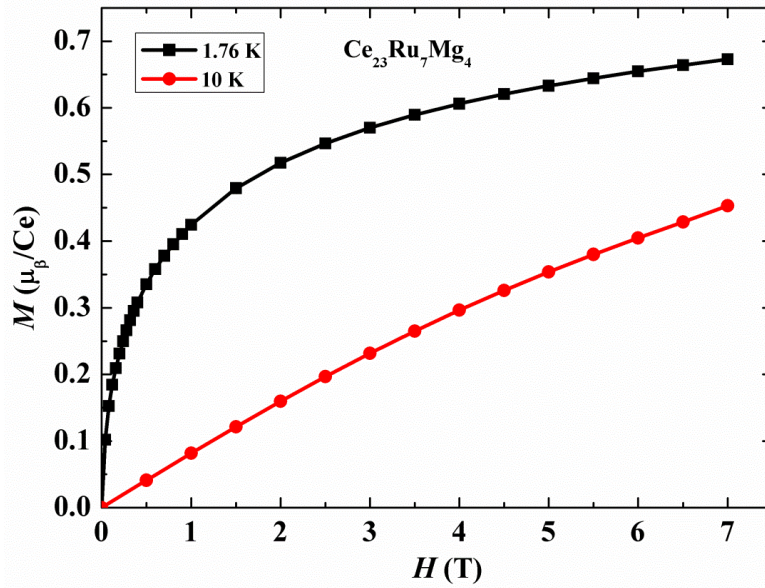


Figure 2. The magnetization for $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ at different temperatures.

The magnetization isotherms taken at 1.76 and 10 K are shown in figure 2. The magnetization M increases near linearly with applied magnetic field at temperature $T = 10 \text{ K}$, whereas at $T = 1.76 \text{ K}$ on the other hand a strong curvature in $M(H)$ is found. It is noted that the magnetization is especially strongly curved and field dependent in the $\leq 2 \text{ T}$ region. The value of magnetic moment measured at 1.76 K in the upper field limit of 7 T amounts to $0.67 \mu_B/\text{Ce}$ which is much reduced from the theoretical value for the saturated moment of free Ce^{+3} ion $\mu_s = 2.14 \mu_B$. It can be accredited to crystal field splitting of the $J = 5/2$ ground state and owing to the fact that not all Ce atoms are in a stable trivalent state similar to $\text{Ce}_{23}\text{Ni}_7\text{Mg}_4$ [11]. This has also been observed in $\text{Ce}_{23}\text{Ru}_7\text{Cd}_4$ [12]. The magnetic moment value $0.59 \mu_B/\text{Ce}$ at 8 T and 2.5 K was obtained by Linsinger et al. [9].

The temperature dependence of the heat capacity $C_p(T)$ of $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ in the temperature range 1.94–245 K in zero magnetic field is shown in figure 3. The heat capacity measurement of $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ shows an upturn starting at about 2 K shown in inset of figure 3. The sharp peak with a maximum at about 2 K in zero field, we ascribe to the transition into a magnetically ordered phase. The appearance of the magnetic phase transition is consistent with measured susceptibility where a pronounced maximum is present near this temperature shown inset in figure 1.

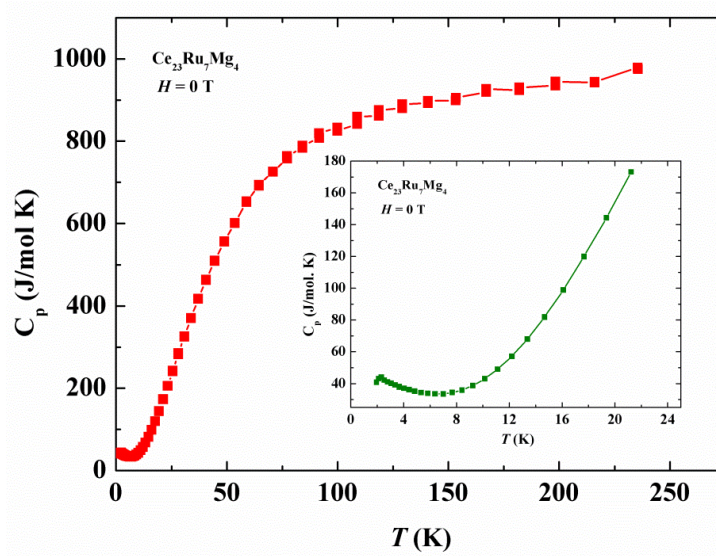


Figure 3. Temperature dependence of the heat capacity of $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ up to 245 K in zero magnetic field. Inset: expanded low- temperature part of the heat capacity up to 22 K.

The value of the electronic coefficient of the specific heat γ has been taken as the extrapolation of the linear part of the C_p/T vs. T^2 curves at low temperatures indicated by red line displayed in Fig. 4. The upturn in C_p/T vs. T^2 observed at low temperatures is related with the magnetic phase transition in low field. The obtained electronic coefficient of specific heat γ value is $127 \text{ mJ/Ce-mol}\cdot\text{K}^2$ for $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$. This value of γ suggests that there may be strong electronic correlations present in $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$. One origin may be the Kondo effect produced by an on-site hybridization between f and conduction electrons. As a precursor effect to the lower-lying magnetic phase transition however, an enhanced value of γ may also result from short-range correlations immediately above T_N . The value of θ_D obtained from the simplified Debye model is approximately 175 K.

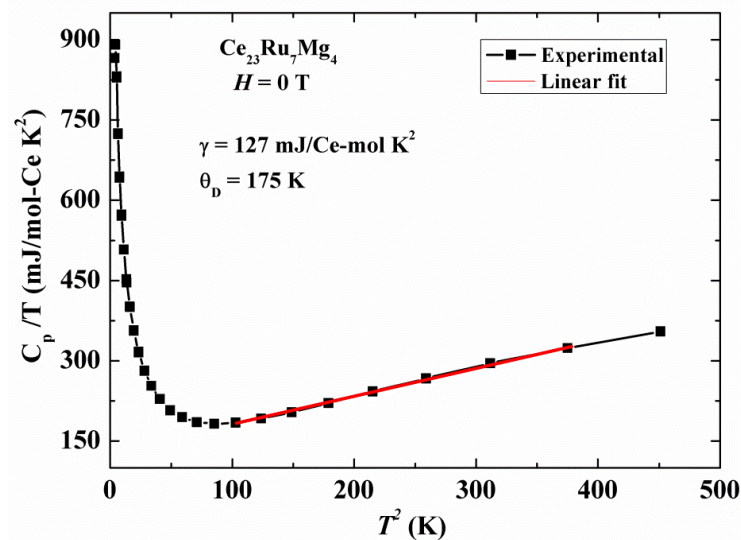


Figure 4. Temperature dependence of the heat capacity displayed as C_p/T vs T^2 .

The temperature dependence of the heat capacity of $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ in zero and 5 T applied magnetic field shown in figure 5. It is shown that with applied magnetic field the peak is flattened and lifted towards higher temperature.

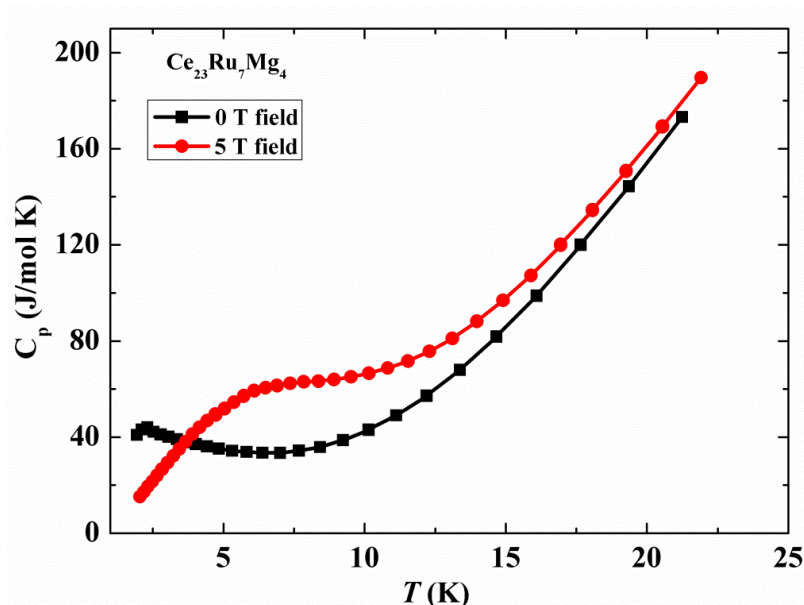


Figure 5. Temperature dependence of the heat capacity in zero and 5 T applied magnetic field.

4. Conclusion

The magnetic susceptibility and heat capacity of $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$ compound have been studied above room temperature to very low temperature. The magnetic susceptibility exhibits a distinct anomaly at 1.2 K with paramagnetic to antiferromagnetic phase transition. The heat capacity measurement shows a maximum at about 2 K which is almost consistent with the measured susceptibility where strong maximum is present near this temperature. Magnetic susceptibility with paramagnetic to antiferromagnetic phase transition and heat capacity measurements suggests that there may be strong electronic correlations present in $\text{Ce}_{23}\text{Ru}_7\text{Mg}_4$. The heat capacity measurements provide for the electron specific heat coefficient $\gamma = 127 \text{ mJ/Ce-mol}\cdot\text{K}^2$ which originates from an appreciable amount of f – electron involvement in the conduction electrons and the resulting strongly hybridized state.

Acknowledgements

JCD acknowledges the NRF (Grant no.85115) for an innovation Postdoctoral Research Fellowship, supervised by AMS. AMS thanks the SA-NRF (78832) and the URC of UJ for financial assistance.

References

- [1] Paschen S, Godart C and Grin Y 2011 in Complex metallic alloys: fundamentals and applications, Jean-Marie Dubois; Esther Belin-Ferré (eds.), Weinheim: Wiley-VCH Verlag
- [2] Linsinger S, Tuncel S, Hermes W, Eul M, Chevalier B and Pöttgen R 2009 Z. Anorg. Allg. Chem. **635** 282
- [3] Kainer K U (Ed.) 2004 Magnesium, Proceedings of the 6th International Conference on Magnesium Alloys and their Applications, Wiley-VCH, Weinheim
- [4] Rodewald U Ch, Tuncel S, Chevalier B and Pöttgen R 2008 Z. Anorg. Allg. Chem. **634** 1011
- [5] Rodewald U Ch, Chevalier B and Pöttgen R 2007 J. Solid State Chem. **180** 1720
- [6] Solokha P, De Negri S, Saccone A, Pavlyuk V, Marciniak B and Tedenac J-C 2007 Acta Crystallogr. **C63** i13
- [7] De Negri S, Solokha P, Saccone A and Pavlyuk V 2009 Intermetallics **17** 614
- [8] Solokha P, De Negri S, Pavlyuk V and Saccone A 2009 Solid State Sci. **11** 801
- [9] Linsinger S, Eul M, Hermes W, Hoffmann R-D and Pöttgen R 2009 Z. Naturforsch. **64b** 1345
- [10] Matar S F, Chevalier B and Pöttgen R 2012 Intermetallics **31** 88
- [11] Tuncel S, Hermes W, Chevalier B, Rodewald U Ch and Pöttgen R 2008 Z. Anorg. Allg. Chem. **634** 2140
- [12] Tappe F, Hermes W, Eul M and Pöttgen R 2009 Intermetallics **17** 1035

Can Shapiro step subharmonics be “charged”?

H Azemtsa Donfack^{1,2}, **Yu M Shukrinov**^{1,2,3} and **I R Rahmonov**^{2,3,4}

¹Department of Physics, University of South Africa, P.O. Box 392, Pretoria 0003, South Africa

² BLTP, JINR, Dubna, Moscow Region, 141980, Russia

³ Dubna International University of Nature, Society, and Man, Dubna, Moscow region, 141980 Russia

⁴Umarov Physical Technical Institute, TAS, Dushanbe, 734063 Tajikistan

E-mail: adonfh@unisa.ac.za

Abstract. Using the capacitively coupled Josephson junctions with diffusion current (CCJJ + DC) model, we performed a precise numerical study of phase dynamics of intrinsic Josephson junctions under external electromagnetic radiation. We survey the different Shapiro step subharmonics found in these systems. We demonstrate the charging of superconducting layers in bias-current interval corresponding to a given Shapiro step subharmonics and the existence of longitudinal plasma wave along the stack of junctions.

1. Introduction

The system of superconducting layers in high temperature superconductors (HTSC) such as $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ (Bi-2212) represents intrinsic Josephson junctions (IJJs). In such systems, as it is the case in single Josephson junctions, the locking of the Josephson oscillations with frequency ω_J to the frequency ω of external electromagnetic radiation leads to the appearance of Shapiro steps [1, 2] in the current voltage characteristics (IV-characteristics). Many devices in existence exploit this effect, notably voltage standards and terahertz radiation emitters/detectors. Therefore, a detailed study of the Shapiro steps and their subharmonics in the IJJs under different resonance conditions presents important research questions with various potential applications.

An interesting feature of the IJJs is the possibility of longitudinal plasma wave (LPW) propagating along the c-axis [3]. Each superconducting layer is characterized by the order parameter $\Delta_l(t) = |\Delta| \exp(i\theta_l(t))$ with the time dependent phase $\theta_l(t)$. The thickness of superconducting layers (about $\sim 3 \text{ \AA}$) in an HTSC is comparable with the Debye length r_D of electric charge screening. Therefore, there is no complete screening of the charge in the separate layers, and the electric field induced in each JJ penetrates into the adjacent junctions. Thus, the electric neutrality of superconducting layers is dynamically broken and, in the case of the ac Josephson effect, a capacitive coupling appears between the adjacent junctions [3]. The absence of complete screening of charge in the superconducting layer leads to the formation of a generalized scalar potential Φ_l of the layer, which is related to the charge density Q_l in the superconducting layer as follows [3, 4]: $Q_l = -\frac{1}{4\pi r_D^2} \Phi_l$.

The existence of a relationship between the electric charge Q_l of the l -th layer and the generalized scalar potential Φ_l of this layer reflects a non-equilibrium nature of the ac Josephson effect in layered HTSC [4]. In this case, the diffusion contribution to the quasiparticle

current arises due to the generalized scalar potential difference. The latter is taken into account in the capacitively coupled Josephson junction model with diffusion current (CCJJ+DC model)[5]. At $\omega_J = 2\omega_{LPW}$ (ω_J and ω_{LPW} are the Josephson and longitudinal plasma wave frequencies, respectively) the parametric resonance is realized: the Josephson oscillations excite the longitudinal plasma wave. The charge in the superconducting layer at parametric resonance can have complex oscillations depending on the number of junctions in the stack, coupling and dissipation parameters and boundary conditions. Fourier analysis [6] of the temporal dependence of the charge in a superconducting layer shows a spectrum of frequencies, in particular, ω_{LPW} , ω_J , and their combinations.

In this paper, we present the results of the investigation of the effects of electromagnetic radiation on the phase dynamics of the intrinsic JJs and the temporal oscillations of the electric charge in superconducting layers in HTSC. For fixed parameters of JJs and parameters of simulations, we have studied the manifestation of the Shapiro step subharmonics in IV-characteristics and demonstrated their “charging”, i.e. the charging of superconducting layers in the corresponding bias current intervals at fixed amplitude of external radiation. To escape the complexity related to the overlapping of the SS subharmonics, we consider here a case of small radiation amplitudes.

2. Model and method

To investigate the phase dynamics of the IJJ, we use the one-dimensional CCJJ+DC model [5] with the gauge-invariant phase differences $\varphi_l(t)$ between S-layers l and $l+1$ in the presence of electromagnetic irradiation described by the system of equations:

$$\begin{cases} \frac{\partial \varphi_l}{\partial t} = V_l - \alpha(V_{l+1} + V_{l-1} - 2V_l) \\ \frac{\partial V_l}{\partial t} = I - \sin \varphi_l - \beta \frac{\partial \varphi_l}{\partial t} + A \sin \omega t + I_{noise} \end{cases} \quad (1)$$

where t is the dimensionless time normalized to the inverse plasma frequency ω_p^{-1} , $\omega_p = \sqrt{2eI_c/\hbar C}$, C is the capacitance of the junctions, $\beta = 1/\sqrt{\beta_c}$, β_c is the McCumber parameter, α gives the coupling between junctions [3], ω and A are the frequency and amplitude of the radiation, respectively. To find the IV-characteristics of the stack of the intrinsic JJ, we solve this system of nonlinear first order differential equations (1) using the fourth order Runge-Kutta method. In our simulations we measure the voltage in units of $V_0 = \hbar\omega_p/(2e)$, the frequency in units of ω_p , the bias current I and the amplitude of radiation A in units of I_c .

To calculate the voltages $V_l(I)$ at each I , we simulate the dynamics of the phases $\varphi_l(t)$ by solving the system of equations (1) using the fourth-order Runge-Kutta method with a step in time T_p (a schematic of the numerical procedure and additional parameters of simulation are shown in figure 1 of Ref. [7]). As a result, we obtain the temporal dependence of the voltages in each junction at a fixed value of bias current. So, we can calculate the temporal dependence of the charge in each layer as well as through the voltage difference in the neighboring junctions (see below). After completing the calculations for bias current value I , the current value is increased or decreased by a small amount of δI (bias current step) and the voltages in all junctions at the next point of the IV-characteristics is calculated. So actually, the time dependence of voltage in each junction or the charge in each layer consists of intervals at each fixed current value. We use the distribution of phases and voltages achieved at the previous point of the IV-characteristics as the initial distribution for the current point. The average of the voltage \bar{V}_l is given by

$$\bar{V}_l = \frac{1}{T_f - T_i} \int_{T_i}^{T_f} V_l dt \quad (2)$$

where T_i and T_f determine the interval for the temporal averaging. In our simulations, we use $T_i = 2^3T$, $T_f = 2^{10}T$. The time step is $\Delta T = T/(2 \times 160)$. Here the factor T is $T = 2\pi/\omega(2/3)(3/5)(4/7)$. This choice is not unique but necessary for the stability of our calculations.

To study time dependence of the electric charge in the S-layers, we use the Maxwell equation $\text{div}(\epsilon\epsilon_0\vec{E}) = Q$, where ϵ and ϵ_0 are relative dielectric and electric constants, respectively. The charge density Q_l (in the following text - charge) in the S-layer l is proportional to the difference between the voltages V_l and V_{l+1} in the neighbor insulating layers

$$Q_l = Q_0\alpha(V_{l+1} - V_l), \quad (3)$$

where $Q_0 = \epsilon\epsilon_0 V_0/r_D^2$. An estimate, taking the parameters of the IJJ, gives a value of about $2.4 \times 10^{-16}C$. This value of charge is not significantly high, but it makes for an interesting physics. In our numerical calculations, we have used the coupling parameter $\alpha = 0.05$, the dissipation parameter $\beta = 0.2$ and periodic boundary conditions. We note that the qualitative results are not very sensitive to these parameters and boundary conditions. The details of the model and simulation procedure are presented in Ref. [8].

3. Results and discussion

3.1. Shapiro step subharmonics in IV-characteristic

As mentioned earlier, the Shapiro step subharmonics appear when Josephson junctions are exposed to electromagnetic radiation. This happens when the resonance condition $q\omega_J = p\omega$ is fulfilled. As a result, we expect to find steps on the IV-curves at voltages $V_{p,q}$ such that,

$$V_{p,q} = N\omega_J = \left(\frac{p}{q}\right)N\omega, \quad (4)$$

Integers ratio p/q correspond to the main Shapiro steps, whereas non-integer ratio p/q give the subharmonics.

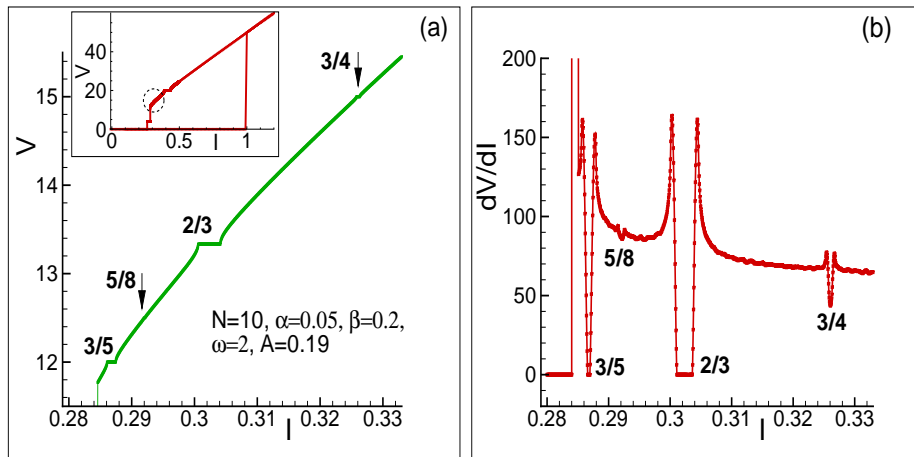


Figure 1. (Color online) (a) Demonstration of subharmonics in IV-characteristic of the stack of 10 IJJs with $\beta = 0.2$, $\alpha = 0.05$ under external radiation at frequency $\omega = 2$ and amplitude $A = 0.19$. Inset shows the total IV-curve with dashed circle indicating the enlarged part of IV-characteristic in main figure; (b) Manifestation of the subharmonics on the dependence of the differential resistance dV/dI versus bias current I for the same stack.

The inset to figure 1(a) shows the total IV-characteristic of the stack of 10 IJJs with $\beta = 0.2$, $\alpha = 0.05$ under external radiation at frequency $\omega = 2$ and amplitude $A = 0.19$. We see the

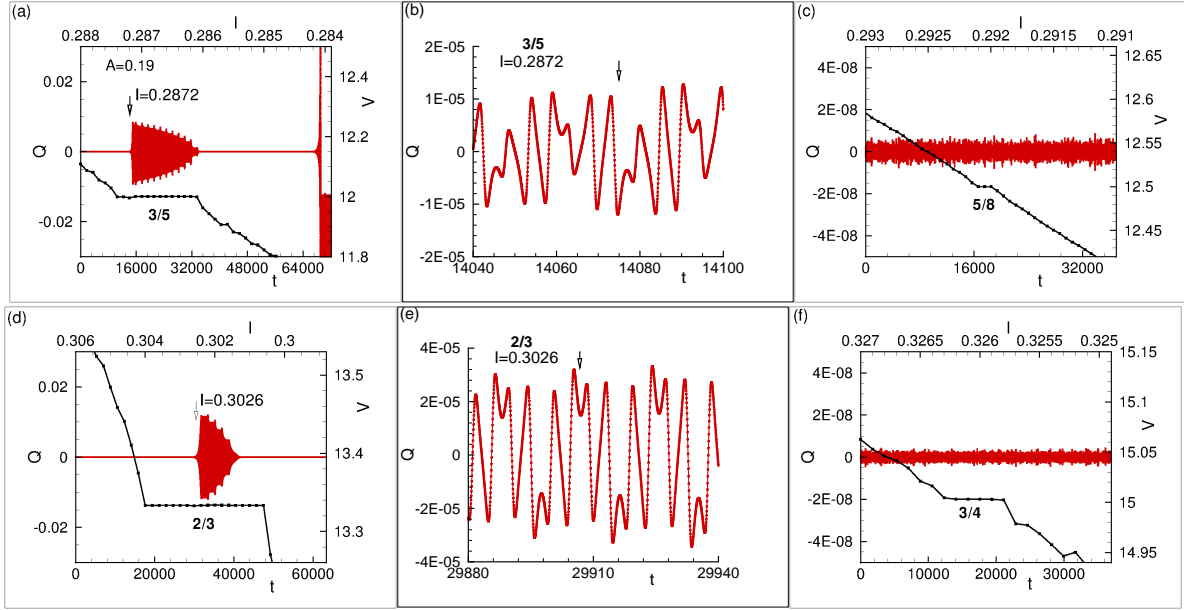


Figure 2. (Color online) Charge-time dependence (top curve) and corresponding parts of IV-characteristics (bottom curve/dotted lines) around Shapiro step subharmonics: (a) $3/5$, (c) $5/8$, (d) $2/3$ and (f) $3/4$. Panels (b) and (e) enlarge the parts of charge-time curves indicated by arrows in (a) and (d). They demonstrate the character of charge oscillations in the superconducting layers at chosen values of bias current.

Shapiro step at $V = 20$ and some inner branches before transition to the zero voltage state. Part of the IV-curve marked by circle is enlarged in the main figure. It demonstrates the Shapiro step subharmonics $2/3$ and $3/5$ observed and the corresponding values of voltage. We note that these steps belong to the third level of continued fractions $N - 1/(n + (1/m))$ with $N = 1$ and $n = 2$ [9]. We show also in figure 1(b) the differential resistance dV/dI versus bias current I to demonstrate the subharmonics $3/4$ and $5/8$ which are not clearly visible in the IV-characteristics. Their positions in IV-characteristic are indicated by arrows. They belong to the continued fractions $N - 1/(n \pm (1/m))$ with $N = 1$, $n = 3$ and sign plus and minus, respectively.

3.2. “Charging” of Shapiro steps subharmonics

A charging of superconducting layers in the bias current interval corresponding to the Shapiro step at double resonance conditions was previously demonstrated [10]. A question concerning the “charging” of Shapiro step subharmonics was not investigated. Here, we attempt this question.

For the current intervals corresponding to the Shapiro steps subharmonics in figure 1, we compute the electric charge in each superconducting layer using Eq. 3. The charge versus time graphs for these current intervals and corresponding parts of the IV-characteristics (black curves) can be found in figure 2.

We observe that the stack of IJJ in the current interval corresponding to the SS subharmonics $3/5$ and $2/3$ has noticeable charge oscillations. The character of charge oscillations is demonstrated in the panels (b) and (e) for the parts of charge-time curves indicated by arrows in (a) and (d). We see the difference in the dynamics of charge oscillations which reflects the different wavelength of realized longitudinal plasma wave in both cases. The SS subharmonics $5/8$ and $3/4$ do not demonstrate “charging”: the amplitude of charge oscillations is in the order

of 10^{-8} (i.e. on the noise level of our simulations).

3.3. Longitudinal plasma waves

The effect of the increase of the external radiation amplitude on the LPW created at the radiation related parametric resonance was investigated in Ref. [10]. A remarkable change in the longitudinal plasma wavelength was demonstrated. Here we show that LPWs with different wavelength can be realized in the stack in the bias current intervals corresponding to the Shapiro step subharmonics.

Figure 3 demonstrates the longitudinal plasma waves created in the stack of 10 IJJs in current intervals corresponding to the “charged” regions of 3/5 and 2/3 steps. We see the realization of LPW with wavelength $\lambda = 5$ in case of Shapiro step subharmonic 3/5 and $\lambda = 2$ in case of 2/3. The effect of the increase of radiation amplitude on the wavelength of the LPW will be studied somewhere else.

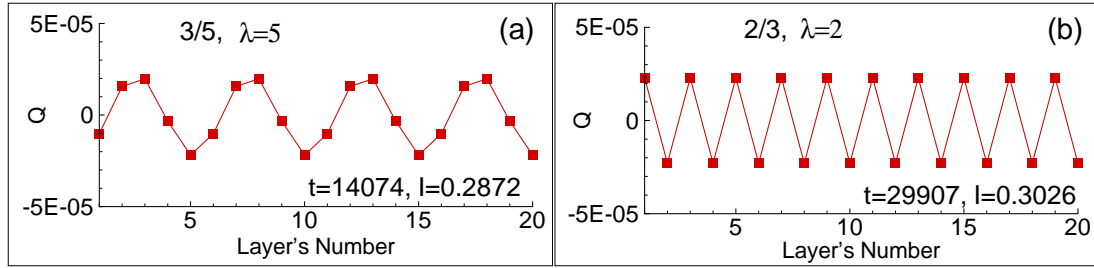


Figure 3. (Color online) Demonstration of the longitudinal plasma waves created in the stack of 10 IJJ: (a) Distribution of charge along the stack at fixed time in the current interval of step 3/5; (b) The same for the step 2/3.

3.4. Conclusions

We have demonstrated the charging of superconducting layers in bias current intervals corresponding to Shapiro step subharmonics. The existence of the longitudinal plasma wave along the stack of junctions in these intervals was illustrated. We found an interesting feature: this “charging” of subharmonics is not regular. Particularly, the step 5/8 which is placed between the “charged” steps 3/5 and 2/3, is not “charged”. This phenomenon requires additional investigation.

Acknowledgments

We thank A. E. Botha, M. R. Kolahchi for detailed discussion of this paper and D. V. Kamanin, L. M. Lekala and the JINR-SA agreement for the support of this work. H.A.D acknowledges the support of the NRF-JINR, grant number 88620. Y.M.S acknowledges the support of the Visiting Researcher Programme of UNISA.

References

- [1] Shapiro S 1963 *Phys. Rev. Lett.* **11** 80
- [2] Tinkham M 1996 *Introduction to Superconductivity*, 2nd edition (New York: McGraw-Hill)
- [3] Koyama T and Tachiki M 1996 *Phys. Rev. B* **54** 16183
- [4] Ryndyk D A 1998 *Phys. Rev. Lett.* **80** 3376
- [5] Shukrinov Yu M, Mahfouzi F and Seidel P 2006 *Physica C* **449** 62
- [6] Shukrinov Yu M, Mahfouzi F and Suzuki M 2008 *Phys. Rev. B* **78** 134521
- [7] Shukrinov Yu M, and Gaafar M A 2011 *Phys. Rev. B* **84**, 094514
- [8] Shukrinov Yu M, Mahfouzi F, Pedersen N F 2007 *Phys. Rev. B* **75** 104508
- [9] Shukrinov Yu M, Medvedeva S Yu, Botha A E, Kolahchi M R, and Irie A 2013 *Phys. Rev. B* **88** 214515
- [10] Shukrinov Yu M, Rahmonov I R and Gaafar M A 2012 *Phys. Rev. B* **86** 184502

Magnetic and Electronic studies on $\text{Cr}_{100-x}\text{Ir}_x$ alloy single crystals

P R Fernando, A R E Prinsloo and C J Sheppard

Department of Physics, University of Johannesburg, PO Box 524, Auckland Park, 2006

Author e-mail address: alettap@uj.ac.za

Abstract. An investigation of the physical properties of $\text{Cr}_{100-x}\text{Ir}_x$ alloy single crystals, with $x = 0.07, 0.15, 0.20$ and 0.25 , were previously used to establish the magnetic phase diagram of $\text{Cr}_{100-x}\text{Ir}_x$ alloy system around triple point concentration where the various magnetic phases co-exist. The present study extends these results by considering the temperature (T) dependence of the Seebeck coefficient (S), specific heat (C_p) and Hall coefficient (R_H) measurements, in addition to the electrical resistivity (ρ). Well defined anomalies were observed in the $\rho(T)$ and $S(T)$ curves of all the samples, except $x = 0.25$ alloy in $S(T)$ curve which is above 380 K ($T_N = 391.5$ K). The anomaly is the effect of an antiferromagnetic to paramagnetic phase transition on heating through the Néel temperature (T_N). Contrary to what is normally expected, it is noted that the anomaly related to T_N is more prominent in the $\rho(T)$ curves than in the $S(T)$ curves. R_H measurements carried out from 360 K down to 2 K in a constant magnetic field of 6 T, shows only weak anomalies at T_N for certain samples. The electronic Sommerfeld coefficient (γ), obtained from the low temperature C_p/T versus T^2 data, fits in well with the γ versus electron-to-atom (e/a) ratio curve previously published for certain Cr alloys.

1. Introduction

The electron and hole Fermi surfaces of pure Cr and its alloys nest on cooling through the Néel temperature (T_N), resulting in the formation of the spin-density-wave (SDW) when the antiferromagnetic (AFM) phase is entered. This nesting decreases the energy of the system through electron-hole pair condensation and results in the appearance of SDW energy gaps at the Fermi surface in certain directions in k -space [1].

The first parameter of importance in influencing the formation of the SDW is the effect of electron concentration on the area of the electron and hole Fermi surface sheets that nest [1]. The nesting area, and concomitantly the stability of the SDW state, depend on the electron concentration per atom (e/a) which can easily be tuned by alloying Cr ($e/a = 6$) with group-8 non-magnetic transition metals to increase or decrease the electron concentration [1].

Cr alloys with group-8 non-magnetic transition metals Ru, Os, Rh, Ir and Pt show large anomalies of magnetic origin at the phase transition temperatures [1]. The addition of very small amount of these impurities is enough to introduce phase changes in the magnetic state in the Cr alloys. Furthermore, the magnetic phase diagrams of the Cr alloys with group-8 non-magnetic transition metals contain the commensurate (C) SDW phase, the incommensurate (I) SDW phase, the transverse (T) incommensurate (I) SDW, longitudinal (L) ISDW phases as well as the paramagnetic (P) phase. The

magnetic phase diagrams of these alloys contain a triple point, where the ISDW, CSDW and P phases co-exist [1]. For impurity concentrations (x) below the triple point concentration (x_L), the LISDW, TISDW and P phases are observed, while for $x > x_L$ CSDW phase is observed. The spin-flip transition (T_{SF}) and the Néel transition (T_N) are known as first order phase transitions and the ISDW-CSDW phase transition (T_{IC}) are second order in nature [1, 2].

The Cr-Ir single crystals used in this study were also previously investigated using electrical resistivity, ultrasonic attenuation, thermal expansion and magnetoelastic measurements [3-5]. In order to extend these investigations and draw new correlations between the various physical properties, electrical resistivity (ρ), Seebeck coefficient (S), specific heat (C_p) and Hall coefficient (R_H) measurements were done in the present study.

2. Experimental

The $\text{Cr}_{100-x}\text{Ir}_x$ single crystals, with $x = 0.07, 0.15, 0.20$ and 0.25 , were prepared using the floating-zone technique [3-5] and previously investigated using different measurement techniques [3-5]. In the present study electrical resistivity $\rho(T)$ in the temperature (T) range $2 \text{ K} \leq T \leq 630 \text{ K}$, Seebeck coefficient $S(T)$, specific heat $C_p(T)$ measurements were done in the temperature range $2 \text{ K} \leq T \leq 380 \text{ K}$ and Hall coefficient (R_H) were carried out in the same temperature range as $C_p(T)$ and at a constant magnetic field of 6 T using a Quantum Design (QD) Physical Property Measurement System (PPMS). All these measurements were performed along the cubic [100] direction. The electrical resistivity measurements were done during heating and cooling.

3. Results and Discussion

Figure 1(a) shows electrical resistivity (ρ) as a function of temperature (T) in the range $2 \text{ K} \leq T \leq 630 \text{ K}$. Well defined anomalies in the form of clear minima followed by prominent domes are observed near Néel temperature (T_N) on cooling through the ISDW-P region. T_N and the size of the dome increase with increasing the Ir concentration, x . This anomalous behaviour is associated with the formation of the SDW on entering the antiferromagnetic phase on cooling through the Néel temperature (T_N) [1]. T_N is often defined for Cr and its dilute alloys as the temperature associated with the minimum in $d\rho/dT$ accompanying the magnetic phase transition [1]. This definition was also used in the present study and an example is shown in Figure 1(b). Present results are in good agreement with previous results of the electrical resistivity measurements on single and polycrystalline $\text{Cr}_{100-x}\text{Ir}_x$ alloys [5-10]. Thermal hysteresis was observed during heating (T_{IC}) and cooling (T_{CI}) in $\text{Cr}_{99.80}\text{Ir}_{0.20}$ single crystal, but the temperature range varies slightly from that previously observed [5].

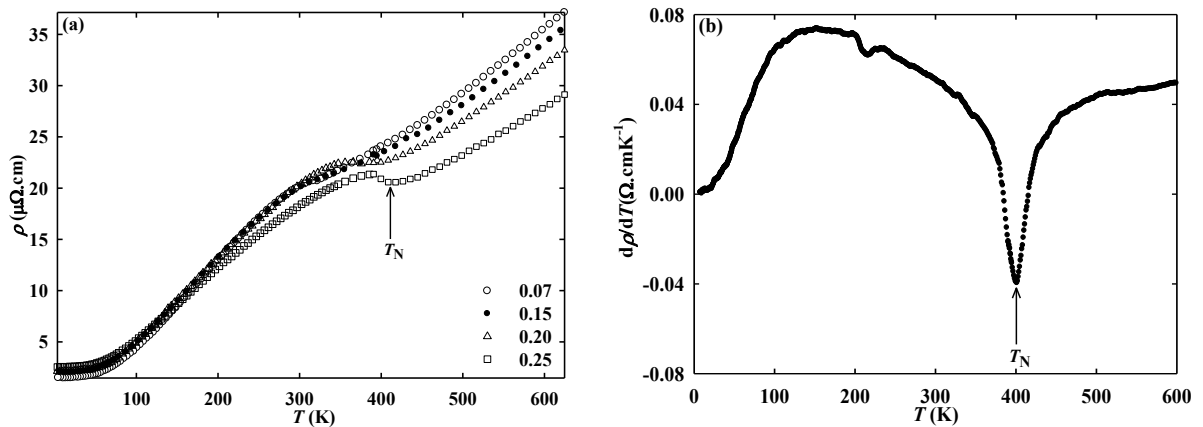


Figure 1. (a) Electrical resistivity (ρ) as a function of temperature (T) in the range $2 \text{ K} \leq T \leq 630 \text{ K}$ for the $\text{Cr}_{100-x}\text{Ir}_x$ single crystals. Symbols for concentrations are indicated in the legend. (b) $d\rho/dT$ as a function of temperature (T) for $x = 0.25$ sample. The minimum is used to indicate the Néel temperature, T_N .

The Seebeck coefficient (S) of $\text{Cr}_{100-x}\text{Ir}_x$ alloys were measured as a function of temperature (T) in the range $2 \text{ K} \leq T \leq 380 \text{ K}$. Figure 2 shows $S(T)$ curves for alloys, with $x = 0.07$ and 0.20 . The S versus T curve for $0.07 \text{ at.}\%$ Ir, shows a well-defined anomaly with prominent dome associated with T_N , around the same temperature region in which the electrical resistivity shows similar anomaly. T_N for $0.25 \text{ at.}\%$ Ir lies above the measured temperature range $2 \text{ K} \leq T \leq 380 \text{ K}$ and this cannot be seen in the curve. This T_N was calculated from the derivative of $dS(T)/dT$ accompanying the $S(T)$ magnetic anomaly, which is indicated by an arrow in Figure 2.

The value of T_N , as well as the magnitude of the dome in $S(T)$ increases with x and its position shifts to higher temperatures. The T_N values obtained from $\rho(T)$ and $S(T)$ in the present studies are approximately equal to the previous studies. The formation of dome in $\rho(T)$ and $S(T)$ are due to the formation of the SDW resulting in the formation of additional energy gaps in band structure of Cr-alloys [1, 5]. Using the experimental electrical resistivity and Seebeck coefficient results, the energy gap (ΔE_g), can be calculated using the equation, $E_g = \frac{\pi^2 k_B^2 T}{3e} \left(\frac{\rho}{S} \right)^2 \frac{1}{S}$ [11]. Here ρ is the total electrical resistivity, $\Delta\rho$ is the extrapolated resistivity, ΔS is the extrapolated Seebeck coefficient at temperature T . The quantity $\Delta\rho$ and ΔS are the difference between the maximum value of ρ and S in the antiferromagnetic region and the extrapolated resistivity and Seebeck coefficient value from the paramagnetic region at the same temperature. The Fermi energy, ΔE_g , was calculated for $x = 0.07$, which is equal to 0.1 meV .

The electronic Sommerfeld coefficient, γ , was determined for the single crystals by fitting the low temperature specific heat data to the equation $C_p/T = \gamma + \beta T^2$, where γ is the electronic Sommerfeld coefficient related to density of state, $n(E_F)$, and β is the lattice contribution related to Debye temperature θ_D . The experimental data points in the C_p/T versus T^2 curve was well fitted by least-squares linear fits, which is a straight line in the temperature range $2 \text{ K} \leq T^2 \leq 7 \text{ K}$ for all $\text{Cr}_{100-x}\text{Ir}_x$ single crystals. The inset in Figure 3 shows such a fit for the $x = 0.20$ sample as an example. The calculated gradients for the C_p/T versus T^2 curves for all the single crystals in the range $T^2 \leq 5.5 \text{ K}$ is approximately equal to the gradients obtained for $4 \leq T^2 \leq 400 \text{ K}$.

The values of γ for the $\text{Cr}_{100-x}\text{Ir}_x$ single crystals is compared with previously studied alloys, such as Cr-V (\blacktriangledown) [13] and (\blacktriangle) [14], Cr-Re (\square) [15, 16], Cr-Ru (\diamond) [13, 15] alloys in Figure 4. In general these curves show a gradual increase with increasing the e/a ratio of the alloys and started to level off at, $e/a = 6.35$. Line curves in Figure 4 were plotted by least-squares polynomial fit. The dashed line curves in Figure 4 plotted with the aid of the paramagnetic data points from Cr-V and Cr-Re.

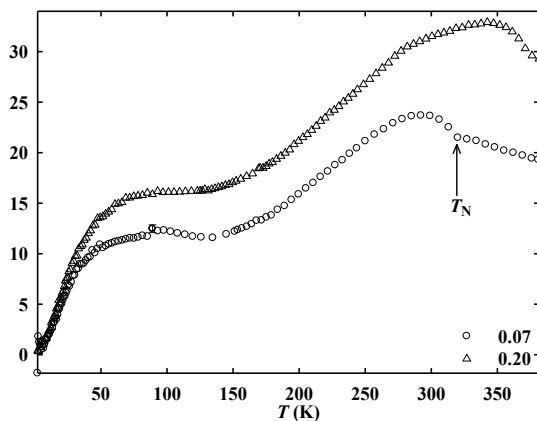


Figure 2. The Seebeck coefficient (S) for $\text{Cr}_{100-x}\text{Ir}_x$ alloys with $x = 0.07$ (\circ) and 0.20 (Δ) as a function of temperature (T) in the range $2 \text{ K} \leq T \leq 380 \text{ K}$.

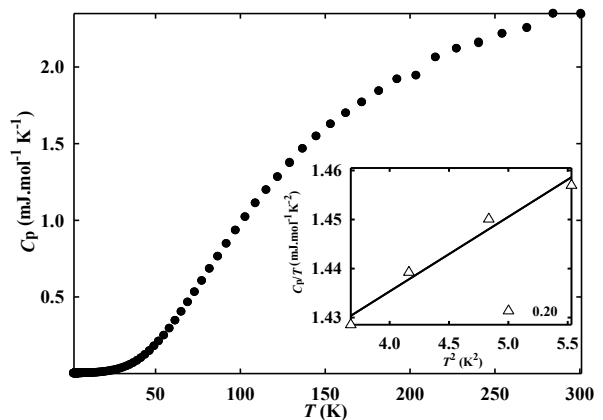


Figure 3. The graph of C_p versus T for $0.25 \text{ at.}\%$ Ir alloy in the temperature range $2 \text{ K} \leq T \leq 300 \text{ K}$. The inset shows C_p/T versus T^2 for $0.20 \text{ at.}\%$ Ir alloy, $T^2 \leq 5.5 \text{ K}$. The experimental data points were well fitted using least-squares linear fits.

The curve shows a smooth change in γ in the PM region for Cr-V, in the low region where $e/a \leq 5.95$ and for Cr-Re, in the high diluent region where $e/a \geq 6.25$. For all the AFM alloys all the experimental γ values are suppressed below those observed in the paramagnetic region. This effect is due to the overlapping of the electron and hole Fermi surface sheet and the reduction in the density of states at the Fermi level in the formation of SDW. The γ values of Cr-Ir single crystals are indicated by large open circles in Figure 4, which have approximately similar γ values at $e/a \approx 6.00$. These γ values lie on the dotted curve, which belongs to the alloys Cr-Ru and Cr-Re in the region where $e/a \geq 6.00$ as shown in Figure 4. Therefore, as expected, the Cr-Ir may have the similar trend of Cr-Ru and Cr-Re in high concentration (x), because these alloys are in the same group [1, 2].

Figure 5 illustrates the Hall coefficient of the $\text{Cr}_{100-x}\text{Ir}_x$ single crystals curves as a function of temperature (T), which shows that the value of $R_H(x)$ at 2 K decreases with increasing the concentration of Ir. $R_H(T)$ curves gradually increase when cooling down to 2 K from 380 K. The single crystal alloys both show a prominent upturn on cooling to $T < T_N$, and the magnitude of the upturn decreases with increasing, x . These effects may be due to an influence of progressive opening up of energy gap due to the SDW in the system. At $T \leq 50$ K the $R_H(T)$ curves levels off, this is attributed to the spin-flip transition (T_{SF}), where the (T) ISDW to (L) ISDW phase transform on cooling [17, 18]. Arrows shown in Figure 5 indicate T_N , T_{SF} and T_{IC} obtained from the work of Martynova *et al.* [3, 5]. The behavioural change in $R_H(T)$ decrease with increasing the concentration, x . This effect may be due to the ISDW to CSDW phase transition around triple point concentration [18] and/or decreasing the number of hole mobility by alloying as well as the domination of the impurity scattering [19].

4. Conclusion

The electrical resistivity $\rho(T)$, Seebeck coefficient $S(T)$, specific heat $C_p(T)$ and Hall measurements $R_H(T)$, of $\text{Cr}_{100-x}\text{Ir}_x$, with $x = 0.07, 0.15, 0.20$ and 0.25 , single crystals were investigated in the present study. Well defined magnetic anomalies in the form of a hump in the electrical resistivity, $\rho(T)$, curves, show similar trend, compared to previous studies [3-5]. T_N values obtained from $\rho(T)$ and $S(T)$ measurements compare well. However, there is no clear indication for a spin-flip (T_{SF}) transition, previously observed in ultrasonic velocity of sound studies in $x = 0.25$ crystals in the present study [3]. The γ versus e/a values obtained fits in well with that previously found for other Cr alloys with group-8 diluents [15]. This study can be further expanded as to including $S(T)$, $R_H(T)$ and magnetic susceptibility $\chi(T)$ measurements as a function of temperature for more single crystals for better comparison.

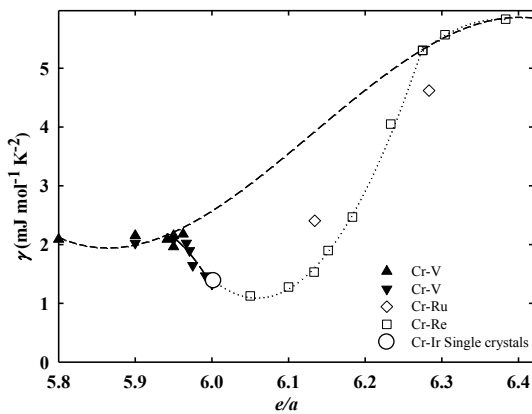


Figure 4. The behaviour of electronic Sommerfeld coefficient (γ) as a function of electron-atom (e/a) ratio for the Cr-V (\blacktriangledown [13] and \blacktriangle [14]), Cr-Re (\square) [16], Cr-Ru (\diamond) [13] poly crystals from previous studies redrawn from Heiniger *et al.*[15] and Cr-Ir single crystals (\circ) from this study. The line curves are guide to eye.

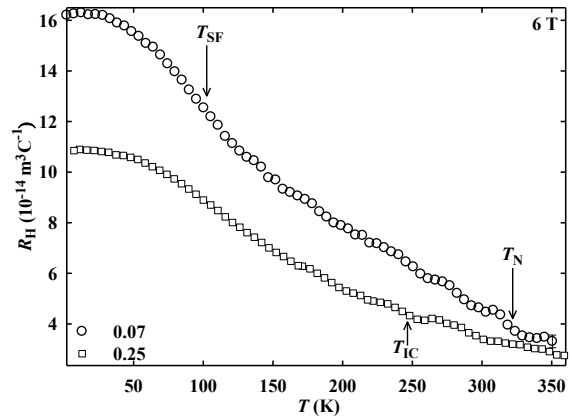


Figure 5. Hall coefficient (R_H) as a function of temperature (T) in the range $2 \text{ K} \leq T \leq 360 \text{ K}$ in a constant applied magnetic field 6 T. The arrows indicate the T_N , T_{SF} and T_{IC} , obtained from the work of Martynova *et al.* [3, 5].

Acknowledgements

This work was supported by the NRF of South Africa under grant numbers 78832 and 61388.

References

- [1] Fawcett E, Alberts H L, Galkin V Y, Noakes D R and Yakhmi J V 1994 *Rev. Mod. Phys.* **66** 25.
- [2] Fawcett E 1988 *Rev. Mod. Phys.* **60** 209.
- [3] Martynova J, Alberts H L and Smit P 1998 *J. Magn. Magn. Mat.* **187** 345.
- [4] Martynova J, Alberts H L and Smit P 1998 *J.Phys.:Condens.Matter* **10** 1173.
- [5] Martynova J, Alberts H L and Smit P 1996 *J.Phys.:Condens.Matter* **8** 10473.
- [6] Fukamichi K and Siato H 1975 *J. Less Common Met.* **40** 357.
- [7] Gopalakrishnan I K, Yakhmi J V and Iyer R M 1984 *J. Magn. Magn. Mat.* **46** 207.
- [8] Arajs S, Rao K V, Åström H U and De-Young F.Tice 1973 *Physica Scripta.* **8** 109.
- [9] Yakhmi J V, Gopalakrishnan I K and Iyer R M 1983 *J. Less Common Met.* **91** 327.
- [10] De Young F.Tice, Arajs S and Anderson E E 1971 *Amer. Inst. Phys. Conf. Proc. Magn and Magn. Mat.* **5** 517.
- [11] Schröder K and Harry T 1968 *Phys. Kondens. Materie* **7** 318.
- [12] Takeuchi J, Sasakura H and Masuda Y 1980 *J. Phys. Soc. Jpn.* **49** 508.
- [13] Heiniger F 1966 *Phys. Kondens. Materie* **5** 285.
- [14] Cheng C H, Wei C T and Beck P A 1960 *Phys. Rev.* **120(2)** 426.
- [15] Heiniger F, Bucher E and Müller J 1966 *Phys. Kondens. Materie* **5** 243.
- [16] Muheim J and Müller J 1964 *Phys. Kondens. Materie* **2** 377.
- [17] Furuya Y 1976 *J. Phys. Soc. Jpn.* **40** 490.
- [18] Sheppard C J, Prinsloo A R E, Alberts H L, Muchono B and Strydom A M 2014 *J. Alloys and Compounds.* **595** 164.
- [19] Shabel B S and Schröder K 1967 *J. Phys. Chem. Solids* **28** 2169.

Combustion Synthesis of Dy³⁺–doped YVO₄ phosphor

KE Foka¹, BF Dejene¹ and HC Swart²

¹Department of Physics, University of the Free State (Qwaqwa), Private Bag x 13, Phuthaditjhaba, 9866, RSA

²Department of Physics, University of the Free State, P.O. Box 339, Bloemfontein, 9300, RSA

fokake@qwa.ufs.ac.za

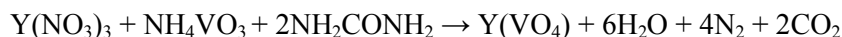
Abstract. Dy doped YVO₄:Dy³⁺ phosphors were produced by the combustion method at 600°C. The structure and optical properties of the powders were investigated using X-ray diffraction (XRD), scanning electron microscopy (SEM), Fourier transform infrared spectroscopy (FTIR) and photoluminescence (PL). The XRD patterns showed the tetragonal phase similar to the standard JCPD file (17-0341). SEM shows that the particle sizes were small and agglomerated, and the size increased with the Dy³⁺ dopant concentration and its shape changed to bulk-like particles. In PL, the emission spectra exhibited a weak band at 663 nm for the ⁴F_{9/2} - ⁶H_{11/2} transition and a peak at 483 nm (blue) for the ⁴F_{9/2} - ⁶H_{15/2} transition and a 574 nm (yellow) peak with higher intensity for the ⁴F_{9/2} - ⁶H_{13/2} transition.

1. Introduction

Yttrium orthovanadate (YVO₄) belongs to the space group ¹⁹D_{4h} [1] and it is an important optical material that has been given considerable attention due to its outstanding characteristics, such as excellent thermal, mechanical and optical properties. In recent years, significant progress has been made on luminescent materials using YVO₄ as host lattice and its emission colors are based mainly on the f–f transition [2]. Eu³⁺ doped YVO₄ nanocrystals have been widely used as an important commercial red phosphor. Their applications include color television, cathode ray tube, fluorescent lamps and plasma display panels. Yttrium orthovanadate activated by trivalent dysprosium (YVO₄:Dy) is a well-known phosphor material with high efficiency. The emission color of the luminescence is close to white because of the yellow (⁴F_{9/2} → ⁶H_{13/2}) and blue (⁴F_{9/2} → ⁶H_{15/2}) emissions of Dy³⁺. There are two basic approaches to generate white light from light emitting diodes (LEDs). One is by mixing light of different colours emitted by several chips called multichip LEDs and the other is to convert the light emitted from a blue or ultraviolet (UV) LED to a longer wavelength light using phosphors, which are called phosphor-converted (pc) – LEDs [3]. In order to produce a phosphor that will produce white light for the LED applications Dy³⁺–doped YVO₄ phosphors were produced by a combustion method at 600°C. There are many methods to prepare YVO₄ phosphors, such as the chemical co-precipitation method [4], sol–gel method [5] and hydrothermal method [6]. Combustion has for a long time been the major source of energy for heating, transportation and production of electricity [7]. The combustion synthesis is very good because it can provide a product without sintering. In this paper, we report on the synthesis of YVO₄ phosphor material doped with different Dy³⁺ concentrations by combustion method. The structure and the luminescence properties of the YVO₄: Dy³⁺ is discussed.

2. Experimental

YVO_4 : Dy^{3+} was prepared by the combustion method. The starting materials were yttrium nitrate $\text{Y}(\text{NO}_3)_3$, ammonium metavanadate (NH_4VO_3), urea (NH_2CONH_2) and dysprosium nitrate $\text{Dy}(\text{NO}_3)_3$. The chemical reaction is as follows:



All the ingredients were mixed according to the stoichiometric ratio in an agate mortar and a pasty solution was formed. The solution was transferred to a crucible and then kept in a furnace maintained at a temperature of 600°C . A combustion process started in a few minutes and a flame was observed. The formation of a foamy powder was observed and a pale yellow powder was obtained. The powder was then characterized by X-ray diffraction (XRD) patterns using a D8 advanced AXS GmbH X-ray diffractometer. The types of chemical bonds were investigated using a Bruker TENSOR 27 Series Fourier transform infrared spectroscopy (FTIR). The morphology was examined by scanning electron microscopy (SEM). The photoluminescence (PL) properties of the phosphor (emission and excitation) were measured using a Carry Eclipse PL spectrophotometer with a 150W xenon lamp as the excitation source.

3. Results and discussion

3.1. XRD study

Figure 1 shows the XRD patterns of YVO_4 phosphor material doped with different concentration of Dy. The XRD patterns show the tetragonal phase and agreed well with standard JCPD file (17-0341) [8]. It is clear from the XRD patterns that no impurity phases were present which shows that all reactants have reacted completely. The average crystallite size determined from the broadening of the peaks (200), (112) and (312) using Scherrer formula was about 31 nm. The calculated lattice parameters were found to be $a=0.711$ nm and $b=0.628$ nm.

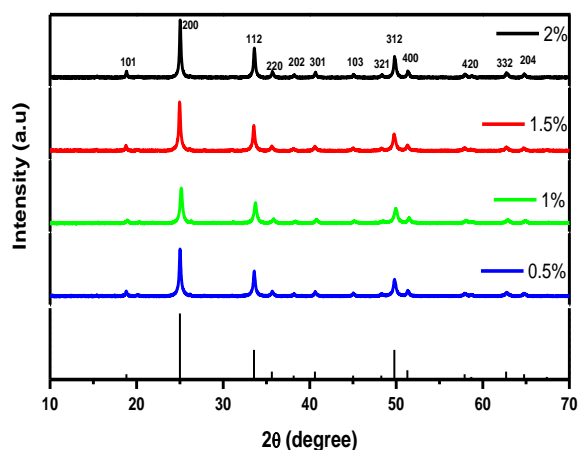


Figure 1. XRD pattern for the YVO_4 phosphors –doped with different concentration of Dy^{3+} as well as the standard JCPD file (17-0341).

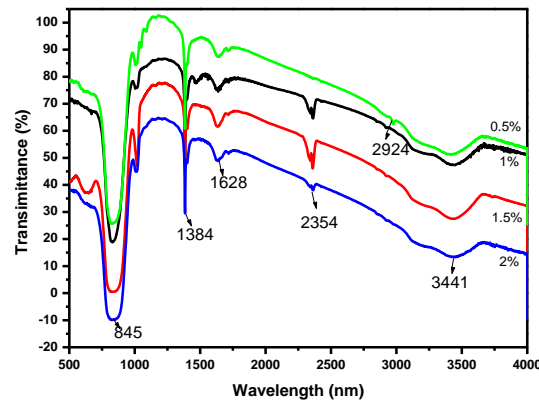


Figure 2. FTIR spectra of YVO₄:Dy³⁺.

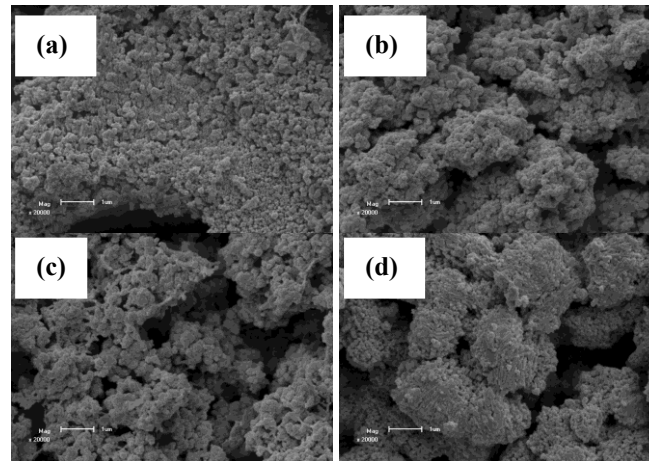


Figure 3. SEM image of the YVO₄:Dy³⁺ doped with (a) 0.5, (b) 1, (c) 1.5 and (d) 2 mol% Dy³⁺ ions

3.2. FTIR study

Figure 2 shows the FTIR spectra prepared for the combustion method at different concentrations of Dy³⁺. The most intense band shows the characteristics of the V-O bonds with strong band ranges from 776-921 cm⁻¹ consisted of a peak at 845 cm⁻¹. A broad band centered at 3441 cm⁻¹ is assigned to the symmetric stretching vibration of H-O-H (H₂O molecules). The strong peak observed at 1384 cm⁻¹ is ascribed to the asymmetrical vibration of the carboxyl groups. It can be seen that the position of the V-O stretching vibration remained the same for all the samples. The two weak bands at 2354 cm⁻¹ and 2924 cm⁻¹ correspond to the symmetrical and asymmetrical stretching vibration modes of the CH₂ group, respectively. The low 0.5% Dy concentration has some influence on symmetrical vibration modes of CH₂ since the band on 2.54 cm⁻¹ is very weak.

3.3. SEM study

SEM has been carried out to study the morphological structure of the YVO₄:Dy³⁺ prepared by the combustion method. The SEM micrographs that are shown in figure 3 show the YVO₄ doped with (a) 0.5, (b) 1, (c) 1.5 and (d) 2 mol% Dy³⁺ ions. In a lower Dy³⁺ concentration the particle size were small and much agglomerated. The particle size seems to increase along with the concentration and its shape changed to bulk-like particles. The surface of the morphology was rough. Table 1 shows the calculated average grain size of the YVO₄ doped with the various concentration of Dy³⁺. The grain size decreased

to 26 nm for 1 mol% of Dy^{3+} , and then increased as the Dy^{3+} concentration increased. When doped with 2 mol% Dy^{3+} , the grain size was found to be 35 nm, which is also indicated in the increase in the intensity of the diffraction peaks, as observed in the XRD analysis.

Table 1: The average crystallite size for the YVO_4 doped with different concentration of Dy calculated by Scherer equation

Dy^{3+} concentration (mol%)	Average grain size (nm)	Experimental error (\pm)
0.5	30	± 4.5
1	26	± 3.7
1.5	30	± 1.2
2	35	± 2.6

3.4. PL study

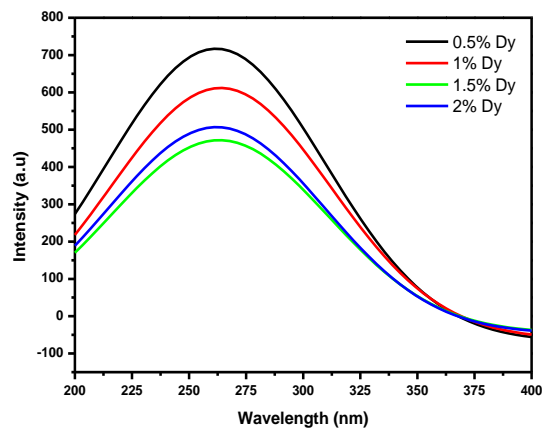


Figure 4. Excitation spectra of $\text{YVO}_4:\text{Dy}^{3+}$

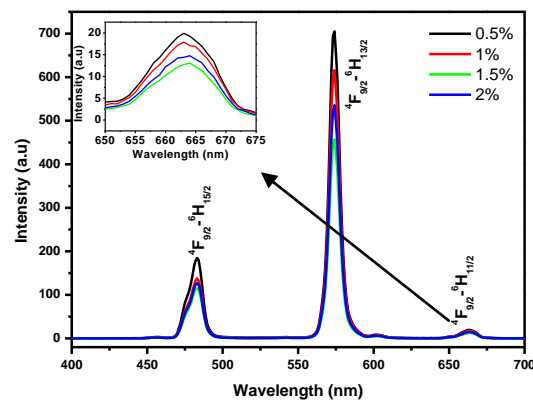


Figure 5. PL emission spectra of $\text{YVO}_4:\text{Dy}^{3+}$.

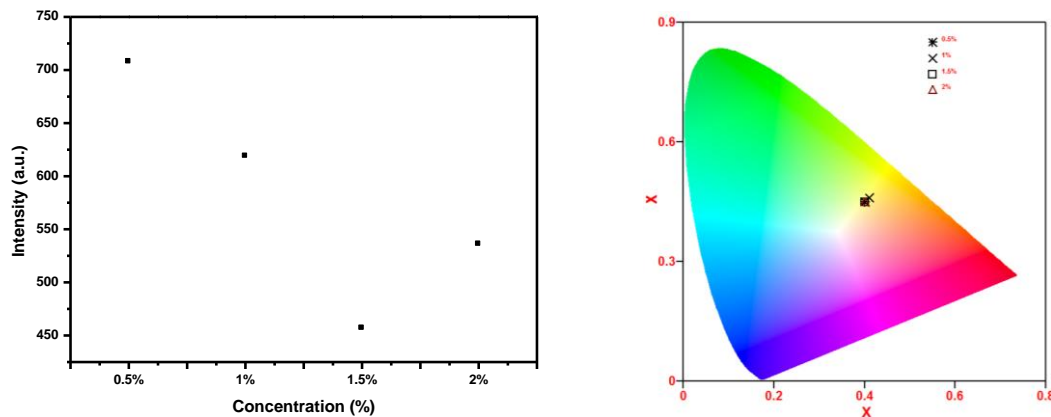


Figure 6. (a) 574 nm PL peak intensity vs concentration graph of the $\text{YVO}_4:\text{Dy}^{3+}$ phosphor, (b) CIE of $\text{YVO}_4:\text{Dy}$.

The PL excitation spectrum of $\text{YVO}_4:\text{Dy}^{3+}$ phosphor monitored at an emission wavelength of 574 nm is shown in figure 4. The excitation spectra showed a broad peak around 257 nm. Figure 5 shows the emission spectra of the $\text{YVO}_4:\text{Dy}^{3+}$ prepared at different concentration of Dy^{3+} . The insert shows the enlargement of the 663 nm peak. The spectra show the characteristics of the emission peaks of Dy^{3+} . The peaks are located at 483 nm (blue) and 574 nm (yellow). The peaks are related to the hypersensitive transition $^4\text{F}_{9/2}-^6\text{H}_{15/2}$ and the $^4\text{F}_{9/2}-^6\text{H}_{13/2}$, respectively. There is a very weak emission band located at 663 nm which correspond to the $^4\text{F}_{9/2}-^6\text{H}_{11/2}$ transition. The intensity of the yellow emission is stronger than that of the blue emission, because Dy^{3+} is located at a site of the D_{2d} , which is deviated from an inverse center in the YVO_4 host [9]. The maximum PL intensity was observed for the lower concentration of 0.5%, figure 6 (a). The PL peak intensity showed a decrease in intensity as the concentration of Dy^{3+} increased due to concentration quenching and then the intensity increased again when the concentration reached 2% as shown in figure 6. The reason is when the concentration is far above the critical concentration the intensity of the Dy^{3+} emission will increase quickly as the concentration increase until it reaches another critical concentration, and then the intensity of Dy^{3+} emission will decrease again for the energy transfer among different Dy^{3+} ions [10].

3.5. CIE chromaticity diagram

The emission color of the $\text{YVO}_4:\text{Dy}^{3+}$ phosphor can be expressed by Commission Internationale de l'Éclairage (CIE) coordinates as indicated in figure 6(b). The chromaticity coordinate for the $\text{YVO}_4:\text{Dy}^{3+}$ with different Dy^{3+} concentrations are very similar. It is clear that the $\text{YVO}_4:\text{Dy}^{3+}$ emitted a more yellowish colour light. The blue part helped to shift it slightly towards the white part of the spectrum. The phosphor can be used as a yellow emitted material in LEDs.

3.6. Lifetime study

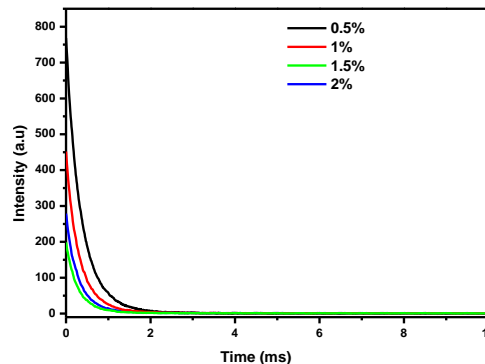


Figure 7. Decay curves of $\text{YVO}_4:\text{Dy}^{3+}$ phosphors with different concentration of Dy^{3+} .

Figure 7 shows the decay curves for the different concentration of Dy^{3+} doped YVO_4 . The logarithmic plots of the curves were fitted by a two order exponential decay equation:

$$I = A_1 \exp(-t/\tau_1) + A_2 \exp(-t/\tau_2).$$

Where I represents the phosphorescent intensity, τ_1 and τ_2 correspond to shorter and longer lifetime constants, and A_1 and A_2 are constants. The parameters for the fitting data are listed in table 2. The phosphorescence decay times as shown in table 2 can be explained by an initial rapid decay and then an intermediate transitional. A decrease in the intensity with an increase in concentration is due to the improvement of the crystallinity as the particle size increased with the concentration [11].

Table 2: Lifetime values obtained after bi-exponential fitting to the data of $\text{YVO}_4:\text{Dy}^{3+}$.

%Dy	τ_1 (ms)	τ_2 (ms)	A_1	A_2
0.5	0.27	0.56	273	493
1	0.26	0.89	20	177
1.5	0.26	0.65	27	267
2	0.28	0.64	47	150

4. Conclusion

$\text{YVO}_4:\text{Dy}^{3+}$ phosphors prepared with different concentration of Dy^{3+} have been synthesized by the combustion method at 600°C . The XRD characteristics confirmed the crystalline single phase present in the sample. In SEM images the particle size seemed to increase along with the concentration of Dy^{3+} . The emission colour of the luminescence is yellow ($^4\text{F}_{9/2} \rightarrow ^6\text{H}_{13/2}$) and approaching the white region due to the blue ($^4\text{F}_{9/2} \rightarrow ^6\text{H}_{15/2}$) emissions of Dy^{3+} and has the potential to be used as a phosphor for pc-LEDs. The intensity of the PL was decreasing while increasing the concentration due to concentration quenching.

Acknowledgements

The author is grateful for the financial support by the National Research Foundation (NRF) and University of the Free State.

References

- [1] Hao W, Huifang X, Qiang S, Tianhu C and Mingmei W 2003 J. Mater. Chem., **13** 1223–1228.
- [2] Deming L, Lizhu T, Jianhui S and Hua Y 2012 Journal of Alloys and Compounds, **512** 361–365.
- [3] Young H, Jong P S K, Ji K, Jong K and Young D 2004 Bull. Korean Chem. Soc. **25** No. 10 1585-1588.
- [4] Wang J, Xu Y, Mirabbos H, Cu Y, Liu H and G 2009 Journal of Alloys and Compounds, **24** 772–776.
- [5] Yu M, Lin J, Wang Z, Fu J, Wang S, Zhang H J and Han Y C 2002 journal Chem. Mater. **14** 2224-2231.
- [6] Liang Y, Chui P, Sun X, Zhao Y, Cheng F and Su K 2013 Journal of Alloys and Compounds, **552** 289–293.
- [7] Aldén M, Alaa O and Mattias R 2011 Gustaf Särner Progress in Energy and Combustion Science, **37** 422-461.
- [8] Zhang H, Fu X, Niu S, Xin Q 2008 J. Of alloys and Compounds, **457** 61-65.

- [9] Yang K, Zheng F, Wu R Li H and Zhang X 2006 journal of Rare earths, **24** 162-166.
- [10] Chen L, Chen K, Lin C, Chu C, Hu S, Lee M and Liu R 2010 J.Comb.Chem, **12** 587-594.
- [11] Meitram N L, Raghumani S N, Sri K S and Rajesh K V 2011 Journal Mater.Chem. **21** 5326-5337.

Pressure effects on the magnetic-electronic behavior of the local moment ferromagnet CeCuSi

G R Hearne¹, G Diguët¹, A M Strydom¹, B Sondezi-Mhlungu¹, K Kamenev²,
F Baudalet³, L Nataf³

¹ Department of Physics, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, Johannesburg, South Africa

² Centre for Science at Extreme Conditions (CSEC) and School of Engineering, University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom

³ Synchrotron SOLEIL, Saint-Aubin, 91192 Gif-sur-Yvette Cedex, France

E-mail : grhearne@uj.ac.za

Abstract. Magnetization measurements to ~10 GPa have been used to monitor both T_C and magnetic susceptibility in CeCuSi. The Ce valence and magnetic ($5d$ moment) state to ~16 GPa at 6 K has been probed by means of L_3 -absorption core-hole spectroscopy; in particular by X-ray absorption near edge spectroscopy (XANES) profiles and X-ray absorption dichroism (XMCD), respectively. Increasing pressure up to 10 GPa increases T_C from 15 K to 30 K. By contrast there is a decreasing XMCD intensity, signalling a collapsing Ce ordered moment. The XANES profiles show that the original $|4f^1\rangle$ white line feature, typical of the localized $4f$ moment, diminishes at the expense of an emerging $|4f^0\rangle$ (electron delocalization) component. The volume dependence of the hybridization between the $4f$ and d conduction-band states enhances the indirect exchange coupling between Ce ions and hence enhances T_C . However there is an attendant $4f$ level broadening resulting in a valence instability, which is disruptive to the stable configuration local moment situation prevailing at low pressure.

1. Introduction

Interest in Ce or U based ternary intermetallics has been ongoing for the last three decades, because they show a variety of exotic magnetic-electronic ground states (e.g., heavy fermion behavior, non-Fermi liquid effects, valence fluctuations, etc). Among the most intensively studied is the equiatomic CeTX (T is a transition-metal, X is a p element) type compounds. Most magnetic CeTX compounds order antiferromagnetically. One candidate, CeCuSi, is among a select number that rather exhibits ferromagnetic ordering at low temperatures; other candidates being CePdX ($X = P, As, Sb$) [1-3]. CeCuSi crystallizes in an ordered hexagonal ZrBeSi-type structure (space group $P6_3/mmc$, number 194) [3, 4]. The ferromagnetic transition in this compound has been established from both C_p specific heat data in which a λ -type anomaly is manifested at $T_C = 15$ K and magnetization measurements in which an ordered moment of $\sim 1 \mu_B$ and stable configuration Ce^{3+} state has been obtained [3, 5, 6].

The hybridization between the localized $4f$ and itinerant d electrons (Ce $5d$ and T $3d$) is readily tuned (to increase) under pressure [7-9]. Consequently new ground states can be stabilized at reduced inter-atomic spacing without the complexity of disorder from doping. Many well known antiferromagnetic CeTX compounds have been the focus of attention in pressure studies in the last

decade. There has been much less done, if any, in elucidating the pressure response of *ferromagnetic* analogs. This work reports on the magnetic-electronic pressure response of CeCuSi, using novel high pressure capabilities in diamond anvil high pressure cells (DACs).

SQUID magnetization pressure measurements of CeCuSi have been used to monitor the pressure evolution of both T_C and the susceptibility in the magnetically ordered state at $T < 15$ K. X-ray absorption spectroscopy (XAS) at the Ce L_3 -edge (involving $2p \rightarrow 5d$ electronic transitions) is feasible under these stringent extreme conditions using specialized DAC methodology [10]. Therefore element specific X-ray absorption near edge spectroscopy (XANES) and X-ray magnetic circular dichroism (XMCD) measurements from these XANES spectra has been used to probe the Ce valence and its $5d$ moment (exchange coupled to the $4f$ moment), respectively; as implemented up to ~ 16 GPa at liquid helium temperatures.

2. Experimental

2.1. SQUID-magnetization pressure measurements

A miniature turn-buckle magnetic DAC (TM-DAC) [11] recently developed for magnetization studies in a small bore SQUID magnetometer (QD : MPMS), has been used for the magnetic pressure studies. This TM-DAC comprised of Cu-Be, with maximal pressure capabilities of 10 - 12 GPa, has sufficiently low background which is smoothly varying over a wide temperature range. Hence this can be readily subtracted from the measured DAC plus sample signal, to reliably reveal the magnetization details of the sample only. Diamond anvils with 800 μm culets have been used in the TM-DAC. The Cu-Be gasket has been pre-indented to a thickness of ~ 100 μm from a starting thickness of 250 μm and a cavity of ~ 400 μm diameter drilled in the center of the indentation.

Well characterized homogeneous and single-phased polycrystalline CeCuSi from the same batch as previous studies has been used for these pressure experiments [4, 5]. A microscopic solid fragment of several μg of the CeCuSi sample was loaded into the cavity of the TM-DAC to ensure $\sim 75\%$ filling factor. Tiny ruby chips were loaded alongside the sample for pressure determination from the ruby fluorescence [12]. Transparent Daphne 7373 mineral oil, sufficiently hydrostatic to several GPa has been used as pressure transmitting medium. The measurement protocol involved cooling the DAC in zero field to about 10 - 20 K above T_C , after which a 30 Oe field was applied and the magnetization (M) measured upon cooling the sample through the transition temperature region. At initial DAC closure the background corrected data revealed the same features of the transition as for the sample measured under ambient conditions outside of the DAC, confirming $T_C = 15$ K, see figure 1(a).

After unloading the DAC from the SQUID, pressure was increased and measured at room-temperature before commencing with the next M - T run at pressure. The pressures obtained from the ruby fragment loaded near the center of the cavity was typically within 5% of that of the ruby near the edge. Moreover, previous tests have indicated that when the TM-DAC is cooled to 2 K, there is a minimal change in pressure from the value locked in at room temperature [11].

2.2. XAS measurements

X-ray absorption measurements at the Ce L_3 -edge (~ 5.7 keV) were performed at the energy-dispersive ODE (Optique Dispersive EXAFS) beamline of Synchrotron SOLEIL. An external field of 1.3 T has been applied along the beam path to impose a magnetization direction (i.e., orientations of majority and minority electron spin bands in the ferromagnet). The magnetic field is switched parallel and anti-parallel to the beam direction which is fully equivalent to changing the helicity of the photons incident on the sample. The dichroic signal (XMCD) is obtained from the difference between the (near edge) XANES spectra for two opposite directions of the external magnetic field.

The absorption of these incident photons at $L_{2,3}$ edges result in corresponding spin-polarized photoelectrons excited from $2p$ core levels, for “detecting” (i.e., populating) majority and minority

spin holes $N_h(m_s)$ in the Ce $5d$ conduction band above the Fermi level E_F ($2p \rightarrow 5d$ electric dipole transitions), where m_s is the electron spin quantum number [13].

Thus electron spin polarized X-ray absorption spectra (or spin dependent absorption coefficients) μ_+ and μ_- are recorded, from which the dichroism signal may be obtained (i.e., $\mu_+ - \mu_-$). A normalized spin-dependent absorption profile $(\mu_+ - \mu_-)/(\mu_+ + \mu_-)$ is computed from the absorption spectra so as to obtain the thickness-independent XMCD for comparison at various pressures.

The beam spot at the sample was $28 \times 42 \mu\text{m}$ (FWHM). Energy calibration was by means of a CeFe_2 reference sample [14]. Pressurization involved an in-house membrane-DAC assembly cooled to $\sim 6 \text{ K}$ in a closed-cycle cryogenic system and pressure adjusted “on-line” at the desired temperature. Perforated diamond anvil methodology was employed; imperative for the high gains in transmission desirable at the $\sim 5.7 \text{ keV}$ L_3 -edge where there is considerable absorption ($> 95 \%$) by conventional diamond anvils. This arrangement entailed mounting miniature diamond anvils of $500 \mu\text{m}$ thickness on bigger diamonds serving as backing plate supports through which holes had been laser drilled.

Powdered sample was loaded into the confining cavity (diam. $\sim 80 \mu\text{m}$) drilled in an inconel gasket pre-indented to a final thickness of $\sim 20 \mu\text{m}$. Silicone oil was loaded as an appropriate pressure transmitting medium. Pressure values were determined using the ruby fluorescence method, from ruby balls loaded in the cavity. Pressure was initially set at room temperature and after the DAC had been cooled to $\sim 6 \text{ K}$ the pressure was again checked in-situ in the beamline for drift.

3. Results and discussion

3.1 Magnetization measurements

Figure 1(a) shows that the precipitous step change in the magnetization typically occurring at the spin alignment transition (T_C) diminishes upon rising pressure, until such transition signatures are considerably reduced at $\sim 10 \text{ GPa}$. The T_C (midpoint) increases monotonically from 15 K at ambient pressure up to $\sim 30 \text{ K}$ at 10 GPa (figure 1(b)). This behavior is nearly replicated by T_C (onset) and shows that spin correlations are onset well above 30 K at the highest pressures. The magnetic susceptibility of the spin aligned state (low field magnetization value at $\sim 9 \text{ K}$, figure 1(c)) decreases non-monotonically and extrapolates to the paramagnetic baseline in the range $10 - 15 \text{ GPa}$. This contrasting behavior of increasing T_C yet collapsing magnetic susceptibility as a function of pressure merits deeper investigation, e.g., of how the pertinent Ce electronic state is responding under pressure at $T \ll T_C$. Therefore element specific XANES and its dichroism (XMCD) at the Ce L_3 -edge under these stringent extreme conditions is considered appropriate (and is perhaps the only means) to directly monitor the Ce valence and probe the Ce moment, respectively.

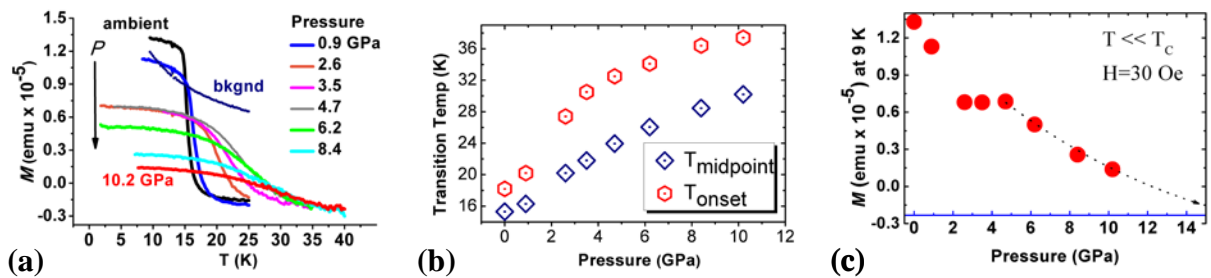


Figure 1 (color online): (a) Background subtracted magnetization-temperature curves as a function of increasing pressure. (b) Pressure evolution of T_C and (c) low field magnetization (susceptibility) at 9 K .

3.2 XANES and XMCD results

L_3 -absorption is a core-hole spectroscopy based on the excitation of $2p$ core electrons into $5d$ valence orbitals. Sensitivity to the $4f$ electrons arises through d - f hybridization and the intra-atomic $4f$ - $5d$ magnetic exchange interaction. The different components of the Ce electronic ground state, i.e., $|4f^0\rangle$ (itinerant), $|4f^1\rangle$ (localized) and $|4f^2\rangle$ (strong hybridization with conduction electrons), make different contributions in the XANES spectrum. This is because the energy attributed to each $|4f^i\rangle$ final state configuration in the $2p \rightarrow 5d$ transition is different, due to their respective screening effects on the $2p$ core hole. For example, in the final state either the core hole is well screened by a $4f$ electron ($|4f^1\rangle$ localized configuration) yielding the low-energy structure at the L_3 edge (so called “white line”) or is poorly screened by the $5d$ conduction band electron ($|4f^0\rangle$ itinerant configuration), the latter giving rise to a white line feature at higher energies (6 – 10 eV) from the $|4f^1\rangle$ peak. The effect of a $|4f^2\rangle$ configuration is more difficult to discern and usually occurs as a low intensity peak in the pre-edge region, where it is masked by core-hole lifetime broadening effects of the $|4f^1\rangle$ main white line [10]. Relative intensities of these spectral features may be used to establish the ground state admixture $c_0|4f^0\rangle + c_1|4f^1\rangle$ [10, 14], where $(c_i)^2$ represents the weight of the individual electronic configurations. The fractional $4f$ occupancy number n_f may be obtained from relative intensities of the $4f^0$ and $4f^1$ contributions in the XANES profile, $n_f = (c_1)^2 / [(c_1)^2 + (c_0)^2]$ [14]. Such an estimation sets an upper limit on n_f , as the contribution from $|4f^2\rangle$ is not accessible and therefore $(c_2)^2$ is not included. The case of $n_f < 1$ (non-integral $4f$ occupancy) signifies a mixed valence or valence fluctuating state of Ce in the compound, that is, electron delocalization.

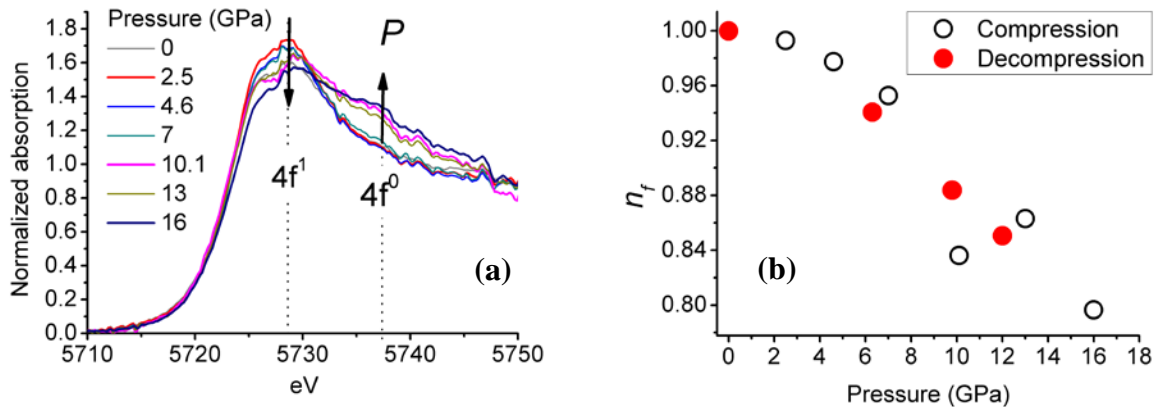


Figure 2 (color online) : (a) Pressure evolution of the XANES profiles at 6 K, upon compression. The positions of the $|4f^0\rangle$ and $|4f^1\rangle$ features have been delineated. (b) The $4f$ level occupancy n_f as estimated from $|4f^0\rangle$ and $|4f^1\rangle$ white lines in the XANES, both for compression and decompression sequences. This is normalized to the assumed full occupancy, Ce^{3+} , in the stable configuration state at ambient pressure.

Figure 2(a) depicts the evolution of the XANES profiles as a function of pressure. These have been normalized to an edge jump of unity. Near ambient pressure the spectrum exhibits a pronounced single white line feature, attributed to the $|4f^1\rangle$ final state configuration anticipated for this local moment ferromagnet. As pressure increases this white line intensity decreases and an additional white line intensity at about 10 eV higher in energy evolves. This hump feature, characteristic of the $|4f^0\rangle$ configuration [10], grows as a function of pressure at the expense of the $|4f^1\rangle$ intensity. The transfer of spectral weight from the $|4f^1\rangle$ to $|4f^0\rangle$ component, conspicuously evident at $P \geq 10$ GPa in figure 2(a), provides here experimental evidence of the $4f$ electron delocalization under pressure. This is representative of a progressive valence transition from Ce^{3+} at ambient pressure [6], towards Ce^{4+} .

The obvious decrease in $4f$ level occupancy plotted in figure 2(b) also exemplifies this electron delocalization. Furthermore the electronic transition is reversible with the stable-configuration state being re-attained upon decompression back to ambient pressure.

The valence instability impacts on the magnetic moment of the Ce ion. This is monitored by the XMCD, obtained from the difference between the XANES profiles associated with right and left circularly polarized incident photons as described in section 2.2. The imbalance of $5d$ unoccupied states in the majority and minority spin channels are probed. The XMCD thus probes the ordered moment (spin and orbital) contribution derived from the $5d$ exchange split majority and minority spin bands [13]. This exchange splitting (relative shift of $5d$ spin-up and down bands) has its origin in the *intra-atomic exchange* field from (i.e., magnetic coupling to) the ordered Ce $4f$ moment [15], as derived from the electronic ground state $c_0|4f^0\rangle + c_1|4f^1\rangle$. Monitoring the L_3 -edge XMCD and associated $5d$ moment is therefore a means of tracking the behavior of the ordered Ce $4f$ moment. Figure 3 shows that there is a monotonically decreasing XMCD intensity as pressure increases, indicative of a collapsing Ce $5d$ ordered moment contribution.

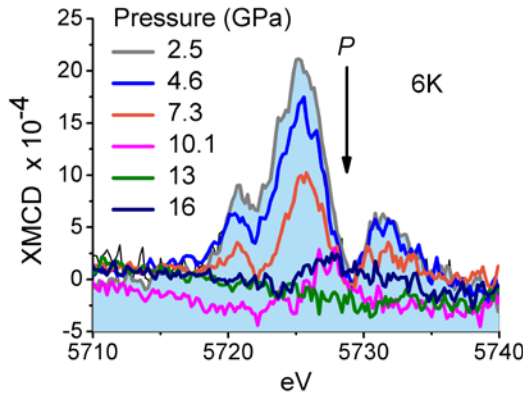


Figure 3 (color online) : XMCD signal (normalized to the absorption edge jump) upon increasing pressure. Signal intensity is proportional to the Ce $5d$ moment contribution. Collapse of the signal is evident at $P \geq 10$ GPa. The intensity is reconstituted upon decompression to ambient pressure, i.e., reversibility obtains as in figure 2(b).

3.3 Discussion

The $4f$ electrons are highly localized and Ce inter-site distances are comparatively much larger. Magnetic coupling between Ce local moments, responsible for spin alignment and the T_C energy scale, is through an *indirect exchange* mechanism involving polarization of the d conduction band electrons; the Ruderman-Kittel-Kasuya-Yosida (RKKY) interaction. The RKKY exchange coupling strength J_{d-f} depends on the hybridization between the (Ce) localized $4f$ electron states with those d band (Cu $3d$ and Ce $5d$) electronic states of comparable energy [9]. The dependence of T_C on J_{d-f} in stable configuration local moment Ce compounds is expected to behave as $T_C \propto (J_{d-f})^2 \times N(E_F)$, where $N(E_F)$ is the density of states at the Fermi energy. The application of pressure generally leads to an increase in the product $|J_{d-f} N(E_F)|$ in such stable configuration systems [9]. This is attributable to the dependence of J_{d-f} on the $d-f$ orbital hybridization V_{d-f} [8]:

$$J_{d-f} \propto \frac{-2}{E_F - \varepsilon_f} \left[\frac{r_f^5 r_d^3}{R^{12}} \right] \quad (1)$$

where ε_f is the energy of the $4f$ level considered to be pinned close to the Fermi level E_F . The term in square brackets, proportional to $|V_{d-f}|^2$, is seen to have strong dependences on the atomic radii r of $4f$ (Ce) and $3d$ (Cu) atoms as well as their inter-atomic separation R . This emphasizes that the RKKY exchange J_{d-f} would have a pressure dependence, via the inter-atomic spacing dependence of the $d-f$ hybridization in $|V_{d-f}|^2$ [8, 16]. The increase in T_C in figure 1 as pressure increases is therefore ascribed to the increased $d-f$ hybridization as R is reduced under pressure.

However, the XANES data of figure 2 suggests progressive delocalization of the $4f$ electron occurs upon increasing pressure. This is compatible with increased d - f hybridization which additionally leads to broadening of the $4f$ level (i.e., ε_f in equation 1) related to such delocalization. Increased partial delocalization is disruptive to the local moment character of the stable configuration Ce^{3+} valence ($|4f^1\rangle$), prevalent in CeCuSi at low pressure [6]. The decrease in magnetic susceptibility in the ordered state in figure 1(c) is thus ascribed to this increasing partial delocalization. XMCD data of figure 3 appears to corroborate this suggestion: notably that the Ce $5d$ moment, resulting from the $4f$ - $5d$ intra-atomic exchange (thus tracking the $4f$ moment) [15], exhibits a progressive diminution as a function of pressure as well.

4. Concluding remarks

In the stable configuration local moment situation at low pressure in CeCuSi, the effect of reducing the unit cell volume beneficiates the RKKY exchange coupling J_{d-f} between Ce moments as mediated by d conduction electrons. This is exhibited in the monotonically increasing T_C upon pressurization. Increasing d - f hybridization in the relation $J_{d-f} \propto |V_{d-f}|^2$ and its strong dependence on inter-atomic spacing between Ce $4f$ and Cu $3d$ atoms rationalizes this enhancement of T_C . On the other hand, the increase in V_{d-f} consequentially broadens the original $4f$ localized level (i.e., band formation occurs). Thus sufficient $4f$ electron delocalization occurs, as seen by L_3 -edge XANES, to the extent that it disrupts the Ce^{3+} stable-configuration moment, $|4f^1\rangle$, as probed by the L_3 -edge dichroism.

Acknowledgements

Project funding derived from NRF (CPRR) and UJ (URC) grants as well as the loan of the perforated-anvils membrane-DAC assembly from Dr J-P Itié' (SOLEIL), is acknowledged with gratitude.

References

- [1] Trovarelli O, Sereni J G, Schmerber G and Kappler J P 1994 *Phys. Rev. B* **49** 15179-83
- [2] Katoh K, Takabatake T, Ochiai A, Uesawa A and Suzuki T 1997 *Physica B* **230** 159-61
- [3] Yang F, Kuang J P, Li J, Brück E, Nakotte H, Boer F R d, Wu X, Li Z and Wang Y 1991 *J. Appl. Phys* **69** 4705 - 7
- [4] Sondezi-Mhlungu B M, Adroja D T, Strydom A M, Paschen S and Goremychkin E A 2009 *Physica B* **404** 3032 - 4
- [5] Sondezi-Mhlungu B M, Adroja D T, Strydom A M, Kockelmann W and Goremychkin E A 2010 *Journal of Physics: Conference Series* **200** 012190 (1 - 4)
- [6] Gignoux D, Schmitt D and Zerguine M 1986 *Solid State Comm.* **58** 559 - 62
- [7] Süllo S, Aronson M C, Rainford B D and Haen P 1999 *Phys. Rev. Lett.* **82** 2963 - 6
- [8] Cornelius A L and Schilling J S 1994 *Phys. Rev. B* **49** 3955 - 61
- [9] Schilling J S 1979 *Adv. in Phys.* **28** 657-715
- [10] Rueff J-P, Raymond S, Taguchi M, Sikora M, Itié' J-P, Baudelet F, Braithwaite D, Knebel G and Jaccard D 2011 *Phys. Rev. Lett.* **106** 186405 (1 - 4)
- [11] Giriat G, Wang W, Attfield J P, Huxley A D and Kamenev K V 2010 *Rev. Sci. Instrum.* **81** 073905 (1-5)
- [12] Chijioke A, Nellis W J, Soldatov A and Silvera I F 2005 *J. Appl Phys.* **98**, 114905 (1 - 9)
- [13] Stöhr J and Siegmann H C 2006 *Magnetism: From Fundamentals to Nanoscale Dynamics* vol 152 (Berlin: Springer)
- [14] Giorgetti C, Pizzini S, Dartyge E, Fontaine A, Baudelet F, Brouder C, Bauer P, Krill G, Miraglia S, Fruchart D and Kappler J P 1993 *Phys. Rev. B* **48** 12732 - 42
- [15] Parlebas J C, Asakura K, Fujiwara A, Harada I and Kotani A 2006 *Phys. Rep.* **431** 1 - 38
- [16] Harrison W A 1983 *Phys. Rev. B* **28** 550 - 9

Probing the magnetic order in $(\text{Cr}_{84}\text{Re}_{16})_{100-y}\text{V}_y$ alloys using neutron diffraction

¹BS Jacobs, ¹ARE Prinsloo, ¹CJ Sheppard, ²AM Venter and ³HE Maynard-Casely

¹Department of Physics, University of Johannesburg, PO Box 524, Auckland Park, 2006, South Africa

²Research and Development Division, NECSA limited, P. O. Box 582, Pretoria 0001, South Africa

³Bragg Institute, ANSTO, Locked Bag 2001, Kirrawee DC NSW 2232, Australia

E-mail address: sjacobs@uj.ac.za

Abstract. This paper reports preliminary results of neutron diffraction investigations of the $(\text{Cr}_{84}\text{Re}_{16})_{100-y}\text{V}_y$ alloy system, with $y = 0, 4.2$ and 6.2 . The previously reported magnetic phase diagram for this alloy system compiled from various physical property studies indicates a possible critical concentration existing at $y_c \approx 10.5$. Amongst others, anomalies were observed in graphs of the Sommerfeld coefficient (γ) as function of y , close to the critical concentration, as well as at $y \approx 4$. The latter warranted further investigation using neutron diffraction in order to shed light on the conjecture that the minima corresponds to the existence of an incommensurate (I) spin-density-wave (SDW) to commensurate (C) SDW transition.

1. Introduction

Neutron diffraction has become indispensable in the investigation of the magnetic structure of elements and alloys. This technique was specifically used by Shull and Wilkinson to elucidate the antiferromagnetic (AFM) structure of pure Cr [1]. Cr and its dilute alloys display magnetic ordering characterized by a spin-density-wave (SDW) that is established through Coulombic interactions between the hole and electron sheets at the Fermi surface, a mechanism referred to as nesting [2]. This nesting decreases the energy of the system through electron-hole pair condensation and results in the appearance of SDW energy gaps at the Fermi surface in specific k -space directions when cooled through the Néel temperature (T_N). The nesting area, and concomitantly the stability of the SDW state, depends on the electron concentration per atom (e/a) which can easily be tuned by alloying Cr ($e/a = 6$) with elements like V ($e/a = 5$) or Mn ($e/a = 7$) to respectively decrease or increase the electron concentration. In pure Cr, the electron sheet is slightly smaller than the hole sheet, resulting in a SDW that is incommensurate (I) with the lattice [2]. Alloying Cr with Re, an electron donor having $e/a = 7$, brings the size of the electron and hole Fermi sheets more in line with each other. The SDW eventually becomes commensurate (C) with the lattice at a Re concentration of about 0.3 at.%. The concentration (c) versus temperature (T) magnetic phase diagram of the $\text{Cr}_{100-c}\text{Re}_c$ alloy system thus exhibits three AFM SDW phases – the longitudinal (L) ISDW, the transverse (T) ISDW and the CSDW with a triple point at $c_t \approx 0.3$ at.% Re where the ISDW, CSDW and paramagnetic (P) phases coexist [2].

Possible quantum critical behaviour (QCB) in the $(\text{Cr}_{84}\text{Re}_{16})_{100-y}\text{V}_y$ alloy system has been reported [3]. One of the significant indicators of QCB in Cr alloy systems is the behaviour of the Sommerfeld coefficient, γ [4,5] that is obtained by fitting the low temperature specific heat (C_p) data to the equation

$C_p = \gamma T + \beta T^3$. A broad deep minimum was observed in the $\gamma(y)$ curve of the $(\text{Cr}_{84}\text{Re}_{16})_{100-y}\text{V}_y$ alloy series at $y \approx 4$, after which it increased continuously on increasing y reaching a maximum at $y \approx 9$. This was followed by a continuous decrease through the critical concentration $y_c \approx 10.5$ and levelling off in the paramagnetic (PM) phase. The decrease in γ through the critical concentration can be attributed to changes in the density of states at the Fermi surface [4]. It was proposed that the minima observed at $y \approx 4$ could correspond to a CSDW-ISDW transition. Neutron diffraction studies were performed to establish this. This study reports on preliminary neutron diffraction investigations into the magnetic phases that exist in the $(\text{Cr}_{84}\text{Re}_{16})_{100-y}\text{V}_y$ alloy system as function of temperature.

2. Experimental

Ternary $(\text{Cr}_{84}\text{Re}_{16})_{100-y}\text{V}_y$ alloys with $y = 0, 4.2$ and 6.2 were prepared by arc melting in a purified argon atmosphere from Cr, Re and V each having mass fractional purities 99.99 %, 99.99 % and 99.8 % respectively. Special emphasis was placed on synthesizing specimens with high homogeneity and metallurgical quality. The alloys were annealed in an ultra-high purity argon atmosphere at 1343 K for seven days and quenched into iced water. The elemental composition and homogeneity were determined using electron microprobe analyses. Electrical resistance (R) was measured over $2 \text{ K} \leq T \leq 390 \text{ K}$, using a standard Physical Properties Measurement System (PPMS) [6]. Resistance measurements at temperatures above 390 K were performed using resistive heating in an inert atmosphere using the standard dc four-probe method and current reversal with Keithley instrumentation. Magnetic susceptibility (χ) was measured over $2 \text{ K} \leq T \leq 390 \text{ K}$ using a SQUID-type magnetometer based on the Magnetic Properties Measurement System (MPMS) platform of Quantum Design [6]. The samples were cooled to 2 K in zero field and the susceptibility measurements subsequently done upon heating in a field of 100 Oe. For the neutron powder diffraction investigations, the alloys were pulverised after annealing to limit the presence of large grains and render more uniform intensity distributions over the Debye Scherer cones. Neutron diffraction measurements were performed over the temperature range $10 \text{ K} \leq T \leq 350 \text{ K}$ on the Wombat powder diffractometer [7] at the Australian Nuclear Science and Technology Organisation (ANSTO) using a wavelength of 0.241 nm and with the primary beam filtered with a pyrolytic graphite filter to reduce higher order wavelength contaminations.

3. Results and discussion

Figure 1(a) shows the X-ray diffraction (XRD) pattern for the $(\text{Cr}_{84}\text{Re}_{16})_{95.8}\text{V}_{4.2}$ sample that was taken with Cu radiation (0.154nm wavelength). Since V neighbours Cr in the periodic table with atomic mass and size close to that of pure Cr, its addition produces the least perturbation on the structure of pure Cr [2]. The diffraction pattern of the $(\text{Cr}_{84}\text{Re}_{16})_{95.8}\text{V}_{4.2}$ sample is well fitted to the XRD pattern of

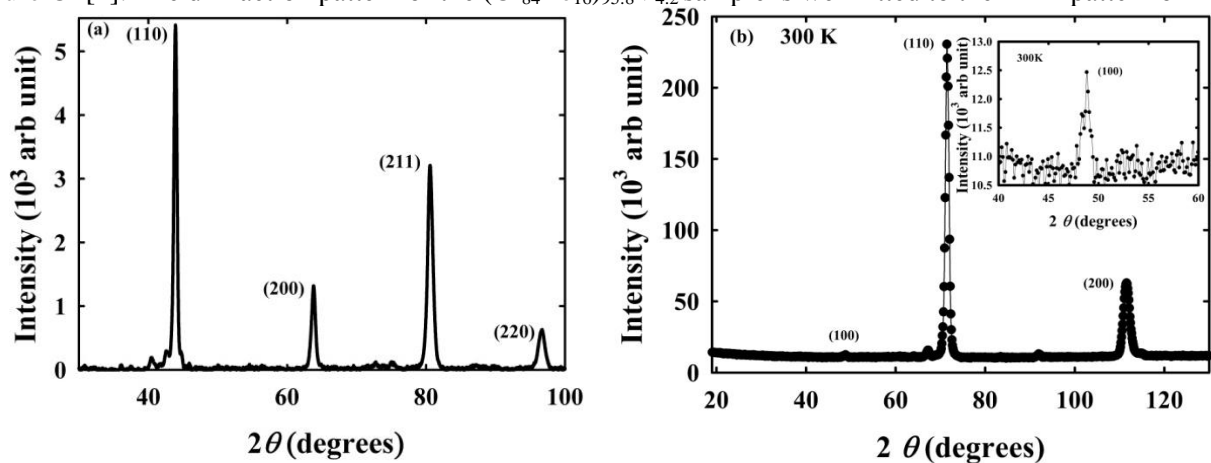


Figure 1: (a) The XRD pattern for the $(\text{Cr}_{84}\text{Re}_{16})_{95.8}\text{V}_{4.2}$ sample with the Bragg peaks indexed and (b) the neutron diffraction pattern for the $(\text{Cr}_{84}\text{Re}_{16})_{95.8}\text{V}_{4.2}$ sample at 300 K with (100) peak in inset.

body centred cubic (bcc) Cr by adjusting the lattice parameter to 0.29210 nm for this alloy from the value of 0.28839 nm for pure Cr. Peak to the left of the (110) peak corresponds to an oxide of Cr. Even though much care is taken during the melting process, the presence of small amounts of oxide is inevitable. Figure 1(b) shows the neutron diffraction pattern for the $(\text{Cr}_{84}\text{Re}_{16})_{95.8}\text{V}_{4.2}$ sample at 300 K with the (100) peak shown in the inset.

The $R(T)$ curves for $\text{Cr}_{84}\text{Re}_{16}$ and $(\text{Cr}_{84}\text{Re}_{16})_{100-y}\text{V}_y$, with $y = 4.2$ and 6.2 , are shown in figure 2. Temperature of the minimum in $dR(T)/dT$ accompanying the $R(T)$ magnetic anomaly is defined as T_N [2]. In general, the size of the anomaly, as well as T_N , decreases with increase in V concentration. The anomalies, which appear as a sudden increase in resistance on cooling through T_N , can be attributed to the Coulombic interactions of electron and hole Fermi surfaces from the nesting process on SDW formation [2].

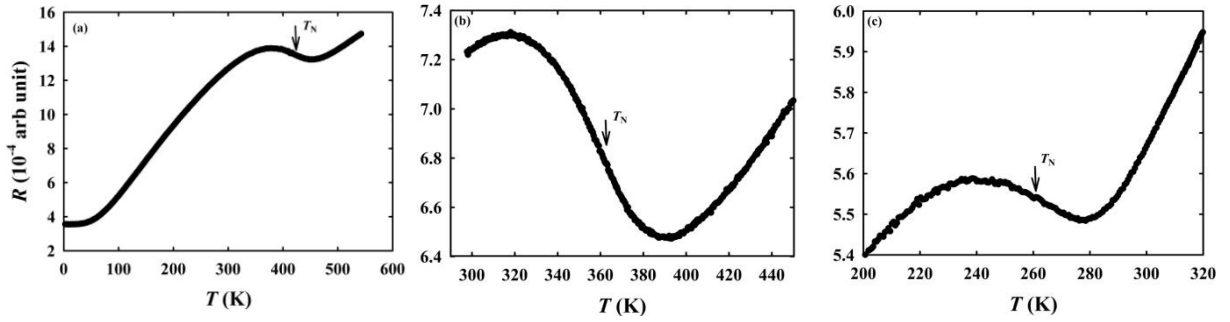


Figure 2: The $R(T)$ curves for the $(\text{Cr}_{84}\text{Re}_{16})_{100-y}\text{V}_y$ alloy system with, with (a) $y = 0$, (b) $y = 4.2$ and (c) $y = 6.2$. Arrow indicates the position of T_N for each alloy.

The temperature dependence of normalized susceptibility for $(\text{Cr}_{84}\text{Re}_{16})_{95.8}\text{V}_{4.2}$ is shown in figure 3 as an example of the typical trends of the samples. Compared with results from other previous studies [8, 9], broad anomalies are seen in the χ - T curves of all the samples of this study. The Néel transition temperature is taken at the point where a kink in the temperature dependence of χ occurs on cooling from the paramagnetic to antiferromagnetic phase. The decrease in χ can be attributed to the decrease in the density of states at the Fermi energy due to the SDW energy gap formation [2]. The broken line in figure 3 serves as a guide to the eye indicating the trend of the curve, should the sample have remained paramagnetic. T_N is taken at the point where the measured values deviate from the broken

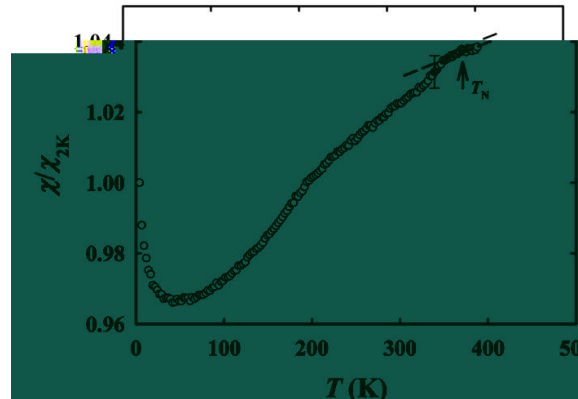


Figure 3. The temperature dependence of normalized susceptibility for the $(\text{Cr}_{84}\text{Re}_{16})_{95.8}\text{V}_{4.2}$ alloy.

line and the T_N obtained in this manner corresponds well with T_N obtained from reported electrical resistivity and thermal transport measurements [3]. The low temperature upturn in the curves is ascribed to the effect of a Curie tail that may arise from the presence of oxide impurities [10].

The neutron diffraction patterns of $\text{Cr}_{84}\text{Re}_{16}$ and $(\text{Cr}_{84}\text{Re}_{16})_{95.8}\text{V}_{4.2}$ alloys are shown in figure 4. Figure 5 shows the diffraction patterns for the $(\text{Cr}_{84}\text{Re}_{16})_{93.8}\text{V}_{6.2}$ alloy. Pure Cr has a magnetic moment of $0.4 \mu_B$ and due to its antiferromagnetic structure, the magnetic moments manifest at the (100) lattice parameter position [11]. The commensurate phase is characterised by the moments located at the corners of the unit cell being parallel to each other but antiparallel to the one at the centre [12]. For the CSDW phase the unit cell lengths of the chemical and magnetic structures are the same. As the moments are collinear to the equivalent {100} planes, the resulting neutron diffraction pattern will have magnetic Bragg peaks at {100}, {111}, {210}, etc. Since the magnetic moment of Cr is weak, only the {100} magnetic peak is observed [12]. The ISDW phase leads to the additional presence of magnetic satellites, offset from the {100} positions at $(1 \pm \delta, 0, 0)$ reciprocal lattice positions [13], whereas the CSDW phase displays no such satellites.

Figure 4(a) shows the neutron diffraction pattern for the $\text{Cr}_{84}\text{Re}_{16}$ mother alloy taken at 300 K and at the highest measured temperature, 560 K (expected $T_N = 420$ K). A clear temperature dependent peak of magnetic origin that dies out is seen at $2\theta = 48.7^\circ$ which corresponds to the (100) position. The second peak at $2\theta = 38^\circ$ is not strongly dependent on temperature and therefore is not magnetic in origin. Within the detection limit of the instrument, since no satellite peaks of magnetic origin are observed, it is concluded that this alloy has a CSDW ordering at room temperature, in correspondence to what is expected from the Cr-Re magnetic phase diagram [2].

The neutron diffraction patterns for the $(\text{Cr}_{84}\text{Re}_{16})_{95.8}\text{V}_{4.2}$ alloy taken at various temperatures are shown in figure 4(b). T_N for this alloy is 370 K. Due to time constraint, a reference paramagnetic diffraction pattern was not recorded. Measurements started at 340 K where a central magnetic peak could be observed at the (100) position. On cooling to 300 K, the (100) intensity grows and in addition a peak appears at approximately $2\theta = 38.6^\circ$ that corresponds to the $(1 + \delta, 0, 0)$ magnetic satellite. On

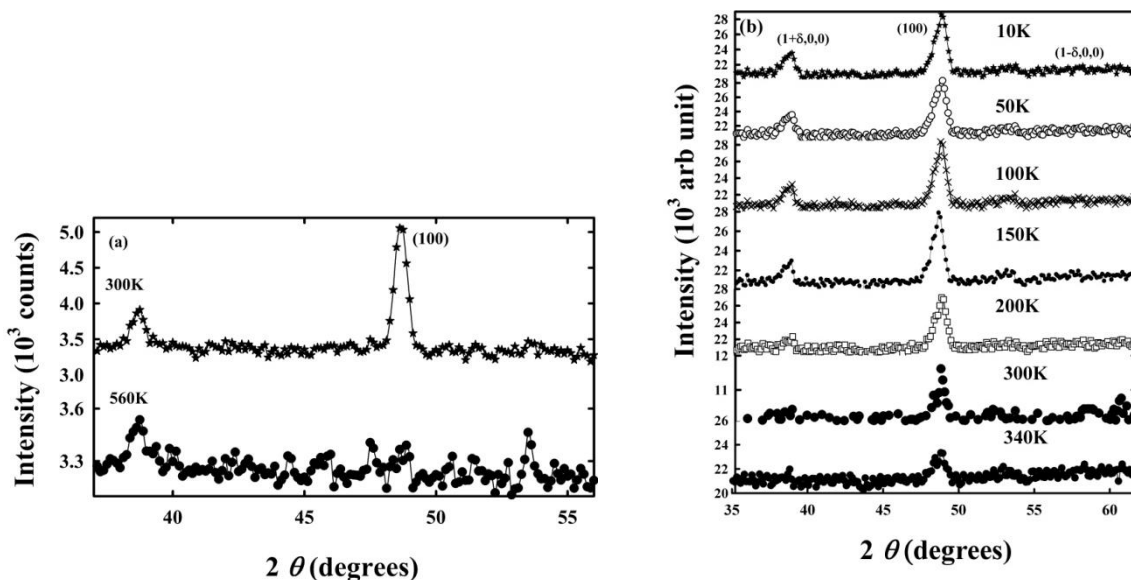


Figure 4. Neutron diffraction patterns of $(\text{Cr}_{84}\text{Re}_{16})_{100-y}\text{V}_y$ alloys, with (a) $y = 0$ at 300 K and 560 K and (b) $y = 4.2$ at various temperatures.

further cooling, the intensities of both of these magnetic peaks increase. The calculated full width at half maximum (FWHM) shows no regular trend. The $(1-\delta,0,0)$ magnetic satellite is expected at the $2\theta = 58.8^\circ$ position but was not observed. Although the satellite peak on the right is observed at certain temperatures, no clear trend is exhibited. These results indicate that the $(\text{Cr}_{84}\text{Re}_{16})_{95.8}\text{V}_{4.2}$ alloy appears to undergo a transition from the CSDW phase at $T > 300$ K to a certain ISDW phase at $T \leq 300$ K, as proposed initially [3].

Figure 5 shows the diffraction pattern for the $(\text{Cr}_{84}\text{Re}_{16})_{93.8}\text{V}_{6.2}$ alloy at various temperatures. T_N for this alloy is 270 K. The diffraction pattern observed at 300 K corresponds to the paramagnetic phase. At 220 K, the central (100) peak is observed indicating the CSDW phase and it appears to remain in the CSDW phase up to 180 K. However, longer measurement times will be required to confirm this. Below 180 K, the existence of a satellite reflection can be detected.

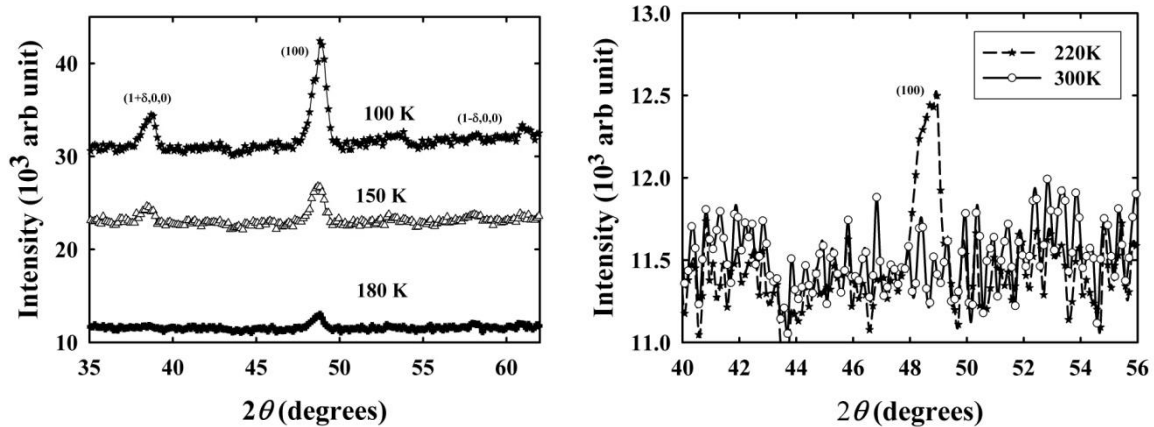


Figure 5. The neutron diffraction patterns of the $(\text{Cr}_{84}\text{Re}_{16})_{93.8}\text{V}_{6.2}$ alloy. The alloy remains in the ISDW phase up to 180 K. At 220 K, the absence of symmetrical satellite peaks indicate CSDW ordering.

The increase in intensity of the central peak and the satellite peak at temperatures below 180 K indicates that the alloy is in the ISDW phase. A symmetric satellite peak on the right hand side is also observed though not very distinct and the variation in its intensity with temperature is not as prominent as the satellite peak on the left. The $(\text{Cr}_{84}\text{Re}_{16})_{93.8}\text{V}_{6.2}$ alloy thus appears to be in the CSDW phase at $T \geq 180$ K, transforming to a ISDW phase at $T < 180$ K.

4. Conclusion.

The present study investigated the possible magnetic origin of the broad minimum observed in the $\gamma(y)$ curve of the $(\text{Cr}_{84}\text{Re}_{16})_{100-y}\text{V}_y$ alloy series at $y \approx 4$. The limited beam time awarded on the Wombat diffractometer at ANSTO was utilised to only probe the magnetic phases of the $\text{Cr}_{84}\text{Re}_{16}$, $(\text{Cr}_{84}\text{Re}_{16})_{95.8}\text{V}_{4.2}$ and $(\text{Cr}_{84}\text{Re}_{16})_{93.8}\text{V}_{6.2}$ alloys. At room temperature, the $\text{Cr}_{84}\text{Re}_{16}$ alloy is in the CSDW phase. Measurements need to be extended to a wider temperature range to further investigate the magnetic ordering in this sample. Preliminary measurements on the $(\text{Cr}_{84}\text{Re}_{16})_{95.8}\text{V}_{4.2}$ alloy indicate satellites around the central (100) peak in the neutron diffraction pattern suggesting that this alloy is in the ISDW phase at $T \leq 300$ K. In the case of the $(\text{Cr}_{84}\text{Re}_{16})_{93.8}\text{V}_{6.2}$ alloy, diffraction patterns indicate the ISDW phase at $T < 180$ K. At the first glance, these preliminary results seem to concur with our initial projection that the minimum in the $\gamma(y)$ curve corresponds to an I-C transition of the SDW state. Measurements need to be extended for the $(\text{Cr}_{84}\text{Re}_{16})_{100-y}\text{V}_y$ series to include temperatures scans for the $y = 0$ sample and higher temperatures for the $y = 4.2$ sample. It is also essential to repeat the preliminary measurements with improved counting statistics. Neutron diffraction study also needs to be extended to include samples having V concentration between 4.2 and 6.2 in order to make concrete conclusions.

Acknowledgements

This work was supported by the NRF of South Africa under Grants 80626, 80928 and the Faculty of Science at the University of Johannesburg.

References

- [1] Shull C G and Wilkinson M K 1955 *Phys. Rev.* **97** 304
- [2] Fawcett E, Alberts H L, Galkin V Y, Noakes D R and Yakhmi J V 1994 *Rev. Mod. Phys.* **66** 25
- [3] Jacobs B S, Prinsloo A R E, Sheppard C J and Strydom A M 2013 *J. Appl. Phys.* **113** 17E126
- [4] Takeuchi J, Sasakura H and Masuda Y 1980 *J. Phys. Soc. Japan* **49** 508
- [5] Yeh A, Soh Y, Brooke J, Aeppli G and Rosenbaum TF 2002 *Nature* **419** 459
- [6] Quantum Design Inc, 6325 Lusk Boulevard, San Diego, USA
- [7] Studer A J, Hagen M E and Noakes T J 2006 *Physica B* **385-386** 1013
- [8] Nishihara Y, Yamaguchi Y, Waki S and Kohara T 1983 *J. Phys. Soc. Japan* **52**(7) 2301
- [9] Arajs A, Kote G, Moyer CA, Kelly JR, Rao KV and Anderson EE 1976 *Phys. Stat. Sol. (b)* **74** K23
- [10] Wagner M J, Dye J L, Pérez-Condero E, Bulgac R and Echegoyen L 1995 *J. Am. Chem. Soc.* **117** 1318
- [11] Bacon G E and Cowlam N, 1969 *J. Phys.* **C2** 238
- [12] Papoular R, Debray D and Arajs S 1981 *J. Magn. Magn. Mater.* **24** 106-110
- [13] Bacon G E 1961 *Acta. Cryst.* **14** 823

Electrical properties of epitaxial $\text{Cr}_{100-x}\text{Co}_x(100)$ alloy thin films on $\text{MgO}(100)$

M B Kadam¹, C J Sheppard¹, A R E Prinsloo¹ and E E Fullerton²

¹Department of Physics, University of Johannesburg, PO Box 524, Auckland Park, 2006

²Center for Magnetic Recording Research, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0401, USA

Email: alettap@uj.ac.za

Abstract. Cr-Co alloys and Cr/Co multi-layered systems have shown considerable promise in practical applications because of its Invar-like properties and enhancement of the SDW, respectively. Comparison between Cr in bulk and thin film forms revealed dimensionality plays an important role in modifying the SDW structure. Epitaxial $\text{Cr}_{100-x}\text{Co}_x$ thin films of thickness (t) \approx 200 nm, with $0 < x < 7.2$, were prepared on $\text{MgO}(100)$ substrates using DC magnetron co-sputtering techniques. The resistivity (ρ) for these samples was measured in the temperature range $2 \text{ K} < T < 395 \text{ K}$ and used to determine the Néel transition temperatures (T_N) for the individual samples. The T_N versus x plot for this sample series shows non-monotonic behaviour and sheds light on the effect of dimensionality on the electrical properties of the Cr-Co alloys.

1. Introduction

In the entire 3d-series, only chromium and manganese exhibits antiferromagnetic ordering [1]. The temperature dependence of the magnetic structure of chromium below the Néel temperature (T_N) is characterized by a longitudinal (L) or transverse (T) spin density wave (SDW), with wave vector \mathbf{Q} , that is incommensurate with the period of the reciprocal lattice [2]. The Néel temperature, $T_N = 311 \text{ K}$, is well established for bulk chromium. Comparison between Cr in bulk and thin film forms indicated that dimensionality plays an important role in modifying the SDW structure [3].

Recent studies revealed that Cr alloys with group-8 magnetic transition metals such as Fe, Co and Ni show magnetic phase diagrams quite different from each other [2]. However, some similarities are observed with individual doping of Fe and Co into a Cr host. Firstly, doping with both metals exhibits a triple point, where the incommensurate (I) spin density wave (SDW), commensurate (C) SDW and paramagnetic (P) phases converge [2]. Secondly, with doping these metals, initially T_N decreases slowly up to certain electron concentration and then increases with increasing electron concentration [2]. Shibata [4] and Moyer *et al.* [5] explained this anomalous behaviour in T_N considering the interaction between local moments and the SDW host. In Cr-Co alloys, Co local moment is strongly coupled to SDW [6]. Hence, unlike in pure Cr, it is expected that dimensionality in Cr-Co alloys should have strong effect on SDW structure. Interestingly, Ge *et al.* [7] reported that at a certain layer thickness, the SDW in Co/Cr multilayers have been enhanced as compared to bulk Cr-Co alloy system

Besides interesting antiferromagnetic properties, Co-Cr alloys have numerous applications. Co-rich Co-Cr alloy thin films have been studied extensively for their applications in a recording and storage media [7, 8]. Ferromagnetic Invar and Elinvar type alloys are widely used in the field of

electromagnetic and precision instrumentation engineering [9]. Their use is often limited due to their ferromagnetism, as these alloys cannot be used in electromagnetic and precision instruments which are used in a static magnetic field. Also, in alternating magnetic field these alloys exhibit magnetostrictive oscillations [9]. Another drawback of these alloys is that the elastic moduli of ferromagnetic Elinvar type alloys are affected remarkably in magnetic fields. However, Cr-Co alloy system is of considerable practical importance because of its both Invar and Elinvar properties along with non-ferromagnetism [9].

The magnetic phase diagram for Cr-Co alloys with lower concentrations of Co is well established [2], but effect of dimensionality on the magnetic phase diagram is yet to be explored. Therefore, taking this into consideration, the present paper studies the effects of thin-film growth and dimensionality on the magnetic phase diagram and electrical properties of Cr-Co alloys.

2. Experimental

Epitaxially-grown $\text{Cr}_{100-x}\text{Co}_x$ thin films of thickness (t) ≈ 200 nm, with $0 < x < 7.2$, were prepared on MgO(100) single-crystal substrates using DC magnetron co-sputtering techniques. Structural studies of these samples were carried using X-ray diffraction (XRD) techniques (Phillips PAN analytical Xpert Pro X-ray diffractometer) in the 2θ range 20 – 100° using $\text{Cu-K}\alpha_1$ radiation ($\lambda = 1.54056\text{\AA}$). The XRD patterns were compared with standard Joint Council of Powder Diffraction Database (JCPDD) files of Cr (06-0694). The electrical resistivity (ρ) measurements for these films were carried out in the temperature range $2\text{ K} < T < 395\text{ K}$ using a Quantum Design Physical Properties Measurement System (PPMS), using the standard four probe technique. For resistivity measurements, the electrical current was directed along the [100] direction. The electrical resistivity measurements were employed to determine Néel temperature for the individual films.

3. Results and Discussion

Figure 1 shows a representative XRD result for the $\text{Cr}_{100-x}\text{Co}_x$ thin films with $x = 1.3$, prepared on a MgO(100) substrate. The result indicates that the sample is indeed epitaxial, exhibiting a single crystallographic orientation, with preferred [100] growth direction. The additional peaks observed in XRD patterns are due to the substrate contribution.

The electrical resistivity $\rho(T)$ curves for $\text{Cr}_{100-x}\text{Co}_x$ thin films, with $0 < x < 7.2$, are shown in Figure 2 (a) to (d). Well defined anomalies in the form of domes are observed in $\rho(T)$ curves of all the thin films. The hump like anomaly in each of the resistivity of the $\rho(T)$ curve of the thin films are attributed to the formation of SDW on entering the antiferromagnetic phase on cooling through the Néel temperature (T_N) [2]. This anomaly observed below T_N , finds its origin in the nesting of electron and hole Fermi surfaces [2], which indicate the decrease in the number of charge carriers available for the conduction below T_N . The actual determination of T_N is discussed further in the text. Generally, exciton formation takes place in Cr-Co alloys during transition from paramagnetic to antiferromagnetic state below T_N . Hence these alloys can be treated as exciton dielectrics [1, 10]. According to Anderson's model of localized states in metals [1, 11], the d-d repulsion energy is higher than the s-s and s-d repulsion energy. Since Co exhibits local moments, a d-shell state on Co impurity atom of spin up is full while spin down is empty. Due to repulsion energy in d-shells an electron with spin down will experience extra repulsion compared to the spin up electron. Thus, the spin up state energy level ($E_{d\uparrow}$) lies below the Fermi level (E_F) and hence transition of electrons to $E_{d\uparrow}$ increases the resistivity over wide temperature range below T_N [1]. From Figure 2, the rise in resistivity below T_N , is found to be more pronounced for the films with $x = 6.4$ and $x = 7.2$ as compared to the films with $x < 6.4$. This indicates the enhancement of SDW in these alloy thin films. Ge *et al.* [7] and Antonoff [12] attributed such behaviour to originate from the interaction between the Co moments and the SDW.

An additional anomaly, the minimum in resistivity below approximately 75 K, is observed for the films with $x = 6.4$ and $x = 7.2$ as seen in Figures. 2 (c) and (d), respectively. An analogous resistivity minimum observed in Cr alloys at low temperature has been investigated [14] and was originally

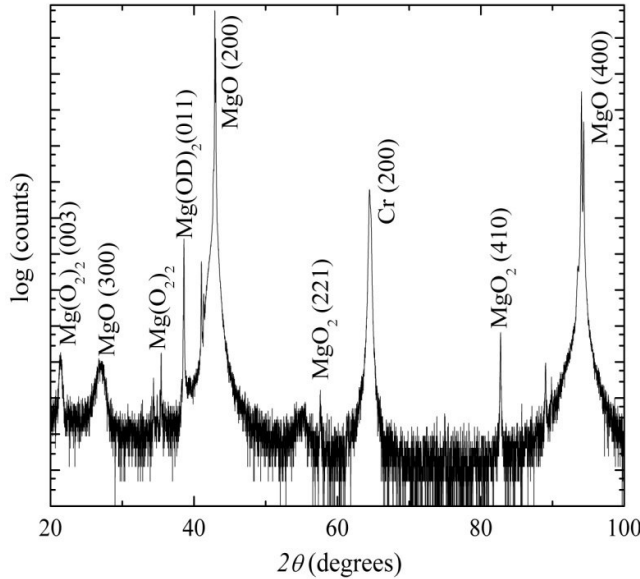


Figure 1. θ - 2θ XRD scan (Bragg-Brentano geometry) in which the scattering vector is maintained normal to the plane of the film, for $\text{Cr}_{100-x}\text{Co}_x$ thin films with $x = 1.3$, prepared on $\text{MgO}(100)$ substrate.

attributed to the Kondo effect [13, 14]. Magnetic susceptibility measurements on Cr-Co alloys system showed that Co have a local moment in Cr alloys [6]. However, the observed susceptibility behaviour of Cr-Co alloys does not satisfy the Curie-Weiss law at low temperatures, indicating that the Co local moment is strongly coupled to the SDW [6]. Generally, the presence of a local moment in metallic host leads to a situation favourable for the Kondo effect [6, 15]. However, in magnetically ordered systems the exchange field suppresses the Kondo state, but interestingly the low temperature minima is pronounced in Cr-Co alloys in their ordered state [6]. Kondorskii *et al.* [1] also suggested that to attribute the low temperature resistivity minimum (T_{\min}) to Kondo-type behaviour the relation, $\rho \sim \ln T$ should hold below T_{\min} . They observed that in bulk Cr-Co alloys with Co impurities in the range 4 at.% to 8 at.%, resistance minimum satisfies $\rho \sim (T)^{1/2}$ relation rather than $\rho \sim \ln T$ relation and discarded the possibility to attribute this resistivity behaviour below T_{\min} to Kondo-like resistance minimum. The $\rho \sim (T)^{1/2}$ dependence of resistivity below T_{\min} , is explained by taking into account the inelastic electron-electron interaction and the elastic scattering of electron by the impurities [1]. Also, Katano *et al.* [13] suggested that the Kondo effect causes the resistivity minimum in Cr-Co alloys with commensurate SDW state. The view that the resistance minimum is Kondo-type behaviour was questioned by Fawcett *et al.* [2].

In the present investigations, the detailed $\rho(T)$ data analysis for films with $x = 6.4$ and $x = 7.2$, revealed that resistivity behaviour below T_{\min} does not obey the $\rho \sim (T)^{1/2}$ nor the $\rho \sim \ln T$ relation up to 2 K. The insets in Figure 2 (c) and (d) shows clear non-linear behaviour in ρ vs $\ln T$ plots for the alloys with $x = 6.4$ and $x = 7.2$. It should be noted that experiments done by Kondorskii *et al.* [1] were carried out on bulk Cr-Co alloys down to 4 K below T_{\min} , while the experiments reported here are done on Cr-Co thin films down to 2 K. Considering the current ρ versus $\ln T$ results it appears that a linear region can be found in the temperature range $44 \text{ K} < T < 12 \text{ K}$ corresponding to the work of Kondorskii *et al.* [1], but it deviates for $T < 12 \text{ K}$.

Boekelheide *et al.* [16] studied the electrical resistivity of Cr thin films and observed the low temperature resistivity minima analogous to that observed in Cr-Co alloy thin films in our experiments. They stated that the minimum was attributed to the impurity resonant scattering (IRS). This conclusion was supported by the following: firstly, the specific heat measurements on the Cr thin films showed large density of states at Fermi level and secondly, the SDW suppressed the spin-flip scattering of the conduction electrons [16]. However, to confirm that T_{\min} in the present Cr-Co films is attributed to IRS specific heat measurements on these films should be done.

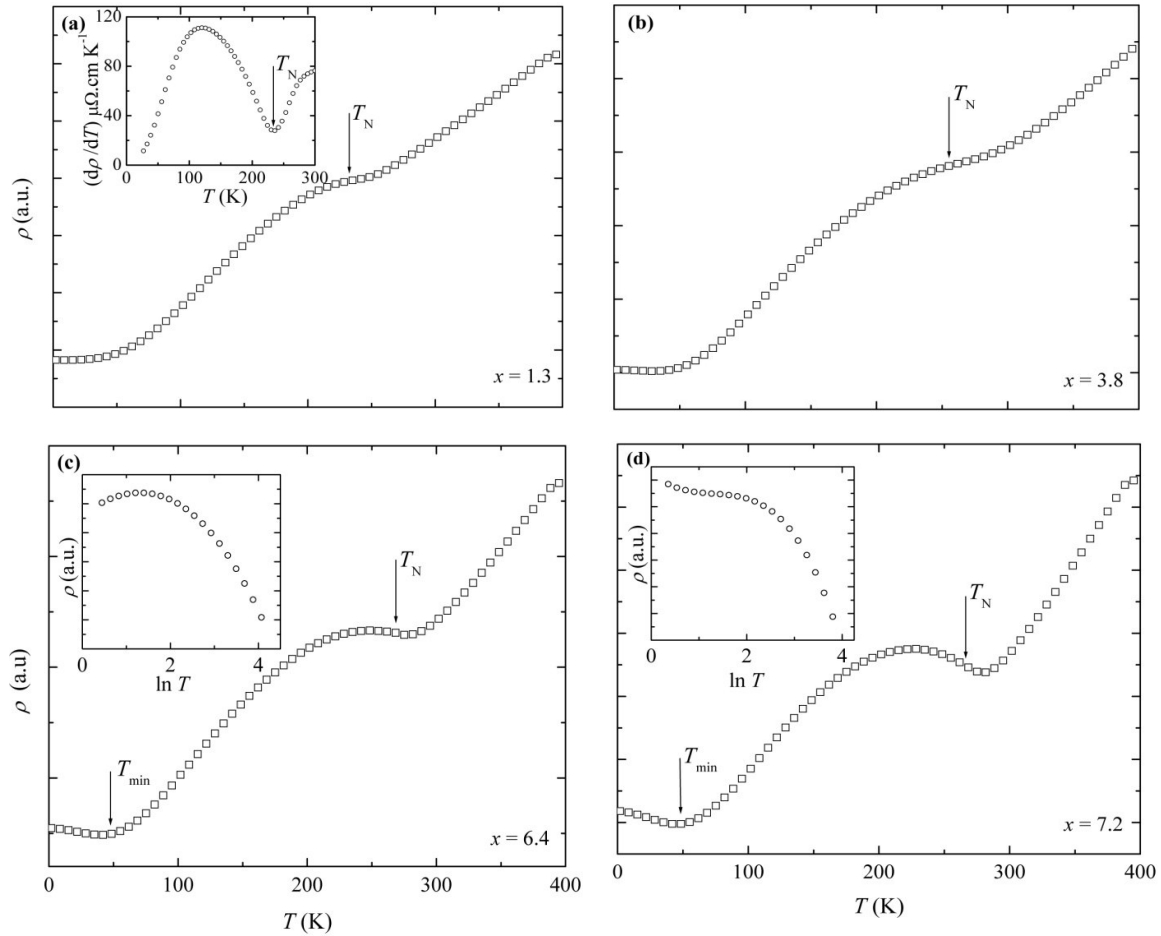


Figure 2. The resistivity (ρ) as a function of temperature (T) for $\text{Cr}_{100-x}\text{Co}_x$ thin films: (a) $x = 1.3$, (b) $x = 3.8$, (c) $x = 6.4$ and (d) $x = 7.2$. The arrows indicate the position of the Néel temperature (T_N) obtained from minimum in $d\rho/dT$ versus T plots for all the samples. Insets show $d\rho/dT$ versus T plot for (a) $x = 1.3$ and the ρ vs $\ln T$ plots for (c) $x = 6.4$ and (d) $x = 7.2$.

In the case of Cr alloys T_N can be defined in one of the two ways: either as the temperature of minimum in $d\rho(T)/dT$ accompanying magnetic phase transition [2] or T_N is defined at the inflection point of the $\rho(T)$ curve. In this study the second approach was followed because of the broadened transition often observed in films. In order to determine the T_N , $d\rho/dT$ versus T plots for all Cr-Co thin films were considered and T_N was taken as a temperature associated with the minima on the corresponding plot. A representative plot used to determine T_N for $\text{Cr}_{100-x}\text{Co}_x$ thin films with $x = 1.3$ is shown in the inset of Figure. 2 (a). The obtained T_N values are denoted by the arrows in the corresponding Figures 2 (a) to (d). The variation of T_N as a function of x for $\text{Cr}_{100-x}\text{Co}_x$ thin films is shown in Figure. 3. It is seen that T_N decreases with doping up to 2.6 at.% Co. Further T_N increases rapidly for doping up to 6.4 at.% Co reaching a maximum and then decreases. This is in agreement with the dependence of T_N non-monotonically on cobalt densities, obtained from neutron diffraction studies on bulk materials [1, 17]. In case of Cr-Co system in bulk form, the triple point (L) is observed for triple point concentration (x_L) 1.3 at.% Co concentration at triple point temperature (T_L) 280 K [2]. However, from Figure. 3, it is seen that in the case of epitaxial Cr-Co thin films, L appears to be

Table 1. Comparison between Néel temperature (T_N) for bulk and thin film form of $\text{Cr}_{100-x}\text{Co}_x$ alloy with different values of x .

$\text{Cr}_{100-x}\text{Co}_x$ alloy with x	T_N , K for Cr-Co alloy in Bulk form [2]	T_N , K for Cr-Co alloy in thin film form
1.3	281	234 ± 2
2.6	319	222 ± 3
2.9	320	233 ± 3
3.8	317	259 ± 2
6.4	290	268 ± 1
7.2	286	258 ± 2

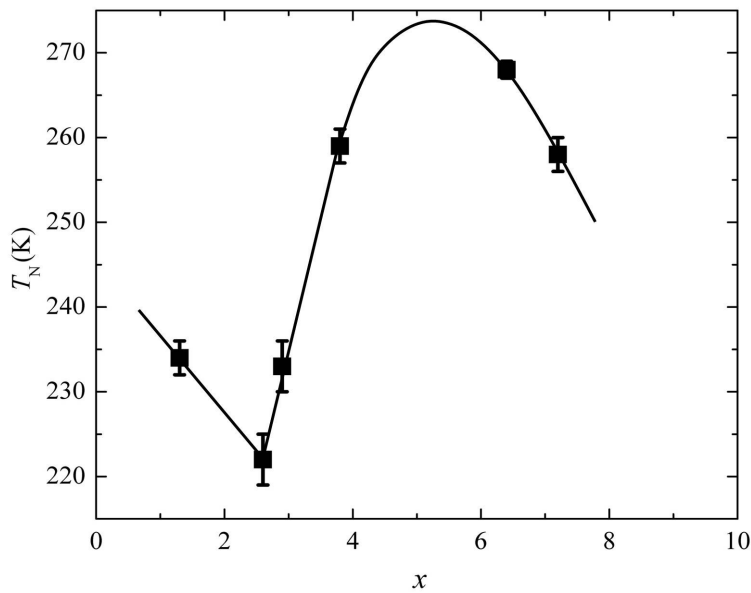


Figure 3. The Néel temperature (T_N) as a function of x for the epitaxial $\text{Cr}_{100-x}\text{Co}_x$ thin films prepared on $\text{MgO}(100)$.

shifted to $x_L = 2.6$ and T_L decreased to 222 ± 3 K. Table 1 shows comparison between Néel temperature (T_N) for bulk and thin film form of $\text{Cr}_{100-x}\text{Co}_x$ alloy with different values of x . From Table 1 it is seen that the T_N for Cr-Co system in epitaxial thin film form is shifted towards the lower temperature than its corresponding value in bulk form. The shift in triple point concentration (x_L), triple point temperature (T_L) and the decrease in T_N in epitaxial thin film form compared to bulk form may be attributed to the dimensionality effects and epitaxial nature of the thin films [7] and need to be investigated further.

The initial decrease in T_N up to $x = 2.6$ in Figure. 3 can be attributed to the contribution due to interaction between the local moments in Co and the Cr SDW [2]. The pairing effect arising from the antiferromagnetic alignment of the impurity spins parallel to SDW polarization and the depairing effect due to spin and potential scattering acts in opposition to each other. Thus, these two effects give opposite temperature dependence of T_N . Thus, under certain conditions the combined effect may give rise to a decrease in T_N [2, 4].

In high-quality thin films, SDW is quantized [18]. It indicates that change in quantization value influence the carrier density in discrete manner. The changes in charge density takes place when the carriers participate in the formation of SDW. During SDW formation, these carriers become localized resulting in the loss of mobile charge carriers. These effects can be more prominently observed by

Hall coefficient measurements [18]. The reliable profiles of the charge type, carrier density and mobility can also be obtained using Hall coefficient measurements. Thus, it is of interest to study the behaviour of R_H in these Cr-Co alloy thin films. Therefore, it is envisaged that this study will further be extended to include Hall measurements.

4. Conclusions

In the present manuscript the effects of epitaxial growth and dimensionality on the magnetic phase diagram and electrical properties of epitaxial $\text{Cr}_{100-x}\text{Co}_x$ thin films with $0 < x < 7.2$, prepared on MgO(100) substrates using DC magnetron co-sputtering are reported. XRD analyses confirmed the epitaxial nature of these monolayers. The electrical resistivity measurements on these $\text{Cr}_{100-x}\text{Co}_x$ thin films revealed that below T_N , the resistivity humps increases with increase in Co concentration giving rise to the dome like anomaly. It is found that the Néel temperature (T_N) for Cr-Co alloy in epitaxial thin film form is lower than the T_N obtained for bulk Cr-Co alloys with similar concentration. The resistivity behaviour of the studied CrCo thin films below T_{\min} is found to be complex but appears to be the result of impurity resonant scattering (IRS). The magnetic phase diagram of Cr-Co alloy system in epitaxial thin film form showed shift in triple point concentration to $x_L = 2.6$ at $T_L = 222 \pm 3$ K as compared to bulk Cr-Co alloy system at $x_L = 1.3$ and $T_L = 280$ K. Future work will focus on Hall measurements in order to understand the charge carrier effects.

Acknowledgements

Financial support from the South African NRF (Grant numbers 80928 and 80626) and Faculty of Science from the University of Johannesburg are acknowledged.

References

- [1] Kondorskii E I, Kostina T I and Trubitsina N V 1983 *Sov. Phys. JETP* **58** (1) 155
- [2] Fawcett E, Alberts H L, Galkin V Y, Noakes D R and Yakhmi J V 1994 *Rev. Mod. Phys.* **66** 25
- [3] Zabel H 1999 *J. Phys. Condens. Matter* **11** 9303
- [4] Shibatani A 1971 *J. Phys. Soc. Jpn.* **31** 1642
- [5] Moyer C A, Kelly J and Araj S 1979 *Phys. Status Solidi* **B 96** 775
- [6] Jayaraman A 1998 *Curr. Sci.* **75** (11)1200
- [7] Ge S and Stearns M B 1988 *J. Appl. Phys.* **63**(8) 4297
- [8] Hadicke E, Werner A and Hibst H 1989 *Textures and Microstructures* **11** 231
- [9] Fukamichi K, Fukuda N and Saito H 1976 *Trans. Jpn. Inst. Metall.* **17** 125
- [10] Zittartz J 1967 *Phys. Rev.* **164** 575
- [11] Anderson P W, 1961 *Phys. Rev.* **124** 41
- [12] Antonoff M M 1980 *J. Magn. Magn. Mater.* **15** 277
- [13] Katano S, Mori N and Nakayama K 1980 *J. Phys. Soc. Jpn.* **48** 192
- [14] Araj S, Dunmyre D R and Dechter S J 1967 *Phys. Rev.* **154** 448
- [15] Kondo J 1964 *Prog. Theor. Phys.* **32** 37
- [16] Boekelheide Z, Cooke D W, Helgren E and Hellman F 2009 *Phys. Rev. B* **80** 134426
- [17] Endoh Y, Ishikawa Y, and Ohno H 1969 *J. Phys. Soc. Jpn.* **24** 263
- [18] Soh Y A and Kummamuru R K 2011 *Phil. Trans. R. Soc. A* **369** 3646

Irradiation-induced improvement in crystalline quality of epitaxially grown InGaN thin films: A preliminary study

Morgan Madhuku^{1*}, Ghulam Husnain², Ishaq Ahmad² and Hamid Saleem²

¹iThemba LABS (Gauteng), Wits 2050, South Africa

²National Centre for Physics, Islamabad 44000, Pakistan

*E-mail: madhuku@tlabs.ac.za

Abstract. Irradiating materials with energetic particles is commonly thought to introduce undesirable phenomena, but recent experiments on ion irradiation of various nanostructures have shown that it can be beneficial and that ion beams may be harnessed to tailor the structure, mechanical, electronic, and even magnetic properties of nano-systems with high precision. However, the structural transformation and physical property changes in III-nitride thin films under ion irradiation have not yet been fully investigated. In this work, we focused on InGaN material for investigation of the effects of Phosphorus (P^+) ion irradiation on its structure, optical and electronic properties. Presently, there are a limited number of reports on InGaN:P in the literature. Hence, it was necessary to investigate the effects of P^+ incorporation into InGaN thin films. Preliminary results from Rutherford backscattering spectrometry and channeling (RBS/C) indicate that the crystalline quality of the InGaN thin films is improved by ion beam irradiation. Irradiating with 0.7 MeV P^+ ions to 1×10^{14} ions/cm² at room temperature reduces the channeling minimum yield (X_{\min}) from about 60% to about 30%, but X_{\min} increases considerably above this dose. Detailed experimental investigations will be carried out to obtain more information on the observed irradiation-induced-improvement in crystalline quality of InGaN thin films.

1. Introduction

The preparation and characterization of advanced optoelectronic material thin films has attracted tremendous scientific attention in recent years. It is well known that the structural, electrical, and optical properties of thin films are strongly correlated with their composition, microstructure, film thickness and impurities [1]. Irradiation of materials with energetic particles, such as ions or electrons, usually spoils their chemical or physical properties as a result of nuclear collisions and electronic excitation. However, ion irradiation may, on the whole, have beneficial effects on the target [2]. Two important examples of positive effects of ion beam irradiation are ion implantation onto semiconductors and the radiation-assisted treatment of cancer [3]. Some of the recent studies on the effects of ion irradiation on properties of materials can be found in Kumar *et al.*, 2014 [4] and Som *et al.*, 2014 [5]. Ion beam irradiation provides an active research field for the low-temperature processing, especially in semiconductor technology. The generation and migration of vacancies induced by ion irradiation make it possible to synthesize and modify materials at temperatures well below those required for thermally activated processes [6]. The evolution of the structure of thin films by ion irradiation is interesting both from theoretical and experimental points of view. In this preliminary study, we investigate the effects of Phosphorus (P^+) ion irradiation on the crystalline

quality of epitaxial InGaN/GaN thin films grown on sapphire substrates. Rutherford backscattering combined with channeling(RBS/C) was utilized to characterize the as-grown and the ion-irradiated InGaN/GaN samples. Future studies will look into structural, electrical and optical property changes after P^+ ion irradiation using additional and/or complementary techniques like transmission electron microscopy (TEM), X-ray diffraction (XRD), four-probe electrical method, and spectroscopic ellipsometry (SE).

2. Background

2.1. Group-III Nitride Materials

Group III-nitride semiconductors, aluminium nitride (AlN), gallium nitride (GaN), indium nitride (InN) and related alloys are obtained by combining group III elements (Ga, Al, In) with the group V element, N [7]. The III-nitrides are an unparalleled material system with many prominent features such as wide bandgap (0.7 eV for InN, 3.4 eV for GaN, and 6.2 eV for AlN), high mechanical and thermal stability, large piezoelectric constants, and excellent electro-optical properties [8]. The bandgap energy of a semiconductor is an essential parameter which influences its transport and optical properties including many other phenomena [9]. The energy gap of ternary III-nitride alloys like indium gallium nitride (InGaN), aluminum gallium nitride (AlGaN), or indium aluminum nitride (InAlN), can be adjusted to conform to light emission in the whole visible spectrum and into the deep ultraviolet (UV) region. In principle, the bandgap energy of the alloys can be varied continuously from 0.7 eV (pure InN) to 6.2 eV (pure AlN).

In contrast to conventional semiconductors like silicon (Si) or gallium arsenide (GaAs) which have a diamond or zinc-blende (zb) structure with cubic symmetry, the group III-nitrides can crystallize in a hexagonal wurtzite (wz) or in a cubic zb crystal structure [8]. The thermodynamically stable crystalline structure of the III-nitrides is the hexagonal wz crystal structure [10], which are partially ionic solids due to large differences in the electronegativity of the group-III metal cations (Al = 1.18, Ga = 1.13, In = 0.99) and nitrogen anions (N = 3.0) [8,11,12].

The III-nitrides have been studied intensively due to the great number of potential applications, and have recently attained an important position in the science and technology of compound semiconductors [13]. GaN and its alloy InGaN have become dominant materials for producing high brightness LEDs and LDs that emit light in the green/blue region of the visible spectrum. InGaN based light emitting diodes are already in use in full colour liquid crystal diode (LCD) displays and traffic lights. All these applications demonstrate the technological relevance of the III-nitride compounds and the reason for these materials to be the subject of an active research field (see, for example, [8,14-17]).

2.2. Ion implantation into InGaN alloys

$In_xGa_{1-x}N$ has the widest direct bandgap range of any compound semiconductor, varying from 0.7 eV (for $x = 1$) to 3.4 eV (for $x = 0$), which can be utilised in optoelectronic device applications from ultraviolet to infrared. This makes InGaN ideal for a wide range of applications, from high-density optical data storage, medical photonics, to communication devices. Recently, there has been a lot of interest in using InGaN in solar photovoltaic (PV) technology [18-21].

Ion implantation is an adaptable technique for introducing an array of doping or compensating impurities into semiconductors. Implantation can also be used to introduce a controlled level of impurities into the lattice for defect studies. It is the technique of choice to dope selective areas of a layer to desired carrier concentrations (for example, to create contact or channel regions) and to form semi-insulating regions for device isolation in such applications as field effect transistors (FETs). As the crystal quality of the group III-nitride materials continues to improve, ion implantation is playing an enabling role in exploring new dopant species and device structures [22]. Presently, some of the ion implantation doping and doping during film growth studies for InGaN that can be found in the literature are for the following impurities: N and F [23], Er [24], and Si and Zn [25]. There are no or limited studies on phosphorus (P)-doped InGaN available in the literature at the moment. However, P

ion doping has been investigated for other wide bandgap materials like ZnS [26,27], ZnSe [28,29], ZnO [30-32], diamond [33-35], and SiC [36,37]. Moreover, phosphorus is a common dopant for n-type Si. Therefore, it was necessary to investigate the effects of P^+ incorporation into InGaN thin films.

3. Experimental techniques

Ion beam analysis (IBA) methods have been used in materials analysis for several decades [38]. With these methods, information about composition, uniformity, impurities, and elemental depth profiles of major, minor, and trace elements of materials is obtained through detection of products from the interactions between energetic ions (typically MeV) and electrons and atoms in the target material. For example, in the particle induced X-ray emission (PIXE) technique [39] the characteristic X-rays of target atoms are detected, and in Rutherford backscattering spectrometry (RBS) [40] and elastic recoil detection (ERDA) [41,42], back scattered and recoiled ions, respectively, are detected.

Four InGaN thin films grown, with a GaN buffer layer, on sapphire by metal organic chemical vapour deposition (MOCVD) were used in this work. Rutherford backscattering spectrometry and channeling (RBS/C) measurements were performed using a collimated beam of He^{++} ions of energy 2.084 MeV delivered by a 5UDH-2 Pelletron Tandem Accelerator at National Centre for Physics (NCP) in Islamabad, Pakistan. The samples were mounted on a high-precision goniometer in a vacuum chamber, so that the orientation of the samples relative to the He^{++} beam could be precisely controlled. A silicon particle detector, with an energy resolution of 25 keV, was placed at a scattering angle of 170° . First, three InGaN samples were irradiated, at room temperature, with P^+ ions of energy 700 keV at ion fluences of 1×10^{14} , 1×10^{15} and 7.75×10^{15} ions cm^{-2} . The fourth sample was not irradiated and taken as a reference/control sample. Then the defects in the crystalline structure produced by P^+ ion irradiation on the InGaN thin films were explored using RBS/C and compared with the as-grown sample.

4. Results and Discussion

Rutherford backscattering and channeling (RBS/C) is a nondestructive quantitative technique used to characterize composition, thickness, interface and crystalline quality of an epilayer as a function of depth [43]. Ion channeling has proved to be successful for strain measurement for a wide range of semiconductor thin films including III-nitrides [44]. When the ion beam is directed along a high-symmetry crystal direction, a phenomenon called “channeling” occurs. The incident ion beam will undergo a series of correlated small angle scatterings, and as a result the backscattering events significantly reduce compared with the backscattering events when the ion beam is randomly directed. X_{min} , which is the ratio of the backscattered yield when the ion beam is aligned with a crystallographic axis (Y_A) to that with the random direction (Y_R), is a measure of the crystalline quality of the film [45].

The RBS/C data was acquired using model RC43 analytical software supplied by the National Electrostatics Corp. (NEC) [46]. First, an RBS random spectrum of the as-grown sample was taken and then fitted with the Rutherford Universal Manipulation Program (RUMP) [47] in-order to determine the thickness and amounts of In, Ga, and N (figure 1). The InGaN thin film and GaN buffer were found to be 950 and 1830 nm thick, respectively. The composition of the epilayer was calculated to be $In_{0.05}Ga_{0.47}N$. The ratio of the backscattered events from the aligned spectrum to that from the random spectrum in the same region close to the surface had minimum yield $X_{min} = 58.43 \%$, indicating that the $In_{0.05}Ga_{0.47}N$ layer had a poor crystalline quality. A X_{min} value of $1 - 2 \%$ indicates perfect crystalline quality of the thin film [48]. Figure 2(a) shows a comparison of the RBS channeling and random spectra and figure 2(b) shows the channeling crystal image of the as-grown InGaN thin film. It is clear from figure 2(a) that the InGaN thin films were of poor crystalline quality because the aligned spectrum is very close to the random spectrum, indicating that the dechanneling rate was high.

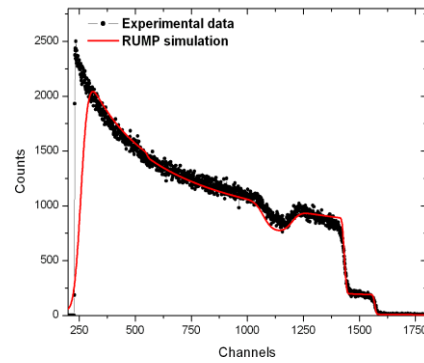


Figure 1. RBS random spectrum of as-grown InGaN sample simulated with RUMP program.

Next, the aligned and random spectra of the three P^+ ion-irradiated samples were acquired and X_{\min} calculated. Figures 3, 4 and 5 below show the aligned and random spectra for each of the three P^+ ion-irradiated InGaN thin film samples with the corresponding channeling crystal image.

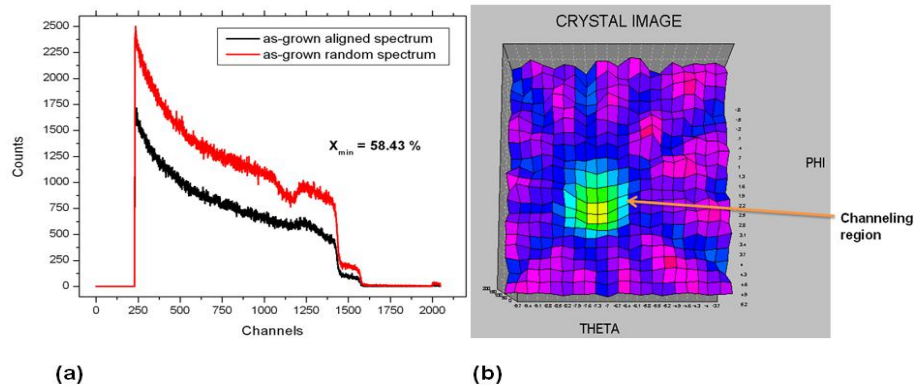


Figure 2. (a) RBS random and channeling spectra, and (b) Channeling crystal image of the as-grown InGaN thin film sample

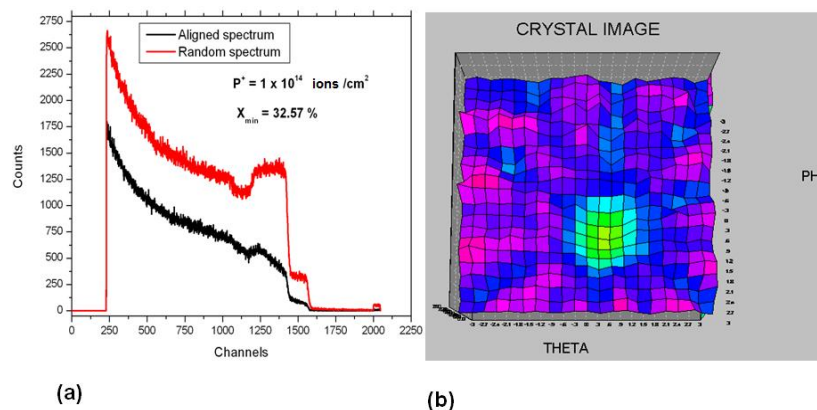


Figure 3. (a) RBS random and aligned spectra, and (b) Channeling crystal image of the InGaN thin film sample at P^+ ion fluence of 1×10^{14} ions/cm²

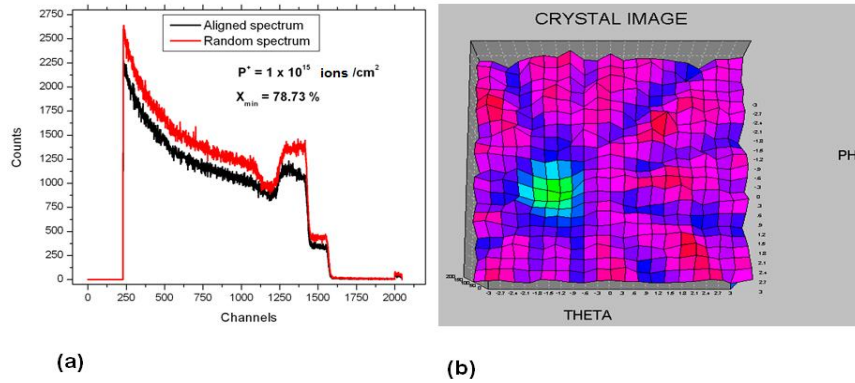


Figure 4. (a) RBS random and aligned spectra, and (b) Channeling crystal image of the InGaN thin film sample at P^+ ion fluence of 1×10^{15} ions/cm²

It is clear from figure 4 and figure 5 that the dechanneling rate increases with the increase in ion irradiation for P^+ fluences of the order of 10^{15} ions/cm². No channeling region was observed at $P^+ = 7.75 \times 10^{15}$ ions/cm² (see figure 5(b)), which was the highest ion fluence reached in this study. However, at $P^+ = 1 \times 10^{14}$ ions/cm² the channeling minimum yield, X_{\min} , is lower than that of the as-grown sample (see figures 2 and 3). This observation suggests that the crystalline quality of the InGaN thin films is improved by low doses of ion irradiation. The same phenomenon has been observed by Takahiro and co-workers on epitaxially grown Ag thin films and Cu films on Si substrates [6,49,50]. Follow up experiments and additional techniques like transmission electron microscopy (TEM) may provide further information on the improvement of crystalline quality induced by ion irradiation.

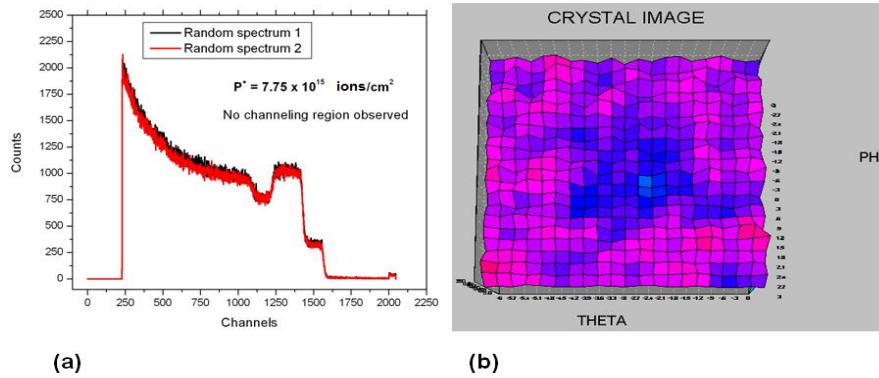


Figure 5. (a) RBS random spectra, and (b) Channeling crystal image of the InGaN thin film sample at P^+ ion fluence of 7.75×10^{14} ions/cm².

Figure 6 shows the variation of the calculated channeling minimum yield, X_{\min} , with P^+ ion fluence for all the four InGaN samples used in this study. As can be seen in figure 6, the channeling minimum yield of the as-grown sample fell from about 60% to about 30% at a P^+ ion dose of 1×10^{14} ions/cm².

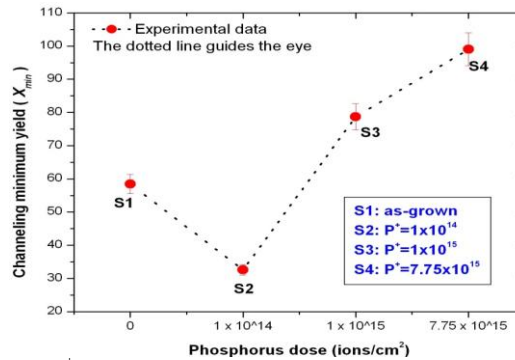


Figure 6. Variation of X_{\min} with P^+ ion fluence. Note the decrease in X_{\min} at $P^+ = 1 \times 10^{14}$

5. Conclusion

In this study, Rutherford backscattering spectrometry and channeling (RBS/C) structural analysis of P^+ ion-irradiated InGaN/GaN thin films grown by MOCVD on sapphire has been presented. We performed RBS/C experiments to find the composition, thickness, crystalline quality and channeling minimum yield, X_{\min} , dependency on ion fluence. Irradiations with 0.7 MeV P^+ ions were carried out to fluences ranging from 1×10^{14} to 7.75×10^{15} ions/cm² at room temperature. The InGaN thin film and GaN buffer were found to be 950 and 1830 nm thick, respectively. The composition of the epilayer was calculated to be $\text{In}_{0.05}\text{Ga}_{0.47}\text{N}$. The crystalline quality of the InGaN thin films was analysed before and after irradiation. Results from RBS/C indicate that the quality of the InGaN thin films is improved by ion irradiation. Irradiation with 0.7 MeV P^+ ions to 1×10^{14} ions/cm² at room temperature reduced X_{\min} from 58.4% to 32.6% at the InGaN surface but X_{\min} increased considerably above this dose. We suggest that the improvement in the crystalline quality of the InGaN thin films could have been brought about by collision-induced atomic re-arrangements at low ion doses.

Acknowledgements

Morgan Madhuku would like to acknowledge the intellectual and material contributions of NCP and TWAS in this work.

References

- [1] Kennedy J, Markwitz A, Trodahl H J, Ruck B J, Durbin S M and Gao W 2007 *J. Electron. Mat.* **36** 472
- [2] Krasheninnikov A V and Banhart F 2007 *Nat. Mater.* **6** 723
- [3] Krasheninnikov A V and Nordlund K 2010 *J. Appl. Phys.* **107** 071301
- [4] Kumar Vinoda, Kumar Vijay, Som S, Purohit L P, Ntwaeaborwa O M and Swart H C 2014 *Journal of Alloys and Compounds* **594** 32
- [5] Som S, Dutta S, Kumar Vijay, Kumar Vinod, Swart H C and Sharma S K 2014 *Journal of Luminescence* **146** 162
- [6] Takahiro K, Takeshima N, Kawatsura K, Nagata S, Yamamoto S and Naramoto H 2002 *Surf. Coat. Technol.* **158–159** 334
- [7] Majid A 2012 Ne and Mn Ion Implantation into AlInN *Ion Implantation* ed Mark Goorsky
- [8] Ambacher O 1998 *J. Phys. D: Appl. Phys.* **31** 2653–2710
- [9] Hill Arlinda 2011 *Ph.D Thesis* Arizona State University
- [10] Orton J W and Foxon C T 1998 *Rep. Prog. Phys.* **61** 1–75
- [11] Kung K and Razeghi M 2000 *Opto-Electronics Review* **8** 201
- [12] Morkoc H (ed.) 2008 *Handbook of Nitride Semiconductors & Devices Vol. 1.* (Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA)
- [13] Razeghi M and McClintock R 2009 *J. Cryst. Growth* **311** 3067
- [14] Strite S and Morkoc H J 1992 *Vac. Sci. Technol. B* **10** 1237
- [15] Pearton S J, Zolper J C, Shul R J and Ren F 1999 *J. Appl. Phys.* **86** 1-78

- [16] Jain S C, Willander M, Narayan J and Overstraeten R Van 2000 *J. Appl. Phys.* **87** 965
- [17] Pearton S J, Abernathy C R, Overberg M E, Thaler G T, Norton D P, Theodoropoulou N, Hebard A F, Park Y D, Ren F, Kim J and Boatner L A 2003 *J. Appl. Phys.* **93** 1
- [18] Cai Xiao-mei, Zeng Sheng-wei and Zhang Bao-ping 2009 *Appl. Phys. Lett.* **95** 173504
- [19] Lai K Y, Lin G J, Lai Y -L, Chen Y F and He J H 2010 *Appl. Phys. Lett.* **96** 081103
- [20] McLaughlin Dirk V P and Pearce Joshua M 2012 *Mat. Sci. Eng. B* **177** 239
- [21] Chun-Feng L, Chung-Chieh C, Min-Hsueh W, Chung-Kwei L and Mau-Kuen W 2013. *Phys. Stat. Sol. A* **210** 1133
- [22] Zolper J C 1997 *J. Cryst. Growth* **178** 157
- [23] Zolper J C, Pearton S J, Abernathy C R and Vartuli C B 1995 *Appl. Phys. Lett.* **66** 3042
- [24] Dahal R, Ugolini C, Lin J Y, Jiang H X and Zavada J M 2010 *Appl. Phys. Lett.* **97** 141109
- [25] Chang S J, Wu L W, Su Y K, Kuo C-H, Lai W C, Hsu Y P, Sheu J K, Chen J F and Tsai J M 2003 *IEEE Trans. Elect. Dev.* **50** 519
- [26] Zhang S, Kinto H, Yatabe T and Iida S 1988 *J. Cryst. Growth* **86** 372
- [27] Ichino K, Yoshida H, Kawai T, Matsumoto H and Kobayash H 2008. *J. Kor. Phys. Soc.* **53** 2939
- [28] Wolf H, Burchard A, Deicher M, Filz T, Jost A, Lauer St, Magerle R, Ostheimer V, Pfeiffer W and Wichert Th 1995 Proceedings of the 18th Intern. Conf. Defects Semicond.: Sendai Japan **Part 1** 309
- [29] Calhoun L C and Park R M 1999 *J. Appl. Phys.* **85** 490
- [30] Heo Y W, Ip K, Park S J, Pearton S J and Norton D P 2004 *Appl. Phys. A* **78** 53
- [31] von Wenckstern H, Benndorf G, Heitsch S, Sann J, Brandt M, Schmidt H, Lenzner J, Lorenz M, Kuznetsov A Y, Meyer B K and Grundmann M 2007 *Appl. Phys. A* **88** 125
- [32] Dongqi Yu, Lizhong Hu, Shuangshuang Qiao, Heqiu Zhang, Song-En Andy Len, L K Len, Qiang Fu, Xi Chen and Kaitong Sun 2009 *J. Phys. D: Appl. Phys.* **42** 055110
- [33] Gheeraert E, Koizumi S, Teraji T, Kanda H and Nesladek M 2000 *Diamond Relat. Mater.* **9** 948
- [34] Koizumi S, Teraji T and Kanda H 2000 *Diamond Relat. Mater.* **9** 935
- [35] Hiromitsu Kato, Toshiharu Makino, Satoshi Yamasaki and Hideyo Okushi 2007 *J. Phys. D: Appl. Phys.* **40** 6189
- [36] Khemka V, Patel R, Ramungul N, Chow T P, Ghezzi M and Kretchmer J 1999 *J. Electron. Mater.* **28** 167
- [37] Handy E M, Rao M V, Holland O W, Jones K A, Derenge M A and Papanicolaou N 2000 *J. Appl. Phys.* **88** 5630
- [38] Tesmer J R (ed) and Nastasi M A (ed.) 1995 *Handbook of modern ion beam materials analysis* (Pennsylvania: Materials Research Society)
- [39] Johansson S A E (ed), Campbell J L (ed) and Malmqvist K G (ed) 1995 *PIXE spectrometry* (New York: John Wiley & Sons Inc.)
- [40] Chu W K, Mayer J W and Nicolet M A 1978 *Backscattering spectrometry* (New York: Academic Press)
- [41] Doyle B L and Peercy P S 1979 *Appl. Phys. Lett.* **34** 811
- [42] Arnold Bik W M and Habraken F H P 1993 *Rep. Prog. Phys.* **56** 859
- [43] Husnain G, Tian-Xiang C, Tao F and Shu-De Y 2010 *Chin. Phys. B* **19** 087205
- [44] Alves E, Pereira S, Correia M R, Pereira E, Sequeira A D and Franco N 2002. *Nucl. Instrum. Methods Phys. Res. B* **190** 560
- [45] Wu M F, Chen ChangChun, Zhu DeZhang, Zhou Shengqiang, Vantomme A, Langouche G, Zhang B S and Yang Hui 2002 *Appl. Phys. Lett.* **80** 4130
- [46] Pollock T J, Haas T A and Klody G M 2009 *AIP Conf. Proc.* **1099** 323
- [47] Doolittle L R 1985 *Nucl. Instrum. Methods Phys. Res. B* **9** 344
- [48] Zhi-Bo D, Kun W, Sheng-Qiang Z, Tian-Xiang C and Shu-De Y 2007 *Chin. Phys. Lett.* **24** 831
- [49] Takahiro K, Nagata S and Yamaguchi S 1996 *Appl. Phys. Lett.* **69** 2828
- [50] Takahiro K, Kawatsura K, Nagata S, Yamamoto S and Naramoto H 2004. *J. Appl. Phys.* **96** 4166

The effect of silver (Ag) dopant on the structural and optical properties of sol gel prepared CdO nanoparticles

MW Maswanganye*, RV Makgobela*, KE Rammutla^{*1}, TE Mosuang* and BW Mwakikunga[#]

* Department of Physics, University of Limpopo, P/Bag x1106, SOVENGA, 0727, South Africa

[#]DST/CSIR National Centre for Nano-Structured Materials, P O Box 395, Pretoria 0001, South Africa

E-mail: erasmus.rammutla@ul.ac.za

Abstract. Samples of undoped CdO and 5% Ag doped CdO nanoparticles were prepared by sol-gel method and annealed for 1h at various temperatures ranging from 100 – 500 °C. The prepared samples were characterized using XRD, UV-Vis, Photoluminescence and SEM. The results showed that for 5% Ag-CdO, the nanoparticles were completely crystallized at the annealing temperature of 400 °C. Almost similar lattice parameters were obtained for both the undoped and Ag doped CdO nanoparticles. Ag dopant appears to reduce the average grain size of CdO, although it increased as the annealing temperature was raised. Introduction of Ag in CdO resulted in the band gap widening and an increase in the recombination rate.

1. Introduction

Nanocrystalline oxides have attracted much attention recently due to their technological importance in solar cells, chemical sensors and liquid crystal displays [1]. CdO is an n-type semiconductor with an FCC structure and has a band gap of between 2.2 eV and 2.7 eV [2]. It has promising applications in a wide range of fields such as solar cells, transparent electrodes, phototransistors and gas sensors [3]. Previous studies have shown that when a dopant has an atomic radius smaller than that of Cd²⁺ (0.97Å), the electrical conductivity of the doped CdO increases and the lattice unit cell compressed [4]. Impurities of high ionic radii are expected to decrease the impurity diffusion towards the p–n junction and are therefore important in the fabrication of solar cells and diodes [5]. In this contribution, focus will be on the effect of doping CdO nanoparticles with Ag (1.26Å) which is a high ionic radius impurity in order to understand the structural and optical properties.

2. Experimental

2.1. Sample preparation

The undoped and 5% Ag doped CdO nanoparticles were prepared using the precursors cadmium acetate dihydrate, 2-methoxyethanol, silver nitrate and ethanalamine. For the undoped CdO, cadmium acetate dihydrate (0.5M) solution was dissolved in 2-methoxyethanol (0.5M) at a constant magnetic stirring for 10 min. For the Ag doped CdO, cadmium acetate dihydrate (0.5M) and 5 wt.% silver

¹ To whom any correspondence should be addressed.

nitrate solutions were dissolved in 2-methoxyethanol (0.5M) with a constant magnetic stirring for 10 min and then the ethanolamine (0.5M) solution was added drop by drop into both the undoped and the doped solutions separately with continuous stirring. The final solutions were stirred for 2 h at room temperature. The solutions were then allowed to age for 24 h and then filtered to dry. The precipitates for both the undoped and doped CdO were annealed at different temperatures (100, 200, 300, 400 and 500 °C) for 1h.

2.2. Sample characterization

X-Ray powder diffraction (XRD) results were obtained using a Philips Analytical X-Ray B.V diffractometer using CuK α ($\lambda = 0.15405\text{nm}$) radiation. The average grain size was calculated using the following familiar Scherrer equation:

$$\beta_{(hkl)} = \frac{K\lambda}{D \cos \theta} ,$$

where $\beta_{(hkl)}$ is the full width at half maximum, D is the grain size, K is a constant λ is the wavelength and θ is the diffraction angle from the XRD pattern.

The strains (ϵ) were determined using the following equation:

$$\epsilon = \frac{\beta_{(hkl)} \cos \theta}{4 \sin \theta} .$$

Ultraviolet-visible (UV-Vis) absorption spectra were recorded using a Perkin-Elmer Lambda 750S UV/Vis spectrometer. Photoluminescence (PL) measurements were carried out using Perkin-Elmer LS 55 Fluorescence spectrometer. The scanning electron microscopy (SEM) images were captured using a JEOL-JSM-7500F microscope at a scale of 100nm and a magnification of $\times 75$.

3. Results and discussions

3.1. X-ray diffraction

The XRD patterns of the 5% Ag doped CdO samples calcined at 100, 200, 300, 400 and 500 °C are shown in figure 1. The patterns of the undoped CdO sample preheated at 400 °C is included for comparison purposes. As can be observed in the figure, the samples heated at 100 and 200 °C do not show any peaks associated with CdO. CdO peaks only start to appear at 300 °C although at this temperature there are still peaks associated with the precursors. The sample completely crystallises to CdO at 400 °C. This behavior was also observed for the undoped sample. The XRD profile of the sample calcined at 400 °C displays only three main CdO peaks at $2\theta = 33.07^\circ$, 35.34° and 55.31° associated with the plane (111), (200) and (220); respectively. This is consistent with the values in the standard card (JCPDS No 05-640) as well as the work by Dakhel AA [2].

The average grain sizes, lattice parameters and strains for undoped CdO and Ag doped CdO samples are shown in table 1. A comparison of the average grain size of the undoped sample (~ 51 nm) and that of the doped sample (~ 40 nm), both preheated at 400 °C shows that the introduction of Ag does reduce the grain growth. The results of the doped sample show that the average grain sizes increase as the temperature increases. This was also observed for undoped CdO [6]. The microstrain is higher for the sample prepared at low temperature and decreases as the temperature increases indicating that the broadening is due to grain size.

The lattice parameters were determined from the interplaner spacing equation and Bragg's law using the X-ray diffraction peaks. As seen from the table, the addition of Ag dopant caused no significant changes in the lattice parameters indicating that the introduction of Ag in CdO nanoparticles did not distort the structure.

3.2. UV-Vis spectroscopy

The UV-Vis absorption spectra of undoped CdO preheated at 400 °C as well as Ag-doped CdO at 400 and 500 °C are shown in figure 2. From the figure, one can observe the absorption excitonic peaks at ~ 557 nm for undoped CdO and ~ 475 nm and ~ 508 nm for the Ag doped CdO samples heated at 400 and 500 °C, respectively. As can be seen from the UV-vis spectra, the excitonic peaks are very broad.

A similar behaviour was observed by J. Liu *et. al.* on pure CdO prepared using thermal decomposition reaction [6]. The corresponding band gap energies of absorption maximum for the samples were calculated using the following equation:

$$E = \frac{hc}{\lambda},$$

where E is the absorption band gap, h is the plank constant, C is a speed of light, λ = absorption excitonic peak wavelength.

The corresponding band gaps obtained were 2.23 eV for the undoped sample preheated at 400 °C and 2.61 eV and 2.44 eV for Ag-CdO preheated at 400 and 500 °C; respectively. The reported value [7] of the direct band gap of bulk CdO is 2.3 which is comparable to the present results.

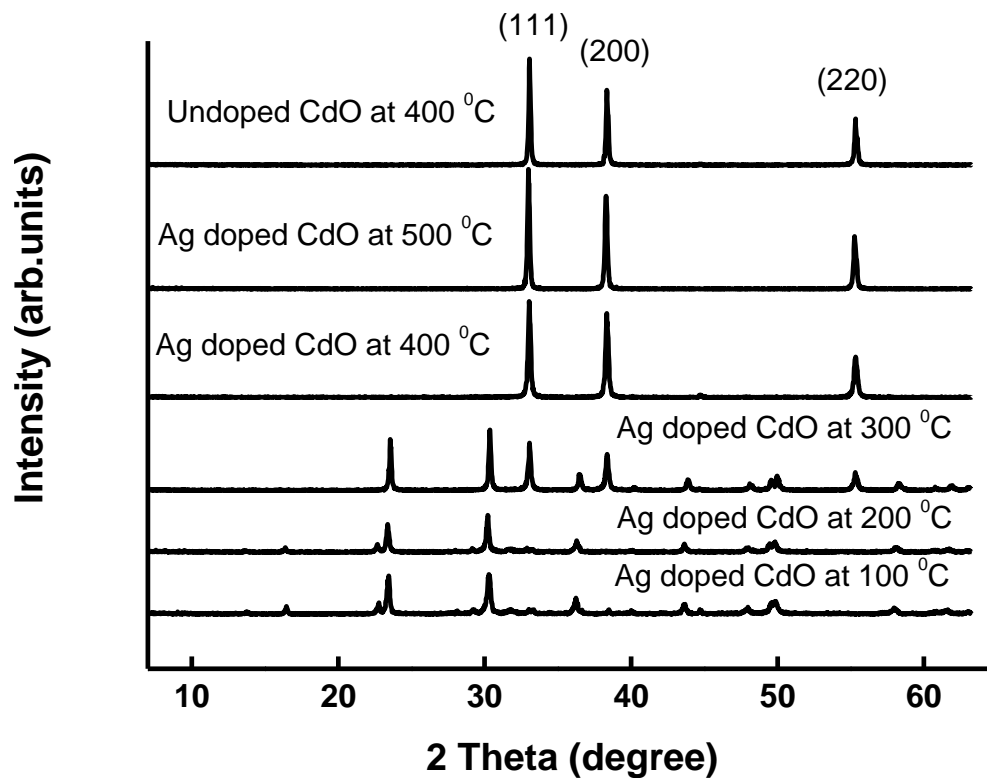


Figure 1. XRD patterns for undoped CdO annealed at 400 °C and 5% Ag doped CdO annealed at 100 °C, 200 °C, 300 °C, 400 °C and 500 °C.

Table 1. Average grain sizes (D), Lattice parameters, energy band gaps and strain (ϵ)

Sample	Temperature °C	Lattice parameters		D (nm)	Band gap (eV) ± 0.03	E ($\times 10^{-3}$)
		a = b = c (Å)	Volume (Å) ³			
Undoped CdO	400	4.690	103.182	51.19	2.23	1.948
5%Ag-CdO	400	4.691	103.228	40.03	2.61	2.490
5%Ag-CdO	500	4.699	103.757	47.03	2.44	2.108

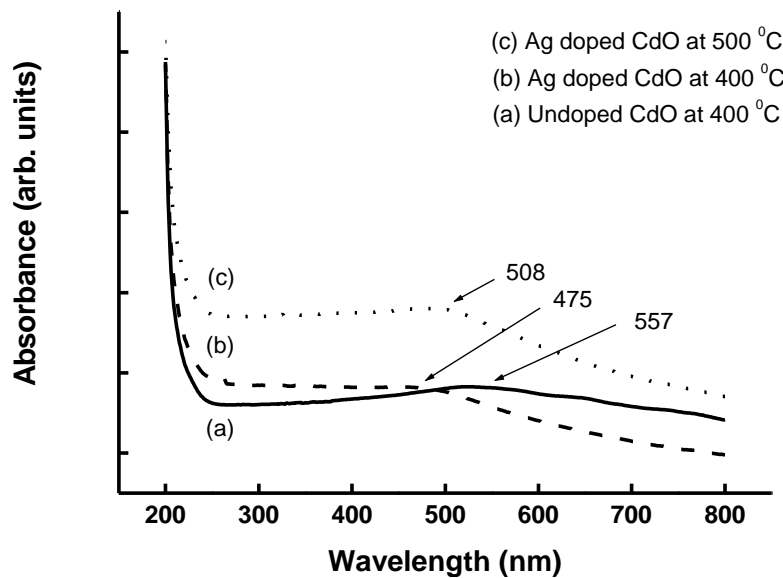


Figure 2. UV-Vis absorption spectra for Undoped CdO annealed at 400°C and 5% Ag doped CdO annealed at 400 °C and 500 °C.

The band gap of Ag doped CdO appears to have blue shifted from that of the undoped sample. This blue shift or band gap widening can be explained in terms of the Moss–Burstein (M–B) effect [8, 9]. According to this M-B theory, the optical absorption edge of a degenerate n-type semiconductor is shifted towards higher energy by the amount which is proportional to the free-electron density.

3.3. Photoluminescence

Previous reports have indicated that bulk CdO does not exhibit luminescence [10]. In the present work the photoluminescence (PL) of pure and Ag doped nanocrystalline CdO were detected. The PL emission spectra of the undoped sample preheated at 400 °C as well as those of Ag-doped CdO preheated at 400 and 500 °C are shown in figure 3. All the spectra were measured using an exciting wavelength of 325 nm. The PL spectra are dominated by the green emission peak around 410 nm. This is comparable to the result obtained for CdO film prepared by spray pyrolysis [11] in which the peak was found at 413 nm. This peak is ascribed to an exciton bound to a donor level. The PL intensity of Ag-doped CdO is higher than that of the undoped sample. Since the PL emission mainly results from the recombination of excited electrons and holes, the higher PL intensity indicates that an introduction of Ag increases the recombination rate. The intensity further increases as the annealing temperature increases.

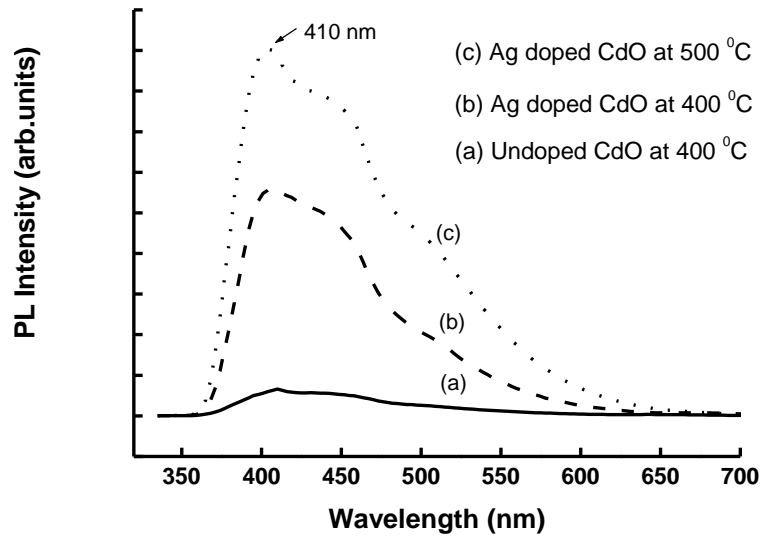


Figure 3. Photoluminescence spectra for Undoped CdO annealed at 400°C and 5% Ag doped CdO annealed at 400 °C and 500 °C.

3.4. SEM

The SEM micrographs are shown in figure 4. The morphologies of the powders are almost spherical in shape. The micrographs also reveal that the particles are agglomerated. Agglomeration could be due to strong hydrogen bonding in the gel network, which is then difficult to remove in the subsequent stages. At very high temperatures, the degree of crystallinity as well as agglomeration further increases considerably. The results are consistent with the XRD results in that as the temperature increases particle size increases and crystallinity improves as indicated by the broadening of the XRD peaks.

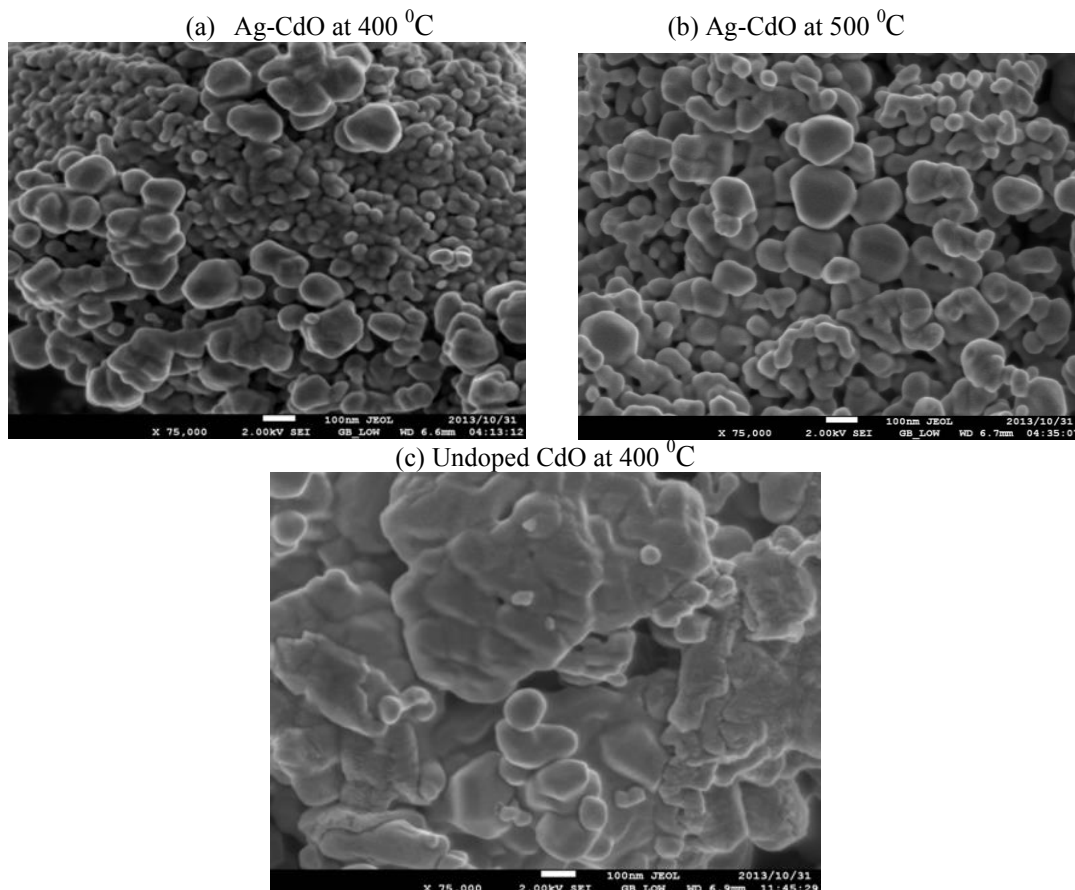


Figure 4. SEM micrographs of Ag-CdO (a) and (b) as well as the undoped CdO (c).

4. Conclusion

XRD results revealed that both the undoped and 5% Ag doped CdO nanoparticle samples had a cubic crystalline structure (FCC structure). Introduction of Ag dopant in CdO had no significant effect on the lattice parameters of CdO nanoparticles. The strains decreased with an increase in annealing temperature indicating that peak broadening is due to grain size. The average grain sizes increased as the annealing temperature was increasing. The introduction of Ag in CdO resulted in the blue shifting of the band gap from that of the undoped sample. The PL intensity of Ag-doped CdO is higher than that of the undoped sample indicating that an introduction of Ag increases the recombination rate. The SEM results confirmed what is observed with XRD in that the average grain sizes of the samples increased with an increase in the annealing temperature.

Acknowledgments

We would like to thank NRF, IBSA and University of Limpopo for financial support and CSIR's Center for Nanostructured Materials for experimental resources.

References

- [1] Jayakrishnan R and Hodes G 2003 *Thin Solid Films* **440** 19
- [2] Dakhel AA, 2008 *Semiconductor Science Technology* **23** 6
- [3] Murali KR, Kalaivanan A, Perumal S, Neelak N and Pillai NJ, 2010 *Journal of Alloys and Compounds* **503**, 350
- [4] Dakhel AA 2012 *Journal of Alloys and Compounds* **539** 26
- [5] Deokate RJ, Salunkhe SV, Agawane GL, Pawar BS, Pawar SM, Rajpure KY, Moholkar AV and Kim JH 2010 *Journal of Alloys and Compounds* **496** 357
- [6] Liu J, Zhao C, Li Z, Yu L, Li Y, Gu S, Cao A, Jiang W, Liu J and Yang C 2011 *Advanced Materials Research* **228** 580
- [7] Skinner AJ and LaFemina JP 1992 *Physical Review B* **45** 3557
- [8] Ferro R, Rodriguez JA, Vigil O and Morales-Acevedo A 2001 *Material Science and Engineering B* **87** 83
- [9] Yoffe AD 1993 *Advances in Physics* **42** 173
- [10] Dong WT and Zhu CS 2003 *Optical Materials* **22** 227
- [11] Seo DJ 2004 *Journal of the Korean Physical Society* **45** 1575

Electronic structures of oxygen adsorption on {110} nickel-rich pentlandite ($\text{Fe}_4\text{Ni}_5\text{S}_8$) mineral surface

P.P Mkhonto¹, H.R Chauke¹ and P.E Ngoepe¹

¹Materials Modelling Centre, University of Limpopo, Private Bag x1106, Sovenga, 0727, South Africa

E-mail: peace.mkhonto@gmail.com

Abstract. Pentlandite ($\text{Co,Fe,Ni}_9\text{S}_8$) is the most abundant iron-nickel sulphide ore containing mineral and has a wide range of applications in industries. The mineral is of commercial importance and can be extracted using floatation processes; one of the processes is by oxidation. This process plays a significant role in forming air bubbles that float the pentlandite mineral. Despite reports that oxidation tend to depress the sulphide minerals, it however, increases the pulp potential during floatation. The present study investigate the clean and oxidised nickel-rich {110} pentlandite surface using *ab-initio* density functional theory (DFT). The Bader analysis have been used to evaluate clean and oxidised surface and suggests that Fe and Ni have 3+ and 2+ (clean) and 4+ and 3+ (oxidised) oxidation state, respectively. Furthermore, when oxygen is adsorbed on the (fcc-hollow site or on Fe-top site) and on Ni-top site, it was found that the surface oxidises as Fe-O-Fe and Ni-O-O, respectively. Oxidation had also shown preferential of iron and we noted a charge transfer from the metals to the oxygen molecule. We also observed that the oxygen (O1) coordinated to the Fe/Ni increases the states of σ_s and π_p bonding orbitals with no π_p^* antibonding orbital present. Furthermore, the oxygen (O2) coordinated to O1 occupies the π_p^* antibonding orbital. The σ_p^* antibonding peak is observed to move closer to the Fermi energy, where on Fe-top site adsorption the peak reside just above the Fermi energy while on Ni-top site adsorption is half occupied.

1. Introduction

The oxidation of pentlandite, naturally and during floatation is an important process to understand extraction of mineral ore. As there is a growing demand for nickel [1], pentlandite, in particular ($\text{Fe, Ni}_9\text{S}_8$) is a principal source of nickel [2]. Pentlandite mineral occurs as intergrowth or in solid solution with pyrrhotite [3] and this makes extraction of mineral difficult. The separation of minerals can be achieved by the use of organic collectors which form important role in rendering the mineral hydrophobic [4].

According to literature, floatation is made easy by creating the rising current that acts like 'hot-air balloons' as such providing the necessary buoyancy to carry selected minerals to the pulp surface. Floatation is the process by which air is bubbled through the slag in a floatation cell. The bubbled air usually oxidises the minerals and also create air bubbles [5]. The alteration, in particular the oxidation both in ore deposits and during the extraction, is an important process used during extraction of mineral ore [3]. In this regard it is of paramount importance to investigate the oxidation process. This will provides information on the chemistry of pentlandite surfaces that may be applicable to the separation of pentlandite mineral. In this study we report on DFT

computational investigation of the adsorption strength, bonding behavior, charge transfer and interaction of oxygen molecule with the {110} surface of nickel-rich pentlandite ($\text{Fe}_4\text{Ni}_5\text{S}_8$) mineral.

2. Computational methodology

In order to investigate the surface- O_2 reaction on nickel-rich pentlandite mineral we perform *ab initio* quantum-mechanical density functional theory [6; 7] calculations and analyze the density of states, Bader charges and charge density difference. Bader analysis is necessary for calculating and assigning electronic charges on individual atoms within the system. This is based upon a grid of charge density values where only steepest ascent trajectories confined to the grid points are used to identify the Bader regions [8].

We use the plane-wave (PW) pseudopotential method with Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional [9] in spin-polarized condition, using VASP code [10]. A {110} surface slab composed of four layers of atoms separated by a vacuum slab of 20 Å was used to mimic the interaction of the adsorbate (O_2) with the repeating upper slab. The ultrasoft pseudopotential is used with a plane-wave basis set, truncated at a kinetic energy of 400 eV since this was found to be sufficient to converge the total energy of the system. Brillouin zone integrations are performed on a grid of 5x5x1 k-points. This is chosen according to the scheme proposed by Monkhorst and Pack [11]. We use the Methfessel-Paxton smearing of $\sigma = 0.2$ eV. Different termination were sampled and only the less reactive (low surface energy) for {110} surfaces were considered. The surface stabilities for different termination are determined by their surface energy, calculated using equation 1:

$$E_{\text{surface}} = \left(\frac{1}{2A}\right) [E_{\text{slab}} - (n_{\text{slab}})(E_{\text{bulk}})] \quad (1)$$

where E_{slab} is the total energy of the cell containing the surface slab, n_{slab} is the number of atoms in the slab, E_{bulk} is the total energy per atom of the bulk and A is the surface area. A low positive value of E_{surface} indicates stability of the surface termination. During adsorption the bottom two layers are kept frozen and the top four layers are allowed to interact with the oxygen molecule. The strength of interaction of the surface with the adsorbate is shown by the adsorption energy, calculated by equation 2:

$$E_{\text{adsorption}} = \left(\frac{1}{n}\right) [E_{\text{slab}} - (E_{\text{system}} + nE_{\text{adsorbate}})] \quad (2)$$

where E_{system} is the energy of the surface slab with adsorbate, E_{slab} is the energy of the surface slab as above, n is the number of adsorbate adsorbed on the surface and $E_{\text{adsorbate}}$ is the energy of the isolated adsorbate molecule. A negative value shows a strong interaction between the adsorbate and the surface, whereas a positive value reveals the opposite.

3. Results and discussion

The nickel-rich pentlandite ($\text{Fe}_4\text{Ni}_5\text{S}_8$) {110} surface is shown in figure 1, and the designated atoms as indicated. The surface was cleaved from the optimised bulk pentlandite structure with space group of *Fm-3m* (225) [12].

Table 1. Calculated surface energy ($\text{eV}/\text{\AA}^2$) and adsorption energies (eV) of oxygen molecule adsorbed on {110} nickel-rich pentlandite mineral surface metals.

Surface	Surface energy	Adsorption energies		
{110}	$E_{(\text{surface})}$	Fcc-hollow	Ni-top	Fe-top
	0.061	-1.891	-0.040	-1.902

In order to describe the adsorption of oxygen on the surface, we consider three distinct adsorption trajectories as shown in figure 1(f). The three adsorptions labelled as fcc-hollow, Fe-top and Ni-top. Their calculated surface and adsorption energies are given in table 1. The adsorption strength of oxygen molecule on the surface is found to be stronger on the Fe atoms. We observed that O₂ adsorption on the Fe-top and fcc-hollow site forms bridging with the iron atoms (Fe-O-Fe), and their adsorption energy is slightly more exothermic on Fe-top than fcc-hollow site. It clearly suggests that energy (0.011 eV) is lost in the chemisorption process of fcc-hollow site. Now, considering the adsorption on Ni-top, we observed formation of a superoxo bonding giving rise to adsorption energy of -0.040 eV. It is evident that the adsorption energy of Fe is more spontaneous than on Ni, this suggests the preferential oxidation of Fe. The oxidation behaviour is in line with that reported by Merape *et al.* [13], where the oxygen molecule prefers the iron.

Figure 1(a) and (b) show the un-relaxed and relaxed surfaces, respectively. We observed that after relaxation, the slab changes from four layers (figure 1(d)) to six layers (figure 1(e)), where layer1 relaxes into three layers as shown in figure 1(e). In order to predict the bonding behavior and charge state of the surface, we consider the spin-polarized density of states (DOS) in figure 1(c) and Bader analysis in table 2. Note that the DOS are plotted only for the top most metals (i.e. Fe and Ni) atoms. The DOS shows that this material has a metallic characteristic since there is no band gap observed at the Fermi energy (E_F). The PDOS of Ni atoms show only one sharp peak at the VB and less contribution at E_F, while the PDOS of Fe has broader peaks around E_F. More importantly, we note that the Fermi energy cut the top of Fe d-orbital (high states at E_F).

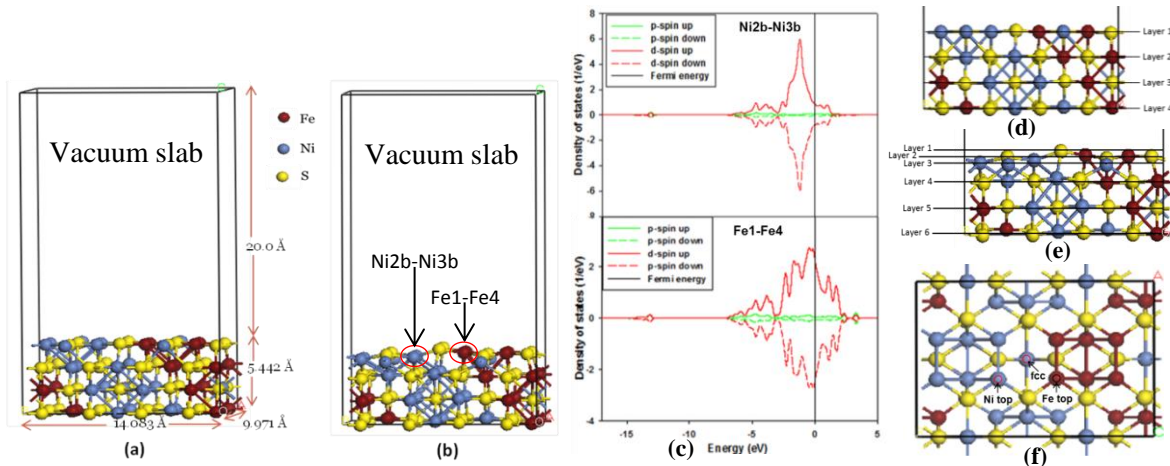


Figure 1: The {110} surface (a) un-relaxed and (b) relaxed stable termination surface, (c) PDOS of the top most clean Fe/Ni atoms, (d) number of layers on the relaxed surface and (e) the three different adsorption geometries investigated: fcc-hollow, Fe-top and Ni-top.

Table 2: The Bader analysis for Fe-O₂ and Ni-O₂ on {110} surfaces: Atom, Bader charge and oxidation state.

Surface	Atom	Clean surface		Fe-top adsorption		Ni-top adsorption	
		Bader charge (e)	Oxidation state	Bader charge (e)	Oxidation state	Bader charge (e)	Oxidation state
{110}	O1	-	-	6.483	-0.48e	6.253	-0.25e
	O2	-	-	6.150	-0.15e	6.118	-0.12e
	Fe1	7.575	+0.43e	7.281	+0.72e	-	-
	Fe4	7.505	+0.51e	7.292	+0.71e	-	-
	Ni2b	9.593	+0.41e	-	-	9.361	+0.64e

The clean iron atoms are observed to have different oxidation states (charges) suggesting that the iron atoms are not charge ordered and there is alternation of the charges. This suggests that the presence of nickel atoms at the octahedral position breaks the charge symmetry of the iron atoms to a small degree, such that the iron with $\text{Fe}^{0.43+}$ has 0.08 electrons less than the $\text{Fe}^{0.51+}$. Such observations of the charge symmetry breakage have been observed on doping of bulk FeS with Ni [14]. The Bader charges for the clean surface suggested oxidation states of Fe^{3+} and Ni^{2+} (table 2). From covalent materials, by analogy it can be deduced that there is a high level of covalent nature predicted for the Ni-S and Fe-S bonds in $\text{Fe}_4\text{Ni}_5\text{S}_8$ pentlandite.

From the calculated Bader analysis of adsorbed system we noted that oxygen (O2) accepts less charges compared to the surface bonding oxygen (O1) (table 2). The charge loss on Ni and Fe atoms increases their oxidation states, suggesting a transformation from Ni^{2+} to Ni^{3+} and Fe^{3+} to Fe^{4+} oxidation states. The oxidation state of Ni^{3+} only occurs with fluoride and oxide [15], thus suggest a Ni^{3+} .

Table 3. The relaxed bond lengths (R , in Å), bond angles (θ , in deg.). The theoretical/ Experimental values are shown in parenthesis for comparison.

Bonds	Oxidation		
	Fcc-hollow site	Fe-top site	Ni-top site
$R(\text{O1-O2})$	1.316	1.316	1.284 (1.284) ^[16]
$R(\text{Fe1-O1})$	1.818 (1.880) ^[17]	1.815 (1.880) ^[17]	-
$R(\text{Fe4-O1})$	1.818 (1.880) ^[17]	1.817 (1.880) ^[17]	-
$R(\text{Ni2b-O1})$	-	-	1.801 (1.728) ^[16]
$\theta(\text{Fe1-O1-Fe4})$	84.68° (81.9) ^[18]	84.96° (81.9) ^[18]	-
$\theta(\text{Ni2b-O1-O2})$	-	-	128.78° (125.3) ^[16]

An interesting observation is that the bond length of Fe1-O1 is shorter compared to Fe4-O1, suggesting that Fe1 losses more charges. These charge transfer is also complimented by the charge density difference (figure 2(f and g)) where we note that a positive electronic cloud that accepts electrons is on oxygen molecule while the negative electronic cloud that donate is on the Ni/Fe atoms. In figure 2(d and e), the relaxed system shows a superoxide or superoxo bonding of oxygen molecule on both Ni and Fe atoms. Superoxo or superoxide is a compound that contains the superoxide anion with the chemical formula O_2^- . Superoxide anion is particularly important as the product of the one-electron reduction of dioxygen O_2 , which occurs widely in nature [19]. The bond distances found (table 3), are in agreement with previous studies and confirmed formation of a superoxo bonding. The cylindrical shape on the oxygen (O2) atom also suggests a superoxo bonding.

Now we present and discuss the spin-polarized density of states (DOS) of oxidised surface (figures 2(a) and (b)). The charge density difference between the oxidised surface system and the clean surface plus the oxygen molecule are visualised using the VESTA software [20]. In order to confirm the charge transfer during adsorption we compute the Bader analysis (table 2). These calculated properties are crucial since they depict the nature of bonding for the adsorbate on the surface as it is not easy to clearly define the bond order from structural analysis. Note that here we only show the DOS and Bader analysis for adsorption on Fe-top and Ni-top, since the fcc-hollow site adsorption showed similar behaviour as the adsorption on Fe-top.

Figure 2(a) and (b) show the DOS for the two adsorption sites, i.e. Fe-top and Ni-top sites. Now, we consider the Fe-top adsorption, the DOS clearly shows that the Fe d-orbital transfer electrons to both bonding and antibonding of the O_2 orbitals, in particular we observe that the LUMO π_p^* antibonding orbital is occupied below the E_F . The Fe PDOS show sharp d-orbital peaks below the E_F compared to the clean surface. Furthermore, the peaks above the E_F merge to form

one peak that has very small splitting peaks. This suggests that electrons are being donated to the oxygen atoms.

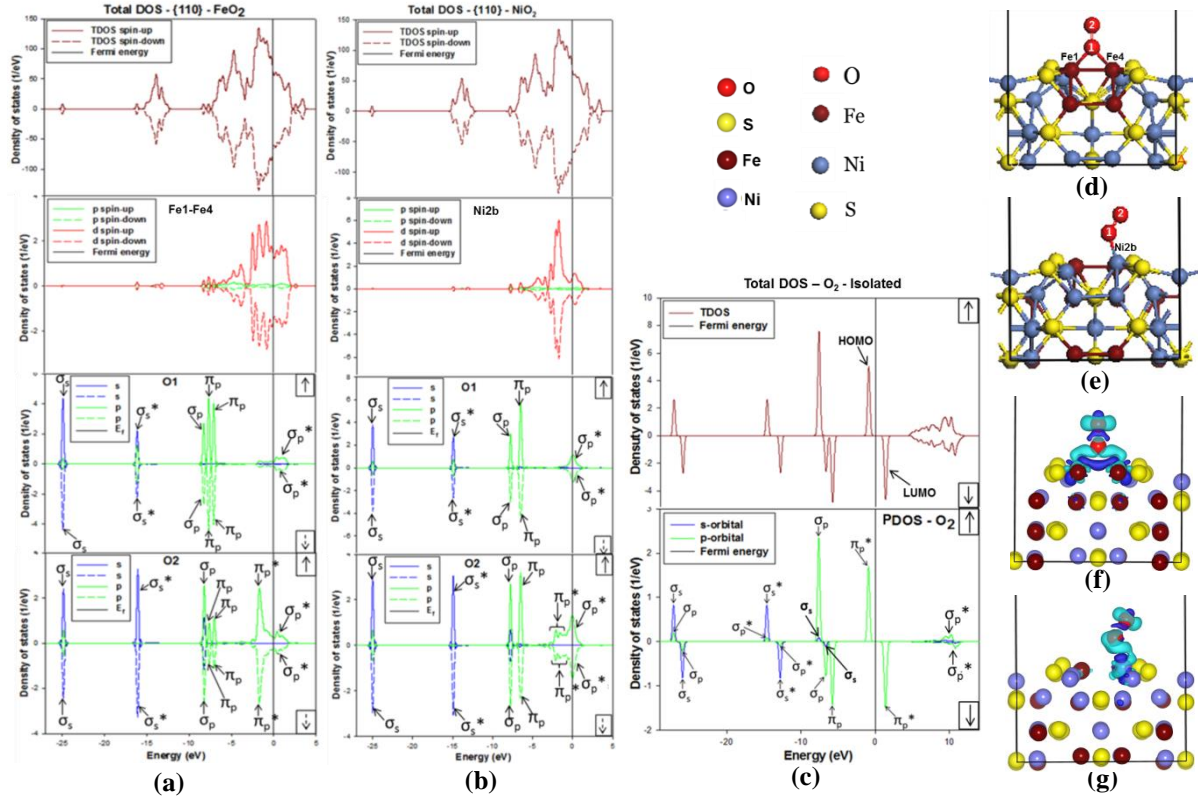


Figure 2: The total density of states (tDOS) and PDOS: (a) adsorption on Fe-top site, (b) adsorption on Ni-top site. Side view of relaxed surface with oxygen molecule: (c) energy levels of the valence orbitals of isolated O_2 molecule, (d) Fe-top site adsorption and (e) Ni-top site adsorption. The sketch of isosurface charge density difference: showing electron transfer between surface metals (Ni/Fe) atoms and oxygen atoms on the surface: (f) Fe-top site (isosurface level = $0.004 \text{ e}\text{\AA}^{-3}$) and (g) Ni-top site (isosurface level = $0.003 \text{ e}\text{\AA}^{-3}$). Cyan represents positive electronic clouds which accept electrons and blue stands for negative electronic clouds which donate electrons.

The PDOS of O_2 shows that σ_s and π_p bonding orbitals and π_p^* antibonding orbitals accept electrons during bonding of the O_2 on the Fe. It is worth noting that O1 is the bridging (bonding) oxygen atom and we observed that it occupies more of the σ_s and π_p bonding orbitals since an increased state are noted with no π_p^* antibonding present below the E_F (figure 2(a) and (b)). The O2 is noted to have the π_p^* antibonding orbitals occupied just below the E_F , suggesting that electrons prefer the outer most shell. Interestingly, we note that the σ_p bonding orbitals on both O1 and O2 is similar to that of the isolated O_2 . Furthermore, we observe that the σ_p^* antibonding orbital moves closer and reside just above the E_F and mix with the d-orbital antibonding.

For Ni-top adsorption (figure 2(b)), the PDOS clearly shows that the sharp d-orbitals peak split into two peaks. Moreover, we noted a formation of a small peak just above the E_F . However, the peak is not completely shifted/formed above the E_F , suggesting that the σ_p^* orbital is not fully occupied and as such show hybridisation of the antibonding d-orbitals and σ_p^* orbital. The PDOS of O_2 molecule (figure 2(b)) of Ni-top shows similar behaviour as Fe-top, however, the difference in this case (Ni-top) is that the π_p^* antibonding orbital has broad peaks, thus reducing their states

height. Moreover, we noted that the σ_p^* orbital on is half occupied at the Fermi energy and merges with the π_p^* orbital.

4. Conclusions

The adsorption energy and bonding nature of oxygen molecule on the {110} surface for flotation purpose was investigated using DFT. We observed that the oxidation favoured the formation of superoxo on Ni-top site and a bridging on iron on Fe-top and fcc-hollow site. Furthermore, we observed a preferential oxidation of iron on fcc-hollow site adsorption. The DOS revealed that the oxygen (O2) not attached to the metal surface has the π_p^* orbital occupied, while the bonding orbitals on O1 have increased states compared to the antibonding orbitals. Interestingly, Bader analysis revealed the amount of charge transferred and those accepted; in which case we noted that O1 accept more charges than O2. The loss of charges on the metals suggested transformation from Ni^{2+} to Ni^{3+} and from Fe^{3+} to Fe^{4+} oxidation states. This study showed how oxygen reacts with iron/nickel naturally and during flotation which is an aspect that may be useful in the floatation of pentlandite mineral.

Acknowledgements

The calculations were performed at the Materials Modelling Centre (MMC), University of Limpopo. We acknowledge the Centre for High Performance Computing (CHPC) for computing resources. Thanks to the South African Mineral to Metals Research Institute (SAMMRI) for financial support.

References

- [1] Ngobeni WA and Hangone G 2013 The effect of using sodium di-methyl-dithiocarbamate as a co-collector with xanthates in the froth flotation of pentlandite containing ore from Nkomati mine in South Africa. *Miner. Eng.* **54** 94–99.
- [2] Borodaev Yu, Bryzgalov S I A, Mozgova N N and Uspenskaya T 2007 Pentlandite and Co-Enriched Pentlandite as Characteristic Minerals of Modern Hydrothermal Sulfide Mounds Hosted by Serpentinized Ultramafic Rocks (Mid-Atlantic Ridge) *Vestnik Moskovsk Universiteta Geologiya* **62** 85–97.
- [3] Rajamani V and Prewitt C T 1973 Crystal chemistry of natural pentlandites *Can. Mineral.* **12** 178–187.
- [4] (n.d.). Retrieved January 13 2013 from http://www.chem.mtu.edu/chem_eng/faculty/kawatra/Flotation_Fundamentals.pdf
- [5] (n.d.). Retrieved october 20 2012 from www.jmeech.mining.ubc.ca/MINE290/Froth%20Flotation.pdf
- [6] Hohenberg P and Kohn W 1965 Inhomogeneous electron gas *Phys. Rev. B* **136** 864–71.
- [7] Kohn W and Sham L J 1965 Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140** 1133–38.
- [8] Bader R F W 1994 *Atoms in Molecules: A Quantum Theory* London: Oxford University Press.
- [9] Heyd J, Scuseria G E and Ernzerhof M 2003 *J. Chem. Phys.* **118** 8207.
- [10] Kresse G and Furthmüller J 1996 Efficient iterative schemes for ab-initio total-energy calculations using a plane-wave basis set *Phys. Rev. B*, **54** 11169–186.
- [11] Monkhorst H F and Park J D 1976 Special points for Brillouin-Zone integrations. *Phys. Rev. B* **13** 5188–92.
- [12] Geller S 1962 *Acta. Cryst.* 1195.
- [13] Merape G 2010 *Fundamental Electrochemical Behaviour of Pentlandite* Pretoria: University of Pretoria.
- [14] Devey A J and de Leeuw N H 2009 *Computer Modelling Studies of Mackinawite, Greigite and Cubic FeS* University College London Department of Chemistry.
- [15] Housecroft C E and Shape A G 2008 *Inorganic chemistry (3rd ed)* Prentice Hall.

- [16] Gutsev G L, Rao B K and Jena P 2000 Systematic Study of Oxo, Peroxo, and Superoxo Isomers of 3d-Metal Dioxides and Their Anions. *J. Phys. Chem. A* **104** 11961–971.
- [17] Pearce C I, Henderson C M B, Telling N D, Patrick R A D, Vaughan D J, Charnock J M, Arenholz E, Tuna F, Coker V S and Van der Laan G 2010 Iron site occupancies in magnetite-ulvöspinel solid solution: A new Approach using XMCD *Am. Mineral.* 1–37.
- [18] Wang L 2000 Photodetachment photoelectron spectroscopy of transition oxide species. **10** 854–957.
- [19] Sawyer D T (n.d.) Superoxide Chemistry McGraw-Hill.
- [20] Momma K and Izumi F 2013 October VESTA: a Three-Dimensional Visualization System for Electronic and Structural Analysis, *National Museum of Nature and Science 4-1-1 Amakubo, Tsukuba Ibaraki 305-0005 Japan and National Institute for Materials Science 1-2-1 Senge, Tsukuba Ibaraki 305-0047 Japan* 1-172.

Influence of zinc acetate concentration in the preparation of ZnO nanoparticles via chemical bath deposition

F V Molefe¹, L F Koao¹, B F Dejene¹, H C Swart²

¹ Department of Physics, University of the Free State (Qwaqwa campus), Private Bag X13 Phuthaditjhaba, 9866, South Africa

² Department of Physics, University of the Free State, P.O. Box 339, Bloemfontein, 9300 South Africa

volksfmv@gmail.com

Abstract. Zinc acetate (ZnAc) concentration was varied to prepare ZnO using the chemical bath deposition method (CBD). X-ray diffraction (XRD), scanning electron microscopy (SEM), ultraviolet visible (UV-vis) and photoluminescence (PL) systems were employed to investigate the influence of ZnAc on the structural, morphological, optical and photoluminescence properties. The obtained crystal structure from XRD analysis was the hexagonal wurtzite structure. It was observed that at higher molar concentration of ZnAc the structure changed to mixture of ZnO and ZnAc. The estimated average crystallite size was 23 nm. There was no effect on the crystallite size as ZnAc concentration was varied. SEM images shows non-uniform agglomerated flower like structures at lower concentrations, at the highest ZnAc concentration flowerlike structures are observed to grow into platelets-like structure. There was no effect on the absorption edges and the energy band gap as ZnAc concentration increases in the UV-vis measurements. PL measurements revealed a slight red-shift in the visible region with an increase in ZnAc concentration.

1. Introduction

Zinc oxide (ZnO) is a metal oxide semiconductor material with wide direct band gap energy equal to 3.37 eV [1]. It is an n – type semiconductor which has a moderate high exciton binding energy of 60 meV at room temperature [2]. This large exciton binding energy indicates that efficient excitonic emission in ZnO can persist at room temperature and higher [3]. From the first half of the 20th century ZnO has been under great study even to-date it is still drawing attention due to its unique structure and morphology [4]. This wurtzite n – type semiconductor has potential applications because of various properties that include good transparency, high electron mobility, and strong room temperature luminescence [5]. Simply because ZnO has good optical properties it can be used in solar cells and light emitting devices [6].

Recently nanotechnology is playing a major role in optimizing metal oxide for various applications in the fields such as electronics, material science and optics [7]. Xu *et al* showed that ZnO has different morphologies ranging from the hexagonal shape up to the rhombohedron shape which can be obtained by different methods of preparation [8].

Several techniques have been employed recently to prepare ZnO such as solution combustion method [9]. Three dimensional flower-like morphologies have been synthesized through facile hydrothermal process by Lu *et al* [10] at 95 °C for 24 hrs who obtained well dispersed flower-like ZnO structures. Koao *et al* [11], reported on the properties of flower-like ZnO nanostructures synthesized using the chemical bath deposition (CBD) technique. They found that by varying the amount of ZnAc there was

no change in morphology and crystal structure, due to high concentration of ZnAc that was used. Following the work of Koao *et al* [11], we prepared ZnO at low ZnAc concentration to study its influence on structural, morphological, optical and luminescence properties using CBD technique. The current study confirms the consistency of CBD method for the formation of flower-like morphology. The CBD method is preferred due to its uses at low temperature and it is relatively inexpensive way of preparation hence it provides a good control of particle size and a uniform morphology.

2. Methodology

The ZnO powders were prepared by the chemical bath deposition method. The following starting materials such as ZnAc ($\text{Zn}(\text{O}_2\text{CCH}_3)_2$), Thiourea ($(\text{NH}_2)_2\text{CS}$) and Ammonia (25% NH_3) were used. The sample solutions were prepared by dissolving 0.1, 0.12, 0.14, 0.16, 0.18 and 0.2 M of ZnAc in 60 ml of H_2O , respectively, and 0.4 M of $(\text{NH}_2)_2\text{CS}$ and also 98.8ml of (25% NH_3), were dissolved into 200 ml of deionized water, respectively. Then a magnetic stirrer was used to stir each of the mixtures for 12 hours at room temperature to ensure homogenous distribution of the solution reagents. An equal volume ratio (1:1:1) was considered for each solution in the following order: 60 ml quantity of ZnAc was first added to the beaker which was placed in the water bath, followed by addition of 60 ml of $(\text{NH}_2)_2\text{CS}$ solution while stirring and finally 60 ml of (25% NH_3) solution was also added while continuously stirring. Water bath was maintained to be at a constant desired temperature of 80 °C. The white precipitates were then formed within 30 seconds after the solution was placed inside the water bath. The beakers were removed from the bath after 10 minutes. The solution was allowed to stabilize for 12 hours in the lab. Finally the solution was filtered and the dry precipitate was collected, washed several times with 60 ml of ethanol and 60 ml of acetone to remove the residue and desiccated for a maximum of 72 hours to ensure that they are dried prior to characterization. The powders synthesized were characterized by XRD, SEM, UV-vis spectroscopy and PL.

3. Results and discussion

3.1. Structural analysis

Figure 1(a) presents the XRD patterns of ZnO prepared at various concentrations of ZnAc ranging from 0.1 to 0.2 M. All the XRD diffraction peaks observed between 0.1 to 0.18 M match with hexagonal wurtzite structure JCPDS card no. 36-1451, with calculated average lattice constants $a = 3.245 \text{ \AA}$ and $c = 5.177 \text{ \AA}$. There was no effect on lattice constants as ZnAc concentration increases. At higher ZnAc concentration 0.2 M there is a mixture of ZnO diffraction peaks marked with circles (°) and ZnAc diffraction peaks marked with asterisks (*). The Scherrer's equation [12], was employed to estimate the ZnO mean crystallite size using the full width at half maximum (FWHM) of all the diffraction peaks. The average estimated mean crystallite size of ZnO powders was $\sim 23 \text{ nm}$, and there was no change in the crystallite size when increasing ZnAc concentration. From figure 1(b) it is observed that the (101) peak of the ZnO powders have slightly shifted to higher diffraction angles when the ZnAc concentration was increased. The slight shift of the peak position indicates a compaction of the unit cell, simply because Zn^{2+} ions which are in excess prefer to occupy the interstitial sites. Furthermore, the peak shift may be due to increased ZnAc concentrations which help to enhance mobility of atoms, subsequently resulting in improved quality of ZnO crystals [13].

3.2. SEM analysis

The SEM images of the ZnO nanoparticles prepared using the CBD method by varying ZnAc concentration are shown in figure 2(a) – (d). In figure 2(a) the image reveals highly agglomerated flower-like particles for the samples prepared at the low concentration of ZnAc 0.1 M. The image in figure 2(b) shows a slight increase in size of flower-like particles for 0.14 M concentration. From figure 2(c) we have SEM image of sample prepared at 0.18 M, which reveals clearly defined flower-like particles. There are noticeable voids and pores on the surface resulting from the evaporation of water molecules when drying the samples. With increasing the ZnAc concentration to 0.2 M we observed an alteration in the structure wherein small flower-like structures are grown on platelets like structure as shown in figure 2(d). We assume that the observed platelets (big chunks) are caused by the unreacted ZnAc. The possible reason for this type of growth is attributed to higher super-saturation at 0.2 M ZnAc concentration since the atoms are readily available for random growth [14].

When comparing the results presented from figure 2(a) – (c) it can be seen that the morphology is non-uniform flower-like structures, except for the sample prepared at 0.2 M which also showed ZnAc impurities in the XRD spectrum. The origin of different morphology in figure 2(d) is due to aggregation of flower-like particles when unreacted metastable ZnAc undergoes self-hydrolysis [15].

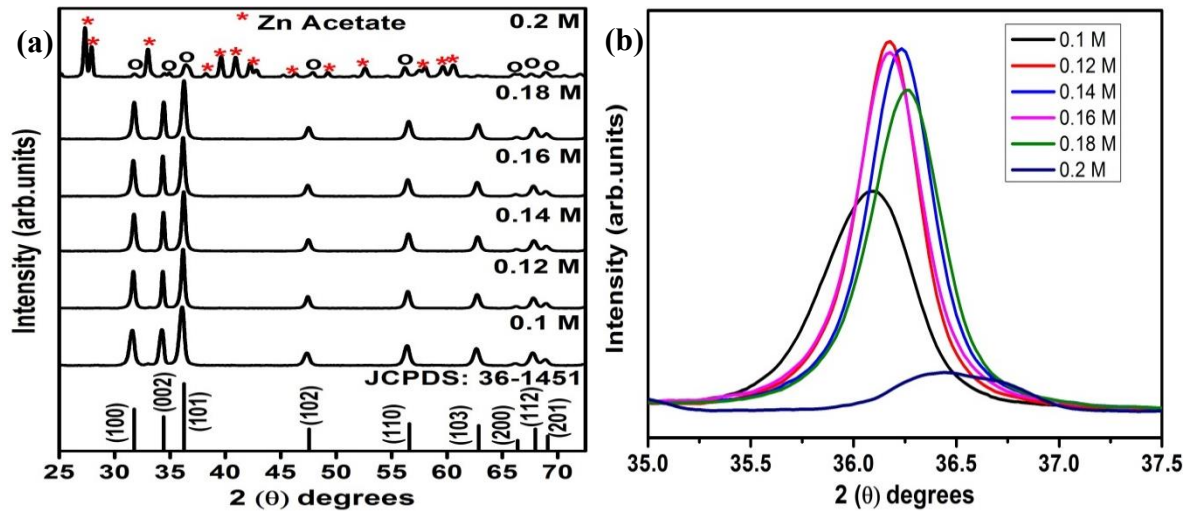


Figure 1. XRD patterns of (a) ZnO nanoparticles prepared at different concentrations of ZnAc using the CBD method as well as the JCPDS standard spectrum, (b) Detail of the (101) peak.

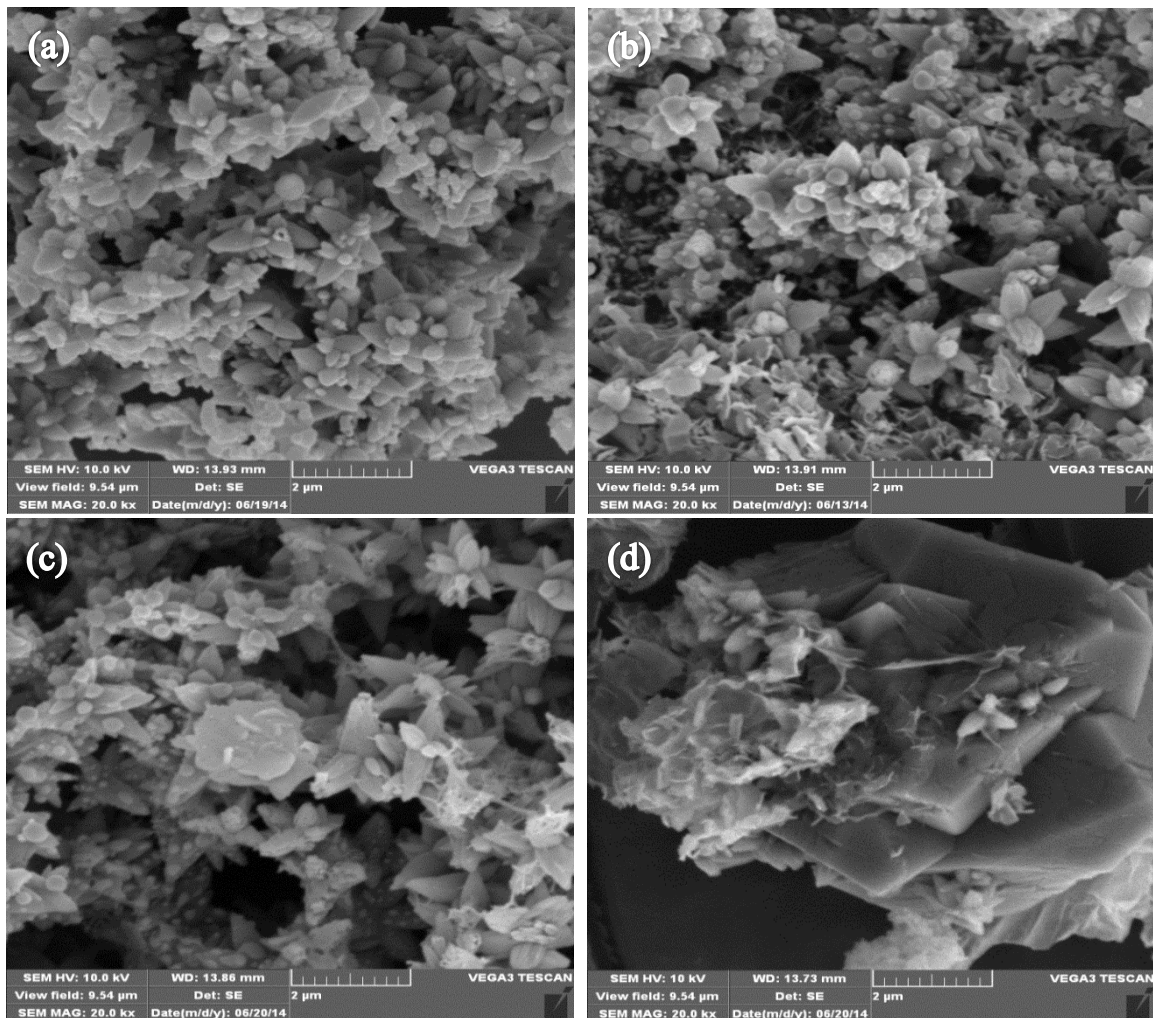


Figure 2. SEM images of ZnO nanoparticles prepared at different concentrations of ZnAc using the CBD method, (a) 0.1 M, (b) 0.14 M, (c) 0.18 M and (d) 0.2 M.

3.3. UV-vis analysis

Figure 3 shows diffuse reflectance measurements of the ZnO nanoparticles in the wavelength range 200 – 800 nm. The percentage reflectance in the visible region increased from ~ 78.5 % at 0.1 M and reached ~ 94.7 % maximum at 0.14 M ZnAc concentration. This indicates that the ZnO absorbs more at low ZnAc concentration. When increasing the ZnAc concentration from 0.1 M to 0.18 M we observed no shift in the absorption band. As illustrated in figure 3 there is a blue shift in absorption band for the sample prepared at 0.2 M ZnAc concentration. This blue shift may be due to ZnAc impurities observed in the XRD spectrum and the formation of platelets like structure observed in the SEM morphology at 0.2 M ZnAc concentration.

The optical band gaps (E_g) for the ZnO powders was calculated from the Kubelka Munk's function $K = (1 - R)^2 / 2R$ [16], by extrapolating the linear portion of the graph to $(Kh\nu)^2 = 0$ as shown in figure 4. It is demonstrated in figure 4 that the optical band gaps (E_g) shows no appreciable change for ZnO powders prepared at 0.1 to 0.18 M ZnAc concentrations. The non-change in band gaps can be because of no change in crystallite size when increasing ZnAc concentration. The average estimated energy band gap for ZnO powders prepared at 0.1 to 0.18 M ZnAc concentrations is ~ 3.20 eV. The inset in figure 4 shows the increase in band gap for the sample prepared at 0.2 M ZnAc, which is equal to ~ 4.60 eV. The increase in band gap energy due to increase in the ZnAc to 0.2 M concentration also relates to ZnAc impurities and structural morphology observed in figure 1 and figure 2(d).

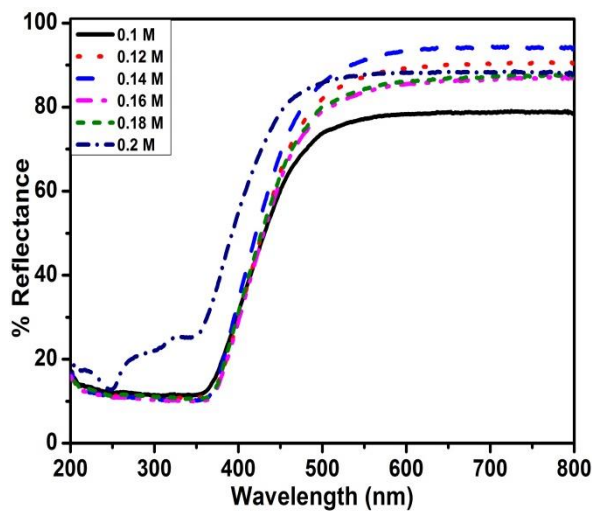


Figure 3. The comparison between diffuse reflectance curves of the ZnO prepared at different concentrations of ZnAc.

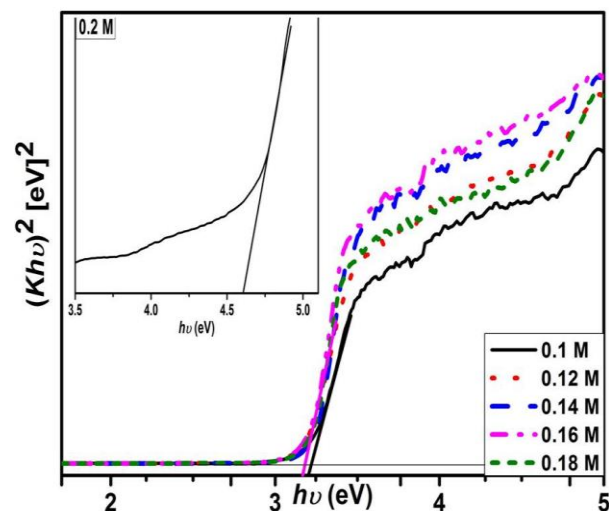


Figure 4. Estimate of the direct energy band gap of ZnO for different concentrations of ZnAc.

3.4. Photoluminescence analysis

Figure 5(a) shows the room temperature PL spectrum of the ZnO flower-like structure prepared at different concentrations of ZnAc. It can be seen that the emission spectrum consists of a broad band in the visible region with two main features. The PL spectrum for the sample prepared at the low ZnAc concentration of 0.1 M exhibits a dominant band emission in the blue region at ~ 450 nm, and a weak green band emission at ~ 528 nm. The luminescence band emission at ~ 450 nm is caused by the electronic transitions from zinc interstitial levels (Zn_i) to the valence band [17]. The observed weak green emission at ~ 528 nm has been attributed by several authors to the oxygen vacancies (V_o) [18], therefore the green emission is due to radiative recombination of photo generated hole with an electron occupying the oxygen vacancy at the local level in the band gap [19]. It is clear that with increasing ZnAc concentration from 0.1 M to 0.18 M the blue emission is quenched while the green emission is enhanced (in proportion), and there is a small red-shift in the green emission. This phenomenon can be clearly seen in the inset in figure 5(a) wherein the spectrum for 0.2 M ZnAc is not included. Further increasing of the ZnAc concentration to 0.2 M enhanced the luminescence

intensity, indicating an increase in the density of defects such as zinc interstitial (Zn_i) and oxygen vacancies (V_O) within the ZnO nanoparticles.

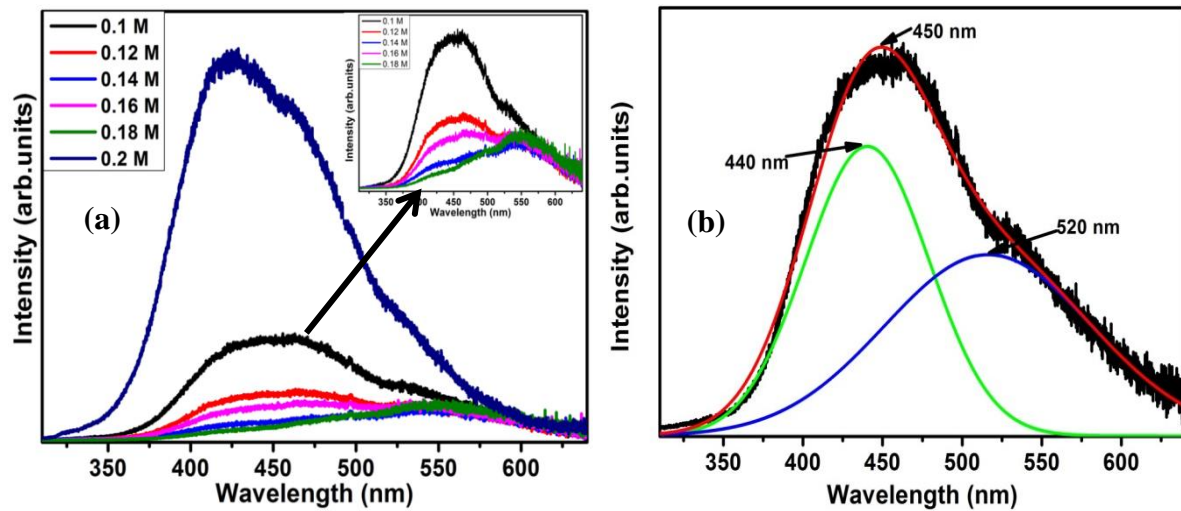


Figure 5. (a) Room temperature PL spectrum for ZnO flower-like structure prepared at different concentrations of ZnAc, (b) De-convoluted spectrum of ZnO sample prepared at 0.1 M ZnAc concentration.

Gaussian fits were only performed on the PL spectrum of ZnO prepared at 0.1 M ZnAc concentration as shown in figure 5(b). The de-convoluted spectrum shows two peaks centred at ~ 440 , and ~ 520 nm. It is well known that a visible emission of zinc originates from intrinsic and structural defects [20]. The blue emission is observed at ~ 440 nm because of surface defects in ZnO, this emission is attributed to Zn interstitials (Zn_i) [21]. The weak green emission observed at ~ 520 nm is well attributed to singly ionized oxygen vacancy in ZnO [22]. The weak green emission at ~ 520 nm typically associated with the recombination of electrons trapped in singly ionized oxygen vacancy (V_O^+) with photo generated holes [23].

4. Conclusion

ZnO nanoparticles were successfully prepared using CBD method by varying molar concentrations of the ZnAc. There was no change in the crystallite size when increasing ZnAc concentration. Not uniform flower-like structure indicated that the morphology is independent from the ZnAc concentration. UV-vis spectroscopy shows no appreciable change in band gap energy of the flower-like structure when the ZnAc concentration increases. PL showed that the emission of the ZnO nanoparticles also depends on the ZnAc concentration.

Acknowledgements

Financial support for M.Sc in nanoscience from National Nanoscience Postgraduate Teaching and Training Programme (NNPTTP) program is highly acknowledged. The authors appreciate the extensive collaboration between UFS, UWC, NMMU and UJ.

References

- [1] Kumar V, Swart H C, Ntwaeaborwa O M, Kroon R E, Terblans J J, Shaat S K K, Yousif A and Duvenhage M M 2013 *Mat. Lett.* **101** 57–60
- [2] Abrarov S M, Yuldashev S U, Kim T W, Kwon Y H and Kang T W 2006 *Optics Commun.* **259** 378–384
- [3] Sun X W, Huang J Z, Wang J X and Xu Z 2008 *Nano Lett.* **8** 1219–1223
- [4] Meaney A J 2010 *On the growth and characterisation of Zinc Oxide* (Dublin City University, Glasnevin, Dublin 9, Ireland) p105
- [5] Awodugba A O and Ilyas A O 2013 *Asian J. Appl. Sci.* **2** 41–44
- [6] Sepet L, Baydogan N, Cimenoglu H, Kayali E S, Tugrul B, Altinsoy N, Albayrak G, Sengel H, Akmaz F and Parlar A 2011 *Def. Diff. Forum* **312–315** 836–841

- [7] Wang Z L 2009 *Mat. Sci. Eng.* **64** 33–71
- [8] Xu L, Guo Y, Liao Q, Zhang J and Xu D 2005 *J. Phys. Chem. B* **109** 13519–13522
- [9] Kumar Vinod, Som S, Kumar Vijay, Kumar Vinay, Ntwaeaborwa O M, Coetsee E and Swart H C 2014 *Chem. Eng. J.* **255** 541–552
- [10] Lu Y C, Wang L L, Xie T F, Chen L P and Lin Y H 2011 *Mater. Chem. Phys.* **129** 281–287
- [11] Koao L F, Dejene F B and Swart H C 2014 *Mat. Sci. Sem. Proc.* **27** 33–40
- [12] Cullity B D 1978 *Elements of X-ray Diffraction* (Addison-Wesley Publishing Company, Inc., London, 1978)
- [13] Manzoor U and Kim D K 2009 *Physica E* **41** 500–505
- [14] Polarz S, Roy A, Lehmann M, Driess M, Kruis F E, Hoffmann A and Zimmer P 2007 *Adv. Fun. Mater.* **17** 1385–1391
- [15] Hu X, Masuda Y, Ohji T and Kato K 2010 *Key Eng. Mater.* **445** 123–126
- [16] Molefe F V, Koao L F, Dolo J J and Dejene B F 2014 *Physica B.* **439** 185–188
- [17] Motaung D E, Kortidis I, Papadaki D, Nkosi S S, Mhlongo G H, Wesley-Smith J, Malgas G F, Mwakikunga B W, Coetsee E, Swart H C, Kiriakidise G and Ray S S 2014 *App. Surf. Sci.* **311** 14–26
- [18] Goswami N and Sharma D K 2010 *Physica E* **42** 1675–1682
- [19] Mohanta A and Thareja R K 2008 *J. Appl. Phys.* **104** 044906
- [20] Yim K G, Cho M Y, Jeon S M, Kim M S and Leem J Y 2011 *J. Korean Phys. Soc.* **58** 520–524
- [21] Dutta S, and Ganguly B N 2012 *J Nanobiotech.* **10** 1–10
- [22] Mazhdi M, Saydi J, Karimi M, Seidi J and Mazhdi F 2013 *Optik.* **124** 4128–4133
- [23] Lin C Y, Wang W H, Lee C-S, Sun K W, and Suen Y W 2009 *Appl. Phys. Lett.* **94** 151909–151909-3

Magnetic susceptibility studies of the $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloy system

B Muchono^{1,2}, C J Sheppard¹, A R E Prinsloo¹ and H L Alberts¹

¹ Physics Department, University of Johannesburg, P.O. Box 524, Auckland Park, Johannesburg 2006, South Africa

² Applied Physics Department, National University of Science and Technology, Box AC939, Ascot, Bulawayo, Zimbabwe

E-mail: alettap@uj.ac.za

Abstract. The magnetic susceptibility as a function of temperature, $\chi(T)$, on a spin-density-wave antiferromagnetic $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloy system in the concentration range $0 \leq x \leq 8.8$ is reported in order to investigate the possibility of quantum critical behaviour in this alloy system. Néel temperatures, T_N , obtained from $\chi(T)$ measurements decrease with Mo concentration and sharply tend towards 0 K at a critical concentration $x_c \approx 4.5$. Antiferromagnetism is suppressed to below 4 K in alloys with $x \gtrsim 4.5$. Alloys in the concentration range $0 \leq x \leq 3.0$ depict an upturn in the $\chi(T)$ curves just above the Néel temperature. The upturn is attributed to local magnetic moments formed around the impurity atoms. The magnetic phase diagram of the alloy system points towards the existence of a quantum critical point at the critical concentration $x = x_c$. The suggestion of quantum critical behaviour in this alloy system from previous electrical resistivity(ρ), Seebeck coefficient (S) and specific heat (C_p) measurements is corroborated in this study.

1. Introduction

Studies of the physical properties of Cr and its alloys around the quantum critical point (QCP) have been a topic of interest in recent years [1-8]. The first experimental investigation of quantum critical behaviour in Cr alloys was done by Yeh *et al.* [8] on a $\text{Cr}_{100-y}\text{V}_y$ alloy system through chemical doping by V solute atoms. Tuning this alloy system through chemical doping across the QCP at the low temperature limit ($T \rightarrow 0$ K) induces an antiferromagnetic (AFM) to paramagnetic (P) phase transition on increasing y [8]. Quantum and not classical fluctuations at $T \approx 0$ K across a QCP cause a de-pairing effect of the electron-hole pairs resulting in an increase in the charge carrier density [3]. This has large effects on several physical properties such as the Hall number, $(qR_H)^{-1}$, and magnetic susceptibility, which both give an indication of the number of charge carriers at the Fermi surface (FS) [8]. Possible quantum critical behaviour was recently reported in the $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloy system through $\rho(T)$, $S(T)$ and $C_p(T)$ measurements [1]. The Néel temperature for this alloy system decreases with x and disappears near a critical concentration $x_c \approx 4.5$ [1]. At the critical concentration the Sommerfeld specific heat coefficient as a function of x , $\gamma(x)$, depicts a peak similar to that observed in the archetypical quantum critical spin-density-wave (SDW) $\text{Cr}_{100-y}\text{V}_y$ alloy system [9].

Local magnetic moments were furthermore reported to exist in Cr doped with non-magnetic elements such as Al [10, 11] and V [12-16]. The $1/\chi$ versus T curves of these alloys follow a Curie-

Weiss (C-W) law just above T_N . The local magnetic moments in these alloys are formed above T_N , unlike those of Cr alloys doped with magnetic elements such as Fe and Mn, which are found to exist in both the P and AFM phases [15, 16].

Investigation of quantum critical behaviour in the $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloy system is extended in this study through $\chi(T)$ measurements as complementary to the previous $\rho(T)$, $S(T)$ and $C_p(T)$ measurements [1]. The present study also investigates the possibility of having local magnetic moments in this alloy system as observed in the $\text{Cr}_{100-z}\text{Al}_z$ [10] alloy system.

2. Experimental

Polycrystalline ternary $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloys in the range $0 \leq x \leq 8.8$ were prepared by repeated arc melting in purified argon atmosphere from 99.999 wt.% pure Cr, 99.999 wt.% pure Al and 99.99 wt.% pure Mo. The alloys were separately sealed into quartz ampoules filled with ultra-high purity argon and annealed at 1300 K for 3 days after which they were quenched into iced water. The composition of Cr, Al, and Mo in each alloy matrix was checked by electron microprobe analyses and the alloys were found to be of good homogeneity [1]. Spark erosion techniques were used to cut and plane the samples into disc shapes of approximately $2 \times 5 \text{ mm}^2$. The magnetic susceptibility measurements were done in the temperature range $4 \text{ K} \leq T \leq 300 \text{ K}$ using the Quantum Design Magnetic Property Measurement System (MPMS). The samples were initially zero field cooled to 4 K and then measurements were done in a constant applied magnetic field of 100 Oe while warming from 4 K to 300 K.

3. Results and discussion

Figure 1 shows representative examples of the $\chi(T)$ curves for the AFM $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ system in the concentration range $0 \leq x \leq 4.3$. The alloys show clear anomalous $\chi(T)$ behaviour in the form of a sharp peak characterised by valley, a systematic upturn and a sudden downturn on cooling through T_N shown by arrows in the figure. The peak is weak and broad for an alloy with $x = 3.8$ (not shown) and becomes smeared out for the alloy with $x = 4.3$. The minimum χ -value of the valley observed above T_N for the alloys in the range $0 \leq x \leq 2.3$ decreases relative to the χ -value at T_N , with an increase in Mo content, but increases on further addition for alloys in the range $3.0 \leq x \leq 4.3$. The downturn observed in these alloys just below T_N is a typical behaviour of the SDW Cr and its alloys [17] and is attributed to a decrease in the density of states at the Fermi energy because of the appearance of a SDW energy gap when the samples are cooled through T_N [17]. On the other hand the downturn observed in the $\chi(T)$ curve just above T_N was also observed in the $\text{Cr}_{100-z}\text{Al}_z$ [10, 11] and $\text{Cr}_{100-y}\text{V}_y$ [12-16] alloys and was attributed to possible local magnetic moments formed around the Al and V impurity atoms, respectively. Just above T_N the curves of these alloys obey a C-W law given by [17]:

$$\chi(T) = \frac{C}{T - T_c}, \quad (1)$$

where T_c and C are the paramagnetic Curie temperature and C-W constant, respectively.

In order to confirm possible C-W behaviour in the present $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloy system $1/\chi$ was plotted against T for alloys in the concentration range $0 \leq x \leq 3.0$ as shown in figure 2. The indicated straight lines represent fits of equation (1) through the data points. The lines give a positive gradient confirming the C-W behaviour in these alloys [16]. It can be concluded that in the present $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloys local magnetic moments seem to be formed in the temperature range $T_N \leq T \leq 190 \text{ K}$ and $T_N \leq T \leq 170 \text{ K}$ for alloys with $x = 0, 0.7$ and 1.1 , respectively, and in the temperature range $T_N \leq T \leq 90 \text{ K}$ for the alloy with $x = 3.0$. Below T_N in figure 2, $1/\chi(T)$ abruptly increases as the C-W behaviour is superseded by antiferromagnetism. Thus, the origin of the upturn at $T > T_N$ for the present alloy system is probably ascribed to the C-W behaviour.

There is also a second upturn, more prominent in figure 1(d), present at low temperatures on the $\chi(T)$ curves of figure 1, presumably arising from upturn effects of a Curie tail [18], connected to oxide impurities [1]. The broken lines in figures 1(a) to (c) were drawn from the P phase and extend into the AFM phase to indicate the behaviour of the curves if they were to remain paramagnetic below T_N .

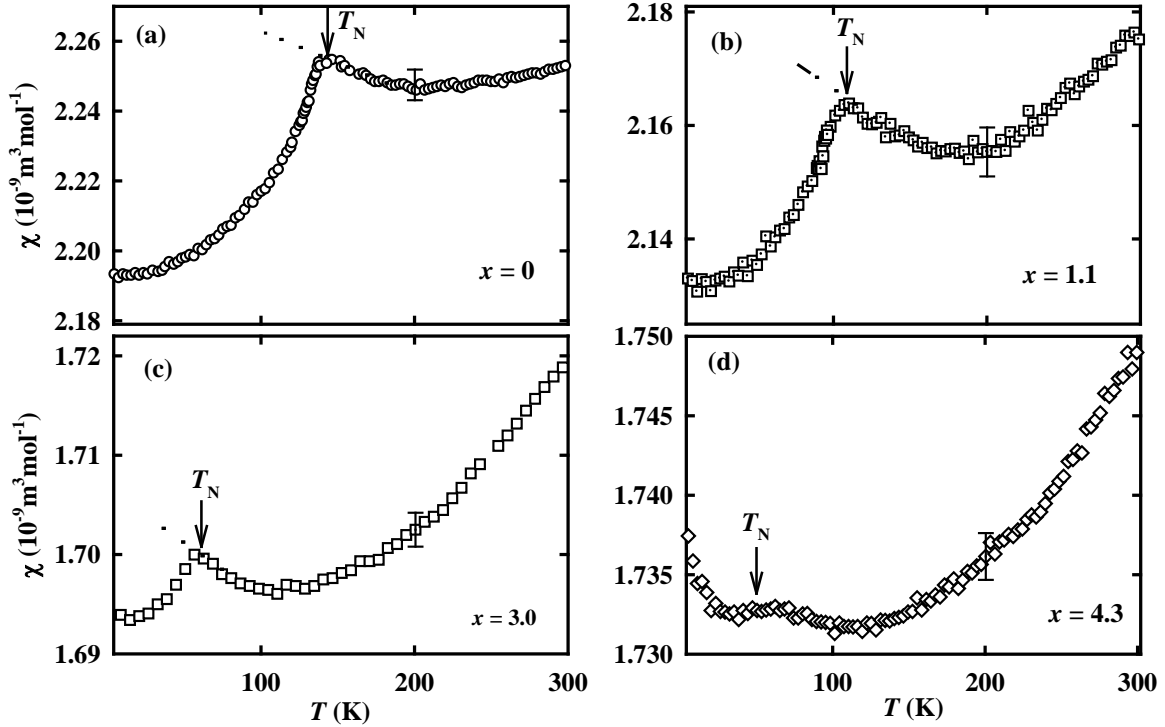


Figure 1. Magnetic susceptibility as a function of temperature, $\chi(T)$, for the $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloys with (a) $x = 0$, (b) $x = 1.1$, (c) $x = 3.0$ and (d) $x = 4.3$. The broken lines in (a) to (c) are lines drawn from the P phase extending into the antiferromagnetic phase to indicate the behaviour of the alloys if they were to remain paramagnetic below T_N . The arrows in (a) to (c) indicate the T_N values obtained from the $\chi(T)$ curves while in (d) the arrow indicates T_N obtained from $\rho(T)$ measurements [1].

The values of T_N were obtained from the temperature where the measured $\chi(T)$ curves deviate from the broken lines. The broad peak anomaly observed on the alloy with $x = 4.3$ makes it difficult to precisely locate T_N . The arrow in figure 1(d) indicates T_N obtained from $\rho(T)$ measurements [1].

Figure 3 shows $\chi(T)$ measurements for the $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloys in the concentration range $4.8 \leq x \leq 8.8$. These alloys did not show an anomalous $\chi(T)$ behaviour associated with SDW effects and were taken to be paramagnetic at all temperatures down to 4 K. $\rho(T)$ and $S(T)$ measurements of these alloys also showed these samples are P down to 2 K [1]. There is a peak observed in the susceptibility data for $x = 6.3$, however, the origin of this peak is unclear at present. It should be noted that $\rho(T)$ and $S(T)$ data for the $x = 6.3$ sample did not show any transitions associated with AFM behaviour at this temperature [1]. The upturn at low temperature observed for alloys with $x = 4.8, 6.3$ and 7.5 is also ascribed to oxide impurity effects, as described above.

Figure 4 shows the magnetic phase diagram of $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloy system obtained from $\chi(T)$ curves of figures 1 and 3. T_N for alloys in the concentration range $0 \leq x \leq 3.8$ decreases with x . Antiferromagnetism disappears for alloys with $x \geq 4.8$. The alloy with $x = 4.3$ shows a slightly higher value of T_N compared to the alloy with $x = 3.8$. The higher value of T_N for this alloy is probably because of an overlap of the Curie tail discussed above and the onset of antiferromagnetism, making it difficult to locate T_N for $x = 4.3$. A similar $T_N(x)$ graph was observed for this alloy system through $\rho(T)$ measurements [1]. The solid line shown in figure 4 is a power law fit through the data points. The fit is of the form:

$$T_N = a(x_c - x)^b, \quad (2)$$

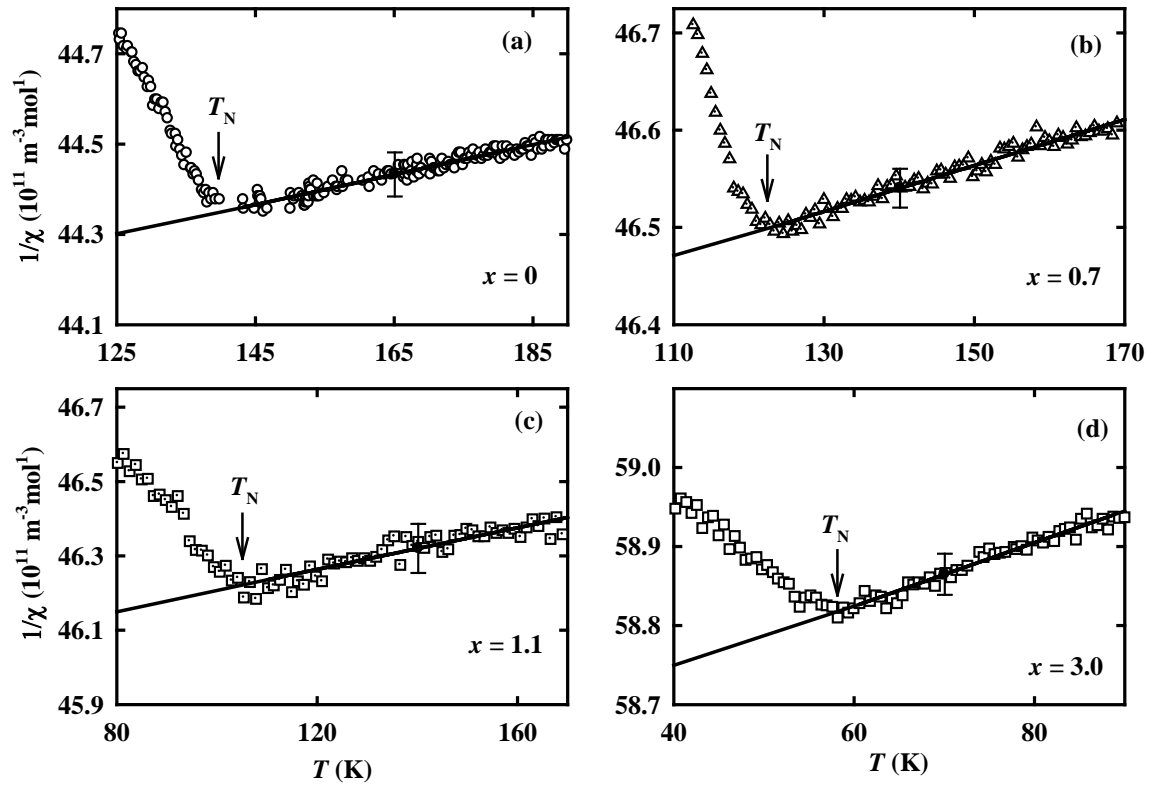


Figure 2. $1/\chi$ as a function of temperature for the $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloys with (a) $x = 0$, (b) $x = 0.7$, (c) $x = 1.1$ and (d) $x = 3.0$. The solid lines are guides to the eye through the data points.

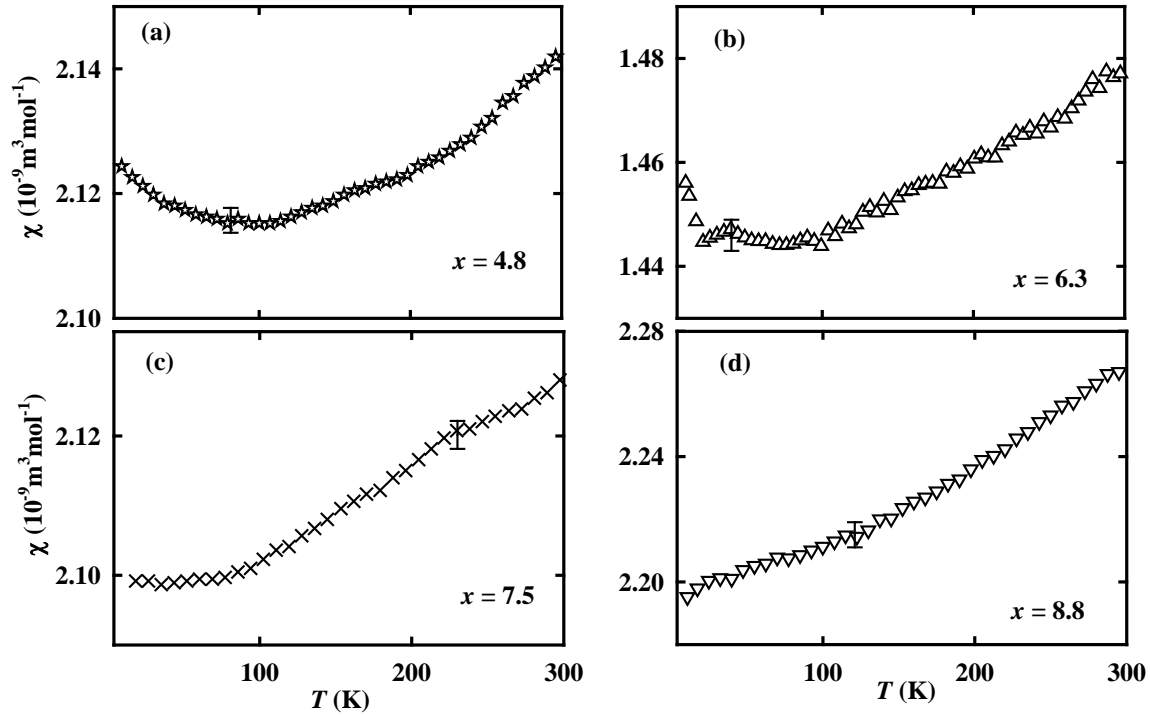


Figure 3. Magnetic susceptibility as a function of temperature, $\chi(T)$, for the $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloys with (a) $x = 4.8$, (b) $x = 6.3$, (c) $x = 7.5$ and (d) $x = 8.8$.

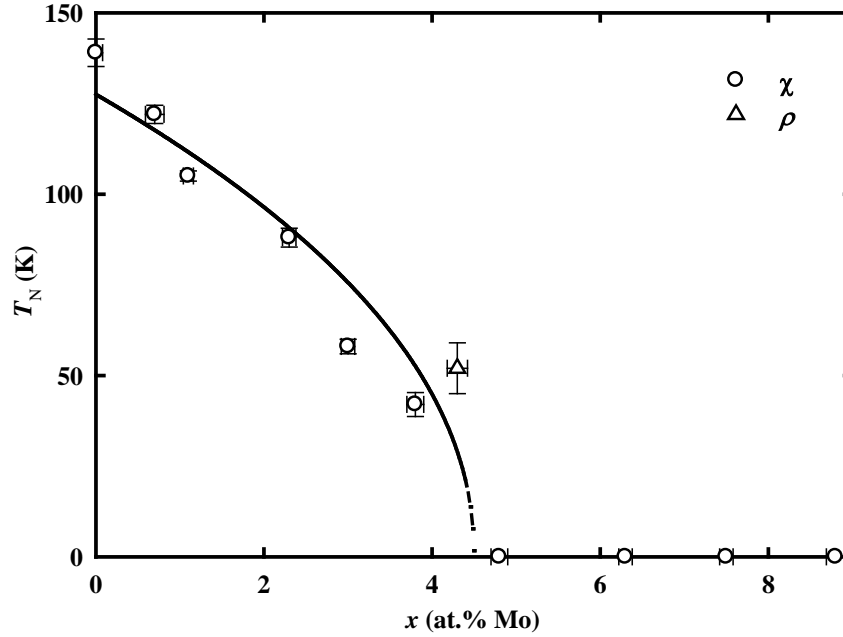


Figure 4. The Néel transition temperature, T_N , as a function of Mo concentration for the $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloy system obtained from (○) the $\chi(T)$ curves of figures 1 and 3, and (Δ) $\rho(T)$ measurements. The solid line is a power law fit through the data points. The error in the data points is represented by the error bars.

where $a = 62 \pm 10$ K is a fitting parameter, $b = 0.5 \pm 0.1$ is a critical exponent and $x_c = 4.50 \pm 0.01$ is a critical concentration where antiferromagnetism disappears. The addition of Mo to the mother alloy $\text{Cr}_{98.4}\text{Al}_{1.6}$ decreases the strength of antiferromagnetism through electron hole pair breaking effects due to electron scattering and a delocalization of the 3-d bands in Cr on introducing the 4-d bands of Mo [17].

The critical exponent b in equation (1) is an important parameter in magnetic systems that approach quantum criticality at low temperatures when an order parameter such as impurity concentration, pressure or magnetic field is varied [19]. The mean field theory for such systems predicts the critical exponent $b \approx 0.5$ [19]. Clearly the value of the critical exponent obtained from the present study is in line with mean field theory. It also compares fairly well with those obtained from electrical resistivity measurements in $T_N \propto (x_c - x)^{0.4 \pm 0.1}$ and $\Delta\rho_0/\rho_0 \propto (x_c - x)^{0.5 \pm 0.1}$ [1] of the same alloy system. These values are also in line with those obtained by Jaramillo *et al.* [3], Lee *et al.* [4] and Reddy *et al.* [5] in the electrical resistivity measurements of pressure tuned Cr and $\text{Cr}_{96.8}\text{V}_{3.2}$ single crystals, and chemical doped $(\text{Cr}_{86}\text{Ru}_{14})_{100-u}\text{V}_u$ polycrystalline alloys, respectively.

4. Conclusion

Magnetic susceptibility studies on the $(\text{Cr}_{98.4}\text{Al}_{1.6})_{100-x}\text{Mo}_x$ alloy system showed that antiferromagnetism in this system is completely suppressed to below 4 K at the critical concentration $x_c \approx 4.5$. Previous $\rho(T)$ results for this alloy system also showed a suppression of antiferromagnetism to below 2 K at the same concentration. At x_c the $\gamma(x)$ for the same alloy system depicts a peak similar to that observed in the archetypical quantum critical spin-density-wave $\text{Cr}_{100-y}\text{V}_y$ alloy system. The value of the critical exponent b of equation (1), obtained from figure 4, is in line with that expected from the mean field theory. Alloys in the concentration range $0 \leq x \leq 3.0$ depict a valley just above T_N indicating the presence of local magnetic moments in these alloys similar to those reported in the $\text{Cr}_{100-z}\text{Al}_z$ and $\text{Cr}_{100-y}\text{V}_y$ alloy systems. The present results and the previous findings point towards the existence of a possible quantum critical point around $x \approx 4.5$.

Acknowledgements

Financial support from the South African National Research Foundation (Grant numbers 80928 and 80626) and the Faculty of Science from the University of Johannesburg are acknowledged.

References

- [1] Muchono B, Sheppard C J, Prinsloo A R E, Alberts H L and Strydom A M 2014 *J. Magn. Magn. Mater.* **354** 222
- [2] Jacobs B S, Prinsloo A R E, Sheppard C J and Strydom A M 2013 *J. Appl. Phys.* **113** 17E126
- [3] Jaramillo R, Feng Y, Wang J and Rosenbaum T F 2010 *PNAS.* **107** 13631
- [4] Lee M, Hussman A, Rosenbaum T F and Aeppli G 2004 *Phys. Rev. Lett.* **92** 187201
- [5] Reddy L, Alberts H L, Strydom A M, Prinsloo A R E and Venter A M 2010 *J. Appl. Phys.* **103** 07C903
- [6] Sheppard C J, Prinsloo A R E, Fernando P R, Venter A M, Strydom A M and Peterson A M 2013 *J. Appl. Phys.* **113** 17E146
- [7] Varella A L S and de Oliveira A J A 2010 *J. Phys.: Conf. Ser.* **200** 012025
- [8] Yeh A, Soh Y A, Brooke J, Aeppli G, Rosenbaum T F and Hayden S M 2002 *Nature* **419** 459
- [9] Takeuchi J, Sasakura H and Masuda Y 1980 *J. Phys. Soc. Jpn.* **49** 508
- [10] Sheppard C J, Prinsloo A R E, Alberts H L, Muchono B and Strydom A M 2014 *J. Alloys and Compounds* **595** 164
- [11] Sousa J B, Amado M M, Pinto R P, Pinheiro M F, Braga M E, Moreira J M, Hedman L E, Åström H U, Khlaif L, Walker P, Garton G and Hukin D 1980 *J. Phys. F: Metal Phys.* **10** 2535
- [12] Fawcett E 1992 *J. Phys.: Condens. Matter* **4** 923
- [13] Hill P, Ali N, de Oliveira A J A, Ortiz W A, de Camargo P C and Fawcett E 1994 *J. Phys. Condens. Matter* **6** 1761
- [14] de Oliveira A J A, de Lima O F, de Camargo P C, Ortiz W A and Fawcett E 1996 *J. Phys.: Condens. Matter* **8** L403
- [15] de Oliveira A J A, Ortiz W A, de Camargo P C and Galkin V Y 1996 *J. Magn. Magn. Mater.* **152** 86
- [16] de Oliveira A J A, Ortiz W A, de Lima O F and de Camargo P C 1997 *J. Appl. Phys.* **81** 4209
- [17] Fawcett E, Alberts H L, Galkin V Y, Noakes D R and Yakhmi J V 1994 *Rev. Mod. Phys.* **66** 25
- [18] Wagner M J, Dye J L, Pérez-Condero E, Bulgas R and Echegoyen L 1995 *J. Am. Chem. Soc.* **117** 1318
- [19] Hertz J A 1976 *Phys. Rev. B* **14** 1165

Residual stress in polycrystalline thin Cr films deposited on fused silica substrates

ZP Mudau¹, ARE Prinsloo¹, CJ Sheppard¹, AM Venter², TP Ntsoane² and EE Fullerton³

¹Department of Physics, University of Johannesburg, PO Box 524, Auckland Park, 2006, South Africa

²Research and Development Division, Necsa Limited, P.O. Box 582, Pretoria 0001, South Africa

³Center for Magnetic Recording Research, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0401, USA

Author e-mail address: cjsheppard@uj.ac.za

Abstract. The Néel temperature in thin film Cr coatings is strongly influenced by dimensionality effects, as well as strain and stress. In an investigation of Cr thin films, with thickness (t) varied between 20 and 320 nm, deposited on fused silica substrates, the T_N values obtained from resistivity measurements indicate an increase with thickness as expected. This study is now extended to investigations of the in-plane stresses in these thin films, using specialised x-ray diffraction $\sin^2\psi$ -method. The in-plane residual strain (ε) present in the coating is determined from the slope of a linear plot through the fractional change in the lattice plane spacing (or Bragg peak position) versus $\sin^2\psi$ plots. Residual stress (σ) are calculated from the ε versus $\sin^2\psi$ data by incorporation the elastic properties of the coating material. The results indicate tensile stresses in all the samples.

1. Introduction

Cr is an archetypical antiferromagnet which forms an incommensurate spin-density-wave (SDW) structure that results from the nesting of the Fermi surfaces [1]. The magnetic transition temperature of bulk Cr is 311 K and has been found to be influenced by applied pressure [1, 2]. For polycrystalline Cr thin films deposited on fused silica substrates by the direct current magnetron sputtering method, a recent study has shown the T_N to be higher than in bulk material and that the T_N increases with the coating thickness (t) [3].

Residual stress in thin films can be present as a result of the deposition and annealing techniques used [4]. Stress in thin films is classified as intrinsic and extrinsic [4, 5]. Intrinsic stress is caused by: i) presence of impurities; ii) voids; iii) partial growth; and v) re-crystallization during deposition [4]. Extrinsic stresses originate from differences in coefficient of thermal expansion (CTE) between the coating and the substrates [4]. The deposition parameters such as evaporation rate, geometry, temperature, and argon sputtering pressure, or impurities in the deposition system can also affect the stress values in both types of stresses [4].

Strain is the fundamental materials parameter that is measured by various techniques [6] based on destructive and non-destructive approaches. In the destructive approaches, strain relaxation is introduced by cutting/slitting actions with the residual strain that existed prior to this action inferred from the extent of the relaxation. The relaxation can be measured by mechanical strain gauges, or laser and optical based dilatometry techniques. Most non-destructive approaches are diffraction based of which x-ray diffraction is most generally used. With this technique, the spacing between atomic planes as an built-in material strain gauge is accurately measured [7]. Stresses are then calculated from the measured strains by incorporation of the elastic properties of the material.

In the x-ray diffraction method the lattice plane spacing, d_{hkl} , is precisely determined for a crystallographic reflection hkl , from the angular-position θ of the measured diffraction peak by application of Bragg's law of diffraction. The lattice plane spacing $d_{hkl} \equiv d_{\phi\psi}$ of crystallites orientated perpendicular to the diffraction vector $\mathbf{L}_{\phi\psi}$, as shown in the geometry of figure 1, is measured at various tilt angles ψ (source/tube geometry always fulfilling the Bragg condition for that reflection). The vector $\mathbf{L}_{\phi\psi}$ along which the strain is measured bisects the angle between the incident and diffracted beam paths. The strain, $\varepsilon_{\phi\psi}$, in the Cr film along $\mathbf{L}_{\phi\psi}$ is defined by [7]:

$$\varepsilon_{\phi\psi} = \frac{d_{\phi\psi} - d_0}{d_0}, \quad (1)$$

where d_0 is the lattice plane spacing in an unstrained Cr thin film. In general defining $\mathbf{L}_{\phi\psi}$ as one of the coordinate axis in figure 1, the strain in the Cr film's coordinates system (\mathbf{S}_i) can be expressed in terms of $\mathbf{L}_{\phi\psi}$ through a second-rank tensor transformation [7]:

$$\varepsilon_{ii} = a_{ik} a_{il} \varepsilon_{kl}, \quad (2)$$

where ε_{ii} is defined along $\mathbf{L}_{\phi\psi}$ and i is usually set to three by substituting the appropriated direction cosines to a_{ik} and a_{il} in the form of ϕ and ψ , ε_{ii} can then be written as (in terms of direction cosines):

$$\begin{aligned} \varepsilon_{\phi\psi} = & \varepsilon_{11} \cos^2 \phi \sin^2 \psi + \varepsilon_{12} \sin 2\phi \sin^2 \psi + \varepsilon_{22} \sin^2 \phi \sin^2 \psi \\ & - \varepsilon_{33} \sin^2 \psi + \varepsilon_{33} + \varepsilon_{13} \cos \phi \sin 2\psi + \varepsilon_{23} \sin \phi \sin 2\psi. \end{aligned} \quad (3)$$

Equation (3) has six unknown variables that can be solved from at least six, but preferably more, independent measurements. Thus ε_{ii} can be solved in the Cr coordinate system (\mathbf{S}_i) [4]. Generally, penetration depth of x-rays into Cr is about 10 μm [8], thus the stress is measured in the surface region of the material only and a biaxial stress condition exists with components \mathbf{S}_1 and \mathbf{S}_2 in the plane of the material with no stress existing perpendicular to the free surface. For this biaxial stress condition it is further assumed that the material is single-phased, randomised crystallite orientations (no texture) and that the grain sizes are substantially smaller than the X-ray gauge volume. Equation (3) can then be simplified to be [5, 7]:

$$\varepsilon_{\phi\psi} = \left(\frac{1+\nu}{E} \right) \sigma_{\phi} \sin^2 \psi - \left(\frac{\nu}{E} \right) (\sigma_{11} + \sigma_{22}), \quad (4)$$

where σ_{ϕ} is the stress in the surface direction \mathbf{S}_{ϕ} shown in Figure 1, E is the Young modulus and ν is the Poisson ratio, σ_{11} and σ_{22} are the principal stress components along \mathbf{S}_1 and \mathbf{S}_2 (shown in figure 1) on the surface of the film.

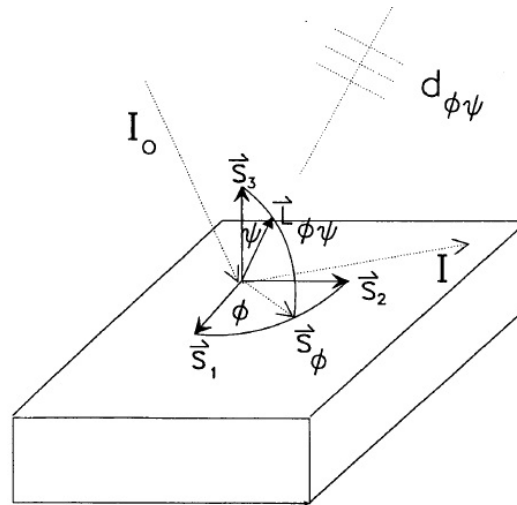


Figure 1: Schematic diagram defining the coordinates system of the sample S_i and the diffraction directions. The diffracted beam I , the incident beam I_o and the normal diffracting planes $L_{\phi\psi}$ lie in the same vertical plane [4].

It is possible to solve the elastic constants empirically, however, the unstrained lattice d_0 value is usually not known. Since $E \gg (\sigma_{11} + \sigma_{22})$ the value of d_0 may be replaced by d_{ϕ_0} , i.e. the value measured at $\psi = 0^\circ$ (perpendicular to the surface). This holds only for the biaxial stress condition applied here. This renders an error of about 0.1% and the values for σ_ϕ , σ_{11} and σ_{22} can be determined within this accuracy. This method thus becomes a differential technique not requiring standard references for calibration of the stress-free lattice spacing [9].

In this paper measurements and results are reported of the strains and stresses in Cr thin films deposited on fused silica substrates. The strains in the thin films were determined using the x-ray diffraction $\sin^2\psi$ technique introduced above.

2. Experimental

The Cr films were prepared using direct current (DC) magnetron sputtering. The films were deposited at a substrate temperature and argon pressure of 973 K and 3 mT, respectively onto fused silica substrates (size: 10 mm² and thicknesses of 0.5 mm). The film thicknesses varied between 20 and 320 nm and were controlled by the deposition times. X-ray diffraction measurements were done on a Bruker D8 Discover diffractometer equipped with a Vantec 500 area detector. The source radiation was Cu K α set at 40 kV and 40 mA in conjunction with a monochromated incident beam 0.8 mm in diameter. The traditional $\sin^2\psi$ method was used to study the residual strain in the material [4]. In strain analyses, changes in the lattice plane spacing are very small (typically 10⁻⁴) which are achieved by selecting reflections with 2θ values larger than 100°. In the case of the Cr polycrystalline films the (310) reflection at 2θ -position 115° was used. Due to the penetration depth of the x-ray beam compared to the coating thickness, the measured strains are averaged through the coating thickness.

3. Results

The fused silica substrates are amorphous and thus do not contribute any coherent diffraction peaks to the diffraction patterns. In figure 2(a) the measured $2\theta_{(310)}$ peak as function of the tilt angle, ψ , is shown for a Cr layer thickness of 160 nm.

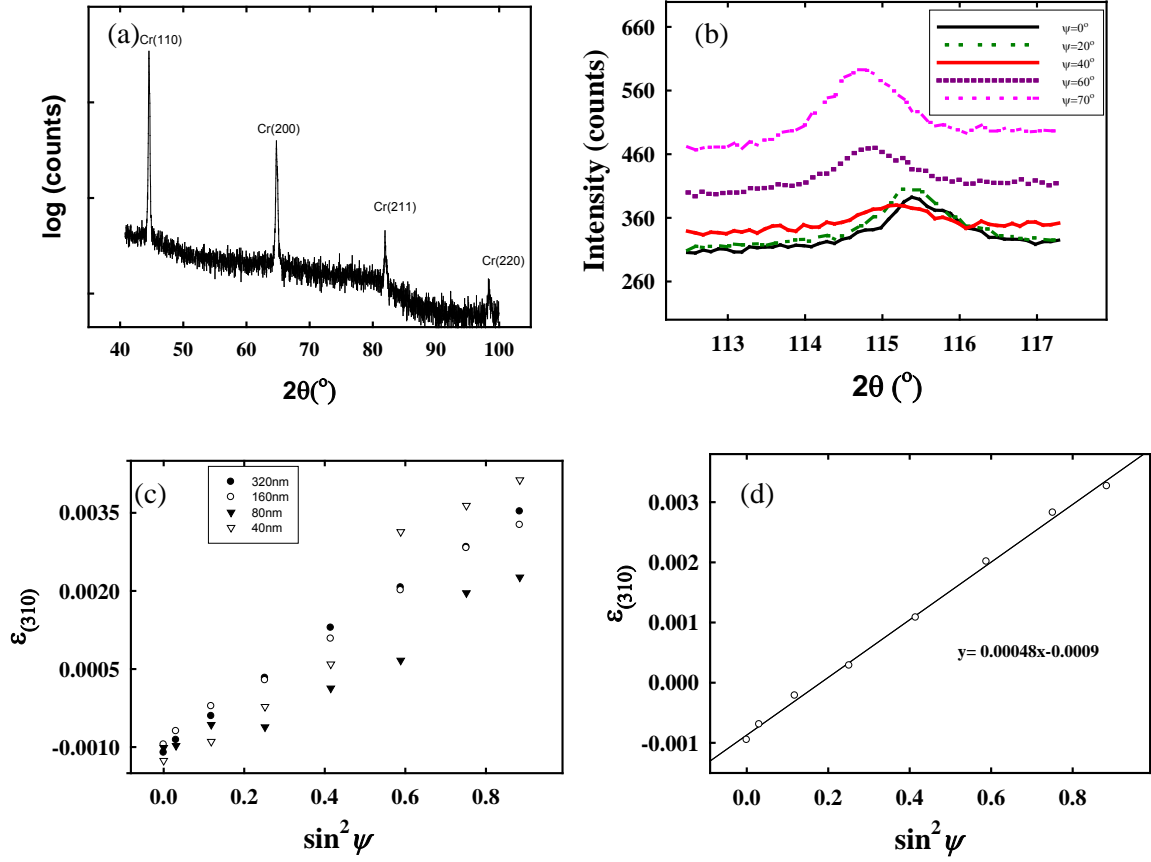


Figure 2: (a) θ - 2θ XRD of the film, for Cr film prepared on fused silica of $t = 320$ nm. (b) $2\theta_{(310)}$ Bragg peak of Cr as function of the tilt angle, ψ , measured on the 160 nm Cr film. (c) Plots of strain (ϵ) as function of $\sin^2\psi$ for different thicknesses of Cr coatings. (d) Strain (ϵ) versus $\sin^2\psi$ for the 160 nm sample showing the linear regression fit, where the gradient was used to calculate the residual stress.

It is clear that the central 2θ -position shifts to lower angles as ψ is increased from 0° to 70° . In figure 2(b) the strain, $\epsilon_{(310)}$, as function of $\sin^2\psi$ is shown for Cr layers with thickness ranging between 40 to 320 nm. At a Cr film thickness of 20 nm no (310) reflection could be observed due to the layer being very thin and not observable within the detection limit of the instrument. The positive slope in the $\epsilon_{(310)}$ as function of $\sin^2\psi$ plots for all Cr film thicknesses indicates that the residual stress in the films are tensile [10] in all cases.

As an example of the typical data analysis procedure, a linear regression fit for $\epsilon_{(310)} = \frac{d_{\phi\psi} - d_0}{d_0}$ as function of $\sin^2\psi$ for the 160 nm thin film is shown in figure 2(c). From equation (4) and assuming that $\sin^2\psi = 0$ at the intercept the unstrained lattice spacing d_0 can be obtained from:

$$d_{\phi 0} = d_0 - \left(\frac{\nu}{E}\right)_{310} d_0 (\sigma_{11} + \sigma_{22}), \quad (5)$$

where the slope of the plot is given by:

$$\frac{\partial d_{\phi\psi}}{\partial \sin^2\psi} = \left(\frac{1+\nu}{E}\right)_{310} \sigma_{\phi} d_0. \quad (6)$$

Using equations (5) and (6) the principal residual stress components σ_{11} and σ_{22} could be calculated. These values are summarized in table 1 and plotted as function of Cr film thickness in figure 3(a). The largest residual stress value is observed in the thinnest Cr coating, from where it decreases sharply to a value 966 MPa for the 80 nm coating followed by a gradual increase to values of 1179 MPa for the 160 nm coating and 1260 MPa for the 320 nm coating.

A qualitative description of the residual stress that develops in coated systems is given. The major contributor is cooling (thermal) stresses associated with the relative differences in their coefficients thermal expansion, α , between the coating (c) and the substrate material (s). As the temperature decreases from the deposition temperature the following scenarios are possible [11]:

- $\alpha_c > \alpha_s$: a tensile stress is generated in the coating,
- $\alpha_c = \alpha_s$: no cooling stress will develop,
- $\alpha_c < \alpha_s$: the resulting cooling stress is compressive.

Since the coefficient of thermal expansion of Cr is larger than that of the fused silica, when the system is cooled from the deposition temperature the Cr will contract more than fused silica, and because of the substantially larger volume of substrate material present, this causes tensile thermal stress [4] in the thin coating. Because the stress values are averaged over the film thicknesses, the result of figure 3(a) indicates that the stresses are influenced by two competing mechanisms: i) the film thickness; ii) the CTE differences. The interactive stress due to the first mechanism is dominant for thicknesses smaller than 80 nm.

Figure 3(b) shows the relationship between σ_{11} as a function of T_N , obtained previously [3]. There is an increase of the T_N with an increase in residual stress values for the samples with a thickness larger than 80 nm. The stress of the 40 nm sample is higher than that of the other samples in the series, this corresponds to an increase seen in the T_N value of the 40 nm sample [3]. It is evident that stress increases the T_N values of the samples series in correspondence to what is expected [2].

4. Conclusions

The residual strain and stress analysis of Cr thin films coated on fused silica substrates, with thickness varying between 40 nm to 320 nm, was done using the x-ray diffraction $\sin^2\psi$ technique. In general, the results reveal a large residual stress when the film is initially deposited, that drops off sharply with thickness, where after the stress values again increase more gradually for film thicknesses larger than 80 nm. As expected, the stresses remain tensile regardless of the film thickness due to the CTE differences between the film and substrate. Overall, there a link is observed between the T_N values and the residual stress values for samples that have thickness larger than 80 nm.

Table 1: Principal stress component results for Cr thin film thicknesses 40 nm to 320 nm deposited on fused silica substrates.

Thickness (nm)	Stress tensors (MPa)	
	σ_{11}	σ_{22}
40	1640 ± 128	1452 ± 129
80	966 ± 50	922 ± 50
160	1179 ± 18	1149 ± 18
320	1260 ± 18	1258 ± 18

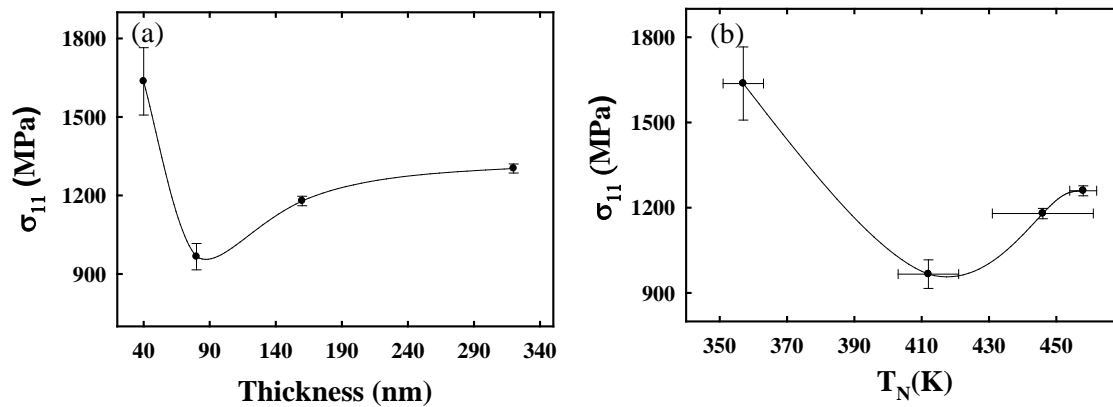


Figure 3: (a) Residual stress (σ_{11}) versus Cr thin film thickness for coatings on fused silica. (b) Plot of residual stress (σ_{11}) versus Néel temperature (T_N). The solid lines in both figure (a) and (b) serve as guides to the eye to show trends.

5. Acknowledgement

The authors wish to thank the National Research Funding of South Africa for financial support towards this study (Grant number 80928 and 806260) and the Faculty of Science from the University of Johannesburg are acknowledged, as well as for the investment made under the National Equipment Programme (NEP, grant 60819) towards the procurement of the Discover instrument and Necsca for its operation and supervision in its use.

References

- [1] Fawcett E, Alberts H L, Galkin V Y, Noakes D R and Yakhmi J V 1994 *Rev. Mod. Phys.* **66** 25
- [2] Zabel H 1999 *J. Phys. Condens. Matter* **11** 9303
- [3] Sheppard C J, Prinsloo A R E, Fernando R R, Mudau Z P, Venter A M and Fullerton E E 2013 *SAIP Conf. Preceedings* Submitted
- [4] Noyan I C, Huang T C and York B R 1995 *Crit. Rev. Solid State Mater.* **20** 125
- [5] Doerner M F and Brennan S 1988 *J. Appl. Phys.* **63** 126
- [6] Jian Lu 1996 *Handbook of measurement of residual stress*, Society for Experimental Mechanics Inc., Ed., The Fairmont Press Inc. ISBN 0-88173-229-X
- [7] Cullity B D 1978 *Elements of X-ray Diffraction* Massachusetts, Addison Wesley p447.
- [8] Van der Voort G and Friel J J 1996 *Developments in Materials Characterization Technologies*, ASM International, Materials Park, OH, p103.
- [9] Chiou S Y and Hwang B H 1998 *J. Phys. D: Appl. Phys.* **31** 349
- [10] Uchida H, Kiguchi T, Saiki A, Wakiya N, Ishizawa N, Shinozaki J and Mizuti N 1999 *J Ceramic Soc. Jpn.* **107** 606
- [11] Oladijo O P, Venter A M, Cornish L A and Sacks N 2012 *Surface and Coatings Tech.* **206** 4725

Influence of solvent casting and weight ratios on the morphology and optical properties of inorganic-organic hybrid structures

T F G Muller^{1,*}, A Ramashia¹, D E Motaung², F R Cummings¹, G F Malgas¹ and C J Arendse¹

¹Department of Physics, University of the Western Cape, Private Bag X17, Bellville 7535, South Africa

²DST/CSIR Nanotechnology Innovation Centre, National Centre for Nano-Structured Materials, CSIR, P. O. Box 395, Pretoria, 0001, South Africa

E-mail: tmuller@uwc.ac.za

Abstract. Organic photovoltaic (OPV) solar cells have the advantage of simple processing, low cost, semi-transparency, high-mechanical flexibility, and light weight. Investigations into OPV solar cells using bulk heterojunction (BHJ) structures have driven extensive and successful efforts to enhance their power conversion efficiencies (PCE). It has been demonstrated that nanoscale morphology is essential for improved transport of charge carriers in the OPV cell containing poly (3-hexylthiophene) (P3HT) and (6,6)-phenyl C₆₁ butyric acid methyl ester (PCBM) structures, and for enhancing its efficiency. Inorganic semiconductors such as zinc oxide (ZnO) have been incorporated into the BHJ structures due to their high carrier mobilities; it is inexpensive and ZnO-nanorod arrays offer an excellent controllable transport path. In this contribution the effect of solvent to control the degree of mixing of the polymer, fullerene and ZnO nanoparticles components into a hybrid inorganic-organic structure, is investigated. Evolution of the domain size, structure and optical properties of hybrid ZnO:P3HT:PCBM thin films spin-coated from different weight is studied, using High Resolution Transmission Electron Microscopy (HRTEM), UV-visible spectroscopy, X-ray diffraction (XRD) and spectroscopic ellipsometry (SE).

1. Introduction

Solution based bulk heterojunction photovoltaic devices incorporate p-type conjugated polymers in combination with fullerenes, with the aim of boosting the efficiencies of organic photovoltaic devices [1]. In such hybrid devices the conjugated polymer may be combined with inorganic semiconductor nanoparticles such as zinc-oxide (ZnO) to take advantage of its high electron mobility [2]. In this contribution we use as active layer regioregular poly (3-hexylthiophene), commonly known as P3HT, in which (6,6)-phenyl C₆₁ butyric acid methyl ester (PCBM) has been blended with a set percentage weight ratio. We further incorporate nanoparticles of ZnO into the blend with the aim of studying the degree of mixing of the polymer, fullerene and ZnO components. Of importance is the domain size,

* To whom any correspondence should be addressed.

structure and optical properties of the thin films when ZnO:P3HT: PCBM is spin-coated from different weight ratios onto substrates.

2. Experimental Procedure

ZnO nanoparticles were prepared by hydrothermal method as discussed before [3]. The photoactive ingredients were prepared by mixing weight percentages of 5 mg P3HT with 5 mg PCBM, and dissolving in dichlorobenzene. The hybrid material was prepared similarly with 5 mg and 10 mg ZnO mixed into the photoactive blend. The solutions were then spin-coated onto Corning 7059 and c-Si wafer substrates at 2000 rpm for 30 s, and dried for 30 min on a hot plate at 50°C in order to evaporate the solvent. The crystal structure of the thin films was studied with a Panalytical X'pert PRO diffractometer with Cu K α radiation ($\lambda = 0.154$ nm) in $\theta - 2\theta$ mode, whereas the crystal structure of the ZnO powder was determined using a Bruker system. ZnO mixed in ethanol, and the hybrid ZnO: P3HT: PCBM in dichlorobenzene were inspected with a FEI Tecnai F20 FEGTEM after drying of the solvent. UV-visible absorption spectroscopy and spectroscopic ellipsometry (SE) were carried out with a Perkin-Elmer Lambda 750 UV-Vis and J. A. Woollam M2000 variable angle ellipsometry (VASE) system respectively. A Cauchy model [4] was used to extract the thickness of the films, and a B-Spline model [5] was used to calculate the optical functions of the thin films, assuming a homogeneous material.

3. Results and Discussion

The crystal structure of the as-synthesized ZnO was investigated with XRD and HRTEM, as shown in figures 1 and 2. The intense (100) and (101) diffraction peaks suggest a highly crystalline material, with $\langle 100 \rangle$ and $\langle 101 \rangle$ the main preferred growth orientations, in a typical hexagonal wurzite structure with average grain sizes of 19.2 nm and 17.1 nm respectively. The d-spacings were calculated using Bragg's Law from the (100) and (101) diffraction peaks and were found to be 0.281 nm and 0.248 nm respectively, with lattice constants $a = 0.3245$ nm and $c = 0.5301$ nm. These findings agree well with the data from the Joint Committee on Powder Diffraction Standards (JCPDS: 36-1451). The HRTEM image shows a random collection of small and larger pyramid-shaped structures intermixed together. No characteristic peaks from other impurities by XRD and Energy Dispersive X-ray Spectroscopy (EDS) were detected, indicating a high purity product.

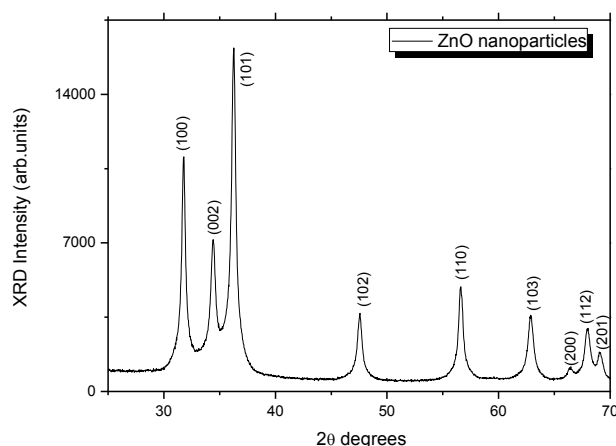


Figure 1. XRD pattern of ZnO nano particles.

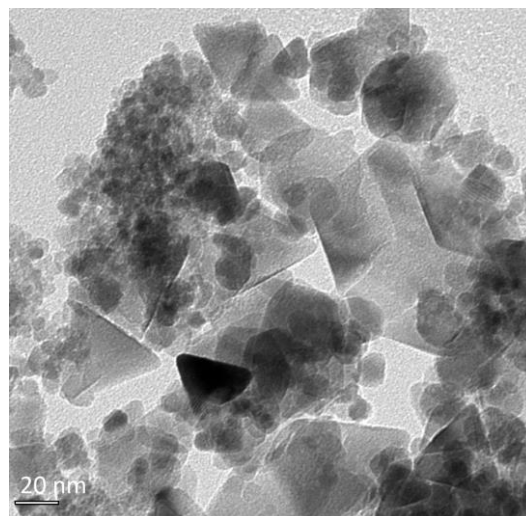


Figure 2. HRTEM micrograph of ZnO nanoparticles.

In order to take advantage of large film roughness to achieve high efficiency for devices it is advisable to spin-cast films of P3HT and PCBM blends prepared with dichlorobenzene as solvent [6]. A donor to acceptor ratio of 1:1 by weight was chosen, as photovoltaic devices utilizing P3HT: PCBM = 1:1 in dichlorobenzene blends gives the best device performance [7]. The XRD patterns of films prepared by spin-coating polymer blends and hybrid inorganic-organic are shown in figure 3. The pristine P3HT film exhibits a (100) peak around $2\theta = 5.4^\circ$, which corresponds to an ordered self-organized lamellar structure with an interlayer spacing formed by parallel stacks of P3HT main chains separated by regions filled with alkyl side chains[8]. It is evident that a peak shift to smaller angle occurs when PCBM is mixed into the polymer, with a hump occurring due to the broadening of the peak; this indicates an interference with the ordering of the P3HT due to the introduction of PCBM into the matrix.

Table 1. Structural and morphological parameters of spin-coated films.

ZnO:P3HT:PCBM	thickness	<100>	
wt ratio	(nm)	2θ position ($^\circ$)	XRD fwhm ratio
0:1:0	84.5	5.39	1.00
0:1:1	44.4	5.35	-
1:1:1	47.6	5.43	1.43
2:1:1	48.3	5.57	1.09

The diffraction peak information for this peak is summarized in table 1, for the P3HT:PCBM blend and hybrid ZnO:P3HT:PCBM thin films. The ratio of the full-width at half-maximum (fwhm) of this peak for the nanocomposite thin films with respect to P3HT, gives an indication of the changes in the crystallite sizes that occur. The change in d-spacing upon adding the PCBM in the polymer matrix, hints that interdigitation or tilting of the side groups occurred, during the formation of ordered aggregates in the solution and this induced crystallinity during film deposition. Upon introduction of ZnO into the polymer blend matrix a further reduction in X-ray intensity is observed when the weight percentage of ZnO is increased, with a shift towards higher 2θ of the (100) peak.

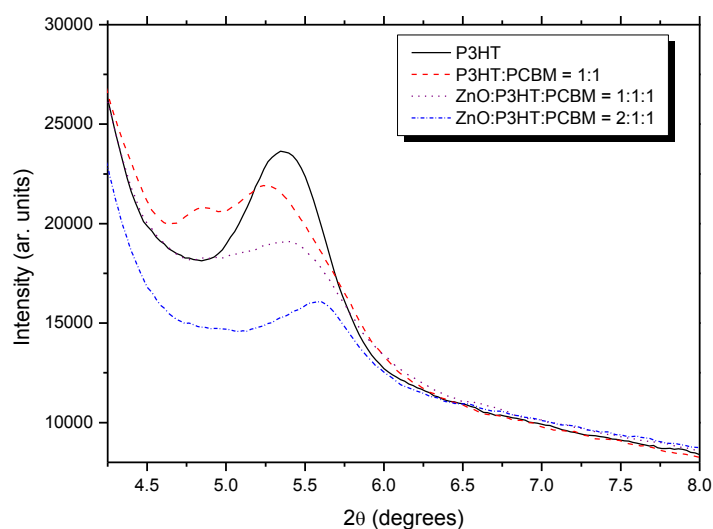


Figure 3. XRD pattern of nano-composites.

The HRTEM image in figure 4 (a) displays the random dispersal of ordered, crystalline regions embedded in amorphous tissue for an equal weight percentage hybrid mixture. In figure 4 (b) a further reduction in the degree of crystallinity, and an increase in the crystallite size is noted when the weight percentage of ZnO is doubled. This is probably due to an excess of inhomogeneously dispersed ZnO in the polymer matrix at higher concentrations, which give rise to the effects of disorder or segregation in the structure. These effects can thus account for the shifts in peak position and the changes in the fwhm of the diffraction peak in the XRD patterns.

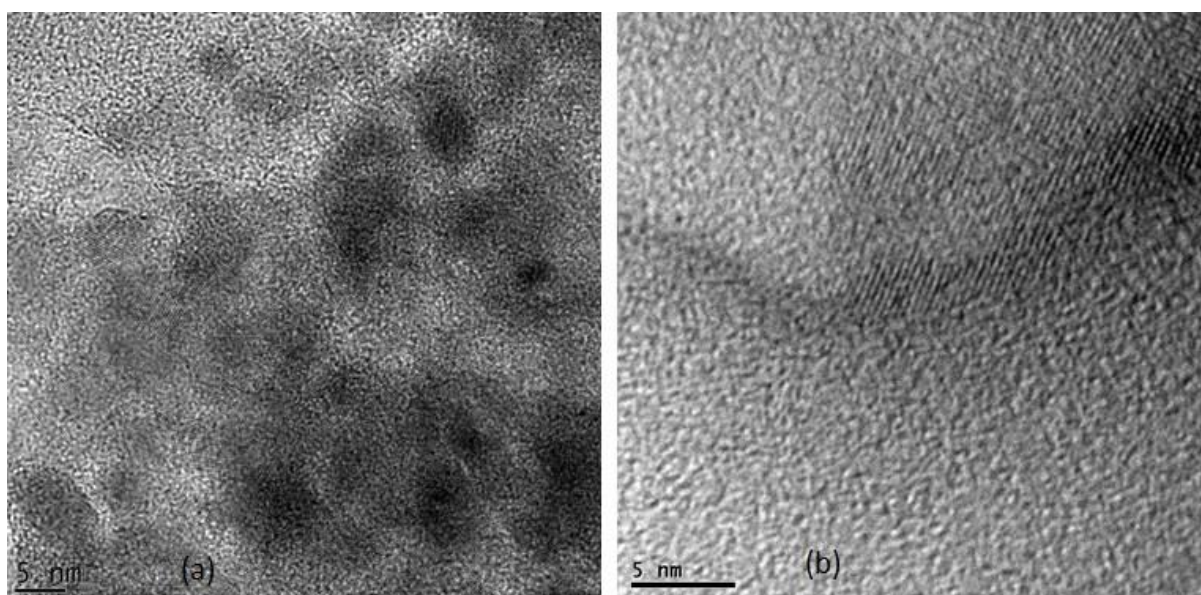


Figure 4. TEM micrographs of ZnO:P3HT:PCBM thin films for (a) 1:1:1 wt. ratio, (b) 2:1:1 wt. ratio.

The morphology of the spin-coated polymer film is thus an important factor in understanding the property of the metal-polymer interface. In the final device the interface between the polymer and the cathode is important as it contributes greatly to the electrical characteristics of the devices [9]. It is expected that the hybrid active layer could enhance the electron transportation to the cathode due to the high electron mobility offered by ZnO nanoparticles present in the active layer, and depending on the shape and size of the ZnO nanostructures the transport rate of charge carriers can be improved and the recombination rates of excitons reduced [10]. The formation of large aggregates (rougher surface) lowers the degree of phase separation between P3HT and ZnO interfaces such that more charge generation can occur; in addition more percolation pathways for charge carriers become available [10].

The contribution from the heterojunction electric field formed between ZnO and organic materials enhance the separation of the electron-hole pairs and hence result in the increase of the photovoltaic property of the hybrid layer [9]. It is also expected that the ZnO: P3HT: PCBM film surface will result in greater light absorption due to the diffuse reflection from the surface of the active layer. UV-vis absorbance is therefore employed to inspect absorption properties of the thin films under study. Figure 5 displays the optical UV-vis absorbance whereas the absorption coefficient obtained from an independent method, spectroscopic ellipsometry, is shown in figure 6.

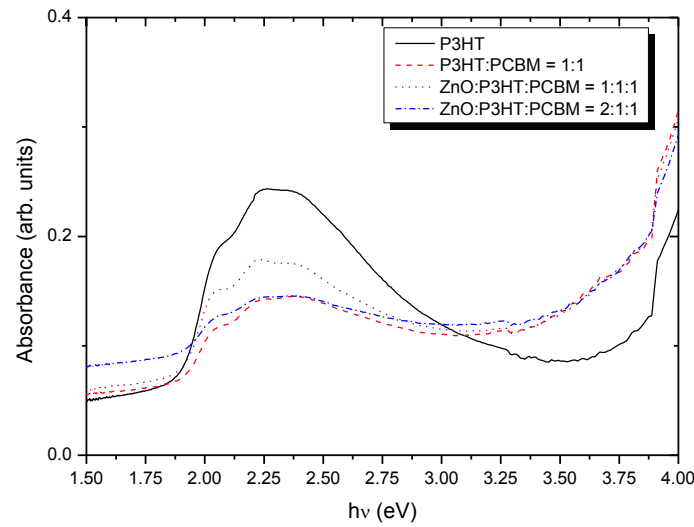


Figure 5. UV-vis absorbance spectra of spin-coated P3HT- based films obtained by UV-vis spectroscopy.

The absorption coefficient was calculated from the extinction coefficient. Regression is performed on the initial parameters inputted by the user; as such the accuracy of the initial values may significantly influence the accuracy of the final results. The spots on the thin films, i.e. the areas measured by UV-Vis and SE may also not necessarily be the same, which can also contribute to a discrepancy between the results obtained by the two methods [7]. It is encouraging however to note that the same features in absorbance and absorption coefficient dispersions can nevertheless be inferred from both figures.

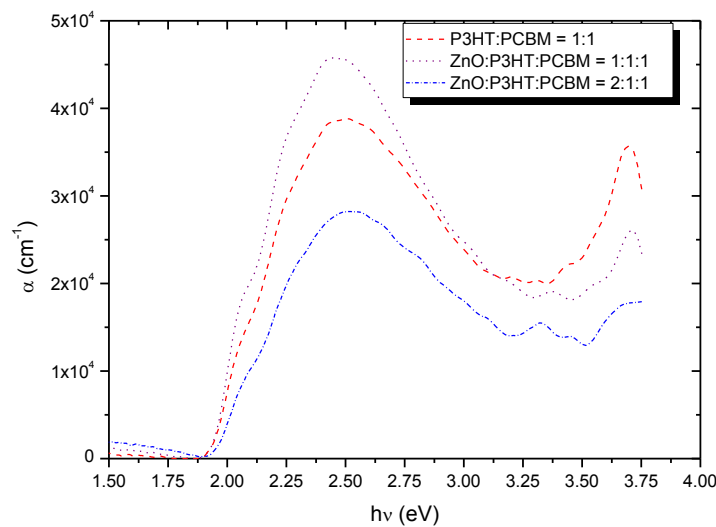


Figure 6. Absorption coefficient of blend and hybrid ZnO:P3HT:PCBM thin films obtained by spectroscopic ellipsometry.

For the pristine P3HT film the absorbance spectrum shows a peak at 2.38 eV, and two humps at 2.04 eV and 2.24 eV respectively. These correspond well with other studies, which ascribe the features to

bands attributed to the 0-0, 0-1, and π - π^* transitions [11]. The absorption wavelength range between 450 nm to 600 nm (2.06 eV - 2.76 eV) is considered to be the peak absorption wavelength band for P3HT. The peak absorption wavelength (energy) experienced a slight blue shift for the P3HT:PCBM blend compared to P3HT; and a concomitant loss of the weak vibronic structures [12]. The shift is attributed to a loss of P3HT polymer-chain stacking and conformational disorder, caused by the mixing of PCBM with the P3HT [12]. For the hybrid mixtures the blue shift is almost reversed completely. However, the maximum absorption intensity changed significantly for the thin films. The blend experienced a significant drop in maximum absorption intensity from pristine P3HT. Further, in both hybrid films the relative contribution from the ZnO absorbance peak at 3.54 eV should increase compared to the π - π^* band of P3HT in the energy range of 2.06 eV to 2.76 eV. Upon introduction of ZnO in equal weight percentage the maximum absorption intensity recovers again, but to less than that of pure P3HT. The quenching of the film absorption when excess ZnO concentration is introduced evidences this. Additional information is further supplied by the absorption coefficient dispersion, which shows an improvement for the 1:1:1 hybrid film over the P3HT:PCBM blend, while the excess ZnO concentration in the 2:1:1 hybrid film causes degradation in the optical properties due to increase in grain size, i.e. increased segregation, as discussed above.

4. Conclusion

It was established that the introduction of pyramid-shaped ZnO nanoparticles into a P3HT:PCBM blend leads to improved absorption of spin-cast films made from the solution, over films made by the P3HT:PCBM blend. The ZnO:P3HT:PCBM = 1:1:1 weight ratio film also exhibits superior absorption properties compared to double the ZnO weight ratio film. It was further noted that spectroscopic ellipsometry is a useful tool to obtain complimentary absorption coefficient information of the spin-cast films, which are notoriously inhomogeneous in thickness across the surface.

Acknowledgements

The authors thank the National Research Foundation of South Africa for financial support.

References

- [1] Yu G, Gao J, Hummelen J C, Wudl F and Heeger A J 1995 *Science* **270** 1789
- [2] Olson D C, Pirijs J, Collins R T, Shaheen S E and Ginley D S 2006 *Thin Solid Films* **496** 26
- [3] Motaung D E, Malgas G F, Arendse C J and Mavundla S E 2012 *Mater. Chem. Phys.* **135** 401
- [4] J. A. Woollam Inc. *Complete Ease™ Data Analysis Manual*, June 15 2008
- [5] Johs B and Hale J S 2008 *Phys. Status Solidi A* **205** 715
- [6] Li G, Shrotriya V, Yao Y and Yang Y 2005 *J. Appl. Phys.* **98** 043704
- [7] Ding Y, Lu P and Chen Q 2008 *Proc. SPIE* **7099** 709919-1
- [8] Zhokavets U, Erb T, Hoppe H, Gobsch G and Sariciftci N S 2006 *Thin Solid Films* **496** 679
- [9] Oh S H, Heo S J, Yang J S and Kim H J 2013 *Appl. Mater. Interfaces* **5** 11530
- [10] Malgas G F, Motaung D E, Mhlongo G H, Nkosi S S, Mwakikunga B W, Govendor M, Arendse C J and Muller T F G 2014 *Thin Solid Films* **555** 100
- [11] Brown P J, Thomas D S, Kohler A, Wilson J S, Kim J S Ramsdale C M, Sirringhaus H, Friend R H 2003 *Phys.Rev. B* **67** 064203
- [12] Beek W J E, Wienk M M and Janssen R A J 2006 *Adv. Funct. Mater.* **16** 1112

Structural characterisation of aluminium and yttrium co-doped tin oxide

JN Ntimane, KE Rammutla¹ and TE Mosuang

Department of Physics, University of Limpopo, P/Bag x1106, SOVENGA, 0727, RSA

E-mail: erasmus.rammutla@ul.ac.za

Abstract. The sol gel technique was successfully used to synthesize the nanocrystals of tin dioxide (SnO_2), co-doped with aluminium (Al) and yttrium (Y). The powders were preheated at different temperatures ranging from 200 to 1000 °C. The effects of temperature and dopants on the structure of tin dioxide nanoparticles were investigated. X-ray powder diffraction (XRD) and Raman spectroscopy were used to probe the structural properties. The average crystallite sizes were found to be in the range between 2.5 – 8 nm in the temperature range studied. Both techniques confirm that at higher temperatures there is a high temperature reaction resulting in the formation of yttrium stannate.

Introduction

SnO_2 is an important semiconducting material due to its dipole forbidden direct band gap of 3.6 eV at room temperature [1]. It has multiple applications and such are in the fields of gas sensing [2], solar energy conversion [3] and catalysis [4]. The multifunctionality of nanostructured SnO_2 is due to its high surface to volume ratio, larger band gap energy and high exciton binding energy. It is also mechanically, thermally and chemically stable [5]. SnO_2 has remarkable resistivity variation in gaseous environment and this property is important in gas sensing technology. Optoelectronic properties of SnO_2 depend on the presence of impurities.

Introducing metal dopants has been found to reduce grain growth and also improve the absorption capabilities of nanocrystalline oxides [6]. Co-doping or introducing suitable amounts of more than one type of metals has recently been found to further improve the photocatalytic properties of metal oxides [7]. In this contribution, aluminium (Al) and yttrium (Y) metal cations have been simultaneously introduced in SnO_2 using sol-gel method. The aim of the study is to investigate the effects of combinational doping on the structural properties of SnO_2 using X-ray diffraction (XRD) and Raman scattering spectroscopy. Their effect on the crystallite growth is also examined.

Experimental details

The sol gel preparation method of Ansari et al. [8] was followed in the present work. To prepare Y and Al doped nanocrystalline tin oxide, appropriate amounts of tin chloride octahydrate and water soluble metal salts (yttrium chloride hexahydrate and aluminium trichloride hexahydrate) were dissolved in 40 ml of distilled water to prepare a 0.1 M solution. To this solution, 0.5 ml of aqueous ammonia was gently added and the resulting white precipitate recovered by evaporation. This was followed by

¹ To whom any correspondence should be addressed.

washing with distilled water to remove the by-product, ammonium chloride. The resulting sample was oven dried at 100 °C for overnight followed by a thorough grinding to make fine nanoparticles. The samples were prepared with dopant concentrations of 10 wt % of Al and 10 wt % of Y. Portions of the sample were heated at various temperatures ranging from 200 to 1000 °C.

XRD patterns of the samples were recorded in an ambient environment using a Panalytical XPERT PRO diffractometer equipped with a Cu K α tube ($\lambda = 1.5406 \text{ \AA}$) monochromatic radiation source operating at 35 kV and 40 mA.

Raman spectra of all the samples were recorded using the HORIBA Jobin Yvon T64000 spectrometer equipped with an Olympus BX-40 microscope. The spectra were excited using the excitation wavelength of 514.5 nm from an argon ion laser.

Results and discussion

The XRD profiles of the Al/Y-SnO₂ preheated at different temperatures are shown in figures 1(a) and 1(b). For the 200 °C sample, three main peaks along (110), (101) and (211) planes are clearly observed indicating that the material is polycrystalline in nature. These diffraction peaks are assigned to tetragonal rutile crystalline phases of tin oxide and are consistent with the values in the standard card (JCPDS 77-0452). As the temperature rises, more peaks associated with SnO₂ become evident indicating the improvement in crystallinity. No characteristic peaks associated with impurities or additional phases are observed for all samples up to 600 °C. However, at 800 °C new peaks at $2\theta = 30^\circ, 36^\circ, 48^\circ$ and 57° associated with the (222), (400), (440) and (622) planes, respectively, start to develop and as can be seen in figure 1(b), at 1000 °C the new peaks are intense. These new peaks are associated with yttrium stannate (Y₂Sn₂O₇) phase [9] suggesting that at higher temperatures tin dioxide reacts with yttrium to form yttrium stannate. In our earlier paper [10], we used extended x-ray absorption fine structure (EXAFS) to study SnO₂ singly doped with Y and the results showed Y preferred the surface layers. The technique also revealed the formation of yttrium stannate at higher temperatures which is consistent with the present work on the double doped sample. SnO₂ singly doped with Al was also investigated [11, 12] and like in the present work no peaks associated with Al metal or alumina were found, indicating that Al substitutes for Sn in the SnO₂. The lattice parameters and crystallite sizes are shown in table 1. The lattice parameters were determined from the interplaner spacing equation and Bragg's law using the X-ray diffraction peaks of the (101) and (110) crystal planes. As seen from the table, the addition of dopants caused no significant changes in the lattice parameters confirming that Y (with a larger ionic radius than the host ion) has not entered into the matrix. The results of the SnO₂ singly doped with Y [10] showed that heating caused the crystallites to grow to about 40 nm at 1000 °C. In the present work, heating to 1000 °C resulted in nanocrystals of ~ 8 nm indicating that double doping resulted in considerable pinning of growth.

Table1: Crystallite sizes and lattice parameters of Al/Y co-doped SnO₂

Sample	Temperature (°C)	Crystallite size (nm)	Lattice parameters (Å)		
			<i>a</i>	<i>b</i>	<i>c</i>
Pure SnO ₂ [13]	Ambient	-	4.738	4.738	3.188
Al/Y - SnO ₂	200	2.6	4.810	4.810	3.158
	400	3.2	4.810	4.810	3.198
	600	4.5	4.747	4.747	3.080
	800	6.0	4.747	4.747	3.065
	1000	8.0	4.706	4.706	3.012

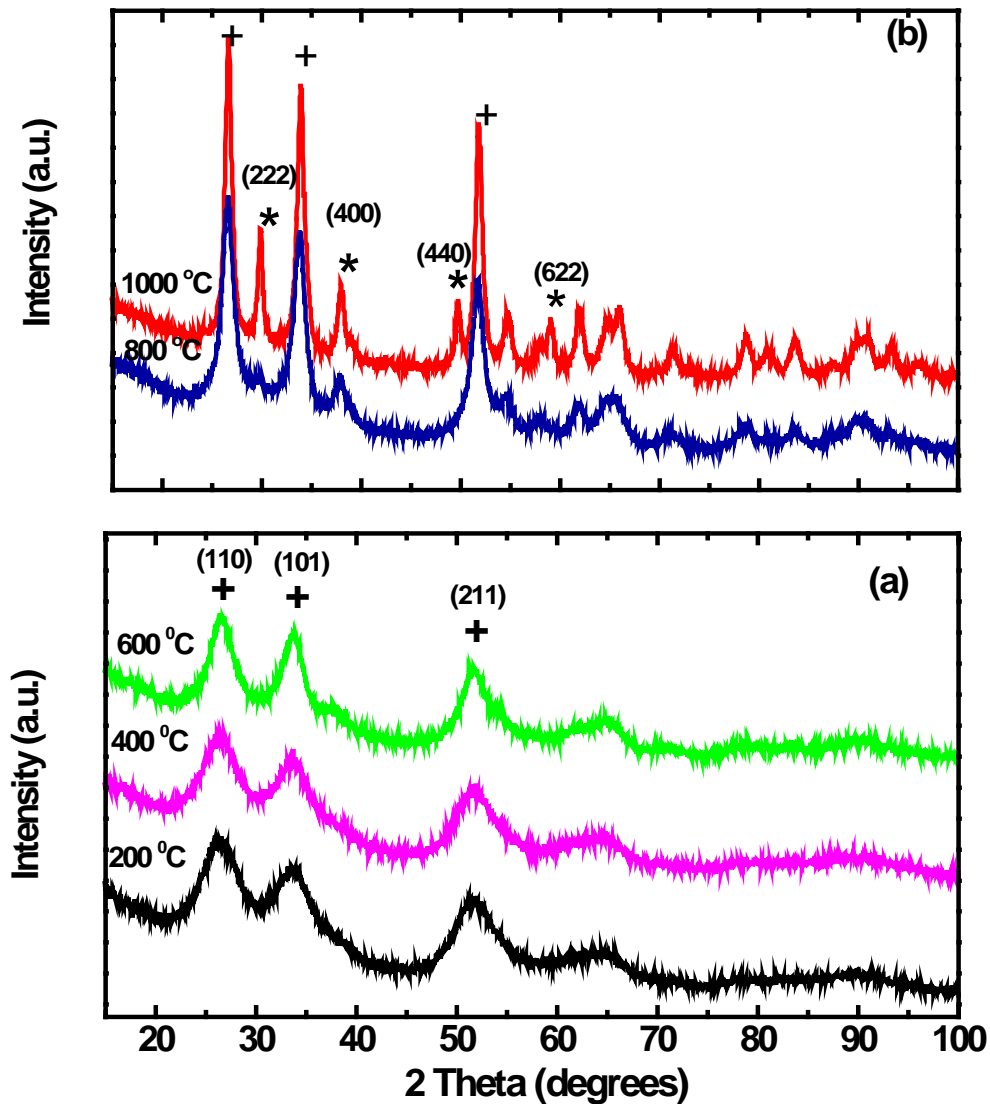


Figure 1: XRD profiles for Y/Al co-doped SnO₂ at (a) 200 °C, 400 °C and 600 °C and at (b) 800 °C and 1000 °C, where peaks associated with SnO₂ are indicated by (+) and those associated with Y₂Sn₂O₇ are indicated by (*).

SnO₂ is expected to display three main Raman bands, i.e. A_{1g} (630 cm⁻¹), B_{2g} (774 cm⁻¹) and E_g (472 cm⁻¹) [14]. Figure 2 shows the Raman spectra of Al/Y co-doped nanocrystalline SnO₂ at various temperatures. The Raman spectrum of the sample heated at 200 °C is similar to the as prepared (not shown) and it displays a broad Raman band at 572 cm⁻¹ associated with the amorphous tin (IV)-hydrous oxide [15]. The 572 cm⁻¹ band is still observed at 400 °C although peaks associated with the SnO₂ start to develop at 630 cm⁻¹ and 774 cm⁻¹. Above 400 °C, the amorphous phase is no longer visible and at 800 °C the three Raman active modes i.e. A_{1g}, B_{2g} and E_g are observed although only the A_{1g} band is clearly visible due the particle size being small. At 1000 °C, all the three bands are clearly visible with the A_{1g} being the most intense. In addition to those associated with SnO₂, more Raman bands at 508 cm⁻¹, 412 cm⁻¹ and 311 cm⁻¹ associated with yttrium stannate (Y₂Sn₂O₇) [16] are observed.

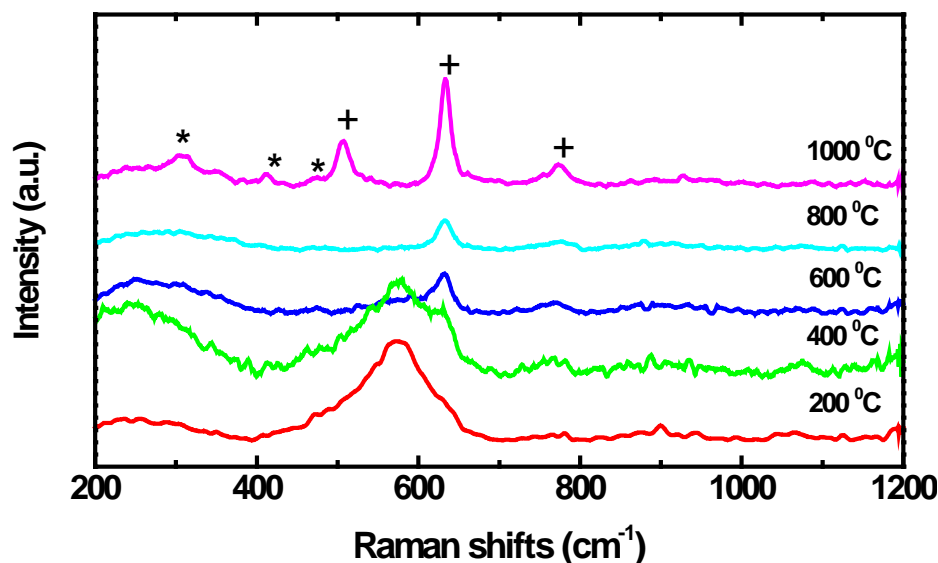


Figure 2: Raman spectra of Al/Y SnO₂ at different temperatures, where peaks associated with SnO₂ are indicated by (+) and those associated with Y₂Sn₂O₇ are indicated by (*).

Conclusions

The sol gel technique was successfully used to synthesize the nanocrystals of tin dioxide (SnO₂) co-doped with aluminium and yttrium. The samples were characterised by XRD and Raman. The Raman results show that at lower temperatures there is an amorphous phase although not clearly shown by XRD. The results reveal that double doping significantly reduces the crystallite growth of SnO₂. There was no evidence of an Al metal or alumina in the sample indicating that Al may have substituted for Sn atom in the SnO₂. Y is probably on the surface layers. Both XRD and Raman results consistently show that at high temperatures SnO₂ chemically reacts with Y to form yttrium stannate.

Acknowledgements The authors greatly acknowledge IBSA, NRF and University of Limpopo for financial support and CSIR for facilities.

References

- [1] Summitt R, Marley JA and Borrelli NF 1964 *J. Phys. and Chem. Sol.* **25** 1465
- [2] Shimizu Y and Egashira M 1999 *MRS Bulletin* **24** 18
- [3] Ferrere S, Zaban A and Gsegg BA 1997 *J. Phys. Chem.* **B 101** 4490
- [4] H Ogawa, Abe A, Nishikawa M and Hayakawa S 1981 *J. Electrochem. Soc.* **128** 2020
- [5] Chopra KL, Major S and Pandya DK 1983 *Thin Sol. Films* **102** 1
- [6] Lajevardi SA and Shahrabi T 2010 *Appl. Surf. Sci.* **256** 6775
- [7] Štengl V, Bakardjieva S and Bludská J 2011 *J. Mater. Sci.* **46 (10)** 3523
- [8] Ansari SG, Boroojerian P, Kulkarni SK, Sainkar SR, Karekar RN and Aiyer RC 1996 *J. Mater. Sci.: Mater. Electron.* **7** 267
- [9] Facer GR, Howard CJ and BJ Kennedy 1993 *Powder Diffraction* **8** 245
- [10] Rammutla KE, Chadwick AV, Harding J, Sayle DC and Erasmus RM 2010 *J. Phys., Conf. ser.* **249** 012054
- [11] Padmavathy R and Sridevi D 2010 *Arch. Phys. Res.* **1(4)** 104
- [12] Xu C, Tamaki J, Miura N and Yamazoe N 1991 *Talanta* **38 (10)** 1169
- [13] Sensato FR, Custodio R, Longo E, Beltrán A and Andrés J 2003 *Catalysis Today* **85** 145
- [14] Yu KN, Xiong Y, Liu Y and Xiong C 1997 *Phys. Rev.* **B 55** 2666
- [15] Ristic M, Ivanda M, Popovic S and Music S 2002 *J. Non-Crystalline Solids* **303** 270
- [16] Li K, Li H, Zhang H, Yu R, Wang H and Yan H 2006 *Materials Bulletin* **41** 19

Structural and optical characterisation of double-doped TiO₂ nanoparticles

OO Nubi¹, KE Rammutla and TE Mosuang

Department of Physics and Geology, University of Limpopo, South Africa

E-mail: olatunbosun.nubi@ul.ac.za

Abstract. With titanium isopropoxide as the precursor, single and double doped nanosized powders of TiO₂ were synthesised by the sol-gel process. The metal dopants used were Ag and Cu at doping levels of 5% (molar weight). The samples were dried at 100 °C in air and then heated at 300°C, 600°C, 900°C and 1100°C for one hour. Structural characterisation of the samples was carried out by X-ray Diffraction (XRD), Raman and Scanning Electron Microscopy (SEM) techniques. The results suggests that the co-doped TiO₂ powders are constituted by both the anatase and brookite phases whereas only anatase is observed in the case of pure and singly doped samples. The co-existence of brookite with anatase in the co-doped sample is thought to be responsible for the enhancement of anatase to rutile transformation. UV-visible measurements were done to study the optical properties of the TiO₂ nanoparticles. Double doping was found to enhance the narrowing of the band gap, compared to single doping.

1. Introduction

Titanium dioxide (TiO₂ or Titania) has been widely studied [1-4] because of its extensive applications in the fields of environment (photocatalysis, sensing) and energy (photovoltaics, hydrogen storage, water splitting, photochromics, electrochromics). It also finds uses as pigments [5], sunscreens [6], ointments [7], paints [8] and toothpaste [9]. Factors that determine the properties of the titania material used in these applications include phase composition, particle size and distribution, morphology as well as porosity [4].

In research and technological applications, the most important polymorphs of titania are the anatase (A), rutile (R) and brookite (B) crystalline phases [10]. Of these, rutile is the most stable while Brookite, being difficult to synthesize, is seldom studied [11]. The metastable anatase and brookite states begin to transform to rutile upon high temperature calcination between 600°C and 700°C. During synthesis, the formation of any particular phase is dependent on the precursors used, method of synthesis and the calcination temperature. It is of significant interest to understand the conditions that affect the transformation from one phase to another as this may affect the properties and performance of devices.

Recently, research studies have explored the effect of doping TiO₂, via chemical synthesis, particularly with metal impurities [12]. Further, considerable enhancement of the photocatalytic activity and other characteristics has been observed for TiO₂ doped with two impurities as compared with single doping [13–16].

¹ To whom any correspondence should be addressed.

In this investigation, the effect of simultaneously introducing two metal dopants, via sol-gel synthesis, on the structural and optical properties of TiO₂ nanopowders, is compared with single-doped species. The Sol-gel method of synthesis affords a way of preparing nanosized materials by a chemical reaction in solution starting with metal alkoxide as a precursor. This route provides the advantages of attaining high chemical purity and precise control over pore structure and dopant concentration [4].

2. Experimental

2.1. Synthesis

In addition to the ‘Undoped/TiO₂’ sample set, “single-doped” TiO₂ nanopowders (TiO₂ doped with either Cu and Ag), namely Ag/TiO₂, Cu/TiO₂, and one “double-doped” (TiO₂ doped with both Cu and Ag) set, Ag+Cu/TiO₂, were synthesised at 5 wt% impurity level. Aldrich’s analytical reagent grade precursors, namely Ti{OCH(CH₃)₂}₄ (Titanium (IV) Isopropoxide), AgNO₃ and CuCl₂, were used to prepare the nanocrystalline TiO₂ samples via the sol-gel formulation. Priming, by first dissolving the dopant salts in water, was followed by adding ethanol to the required amount of each dissolved precursor. The dopant solution was then added, in drops, to the Titanium Isopropoxide, while vigorously stirring for up to 1 hour. The precipitate (xerogel) formed was further diluted with 30 ml water, filtered and left to dry at room temperature for 16 hours. Further drying of the samples was done at 100°C for 1 hour before being ground to powders.

The powdered samples were divided into four so that each could be annealed at temperatures of 300°C, 600°C, 900°C and 1100°C for one hour, thus making a total of 20 samples.

2.2. Characterisation

XRD powder data were recorded in the 2 θ scan range of 10 – 90° with PANalytical X’Pert PRO diffractometer (Cu Ni-filtered K α radiation of $\lambda = 1.540598$ Å, operating at 45kV 40mA). Raman spectra from 50 to 1200 cm⁻¹ were obtained with a Jobin Yvon LabRAM HR 800 UV-VIS-NIR spectrometer equipped with an Olympus Microscope. SEM micrographs of the samples were obtained with a JEOL- JSM 7500F Scanning Electron Microscope. Optical absorption spectra were collected using Perkin Elmer Lambda 750S UV/Vis Spectrometer in the range of 200 – 800 nm, while the Perkin Elmer LS-55 Fluorescence Spectrometer was used for the Photoluminescence (PL) data.

The Scherrer equation, $\tau_{(hkl)} = K\lambda/\beta_{(hkl)} \cos \theta$ was used in estimating the mean size $\tau_{(hkl)}$ of the ordered crystalline domains from $\beta_{(hkl)}$, the line broadening at half the maximum intensity (or the full-width-at-half-maximum, FWHM) in radians. Here, θ is the Bragg angle and λ is the X-ray wavelength (equal to 1.540598 Å for the Cu K α radiation used in this study). By considering the synthesised nanoparticles to be fairly spherically shaped, a value of 0.9 was assumed for the shape parameter (or roundness factor) K [17]. For quantitative comparative analyses, instrumental and strain broadening were neglected and Gaussian fitting of the most intense peaks of the individual phases – (101) and (200) for anatase and (110) and (211) for rutile – were carried out [14], with the aid of the Origin® statistical software package. The phase content of each sample was calculated from the integrated intensities of the respective peaks of anatase (101), rutile (110) and brookite (210) [12]

Many studies [18, 19] have employed direct-transition Tauc plot extrapolations in the investigation of the optical band gap for both thin films and nanostructured TiO₂. This method requires a dimension-related parameter for estimating the absorption coefficients required in the determination of Kubelka-Munk function values. For instance, in the thin-films scenario, the parameter is often the thickness of the film. Determination of this parameter in the case of randomly sized nanoparticles is often not straightforward. Further, extrapolations of a Tauc plot can be somewhat ambiguous, particularly in cases where the linear region is not clearly defined. For these reasons, the Tauc route is not taken in this study. Instead, points on the absorption versus wavelength plot, which gives the sharpest change in absorption, are identified. This was achieved by plotting the absorption sensitivity to wavelength, (dA/d λ), against the wavelength (or photon energy) and selecting the position of the most intense peak. The peak thus obtained provide reasonable estimates of the energy gap between the conduction and valence bands for the samples under investigation.

3. Results and discussion

Nanocrystalline anatase TiO_2 structure indexed to the American Mineralogist Crystal Structure Database (AMCSD) 0019093 were observed for the Undoped/ TiO_2 , as well as the single-doped Ag/ TiO_2 sample at calcination temperatures of 600°C and lower. At the higher temperatures of 900°C and 1100°C, these samples feature patterns that could be indexed to the AMCSD 000173 structure for rutile. This is in line with what is commonly found in literature for the anatase-to-rutile transformation at calcination temperatures above 600°C.

In figure 1 the XRD patterns are given for anatase (A) and rutile (R) samples calcined at 600°C. The presence of the well-formed rutile structure – for the Cu/ TiO_2 and the Ag+Cu/ TiO_2 powders at 600°C – suggests that the phase transition temperature (PTT) in these cases were lower than those of the Undoped/ TiO_2 and Ag/ TiO_2 samples. In particular, at 300°C, the double-doped Ag+Cu/ TiO_2 powder reveals peaks associated with the rutile phase, thus indicating the commencement of the phase change at this temperature. Although doping with Ag did not affect the transition in any appreciable way, the lowering of the transformation temperature by Cu was further enhanced by the inclusion of Ag impurities in the double-doped system.

This observation may be attributed to the presence of the brookite phase (figure 2) in the double-doped Ag+Cu/ TiO_2 sample between the calcination temperatures of 300°C and 600°C. With barely any anatase left untransformed at 600°C, the brookite phase content of the Ag+Cu/ TiO_2 sample was estimated to be 22.7%, the remaining 77.3% being rutile. The transformation of this brookite to rutile was found to be complete at 900°C.

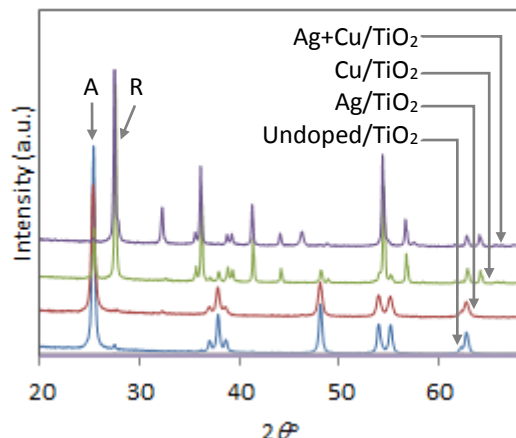


Figure 1. Anatase (A) and rutile (R) peaks in the XRD patterns of samples calcined at 600°C.

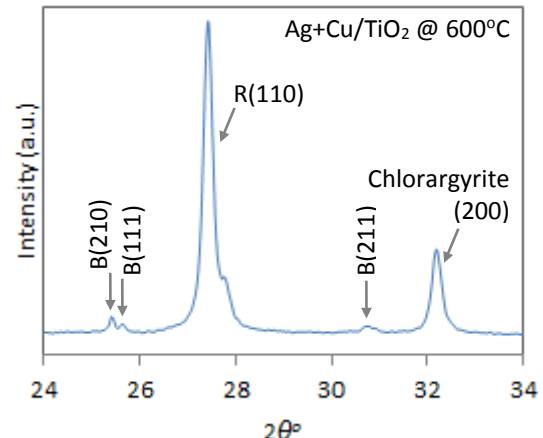


Figure 2. A zoom-in on the brookite (B) peaks present in the XRD patterns of the Ag+Cu/ TiO_2 sample calcined at 600°C.

The SEM micrographs of figure 3 belong to the Undoped/ TiO_2 at 300°C (left) and Ag/ TiO_2 at 900°C (right). Agglomeration of fairly spherically-shaped particulate structure, revealed in the images, are typical of what was observed for the anatase (at 300°C) and the rutile (at 900°C) phases.

Figure 4 displays how the crystallite size varies with increasing calcination temperature for the Undoped- TiO_2 samples. Judging by the 4th-order polynomial fit, the anatase sizes (100°C to 600°C) appear to grow relatively faster than the rutile (900°C and 1100°C). Further work, with smaller temperature intervals, may be needed to confirm this. However, the grains of doped powders are found to be smaller than those of the Undoped/ TiO_2 , (table 1), for the anatase structures, in agreement with other studies that discussed the reduction mechanism [20, 21].

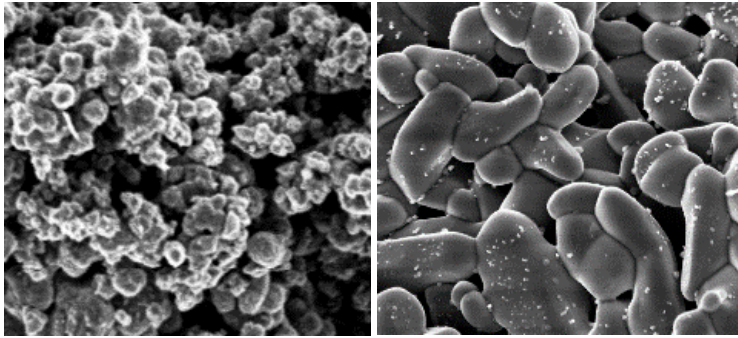


Figure 3. SEM micrographs of Undoped/TiO₂ at 300°C (left) and Ag/TiO₂ at 900°C (right).

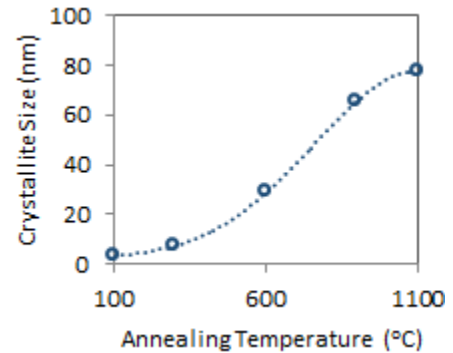


Figure 4. Crystallite size vs. calcination temperature of Undoped/TiO₂.

Table 1. Crystallite sizes and lattice parameters for the anatase and rutile structures.

Sample	Anatase Structure at 300°C			Rutile Structure at 900°C		
	Crystallite size (nm)	Lattice Parameters (Å)		Crystallite size (nm)	Lattice Parameters (Å)	
		a = b	c		a = b	c
Undoped/TiO ₂	7.39	3.79	9.37	65.46	4.59	4.65
Ag/TiO ₂	6.17	3.80	9.09	74.93	4.59	2.96
Cu/TiO ₂	6.32	3.79	9.37	78.53	4.57	2.97
Ag+Cu/TiO ₂	4.65	3.79	9.21	79.69	4.58	2.96

The unit cell parameters (table 1) suggest that the TiO₂ lattice is not deformed by doping when in the anatase phase, whereas the rutile structure reduced in volume by about 36.4% with doping (even though the grain sizes became larger.)

In figure 5 the Raman scans are given for the Undoped/TiO₂ calcined at 300°C (Anatase) and at 900°C (Rutile). The spectrum for Anatase show the six characteristic bands usually identified at 144, 197, 399, 513, 519, and 639 cm⁻¹ with anatase TiO₂ [22, 23]. The Characteristic phonon modes of the rutile TiO₂ structures were also observed for all the samples constituted by the rutile phase. Here, the major peaks were found at 240, 446 and 610 cm⁻¹ while minor peaks are at 818, 707 and 319 cm⁻¹ as expected [24, 25]. The intensity and sharpness of the peaks signify that the samples are highly crystalline and pure. The results of the Raman spectroscopy thus compliment very well the results that were obtained from XRD. A blue-shift of about 6.3 cm⁻¹, for both Cu/TiO₂ and Ag+Cu/TiO₂, was evident in the Raman spectra of the anatase lowest-frequency mode (E_g, due mainly to the symmetric stretching of the O–Ti–O). This blue-shift, displayed in figure 6 for all the powder sets calcined at 300°C, may be ascribed to phonon confinement, non-stoichiometry effects and variations in the dimensions of the nanoparticles [26, 27].

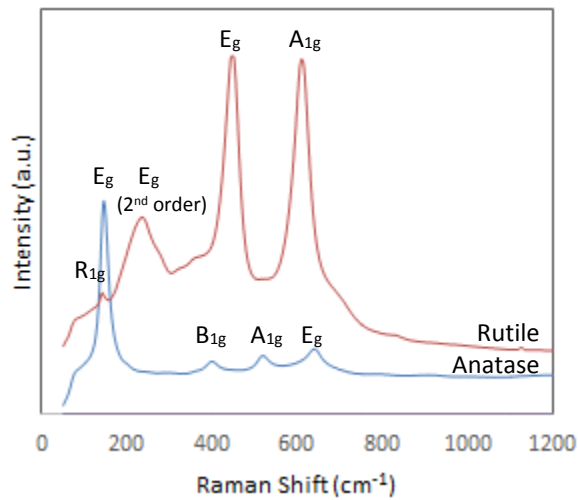


Figure 5. Raman spectra for the Undoped/TiO₂ calcined at 300°C (Anatase) and at 900°C (Rutile).

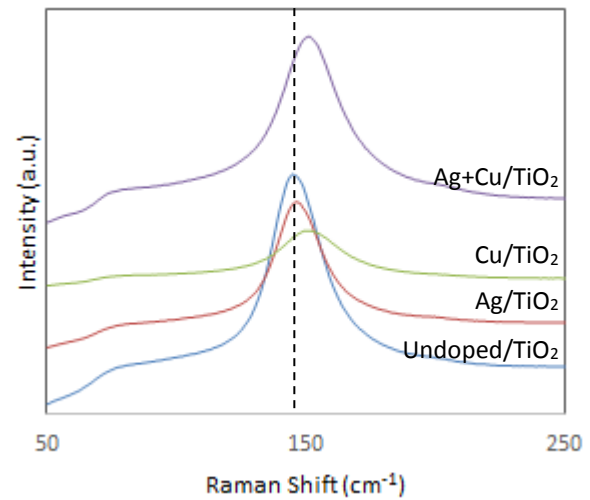


Figure 6. Raman spectra of the E_g mode for samples at calcined at 300°C.

Figure 7 displays the optical absorption spectra of the sample sets at the various TiO₂ phases. Inset is the plot of the wavelength sensitivity of the absorbance versus the photon energy for the Undoped/TiO₂ calcined at 600°C. This yields a sharp peak (or the most rapid change in absorption) for this sample at 3.09 eV. Table 2 shows values for other samples that were derived in a similar manner. The band gap for the single-phase anatase powder Ag/TiO₂ was found to be 3.09 eV as well. The copper-doped (anatase-only) sample, Cu/TiO₂ however showed a red-shift towards a lower energy gap of 2.47 eV. This same improvement in photo absorptions was also observed for the Ag+Cu/TiO₂ sample which contains the mixed phase of anatase and brookite at 300°C. The narrowest band gap of 2.04 eV was observed for the brookite-rutile mix of Ag+Cu/TiO₂ at 600°C. In comparison, all the rutile-only samples (calcined at 900°C and above), doped and undoped, revealed a value of 3.08 eV.

Though others [28] have attributed it to quantum size effects, the shift of the absorbance spectra of Ag+Cu/TiO₂, at 600°C, towards the visible light region may be ascribed to the introduction of some additional energy levels in the host lattice band gap by the impurities. As both Ag and Cu metals form a *p*-type semiconductor, a shift in Fermi energy towards valence band may be realised.

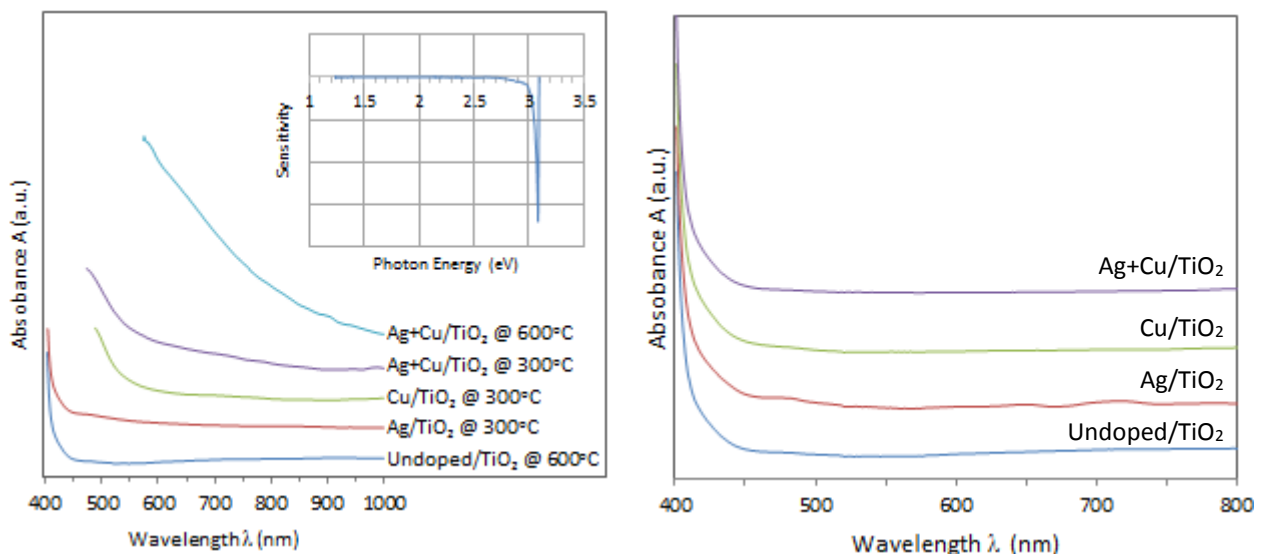


Figure 7. Optical absorption spectra of the sample sets at the various TiO₂ phases. Inset is the plot of the wavelength sensitivity of the absorbance versus the photon energy.

Figure 8. Optical absorption edges of the UV-vis spectra for the various sample sets at 900°C (rutile phase).

Table 2. Band gap of samples at various phase structures.

Sample	Calcination Temperature	Phases Present	Band Gap E_g (eV)
Undoped/TiO ₂	600°C	Anatase	3.09
Ag/TiO ₂	300°C	Anatase	3.09
Cu/TiO ₂	300°C	Anatase	2.47
Ag+Cu/TiO ₂	300°C	Brookite, Anatase	2.47
Ag+Cu/TiO ₂	600°C	Brookite, Rutile	2.04
Ag+Cu/TiO ₂	1100°C	Rutile	3.08

4. Conclusions

Double doping of TiO₂ with Ag and Cu impurities induced mixed phases which included brookite, namely, anatase-brookite (at 300°C) and rutile-brookite (at 600°C) before the transition to rutile TiO₂. The presence of this brookite, though minute, appeared to lower the transition temperature appreciably. Even though doping reduced the grain size of TiO₂, the anatase lattice was not deformed by doping but the rutile shrunk by 36.4%.

In general, single doping with silver did not alter the properties of TiO₂ as much as single doping with copper. For instance, due to phonon confinement, copper was able to introduce a blue-shift towards higher energies, in the Raman E_g mode. Optically, double doping featured a remarkable effect over the copper doped material by reducing the energy gap well into the visible region (608 nm, Yellow) of the electromagnetic spectrum. This could result in greater production of electron-hole pairs under visible light illumination and thus achieving higher photocatalytic efficiencies.

Acknowledgements

The authors would like to express their gratitude to NRF, IBSA and University of Limpopo for financial support and CSIR's Centre for Nanostructured Materials for experimental resources.

References

- [1] Mo S-D and Ching W Y 1995 *Phys. Rev. B* **51** 19
- [2] Hua C, Duoa S, Liua T, Xiang J and Li M 2010 *Appl. Surf. Sci.* **11**, 110
- [3] Colmenares J C, Aramedia M A, Marinas A, Marinas J M and Urbano F J 2006 *Appl. Catal. A-Gen.* **306** 120–7
- [4] Yu J-G, Yu J C, Cheng B, Hark S K and Iu K 2003 *J Solid State Chem.* **174**, 372–80
- [5] Pfaff G and Reynders P 1999 *Chem. Rev.* **99**, 1963
- [6] Salvador A, Pascual-Marti M C, Adell J R, Requeni A and March J G 2000 *J. Pharm. Biomed. Anal.* **22**, 22, 301
- [7] Zallen R and Moret M P 2006 *Solid State Commun.* **137**, 154
- [8] Braun J H, Baidins A and Marganski R E 1992 *Prog. Org. Coat.* **20**, 105
- [9] Yuan S A, Chen W H and Hu S S 2005 *Mater. Sci. Eng. C* **25**, 479
- [10] Murray C B, Norris D J and Bawendi M G 1993 *J. Am. Chem. Soc.* **115** 8706
- [11] Beltrán A, Gracia L and Andrés J 2006 *J. Phys. Chem. B* **110** 23417-23
- [12] Yang P, Lu C, Hua N and Du Y 2002 *Mater. Lett.* **57** 794–801
- [13] Chen D, Jiang Z, Geng J, Wang Q and Yang D 2007 *Ind. Eng. Chem. Res.* **46** 2741-46
- [14] Klug H P and Alexander L E, 1975 *J. Appl. Crystallogr* **8** 573-4
- [15] Yam M C, Chen F, Zhang J L and Anpo M 2005 *J. Phys. Chem. B* **109** 8673
- [16] Stengl V, Bakardjieva S and Bludska J 2011 *J. Mater. Sci.* **46** 3523

- [17] Langford J I and Wilson A J C 1978 *J. Appl. Cryst.* **11** 102-113
- [18] Hung W C, Chen Y C, Chu H and Tseng T K 2008 *Appl. Surf. Sci.* **255** 2205-13
- [19] Segne T A, Tirukkovalluri S R and Challapalli S, 2011 *Int. Res. J. Pure Appl. Chem.* **1** 3 84-103
- [20] Shin H, Jung H S, Hong K S and Lee J K 2005 *J. Solid State Chem.* **178** 15
- [21] Sheng Q R, Cong Y, Yua S, Zhang J L and Anpo M 2006 *Micropor. Mesopor. Mat.* **95** 220
- [22] Ohsaka T, Izumi F and Fujiki Y 1978 *J. Raman Spectrosc.* **7** 321
- [23] Porto S P S, Fleury P A and Damen T C 1967 *Phys. Rev.* **154** 522
- [24] Hardcastle F D and Ark J 2011 *Acad.Sci.* **65** 43
- [25] Katiyars R S, Dawsons P, Hargreaves M M and Wilkinson G R 1971 *J. Phys. C: Solid St. Phys.* **4** 2421–2431
- [26] Zhang W F, He Y L, Zhang M S, Yin Z and Cheng Q 2000 *J. Phys. D: Appl. Phys.* **33** 912–6
- [27] Mazza T, Barborini E, Piseri P, Milani P, Cattaneo D, Bassi A L, Bottani C E and Ducati C 2007 *Phys. Rev. B* **75** 045416
- [28] Chaudhary V, Srivastava A K and Kumar J 2011 *Mater. Res. Soc. Symp. Proc.* **1**.

Synthesis and magnetic-electronic characterization of mixed-valence $\text{LuFe}_2\text{O}_{4-\delta}$: effect of stoichiometry δ

M A Peck^{1,2}, W N Sibanda¹, G Digue¹, C Martin³, E Carleschi¹ and G R Hearne¹

¹Department of Physics, University of Johannesburg, PO Box 524, Auckland Park, 2006, South Africa

²Department of Chemistry, University of Johannesburg, PO Box 524, Auckland Park, 2006, South Africa

³Laboratoire CRISMAT, ENSICAEN, UMR 6508 CNRS, Caen, France

Email: adlipeck@gmail.com

Abstract. Samples of $\text{LuFe}_2\text{O}_{4-\delta}$ of varying oxygen stoichiometry (δ), prepared by solid state synthesis, have been characterised by x-ray diffraction, ^{57}Fe Mössbauer-effect spectroscopy, SQUID magnetometry and transmission electron spectroscopy as well as electron-diffraction analysis. Magnetisation measurements show that the onset of ferrimagnetic ordering is confined to 245-250 K for the synthesised samples. Detailed analysis of Mössbauer spectra measured at variable cryogenic temperature confirm the mixed valence state and reveal the effects of charge and spin frustration of the triangular network of Fe atoms. The electron diffraction analysis demonstrates the existence of charge ordered superstructure (satellites) as well as an inhomogeneous distribution of oxygen throughout the lattice, irrespective of the δ values.

1. Introduction

Multifunctional materials able to combine both magnetism and ferroelectric polarization, termed multiferroics, have strong potential for novel electronic functional devices [1]. This magneto-electric coupling offers manipulation of the electric polarization via a magnetic field and vice versa [2]. Rare-earth ferrites ($R\text{Fe}_2\text{O}_4$) have been shown to exhibit this phenomenon, with a great deal of emphasis towards lutetium ferrite (LuFe_2O_4) since Ikeda *et al.* showed the existence of a mixed-valence superstructure, $\text{Fe}^{2+}/\text{Fe}^{3+}$ charge-ordering (CO), within these compounds [3]. There were initial claims that the CO is intimately linked to ferroelectric properties; however this has now been refuted and the electric polarization is now considered to be an extrinsic effect [4].

A review by Angst [4] highlights that due to the frustrated LuFe_2O_4 lattice system, the triangular network of Fe atoms contains a number of degenerate charge and spin ordered states allowing transitions in both the magnetic and electric ordered dimensions. It has also been reported that oxygen-stoichiometry, $\text{LuFe}_2\text{O}_{4\pm\delta}$, plays a critical role in varying the physical and structural properties of this material. Both excess ($\delta > 0$) [5, 6] and deficient ($\delta < 0$) [7] samples lead to a suppression of the magnetic ordering volume and temperature (T_N) as well as a deformation of the structural lattice. The mechanism of this variation as a function of the absence or excess of oxygen is still not fully understood. In this paper we will report on the synthesis and magnetic-electronic characterization of three oxygen-varied $\text{LuFe}_2\text{O}_{4+\delta}$ samples and compare our results to those reported for excess and stoichiometric samples.

LuFe₂O₄ crystallizes in a rhombohedral lattice, consisting of an alternation of LuO₆ monolayers with triangular networked FeO₅ bi-layers (figure 1(a)), where Fe-O co-ordinates in a trigonal bipyramidal structure as indicated in figure 1(c), characteristic of these rare-earth ferrites [4]. Using high resolution transmission electron microscopy (TEM) and electron diffraction (ED) for samples with $\delta < 0$ Yang *et al.* observed that a structural modulation occurred as well as local distortions in the *b-c* plane and deformation of the FeO₅ polyhedron, presumably affecting magnetic transitions [7]. A similar disruption of the stoichiometric lattice was observed by Hervieu *et al.* for samples where $\delta > 0$, in which case the cationic system expands to accommodate excess oxygen (up to $\delta = 0.5$), resulting in a loss of the rhombohedral layer-stacking [5]. Their results point towards a reversible oxygen-storage application of LuFe₂O₄ as well as the possibility to fine-tune $\delta > 0$. Previous Fe Mössbauer spectroscopy results show that Fe³⁺/Fe²⁺ CO occurs below an onset temperature of $T_{CO} = 320$ K [8] and above 370 K electron-hopping, Fe²⁺ ↔ Fe³⁺, destroys the mixed-valence superstructure (CO melting) [9]. Stoichiometric samples show a paramagnetic to ferrimagnetic transition at $T_N \approx 250$ K [8, 9] whilst off-stoichiometric samples show a suppression of T_N to below 250 K [5-7]. An antiferromagnetic transition also occurs at 175 K for all samples, while only off-stoichiometric samples display spin-glass magnetic behaviour at lower temperatures, $T_{LT} < 175$ K [6, 7].

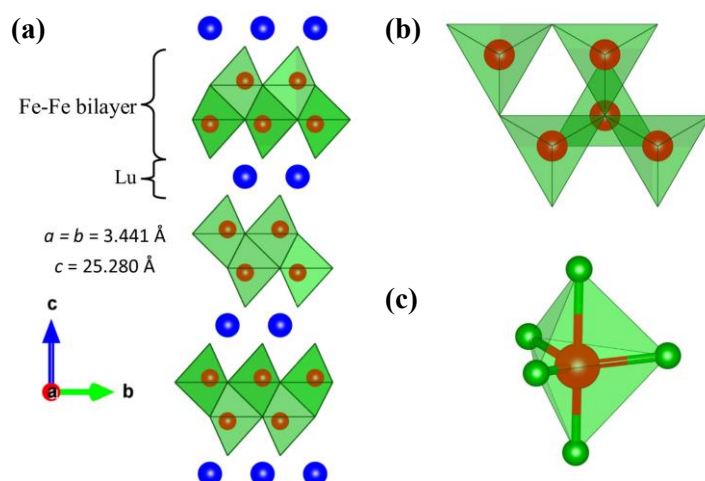


Figure 1 (colour online): (a) Rhombohedral lattice, $R3-m$, of LuFe₂O₄ viewed in the *b-c* plane (b) View down the *c*-axis of a single bilayer showing the triangular network of Fe atoms. (c) Five-fold coordination of Fe-O. Lu atoms are shown in blue, Fe in red and O atoms, at the vertices of the trigonal bipyramid, are shown in green.

2. Experimental

Two LuFe₂O₄ polycrystalline powdered samples, enriched to 10 - 14% ⁵⁷Fe, were prepared by solid-state reaction of starting mixtures according to the stoichiometric ratio 0.50Lu₂O₃: 0.83Fe₂O₃: 0.34Fe. Mixtures were thoroughly ground in an agate mortar, pestled and compressed into 20 mg and 100 mg pellets. The pellets were vacuum-sealed in a quartz ampoule to 3×10^{-6} mbar before heating at 1100 °C for 12 hrs. A third sample with $\delta \approx 0$ (designated MP), prepared on a gram-scale via a similar method, was received from Martin's group at the *Laboratoire Crismat, Caen, France* [5, 10]. This was also characterized and compared to our two synthesized samples.

Structural characterization was performed on synthesized samples by X-ray powder diffraction (XRPD) using a *Phillips X-Pert Pro* instrument with Cu-K α radiation. Magnetic susceptibility measurements were carried out in a SQUID (QD: MPMS) in the temperature range 150 – 300 K. ⁵⁷Fe Mössbauer effect spectroscopy (MES) measurements in transmission mode were conducted using a 10 mCi ⁵⁷Co(Rh) source and all samples were characterized with respect to a 25 μ m α -Fe-foil reference. Low temperature measurements were performed in a top-loading cryostat with the source and sample at the same temperature.

In addition electron diffraction measurements were performed on micron-dimensioned grains in a TEM, similar to those indicated in references [5, 10].

3. Results and discussion

3.1. XRPD

Of the two synthesized samples, the diffraction pattern for the 20 mg sample showed a significant amount of impurities, (indicated by arrows in figure 2(b)) relative to the 100 mg sample. Sample MP's diffraction pattern (not shown here, see reference [5]) showed no discernible impurity phases either. Since sample MP was produced on a gram-scale, a possible inverse proportionality mass-purity effect takes place leading to unreacted material. This can also be observed in the Mössbauer spectrum of the 20 mg sample (figure 2(c)), which shows evidence of remnant Fe_3O_4 and Fe impurity phases (as indicated by arrows) from the reaction sequence, not observed in the MP and 100 mg samples (see figure 3).

The crystal structure was resolved by Rietveld refinement and a best-fit, rhombohedral cell, with $R3\text{-}m$ space group was obtained with parameters $a = b = 3.441 \text{ \AA}$ and $c = 25.280 \text{ \AA}$. This is in good agreement with previous reports [10] and the pattern in figure 2(a) closely matches that of phase-pure samples of references [5] and [9]. The MP sample has already been well characterized in previous reports [5].

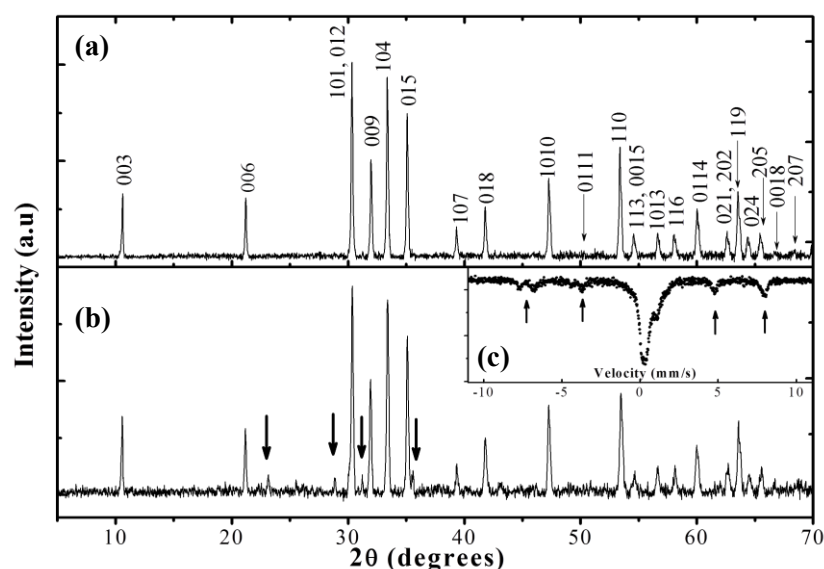


Figure 2 (colour online): XRPD data for (a) 100 mg and (b) 20 mg enriched LuFe_2O_4 pellets. Inset (c) shows the MES spectrum of the 20 mg sample. Arrows indicate impurity phases.

3.2. MES

Mössbauer spectra are shown for the MP and ^{57}Fe -enriched 100 mg samples at room temperature in figure 3 and at 15 K in figure 4. At room temperature the spectra are commonly fitted as a combination of two doublets, corresponding to Fe^{2+} and Fe^{3+} , typical of the mixed-valence state [9]. However, for a best fit it was necessary to include a $\text{Fe}^{3+}/\text{Fe}^{2+}$ electron-hopping (EH) component whilst allowing the isomer shifts and quadrupole splitting to correlate to those of the individual mixed-valence components. The fitting yields an electron exchange rate of $\sim 2 \text{ MHz}$. This likely occurs at CO domain boundaries or may be a result of CO frustration in the triangular network of figure 1(b). The abundance of the EH component in both samples is appreciable ($\sim 20\%$) with the 100 mg sample (figure 3(b)) having a slightly smaller EH frequency in comparison with the MP sample. This is possibly due to the different oxygen stoichiometries in the two samples.

At 15 K (figure 4), where $T \ll T_N$, complex magnetic hyperfine structure is observed. The spectra were both fitted using six sextets allocating three potential sites for Fe^{3+} and two for Fe^{2+} [9], as well as a smeared Fe^{2+} component. This multiplicity of Fe sites is compatible with the known ferrimagnetic structure of the compound and the Fe^{2+} and Fe^{3+} superstructure (CO) that occurs in each sheet of a bi-layer depicted in figure 1(a), as delineated in reference [4]. The model for each spectral component is calculated by solving the static Hamiltonian for mixed magnetic and quadrupole interactions using MOSSWINN [11]. The smeared component (broad linewidth, see last line of table 1) is supposed to

originate from disordered spins (fluctuating or a distribution of B_{hf} values), anticipated to occur in such a frustrated network, see figure 1(b). Longer-range magnetic exchange interactions within the bilayers give rise to the spin aligned (ordered) components. The parameters derived from the fitting of the MES spectra at 15 K are shown in table 1 and confirm that distinct charge states, mixed-valence Fe^{3+} and Fe^{2+} , occur; as compared to parameter values typical of ferrous and ferric iron [12]. It is interesting to note that the disordered spins comprise mainly Fe^{2+} moments.

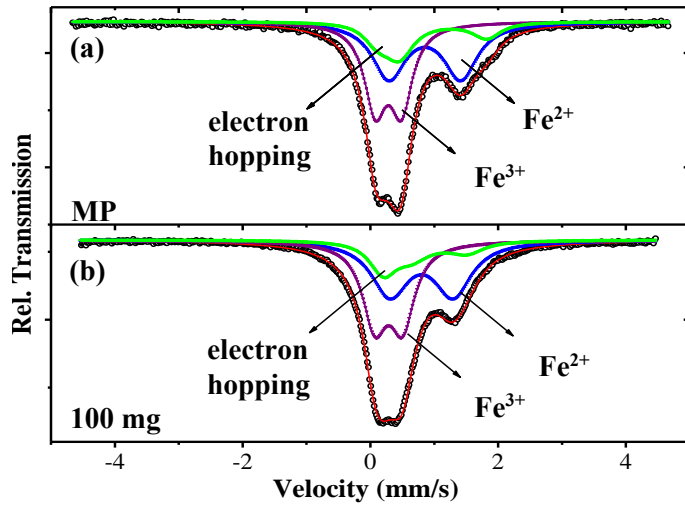


Figure 3 (colour online): Mössbauer spectra at RT for: (a) MP sample and (b) enriched 100 mg pellet.

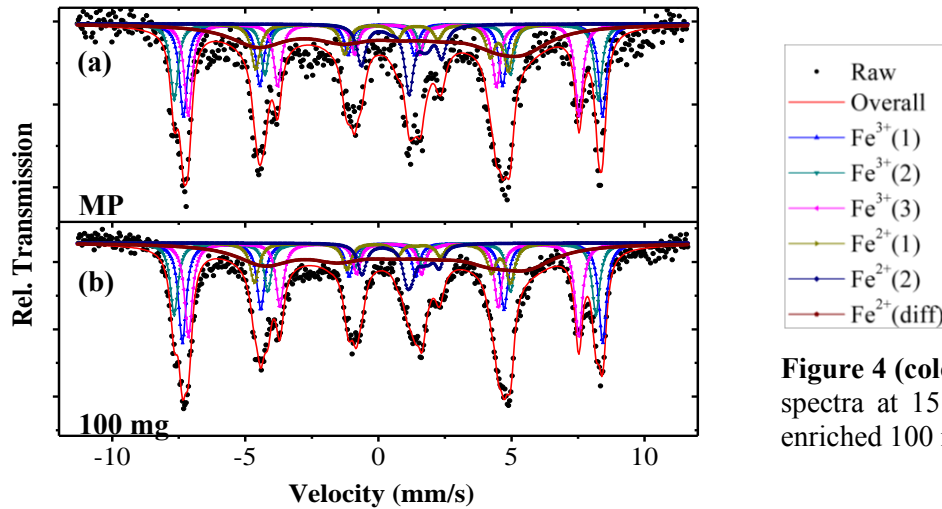


Figure 4 (colour online): Mössbauer spectra at 15 K for (a) MP and (b) enriched 100 mg pellet.

Table 1: MES parameters for the spectrum of the MP sample at 15 K. The isomer shift is δ and Beta is the angle between the electric field gradient V_{zz} and the internal magnetic field at the Fe site, B_{hf} . A very similar tabulation has been obtained for the enriched 100 mg sample.

Site	Γ line width ^a (mm/s)	δ / Fe (mm/s)	B_{hf} (T)	$V_{zz} \times 10^{21}$ (V/m ²)	$ 1/2 \cdot eQV_{zz} $ (mm/s)	Beta (deg)	Abundance (%)
Fe^{3+} (1)	0.35	0.46	49	2.65	0.44	0	18
Fe^{3+} (2)	0.35	0.45	50	-0.33	0.05	0	14
Fe^{3+} (3)	0.35	0.39	45	-7.70	1.28	51	18
Fe^{2+} (1)	0.41	0.93	29	-8.20	1.37	5	10
Fe^{2+} (2)	0.41	1.13	6	-8.90	1.48	26	10
Fe^{2+} (diff)	2.5	1.13	31	-6.50	1.08	0	30

^aFixed parameters

3.3. Magnetic susceptibility

Zero-field cooled magnetic susceptibility measurements on warming are given in figure 5 for the three samples. T_N has been determined from the point of inflection on the high temperature side of the peak. The 20 mg pellet shows a T_N of 247 K as well as a higher magnetic ordering volume, indicative of the impurity phases present. The 100 mg and MP sample show curves with a T_N of 249 and 255 K, respectively. The differences in T_N are due to differences in oxygen stoichiometry δ of the two samples [5-7]. The MP sample has $\delta \sim 0$ and the decrease in T_N in the enriched sample is attributed to a slight oxygen deficiency [7], which is in the process of being quantified by means of TGA measurements. The as received MP sample shows an additional transition at ~ 210 K believed to be a stabilization of one of the degenerate magnetic configurations [13]. This evidently does not occur in the off-stoichiometric sample.

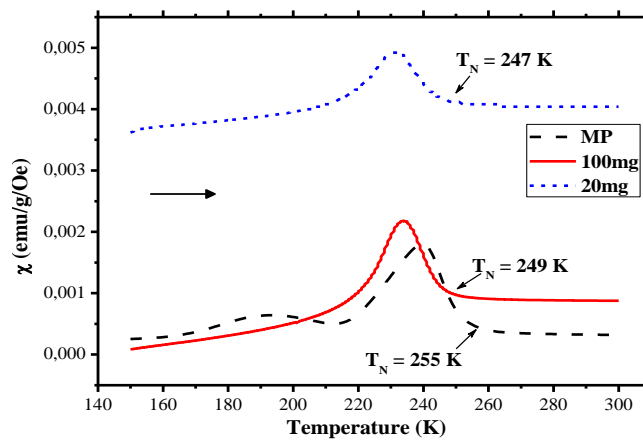


Figure 5 (colour online): Magnetic susceptibility for the ^{57}Fe -enriched 20 and 100 mg samples, as well as for the MP sample.

3.4. TEM and ED

In the numerous papers devoted to this ferrite, ED patterns along the [100] axis are characterized by additional reflections leading to commensurate or incommensurate modulation vectors; all the authors agree that they are the signatures of CO [5, 10]. ED from high resolution TEM studies were taken for the 100 mg ^{57}Fe enriched sample and compared to the as received MP sample.

Along the [100] axis in the MP sample (figure 6(a)), three modulation vectors are observed: \mathbf{q}_2 ($1/3 - \varepsilon$ $1/3 - \varepsilon$ $1/3$), \mathbf{q}_2' by twinning and \mathbf{q}_3 ($1/3$ $1/3$ $1/2$). The point-like satellites are rather intense, however weaker or rather diffuse satellites are also observed due to the $C2/m$ sub-cell of $R3-m$ [5]. Similarly, in the 100 mg enriched sample (figure 6(b)) weak zigzagging diffuse lines are observed (highlighted by the red arrows), which are associated with the \mathbf{q}_2 , \mathbf{q}_2' and \mathbf{q}_3 vectors previously reported. This is similar to the MP samples (likely with short range charge ordered domains); except more intense satellites are present in the MP sample.

Along [010], the ED patterns of the 100 mg (figure 6(c)) and MP samples exhibit either no extra reflections, consistent with the $C2/m$ structure, or there are satellites associated with an incommensurate modulated structure. The latter ones have been shown to be correlated with the formation of “oxidized” domains present in the matrix [5]; thus considered to be a true signature of a local excess of oxygen. The 100 mg sample thus shows an inhomogeneous distribution of oxygen throughout the lattice, as is the case in the MP sample (not shown here).

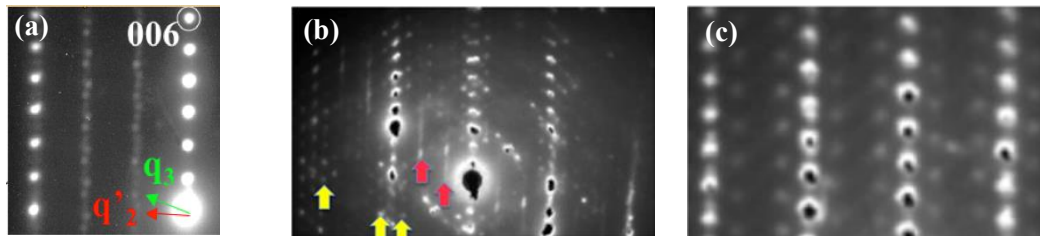


Figure 6 (colour online): ED patterns of (a) MP sample and (b) 100 mg enriched sample, along the [100] axis. Satellites from CO superstructure are emphasized by the red arrows; these are more intense in (a) and much weaker but nevertheless discerned in (b). Yellow arrows indicate unknown point-like satellites. (c) ED pattern for the 100 mg enriched sample along the [010] axis, showing satellites correlated to locally oxidized domains.

4. Conclusions

We are able to prepare small quantities of ^{57}Fe isotopically enriched mixed-valence $\text{LuFe}_2\text{O}_{4-\delta}$ samples in which CO prevails. Such isotopic enrichment is imperative for future envisaged MES studies under high pressure. These are phase pure materials when a minimal amount of ~ 100 mg is synthesized; there are slight deviations of oxygen stoichiometry from ideal $\delta \sim 0$ values. Oxygen stoichiometry may readily be altered by a mild heat treatment (150 - 200 °C) in air [5]. In both stoichiometric and off-stoichiometric samples oxygen is inhomogeneously distributed in the lattice; there being oxidized domains with a local excess of oxygen. The degree of off-stoichiometry does not drastically alter the magnetic-electronic properties as exemplified by the magnetization and Mössbauer characterizations. These results have been compared to those for a well-characterized reference sample ($\delta \sim 0$).

We have conducted a more comprehensive data analysis of the Mössbauer spectra than in the previous literature. The effects of geometrical charge and spin frustration, anticipated for such a triangular networked lattice structure, have been discerned in this detailed analysis by way of a residual electron-hopping component at $T < T_{\text{CO}}$ and disordered spins at $T \ll T_{\text{N}}$, respectively.

References

- [1] Bibes M and Barthélémy A 2008 *Nat. Mater.* **7** 425-6
- [2] Béa H, Gajek M, Bibes M and Barthélémy A 2008 *J. Phys.: Condens. Matter* **20** 434221 (1-11)
- [3] Ikeda N, Ohsumi H, Ohwada K, Ishii K, Inami T, Kakurai K, Murakami Y, Yoshii K, Mori S and Horibe Y 2005 *Nature* **436** 1136 (1 - 3)
- [4] Angst M 2013 *Phys. Status Solidi RRL* **7** 383-400
- [5] Hervieu M, Guesdon A, Bourgeois J, Elkaïm E, Poienar M, Damay F, Rouquette J, Maignan A and Martin C 2014 *Nat. Mater.* **13** 74-80
- [6] Wang F, Kim J, Gu G D, Lee Y, Bae S and Kim Y-J 2013 *J. Appl. Phys.* **113** 063909 (1 - 7)
- [7] Yang H X, Tian H F, Wang Z, Qin Y B, Ma C, Li J Q, Cheng Z Y, Yu R and Zhu J 2012 *J. Phys.: Condens. Matter* **24** 435901 (1 - 7)
- [8] Xu X, Angst M, Brinzari T, Hermann R P, Musfeldt J, Christianson A D, Mandrus D, Sales B, McGill S and Kim J-W 2008 *Phys. Rev. Lett.* **101** 227602 (1 - 4)
- [9] Bang B K, Kouh T and Kim C S 2008 *J. Appl. Phys.* **103** 07E307 (1 - 3)
- [10] Rouquette J, Haines J, Al-Zein A, Papet P, Damay F, Bourgeois J, Hammouda T, Doré F, Maignan A and Hervieu M 2010 *Phys. Rev. Lett.* **105** 237203 (1 - 4)
- [11] Klencsár Z 2013 Mössbauer fitting program: *MOSSWINN* 4.0 <http://www.mosswinn.com/>
- [12] Gütlich P, Bill E and Trautwein A X 2011 *Mössbauer Spectroscopy and Transition Metal Chemistry: Fundamentals and Applications* (Berlin: Springer)
- [13] Christianson A D, Lumsden M D, Angst M, Yamani Z, Tian W, Jin R, Payzant E A, Nagler S E, Sales B C and Mandrus D 2008 *Phys. Rev. Lett.* **100** 107601 (1 - 4)

A Study of Crack Formation and its Effect on Internal Surface Area using Micro-Focus X-ray Computerised Tomography and Fractal Geometry.

T Seakamela¹, G Nothnagel², F C de Beer² and B P Doyle¹

¹Department of Physics, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, South Africa

²South African Nuclear Energy Corporation Limited, P.O. Box 582, Pretoria, 0001, South Africa

E-mail:teffoseakamela@gmail.com

Key words: Micro-focus X-ray tomography; Fractal geometry; Cracks

Abstract. Micro-focus X-ray tomography, as a non-destructive analytical technique, has been applied to obtain a virtual three dimensional image (tomogram) of material at a scale as low as 3 μm . Image analysis software was used to extract quantitative information from 2D planar sections of the tomograms to study the geometry of cracks. The aim of the study was to assess the ability of the technique to determine if crack formation in chosen samples has a fractal character, as this may provide a compact means of characterising the effective surface area of the crack plane. Studies were performed on samples of different materials, viz smooth mortar, large and small bricks and coal. It was found that the cracks, or fracture surface edges, show fractal behaviour for a wide range of length scales but do not give consistent values for different cracks in the same material for the samples studied.

1. Introduction

There is an interest in the nuclear energy industry in the study of the confinement integrity of nuclear waste encapsulation materials, in which case it is important to study materials for the presence of undesirable micro-cracks and pores. In the coal industry, on the other hand, cracks are desirable for gasification of the coal, and methods to characterise their nature and ways to produce them become important. This paper describes the use of micro-focus X-ray tomography (μCT) [1] for the analyses of cracks found within coal, small and large bricks and smooth mortar samples. The μCT technique, with machines such as the MIXRAD facility at Necsa, produces 3D images of a material's internal structure to a best spatial resolution of about 3 μm [2], depending on the sample size (the smaller the sample, the higher the resolution). However, there are similar techniques that are capable of going down to 500 nm spatial resolution such as the nano-focus X-ray CT facility situated at overseas located synchrotron facilities [3].

Fractal methods have been used to study self-similar natural phenomena in many fields of scientific research [4, 5, 6] and have been used in this work to characterize crack formation via details of the crack edge. Fractal structures show self-similarity, which is revealed through a power-law relationship between two variables, and is characterised by a non-integer fractal dimension [7, 8]. In this work we used the compass method as described in Section 2.2.

2. Experimental set-up

The experimental set-up consists of an X-ray source, 2-dimensional high resolution digital X-ray detector and a rotating mechanical sample manipulation stage. Radiographs of the total sample are taken at multiple equally spaced angles between 0 and 360°.

2.1 Sample preparation

In μ CT scanning there is no need for specific sample preparation. A sample such as a piece of coal was formerly cracked in a predetermined process and then analysed using the calibrated and standardized micro-tomography set-up [9]. Figure 1 below shows the samples that were inspected for this study.

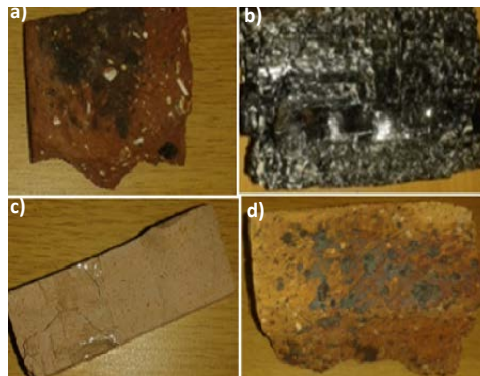


Figure 1: a) Piece of a small brick, b) coal, c) smooth mortar and d) piece of a large brick

2.2 Experimental set-up, image acquisition, reconstruction and analysis

The samples were scanned using the micro-focus X-ray radiography and tomography system (MIXRAD) located at Necsa [9]. The instrument, shown in figure 2, is equipped with a tungsten X-ray source with anode voltage settings ranging between 30 kV and 255 kV, beam current up to 1 mA and a spot size of 3 μ m. A schematic of the process of acquiring 2D radiographic slices, which are then combined into a 3D image via a reconstruction algorithm, is shown on the right.

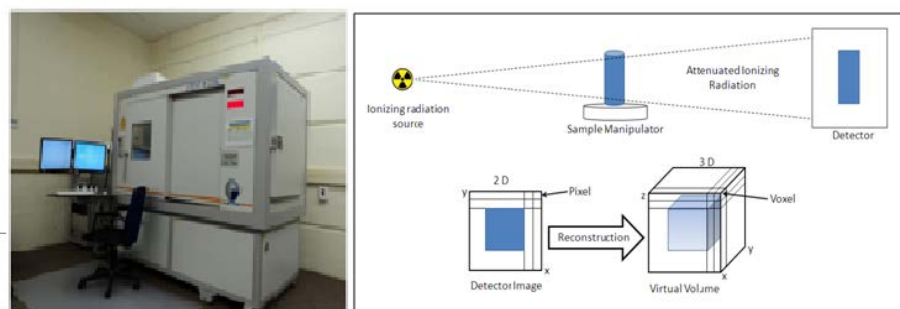


Figure 2: Micro-focus X-ray machine and tomographic process at the MIXRAD facility at Necsa

The X-ray source scanning parameters were set to achieve sufficient contrast in the radiograph to provide a good visualisation of cracks. Due to the size of the samples scanned in this study, the spatial resolution obtained was ≈ 0.022 mm (22 μ m).

After tomographic reconstruction, 3D visualisation by means of VGStudio MAX 2.2 software [9] was performed where 2D slices at appropriate crack location were generated. Image analysis software (Image-J) was then used to study crack properties on these 2D planes of the

virtual 3D object. For this purpose planes that intersect cracks were chosen so as to visualize the line features of the crack across the selected plane. Figure 3 is an example of the image of a crack intersection line on such a surface.

Image-J allows the linear distance between any two points on the plane to be measured in pixels (or calibrated length units) so that the compass method of length measurement along the crack intersection curve can be used to assess self-similarity and fractal nature. In the compass method a fixed scale length is chosen in terms of which the length along a section of a crack line is measured. This is seen in figure 3 where a given scale length N (small line segments between adjacent points), was used to measure the total crack length L_{AB} between pre-determined starting and end points, A and B respectively.

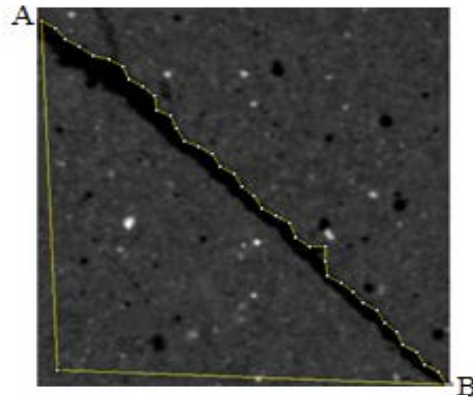


Figure 3: A crack on a plane

The measurement of $L_{AB}(s)$ where s is the scale, is performed for a number of scales values and $\log L_{AB}(s)$ is then plotted against $\log(1/s) = \log(r)$, where r is the resolution. If the geometry of the curve edge is fractal in nature, with ever smaller self-similar structure, the measured length will increase indefinitely (within practical limits), with decrease in scale length, as follows [4]:

$$\log L_{AB}(s) = D_C \log(1/s) + k \quad (1)$$

Where D_C is the so-called compass dimension and k is a constant.

3. Results and discussion

3.1 Smooth mortar sample:

An example of data for a typical measurement on a single crack is shown in table 1 in terms of the quantities defined above and the resulting $(\ln L_{AB})$ versus $\ln(1/s)$ graph is shown in figure 4.

Table 1. Example of smooth mortar data with an instrument resolution of 0.022 pixel/mm.

Scale (Number of pixels)	Number of scale lengths	Scale (mm)	Total Length (mm/pixel)	$\ln(1/s)$ (s in mm^{-1})	$\ln(L_{AB})$ (L_{AB} in mm)
3	98.667	0.066	6.512	2.718	1.874
4	73.00	0.088	6.424	2.430	1.860
8	34.00	0.176	5.984	1.737	1.789
10	27.00	0.220	5.940	1.514	1.782
14	18.16	0.308	5.732	1.178	1.746
16	16.50	0.352	5.808	1.044	1.759
20	13.00	0.440	5.720	0.821	1.744
22	11.55	0.484	5.581	0.726	1.721
24	10.56	0.528	5.573	0.639	1.718

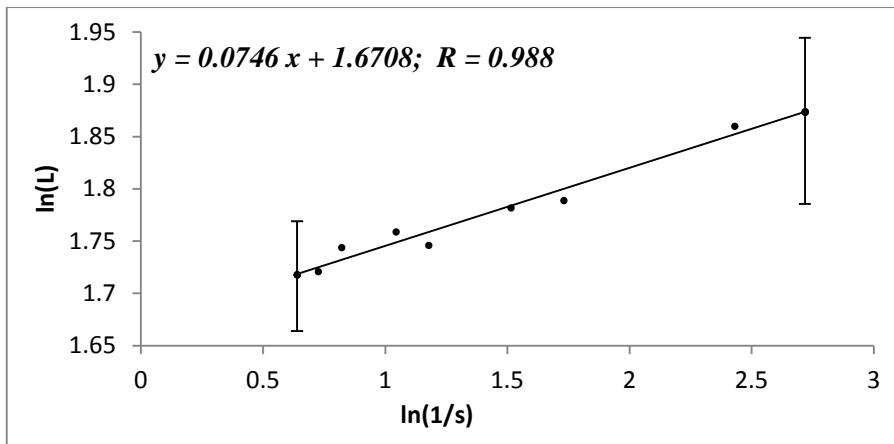


Figure 4: Ln-Ln fractal analysis plot for a crack in the smooth mortar

Similar quantification of cracks was performed in coal, a large and small brick sample were analysed and results are shown in table 2.

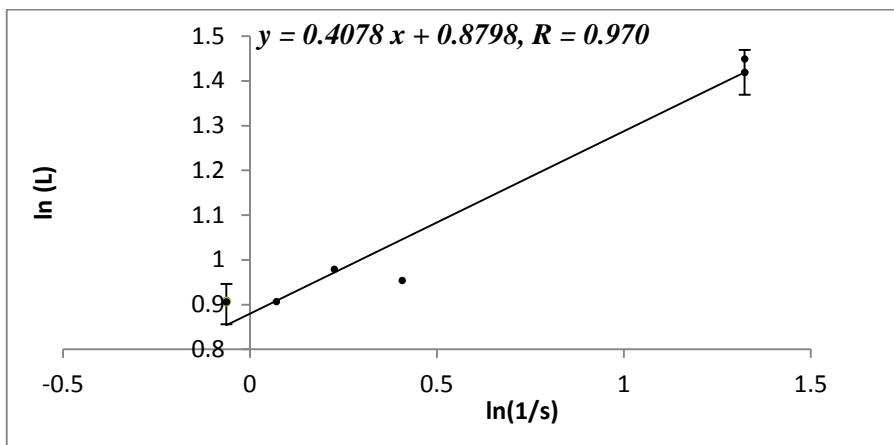


Figure 5: Ln-Ln fractal analysis plot for the coal sample.

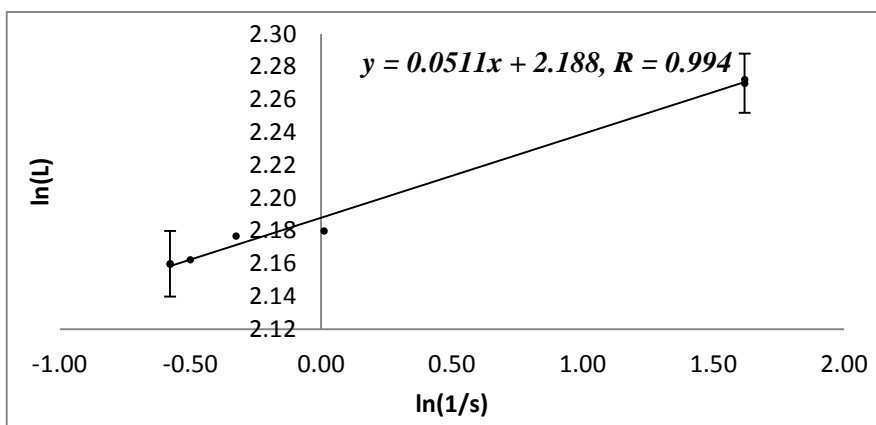


Figure 6: Ln-Ln fractal analysis plot for the large brick sample

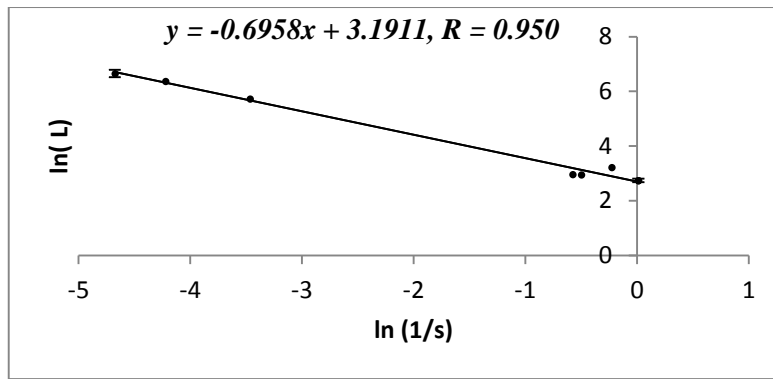


Figure 7: Ln-Ln fractal analysis plot for the small brick sample.

Table 2 contains the summarised results of several cracks in several samples and figure 5 shows the graph for a crack in the coal sample with large fractal dimension. The dimension is large because the crack measured was jagged and sinuated.

Table 2. Tabulation of data from several cracks in several sample types.

Sample description	Fractal dimension	Crack length (mm/ pixels)	Correlation coefficient
Smooth mortar	1.075	5.920	0.987
	1.026	2.575	0.997
	1.020	10.10	0.981
Coal	1.408	2.905	0.970
	1.083	1.534	0.968
	1.049	3.885	0.993
Large brick	1.051	8.946	0.994
	1.025	12.51	0.966
Small brick	1.696	198.6	198.6
	1.015	9.109	9.109
	1.025	9.160	9.160

Figure 8 below shows the crack that was segmented approximately in the middle of the sample. The total volume and surface area of this crack were found to be 49.987 mm³ and 738.863 mm² respectively.

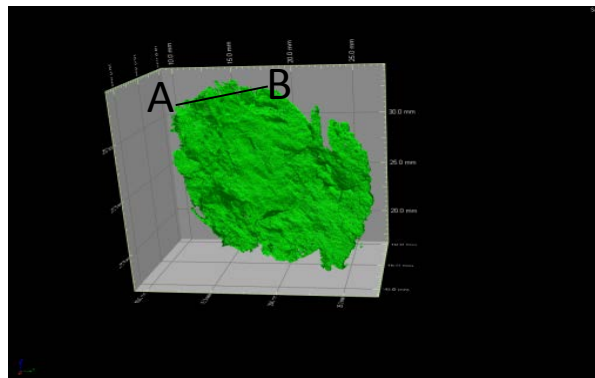


Figure 8: One crack segmented from the coal and displayed in 3D

In figure 4 both the error bar separation and the correlation coefficient indicate a bona fide linear relationship, despite its low value, thereby suggesting a fractal dimension of ~ 1.07 for this crack. Thus we may conclude that the method allows determination of fractal character down to rather low compass dimension values. A crack and thus the material is only fractal when the slope of $\ln(L)$ vs. $\ln(1/s)$ holds for a range of spatial resolutions, thus it becomes clear that the scale can vary in the measurement because it does not affect the slope by changing it to negative. For the small brick sample on figure 7, the slope is negative therefore an absolute of (-0.696) will be taken in and the fractal dimension will become 1.696. The results in table 2 indicate low precision for cracks in similar samples and no clear relationship between the length of crack analysed and the outcome of the fractal analysis. This indicates that it may not be possible to assume the same fractal dimension for all cracks in a given sample size, that the details at this resolution for the different sections of the cracks limit the ability to compare or that other factors such as inhomogeneity of impurities may influence the dimension from crack to crack. On the figure 4 to 7 some data points (outliers) are observed distant away from other observation points. This may be due to variability in the measurement or they may indicate experimental error, the latter are sometimes excluded from the data set. Hence they still contribute to the trend line.

4. Conclusions

Based on the findings of this first study, micro-focus X-ray tomography proves to be a fast technique (no special sample preparation) for the 3D visualization and quantitative analysis of cracks in samples penetrable with X-ray energies up to 70 keV. It was found that the fracture surfaces of the smooth mortar, large and small bricks and coal follow fractal behaviour that is represented by the fractal dimensions recorded in table 2.

Preliminary results show that it may be difficult to rely on the absolute value of the fractal dimension for material characterisation, which does not appear to be the same for all samples of the same material. Further research is needed to assess how much of this is due to inherent measurement limitations of the technique, including plane and crack selection, and how much on real sample properties (such as hardness and impurity distribution).

References

- [1] Van Dalen G, Don A, Nootenboom P and Blonk J C G 2012 *Microscopy and Analysis* S8 (March)
- [2] Landis E N and D T Keane 2010 X-ray Microtomography *Mater. Charact.* **61** 1305
- [3] Bruke O, Brockdorf K, Drews S, Müller B, Donath T, Herzen J and Beckmann F 2008 *Proceedings on the Developments in X-ray Tomography VI* **7078** 16
- [4] Charkaluk E, Bigerelle M and Iost A 1998 Fractals and Fracture *Eng. Fract. Mech.* **61** 119
- [5] Normando P G, Nascimento R S, Moura E P and Vieira A P 2013 *Phys. Rev. E* **87** 043304
- [6] Jiang M Q, Meng J X, Gao J B, Wang X L, Rouxel T, Keryvin V, Ling Z and Dai L H 2010 *Intermetallics* **18** 2468
- [7] Lee K I, Choi S C, Park T N and You D S 1999 *Dentomaxillofac. Rad.* **28** 284
- [8] Kraft R 1997 Fractals and Dimensions *Journal of Agricultural, Nutritional and Environmental Sciences* <http://web2.wzw.tum.de/ane/janes.html>
- [9] Hoffman J W and De Beer F C 2012 *Proceedings of the 18th World Conference on Nondestructive Testing* **18** 16
<http://www.ndt.net/article/wcndt2012/papers/37wcndtfinal00037.pdf>

Charge ordering dynamics under pressure in LuFe_2O_4

W N Sibanda¹, E Carleschi¹, G Diguët¹, C Martin² and G R Hearne¹

¹Department of Physics, University of Johannesburg, PO Box 524, 2006 Auckland Park, Johannesburg, South Africa

²Laboratoire CRISMAT, ENSICAEN, UMR 6508 CNRS, Caen, France

Author e-mail address: wnsibanda@gmail.com

Abstract. The pressure response of charge order behaviour in LuFe_2O_4 has been investigated using pressure as a perturbing thermodynamic variable up to 7 GPa in diamond anvil cell experiments at 300 K. ^{57}Fe Mössbauer spectroscopy was used as the probing technique for magneto-electronic properties. An analysis of the Mössbauer effect spectra indicates that above 3 GPa the charge order has completely collapsed, also evidenced by the change in the Mössbauer spectral line shape. Above 3 GPa the magneto-electronic phenomena can be understood in terms of electron hopping processes, which occur in the time window of the Mössbauer effect (10^{-7} s). As a consequence of this Fe^{2+} and Fe^{3+} valence states are indistinguishable and are described by an intermediate valence state, $\text{Fe}^{2.5+}$. This destabilisation of charge order is compatible with magnetic collapse seen in neutron diffraction experiments. Charge carrier confinement, $\text{Fe}^{2+} \leftrightarrow \text{Fe}^{3+}$, prevails at these high pressures.

1. Introduction

LuFe_2O_4 belongs to a family of hexagonally layered metal oxides, RFe_2O_4 , where R is a rare-earth metal. Recently there has been a surge of interest in LuFe_2O_4 because of multiferroic properties being apparently observed in this system [1]. Multiferroic compounds show a strong coupling between magnetism and ferroelectricity; hence the combination of these two ferroic orders in a material has the potential for the tailor-making of multifunctional devices. Initially LuFe_2O_4 was thought to be the prototypical compound for charge order (CO) based ferroelectricity, where CO refers to a localisation of charge resulting in a mixed-valence superlattice (e.g., $\text{Fe}^{2+} - \text{Fe}^{3+} - \text{Fe}^{2+} - \text{Fe}^{3+} \dots$). Properties reported in this compound due to CO-based ferroelectricity included a large magneto-dielectric response and pyroelectricity near the spin ordering temperature [1]. Recent work has cast doubt on these observations because of the hitherto unknown subtle influence of electrical contacts in generating colossal values of the magneto-dielectric response [2]. This also puts in question the mechanism of charge order driven ferroelectricity. Despite this, there is still considerable interest in LuFe_2O_4 because the magneto-electric coupling is still pronounced.

LuFe₂O₄ has a rhombohedral unit cell (space group, $R\bar{3}m$) and exhibits a layered structure at room temperature, depicted in figure 1. The structure is described by an alternating stacking of [LuO₂]_∞ layers and [Fe₂O₄]_∞ bi-layers along the *c*-axis (figure 1(a)). The unit cell consists of three Fe bi-layers with a 1:1 ratio of Fe²⁺ and Fe³⁺. There is a superstructure arrangement of Fe²⁺ and Fe³⁺ ions within the triangular network of each sheet of a bilayer (figure 1(b)). This particular Fe²⁺/Fe³⁺ CO arrangement has long been thought to lead to charge segregation along the *c*-axis, leading to a net dipole moment and thus the occurrence of ferroelectricity [3, 4]. Lafuerza *et. al.* [5, 6] however have contested this model of an appreciable (ionic) valence separation as well as the occurrence of ferroelectric behavior; although valence segregation does seem to occur.

The corresponding spin values are $S=2$ for Fe²⁺ (d^6) and $S=5/2$ for Fe³⁺ (d^5). The triangular network in the Fe monolayers (each layer is made up of corner sharing FeO₅ trigonal bipyramids, see figure 1(c)) presents a scenario in which the magnetic and charge order interactions are incompatible with the topology of the crystal lattice, hence geometrical CO and spin frustration occur (figure 1(b)). Geometrical frustration can result in a plenitude of degenerate ground states. For example, LuFe₂O₄ undergoes different phase transitions: a 3-D charge order develops at $T_{CO} \approx 320$ K with cooling, while for $320 \text{ K} < T < 500 \text{ K}$ a 2-D CO ground state occurs. 3-D ferrimagnetic ordering sets in at $T_N \approx 240$ K, and a magneto-structural transition is observed at $T \approx 175$ K. Makarova *et. al.* have reported a gradual disappearance of long range magnetic order with pressure up to 3 GPa at 50 K in neutron diffraction measurements [7]. This is thought to be linked to the destabilisation of the magnetic moment and may be correlated with the melting of the CO. The purpose of this work is to investigate the behaviour of the CO in this pressure range.

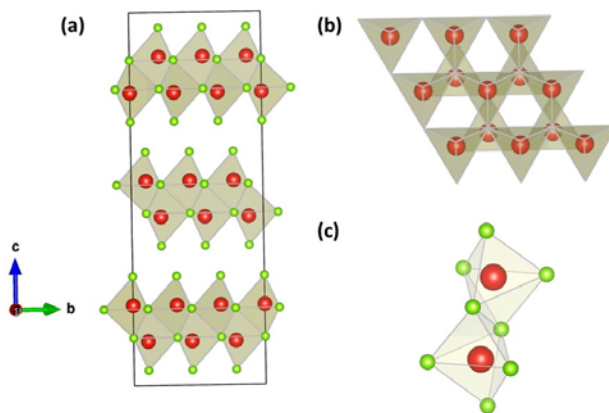


Figure 1 (Colour online): (a) Crystal structure of LuFe₂O₄ at room temperature, showing the $R\bar{3}m$ rhombohedral layered structure atomic arrangement. The unit cell consists of three Fe bi-layers stacked along the *c*-axis. The Fe ion (in red) is coordinated in trigonal bi-pyramidal polyhedra with O ions (in green). The [LuO₂]_∞ layers (not shown) are in between the bi-layers. (b) Projection of the *ab*-plane viewed along the *c*-axis showing the triangular network of Fe ions. (c) Isolated trigonal bi-pyramids sharing a corner. Structure insets generated using VESTA software [8].

To the best of our knowledge Mössbauer pressure measurements on LuFe₂O₄ have not yet been reported. ⁵⁷Fe Mössbauer spectroscopy, which is a nuclear γ -ray spectroscopy, allows probing the valence state of Fe as the only electro-magnetic active ion in the compound. To this end, a combination of high-

pressure methodology and ^{57}Fe Mössbauer spectroscopy has been used in this work to investigate the magneto-electronic behaviour of a highly stoichiometric sample, $\text{LuFe}_2\text{O}_{4+\delta}$ ($\delta=0$). In this present work we show that the CO collapses above 3 GPa in LuFe_2O_4 at 300 K, thereafter only dynamic charge exchange (electron hopping) are observed up to the maximum pressure, 7 GPa, reached in this study.

2. Experimental details

A polycrystalline sample of LuFe_2O_4 (with natural abundance of ^{57}Fe , 2%) was loaded into a Merrill-Basset diamond anvil cell (DAC). Details of its synthesis can be found elsewhere [9]. The sample and a few ruby balls for pressure calibration were loaded into a sample chamber made by drilling a 225 μm hole in a Re gasket pre-indented to 30 μm thickness. Daphne 7373 oil was used as pressure transmitting medium. A Wissel constant acceleration motor was used to scan the velocity range of interest (typically ± 5 mm/s) with a ^{57}Co (Rh) “point” source ($\approx 0.5 \times 0.5$ mm, 10 mCi) at room temperature [10]. A KrCO_2 (1 atm) proportional counter was used to detect the 14.4 keV resonance γ -radiation.

The model used for analysis describes electron hopping relaxation within the framework of the Blume-Tjon theory [11]. In the case of electron hopping processes, the ^{57}Fe nuclei experience a Hamiltonian that relaxes between two eigenstates (with a relaxation rate of ω [rad/s]) characterised by different isomer shift (IS) and quadrupole interaction values. The Hamiltonian is described by:

$$\begin{aligned} H_1 &= \frac{\Delta_1}{6} (3m^2 - I(I+1)) + \delta_1 I \\ H_2 &= \frac{\Delta_2}{6} (3m^2 - I(I+1)) + \delta_2 I \end{aligned} \quad (1)$$

where H_1 and H_2 are the different eigenstates experienced by the ^{57}Fe nuclei. Each eigenstate gives rise to a characteristic quadrupole split doublet with associated isomer shift, quadrupole splitting, QS (Δ) and IS (δ) respectively. Thus H_1 gives rise to Δ_1 and δ_1 , and H_2 to Δ_2 and δ_2 . I and m are the nuclear spin quantum number and magnetic quantum number respectively. The IS and QS values fluctuate simultaneously between characteristic values for Fe^{2+} and Fe^{3+} (i.e. the average valence state is $\text{Fe}^{2.5+}$). The relaxation rate in [rad/s] is involved in the lineshape profile which is derived from a solution of a time dependent Hamiltonian characterizing the fluctuation between the two states (profiles) in Equation 1 [11, 12]. It is therefore obtained as one of the fitting parameters of the spectral profile and then converted to electron hopping frequency [Hz].

The spectra were analyzed using an appropriate Mössbauer spectra fitting program from which the hyperfine interaction parameters and the corresponding relative abundances of spectral components were derived [13]. Thickness broadening, pressure gradients and geometrical effects were taken into account in this analysis. The spectra were evaluated assuming Lorentzian line shapes for the resonant absorption patterns. The reported IS values are relative to α -Fe at 300 K.

3. Results and discussion

Figure 2 below shows a selection of Mössbauer spectra measured at various pressure runs up to 7 GPa at 300 K. The model used in the analysis was first implemented in the ambient spectra measured with the LuFe_2O_4 sample loaded in a conventional 1.7 cm^2 copper clamp holder and a 20 mCi ^{57}Co (Rh) conventional source. This ensures better counting statistics, high signal-to-noise ratio and high resolution obtained by using a conventional source with a narrower line-width; hence reliable hyperfine parameters

could be extracted (see figure 2 (a) for comparison). The parameters obtained were then used as the starting values for the compression measurements.

The fitting model adopted for the sample at ambient pressure and temperature incorporates a fluctuating valence state; a markedly different approach from previous works [14, 15]. The observed Mössbauer spectra were analysed with three sub-spectra; two distinct quadrupole doublets describing Fe^{2+} (in blue) and Fe^{3+} (in green) sites with intensity ratio 1:1 describing the static CO. A third component was added to describe electron hopping between Fe^{2+} and Fe^{3+} ($\text{Fe}^{2+} \leftrightarrow \text{Fe}^{3+}$) (in red). This originates at CO nano-domain boundaries or is a result of the geometrical charge frustration [16-18]. The presence of the right most absorption peaks (higher velocity side) in figure 2 (a) are a signature of Fe^{2+} . Hence we deduce that the CO is already present at room temperature. The disappearance of this absorption peak would therefore indicate CO collapse.

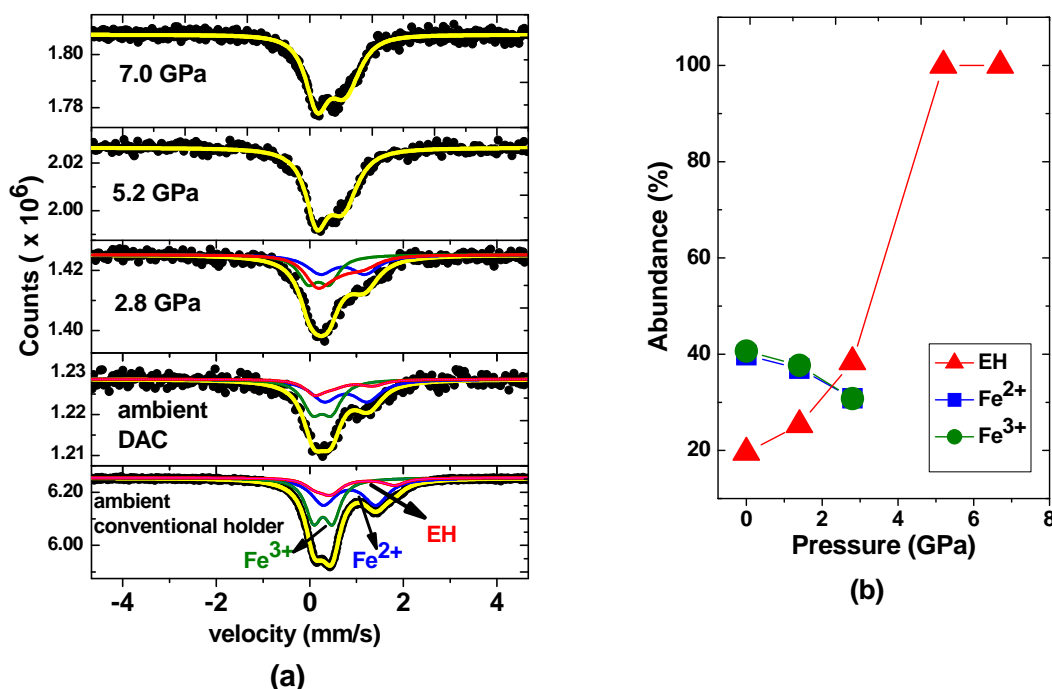


Figure 2 (Colour online): (a) Selected ^{57}Fe Mössbauer spectra of the pressure response of LuFe_2O_4 at 300 K plotted on a restricted velocity scale. The black circles are the data points and the solid line through the data points is the overall fit to the data. With an increase in pressure, at ≈ 3 GPa, the Fe^{2+} signature is no longer easily discernable, indicating the initiation of CO collapse. At $P > 3$ GPa (see 7 GPa spectrum), the Mössbauer line shape is fitted with a component consistent with electron hopping processes only in the MHz regime (indicated by EH). (b) A plot of the EH component abundances as a function of pressure. The increase in the abundance of the EH component is readily apparent. Solid lines through the data points are to guide the eye.

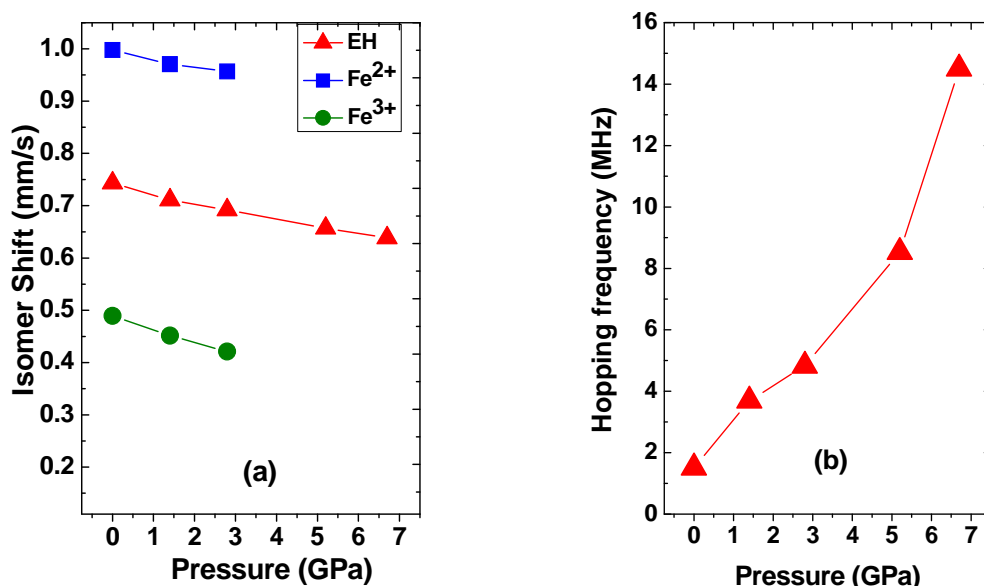


Figure 3 (Colour online): (a) Pressure dependence of the IS. The IS of EH (red triangle) is an average of the corresponding values of the fitted components, Fe^{2+} and Fe^{3+} in equation (1). The IS decreases with increasing pressure. (b) The pressure dependence of the hopping frequency as deduced from the fitting of the MS spectra. Solid lines are to guide the eye in (a) and (b).

The fit results of the ambient pressure spectrum reveal the co-existence of static CO (site-centered CO) of Fe^{2+} and Fe^{3+} in a 1:1 ratio and $\approx 20\%$ abundant sites where electron hopping is occurring (as explained above), in which the hopping frequency is 1.5 MHz, see figure 2 (b) and figure 3 (b). As the pressure increases, the Fe^{2+} signature becomes less discernable, indicating the beginning of CO collapse (see the Mössbauer spectrum at 3 GPa in figure 2 (a)). This is also confirmed by the increase in the abundance of the EH component with pressure (see figure 2(b)). Rouquette *et al.* observed a progressive rhombohedral to orthorhombic structural phase transition of the unit cell from 4–8 GPa [9]. Our Mössbauer analysis therefore infers that the CO collapse and structural phase transition are concomitant. In the range 5–7 GPa the Mössbauer line shape is consistent with electron hopping *only on* a nanosecond time scale (MHz regime), which is within the time (sensing) window of the Mössbauer effect of 100 ns [12, 16].

Figure 3 (a) shows the obtained values of the IS as a function of pressure. The plotted values of the IS (in red) for the EH component are the average of that of the individual components (Fe^{2+} and Fe^{3+}) describing the electron hopping model in equation (1). For all components, the IS decreases with pressure, which is attributed to the s -electron density increasing near the Fe nucleus. In fact, the IS is related to s -electron density, ρ_s , at the nuclear site. The quadrupole splitting (QS) of the components is observed to increase with pressure as expected (not shown). QS reflects deviation from cubic symmetry, an increase in QS is therefore attributed to increased site distortion around the Fe ion when pressure increases.

The pressure dependence of the hopping frequency is plotted in figure 3 (b) and shows a substantial ten-fold increase in the measured pressure range. Pressure decreases the inter-atomic distance and therefore the hopping between Fe sites is expected to be enhanced, consistently with what is observed here. The presence of *only* electron hopping and IS values intermediate between that of Fe^{2+} and Fe^{3+} above 3 GPa infer that the electrons are still confined, perhaps in some sort of dimerised bond between

Fe-Fe pairs in the *ab*-plane of the bi-layers. This would therefore suggest that the compound is still an insulator in this pressure regime, even though the CO is destabilized [19].

4. Conclusions

In summary, we have investigated the stability of the CO in a highly stoichiometric sample of LuFe₂O₄. The CO is destabilised at 3–7 GPa at room temperature. The Mössbauer spectra above 3 GPa indicate *only* electron hopping processes with relaxation rates on the nano-second time scale (2–16 MHz). The presence of electron hopping and IS values intermediate to that of Fe²⁺ and Fe³⁺ suggests charge confinement in the triangular network of the *ab*-plane of the crystal lattice (e.g. formation of dimers, Fe²⁺ \leftrightarrow Fe³⁺). Even though CO has been destabilised at pressure, there still appears to be confinement of the charge carriers. This will have implications for the electrical-transport behaviour. The destabilisation of CO is compatible with the collapse of magnetic order seen in neutron diffraction studies [7].

Acknowledgements

This work was supported by funding from the UJ-URC and NRF (SA), and supply of the sample used in this study from M. Poienar (Caen) is also acknowledged with gratitude.

References

- [1] Subramanian M A, He T, Chen J, Rogado N S, Calvarese T G and Sleight A W 2006 *Adv. Mater.* **18** 1737
- [2] Lafuerza S, Garcia J, Subias G, Blasco J, Conder K and Pomjakushina E 2013 *Phys. Rev. B* **88** 085130
- [3] Ikeda N 2008 *J. Phys.: Condens Matter.* **20** 434218
- [4] Ikeda N, Ohsumi H, Ohwada K, Ishii K, Inami T, Kakurai K, Murakami Y, Yoshii K, Mori S, Horibe Y and Kito H 2005 *Nature* **436** 1136
- [5] Lafuerza S, Subías G, Blasco J, García J, Nisbet G, Conder K and Pomjakushina E 2014 *EPL* **107** 47002
- [6] Lafuerza S, Subías G, García J, Blasco J, G. Nisbet, Conder K and Pomjakushina E 2014 *Phys. Rev. B* **90** 085130
- [7] Makarova O L, Bourgeois J, Poienar M, Mirebeau I, Kichanov S E, André G, Elkaim E, Hanfland M, Hervieu M, Maignan A, Haines J, Rouquette J, Martin C and Damay F 2013 *Appl. Phys. Lett.* **103** 082907
- [8] Momma K and Izumi F 2011 *J. Appl. Crystallogr.* **44** 5
- [9] Rouquette J, Haines J, Al-Zein A, Papet P, Damay F, Bourgeois J, Hammouda T, Doré F, Maignan A, Hervieu M and Martin C 2010 *Phys. Rev. Lett.* **105** 237203
- [10] Takele S and Hearne G R 2001 *Nucl. Instrum. Meth. B* **183** 413
- [11] Tjon J A and Blume M 1968 *Phys. Rev.* **165** 456
- [12] Herber R H and Eckert H 1985 *Phys. Rev. B* **31** 34
- [13] Klencsar Z, MossWinn 4.0Pre Manual, 2013.
- [14] Patankar S, Pandey S K, Reddy V R, Gupta A, Banerjee A and Chaddah P 2010 *EPL* **90** 57007
- [15] Bang B K, Kim S B, Cheong S-W and Kim C S 2007 *Phys. Status Solidi (B)* **244** 4566
- [16] Hearne G R, Sibanda W N, Carleschi E, Pischedda V and Attfield J P 2012 *Phys. Rev. B* **86** 195134
- [17] Angst M, Khalifah P, Hermann R P, Xiang H J, Whangbo M-H, Varadarajan V, Brill J W, Sales B C and Mandrus D 2007 *Phys. Rev. Lett.* **99** 086403
- [18] Litterst F J and Amthauer G 1984 *Phys Chem Minerals* **10** 250
- [19] Diguët G, Hearne G R, Sibanda W N, Carleschi E, Musyimi P, Pischedda V and Attfield J P 2014 *Phys. Rev. B* **89** 035132

Pressure induced charge order collapse in Fe₂OBO₃

W N Sibanda¹, E Carleschi¹, G Diguet¹, V Pischedda², J P Attfield³ and G R Hearne¹

¹Department of Physics, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, Johannesburg, South Africa

²Institut Lumière Matière, UMR5306 Université de Lyon 1-CNRS, 69622 Villeurbanne, Cedex, France

³Centre for Science at Extreme Conditions and School of Chemistry, University of Edinburgh, Erskine Williamson Building, King's Buildings, Mayfield Road, Edinburgh EH9 3JZ, United Kingdom

Author e-mail address: wnsibanda@gmail.com

Abstract. The effect of pressure on the charge order in the mixed-valence Fe₂OBO₃ has been investigated using ⁵⁷Fe Mössbauer spectroscopy at pressures up to 30 GPa in diamond anvil cell experiments. Above 16 GPa only signatures of electron hopping, Fe²⁺ ⇌ Fe³⁺, where “⇌” represents the hopping between the different ionic charge states are discerned. Analysis shows that the hopping frequency is of the order of MHz (~100 nanosecond time scale), that is, in a time window to which the Mössbauer spectroscopy technique is particularly sensitive. The fluctuation signatures persist upon further pressurization to ≈ 30 GPa at 300 K. The frequency lies in the range of 2–4 MHz upon increasing pressure. The charge order collapse initiated by external pressure was found to be reversible upon decompression to ambient conditions.

1. Introduction

Site-centered charge ordering (CO) refers to the spatial localisation of charge carriers on atomic sites with a certain periodicity, resulting in valence electrons localising and giving rise to mixed-valence superstructures within crystal lattices. This may occur in strongly correlated systems as a mechanism to minimize the repulsive energy between valence electrons at the expense of their kinetic energy [1]. The condensation of charge into distinct states or valences (also called a Wigner crystal [2]) invokes the concept of ionic charge order, in which valence states take on integer values. Until recently the concept of ionic charge order in metal oxides has been under doubt because valence separation into distinct charge states does not take on integer values, even in the case of well-studied Fe₃O₄ (magnetite), which is the first CO compound to be discovered. Recent work has then found evidence of a superstructure in iron oxoborate, Fe₂OBO₃, with a valence separation, Fe^{(2+δ)+} — Fe^{(3-δ)+}, near to integral values, where δ ≤ 0.2 [3]. This makes Fe₂OBO₃ the archetypal ionic CO compound.

Fe₂OBO₃ belongs to a family of homometallic oxoborates with a warwickite crystal structure. The structure is characterised by one-dimensional ribbons of edge sharing FeO₆ octahedra, stacked along the *a*-axis direction. The ribbons are connected by corner sharing octahedra and trigonal-planar BO₃

groups, defining non-equivalent crystallographic sites (referred to as Fe (1) and Fe (2)) in the FeO_6 octahedra (see figure 1 (a)).

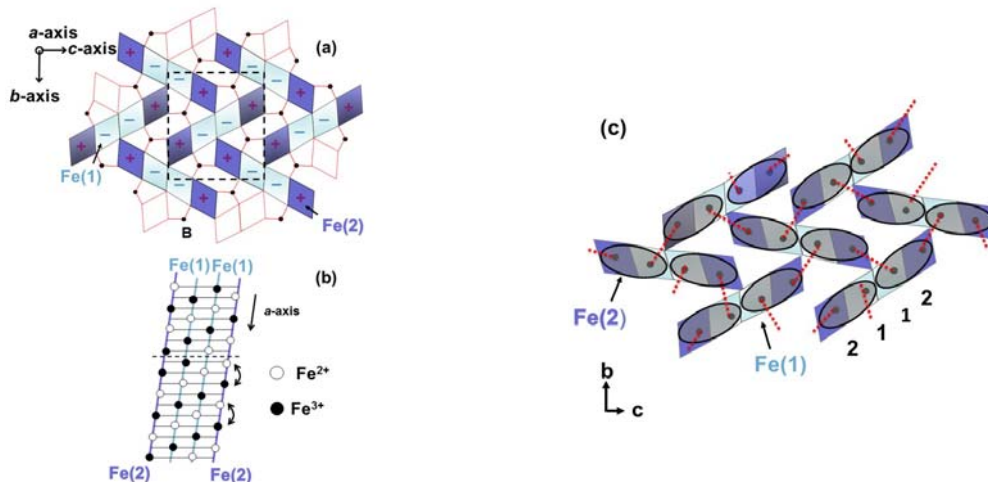


Figure 1 (Colour online): (a) Crystal structure of Fe_2OBO_3 adapted from Attfield et. al. [1]. In the b - c plane this can be represented as Fe (2)-Fe (1)-Fe (1)-Fe (2) edge-sharing octahedra within a ribbon, where Fe (1) and Fe (2) indicate structurally distinct iron octahedral sites that make up such four-chain ribbons. The “+” and “-” signs refer to antiparallel magnetic moments in the magnetically ordered phase. (b) Ribbon in the CO phase comprised of chains of four edge-sharing FeO_6 octahedra that extend infinitely along the crystallographic a -axis. (c) Dimerisation in the orthorhombic structure of Fe_2OBO_3 , depicted by the ellipses. The red dotted lines show the possible paths in which electron hopping can occur.

Fe_2OBO_3 undergoes an orthorhombic-monoclinic structural phase transition at $T_{\text{CO}} = 317$ K, with a concomitant charge delocalization. These structural and electronic changes are attributed to site-centered CO, manifested as a Fe^{2+} and Fe^{3+} mixed-valence state [1, 4]. In this scenario the CO involves the localisation of the extra down-spin electron in a t_{2g} orbital of the Fe^{2+} cation ($t_{2g}^4 e_g^2$). The CO in this case is stabilised in the monoclinic structure.

^{57}Fe Mössbauer spectroscopy has shown that below 270 K (monoclinic unit cell, $P2_1/c$), the Fe^{2+} and Fe^{3+} are equivalently distributed in crystallographically distinct sites. Above 400 K (orthorhombic unit cell, $Pm\bar{c}n$), the semi-valent state $\text{Fe}^{2.5+}$ representing dynamic charge carriers or electronic hopping is present exclusively. In between 270 K – 400 K, there is co-existence of Fe^{2+} , Fe^{3+} and $\text{Fe}^{2.5+}$ [5]. In the temperature range 270 K – 400 K, there exists the so-called intermediate phase [6], which is characterized by an incommensurate CO super structure due to the geometrical frustration occurring in the system. The intermediate phase was shown to undergo resistive switching where application of an electric field induces a change in the resistance. This ability to tune the resistance is thought to be related to mobility of charge carriers in this temperature region [7].

An analogy to this temperature-dependent study is our recent x-ray powder diffraction pressure-dependent study of the structural properties of Fe_2OBO_3 [8]. Analysis of the in-situ XRD data indicates that, starting at ≈ 6 GPa, the ambient pressure $P2_1/c$ monoclinic structure gradually transforms into the high pressure $Pm\bar{c}n$ orthorhombic phase, and at 22 GPa the phase transition is

complete. Moreover, the unit cell was found to be more compressible along the a -axis (10% over the pressure range up to 30 GPa) compared to the other directions in the unit cell which have a 3% compressibility over the same pressure range. The change in volume for the $P2_1/c \rightarrow Pmcn$ phase transition is limited to 1% and this therefore suggests the bulk moduli of the two structures are commensurate.

Our previous ^{57}Fe Mössbauer spectroscopy studies involving the application of external pressure (P) on an Fe_2OBO_3 have led to an extension of the magneto-electronic temperature-pressure phase diagram (not shown in this publication; the detailed phase diagram can be found in ref. [9]). On the basis of the analysis of the Mössbauer data it was revealed that (i) the site-centered CO collapse is initiated at ≈ 11 GPa at 300 K; at $P \geq 16$ GPa, *only* electron hopping processes on the time-scale of the Mössbauer effect (10^{-7} s) were discerned (this was typified by asymmetric relative intensities and hyperfine parameters values intermediate to that of Fe^{2+} and Fe^{3+} , that is, the intermediate valence state $\text{Fe}^{2.5+}$); (ii) suppression of CO and magnetic ordering below 115 K at $P \geq 16$ GPa. Thus in the CO phase the magnetic transition temperature T_M does not increase or rise beyond the ambient pressure value of $T_M = 155$ K.

This study is a continuation of our previous work [9] with a particular focus on the pressure range 16–30 GPa so as to make a quantitative assessment of the electron dynamics. We present a detailed analysis of the high pressure electronic state of Fe in Fe_2OBO_3 using ^{57}Fe Mössbauer spectroscopy as a probing technique of the Fe ion valence. We show that electron hopping frequency increases doubles, from 2 to 4 MHz in the pressure range 16–30 GPa. This, less than one order of magnitude, increase in the hopping frequency is also reflected in the Mössbauer line shape that shows no major discernable change over an extensive pressure range spanning 14 GPa. The high pressure orthorhombic (space group, $Pmcn$) phase is found to be reversible to the ambient monoclinic (space group, $P2_1/c$) phase, where the CO is stabilised.

2. Experimental details

A sample of natural ^{57}Fe abundance (2%) was used in this work. Details of the synthesis can be found in a previous work [4]. The composition and magnetic properties of the samples at ambient pressure were confirmed by conventional powder x-ray diffraction (XRD) and Mössbauer spectroscopy (MS). For MS studies the sample was encapsulated in a 250 μm cavity of a pre-indented Re gasket of thickness 30 μm in a Merrill-Basset diamond anvil cell (DAC). Re also served as a collimator for the 14.4 keV Mössbauer resonance γ -rays. Daphne 7373 oil was used as the pressure transmitting medium (PTM) and ruby fluorescence served as a manometer.

Mössbauer studies were carried out with a 10 mCi ^{57}Co (Rh) point source in transmission geometry. The typical collection time of a single spectrum was ≈ 72 hours. All the spectra were analysed using an innovative fitting model, as described in the following section [10]. Hyperfine parameters were obtained by fitting theoretical sub-spectral curves to experimental data with Lorentzian line shapes to de-convolute Fe phases in the spectrum. Prior to analysis each spectrum was folded with its mirror image and, where necessary, adjacent channels subsequently added. The purpose of this was to remove geometrical base-line distortions and to reduce the signal-to-noise ratio in the final data set to be used for analysis.

3. Results and Discussion

Room temperature Mössbauer spectra of Fe_2OBO_3 in the range 16–30 GPa are shown in figure 2 (a). Their line shape is indicative of electron hopping processes occurring within the time sensing window

of the Mössbauer effect, $\sim 10^{-7}$ s [11-13]. There is a marked disappearance of the right most resonant absorption lines (high velocity side) in the ambient pressure spectra with increasing pressure, see figure 2 (a) and (b) for a visual comparison. Marked changes in spectral features were rationalized as an electronic hopping process between sites of different valence state; in the case of Fe_2OBO_3 , $\text{Fe}^{2+} \leftrightarrow \text{Fe}^{3+}$ in the high pressure regime [9], indicative of a complete CO collapse.

The fitting model adopted for the fluctuating valence state is markedly different from previous works [5, 8]. The model implemented here describes electron hopping relaxation in a paramagnetic powder using the theory proposed by Blume [11, 12]. From the fitting parameters characterising the lineshape, we were able to extract the isomer shift and primarily the hopping frequency of electrons in this dynamical process. The isomer shifts quoted in this work for the $\text{Fe}^{2.5+}$ (electron hopping) are an average of the Fe^{2+} and Fe^{3+} valences.

Each spectrum was fitted with a single hopping component, which has been superimposed to the corresponding experimental spectrum in figure 2 (a) (red solid line). There are two possible routes for the electron hopping to occur; inter-ribbon, between Fe^{2+} and Fe^{3+} sites on adjacent ribbons, or intra-ribbon, within the same ribbon (see figure 1(c)). Hence the MS spectrum can be expected to be composed of two components. Our approach of using a single fitting component offers the safest route because MS cannot easily discriminate between different hopping paths, involving similar hopping frequencies, but can only reveal the valence state as being static or fluctuating (intermediate between two states) when the electron processes are within the time window of the Mössbauer effect (10^{-7} s).

It is apparent that there are no marked or drastic changes in the spectral line shape over the pressure range spanning a range of 14 GPa, indicating that no new electronic phase change occurs and that the hopping persists up to 30 GPa.

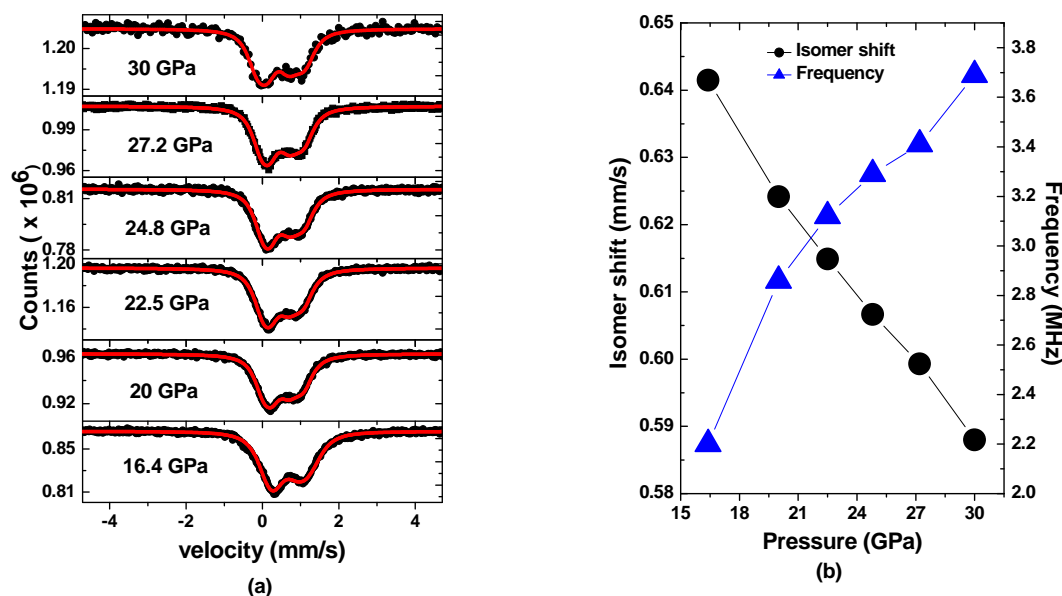


Figure 2 (Colour online): (a) Mössbauer spectra recorded at pressure from 16.4 GPa to 30 GPa. The solid red lines represent the fit to the experimental data (in black). (b) IS and frequency plotted as function of pressure in the range 16-30 GPa. The values of the IS are quoted with respect to α -Fe. The black coloured circles are data points and the lines joining the points are intended as a guide to the eye.

Figure 2 (b) shows a combined plot of the isomer shift (IS) and hopping frequency as a function of pressure, as retrieved from the fitting to the experimental data. The values of the IS are intermediate between those normally associated with Fe^{2+} and Fe^{3+} . This supports the notion that the charge carriers are confined in a lattice dimerised bond, between Fe^{2+} - Fe^{3+} pairs, see figure 1 (c). Our previous work on the structural properties of Fe_2OBO_3 shows that the unit cell is more compressible along the a -axis compared to other axes [8]. It is therefore expected that such a lattice distortion will favour a spontaneous Peierls-type dimerisation along the a -axis compared to other crystallographic directions and such paths are conducive to charge transfer, that is, electron hopping. The values of the IS also decrease with pressure, which is attributed to the s -electron density increasing near the Fe nucleus. The IS is a function of s -electron density, ρ_s , at the nucleus.

The pressure dependency of the hopping frequency shows a 70% increase between 16 GPa and 30 GPa. Application of pressure causes the inter-atomic distances to decrease and hence an increase in pressure is expected to promote the $\text{Fe}^{2+} \leftrightarrow \text{Fe}^{3+}$ electron exchange, from the increase in orbital overlap. Resistivity studies done by our group infer that there is a site-centered to bond-centered insulator transition in Fe_2OBO_3 above 16 GPa, i.e. a Wigner crystal to a dimer-Mott insulating state transition [8].

The high pressure CO collapse was found to be reversible upon releasing pressure to ambient conditions, see figure 3 below. Distinct signatures of Fe^{3+} and Fe^{2+} indicating CO are again retained at ≈ 0.3 mm/s and ≈ 2 mm/s respectively. This therefore means the high pressure orthorhombic phase is reversible to the monoclinic low pressure phase.

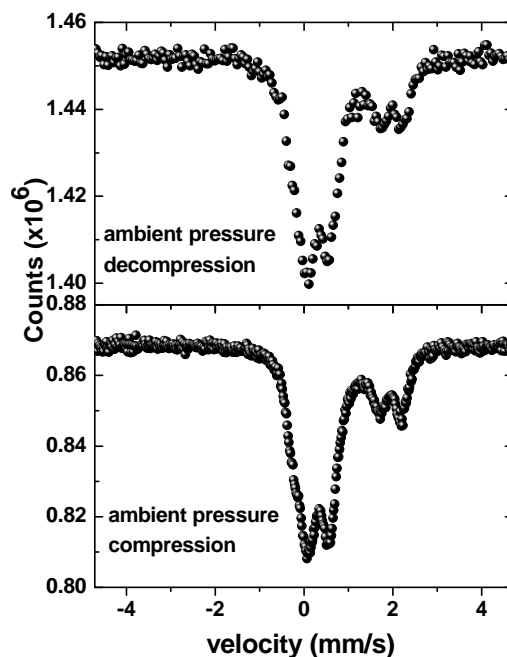


Figure 3. The ambient pressure spectra for the compression and decompression runs. The very good similarity between these two spectra shows that the low pressure monoclinic ($P2_1/c$) site-centered CO phase is recovered after decompression.

4. Conclusions

In summary, we used ^{57}Fe Mössbauer spectroscopy at room temperature to probe the nature of site-centered CO in Fe_2OBO_3 by applying external pressure. Our analysis of the MS spectra above 16 GPa, where such site-centered CO is disrupted, show an average valence of $\text{Fe}^{2.5+}$ as the IS values are intermediate to that normally associated with Fe^{2+} and Fe^{3+} , indicating electron hopping relaxation. The electron hopping persists up to 30 GPa, on the time-scale of the Mössbauer effect, with frequencies on the order of MHz (~ 100 nanoseconds). The site-centered CO disruption was found to be reversible, with Wigner crystallisation (CO) being re-established upon decompression to ambient conditions.

Acknowledgements

The authors thank the UJ-URC and NRF (SA) for financial support. Funding also from EPSRC and the Royal Society (UK) is also acknowledged with gratitude.

References

- [1] Attfield J P, Bell A M T, Rodriguez-Martinez L M, Greneche J M, Cernik R J, Clarke J F and Perkins D A 1998 *Nature* **396** 655
- [2] Wigner E 1934 *Phys. Rev.* **46** 1002
- [3] Angst M, Khalifah P, Hermann R P, Xiang H J, Whangbo M-H, Varadarajan V, Brill J W, Sales B C and Mandrus D 2007 *Phys. Rev. Lett.* **99** 086403
- [4] Attfield J P, Bell A M T, Rodriguez-Martinez L M, Greneche J M, Retoux R, Cernik R J, Clarke J F and Perkins D A 1999 *J. Mater. Chem.* **9** 205
- [5] Douvalis A P, Papaefthymiou V, Moukarika A, Bakas T and Kallias G 2000 *J. Phys.: Condens. Matter* **12** 177
- [6] Angst M, Hermann R P, Schweika W, Kim J-W, Khalifah P, Xiang H J, Whangbo M-H, Kim D-H, Sales B C and Mandrus D 2007 *Phys. Rev. Lett.* **99** 256402
- [7] Akrap A, Angst M, Khalifah P, Mandrus D, Sales B C and Forró L 2010 *Phys. Rev. B* **82** 165160
- [8] Diguët G, Hearne G R, Sibanda W N, Carleschi E, Musyimi P, Pischedda V and Attfield J P 2014 *Phys. Rev. B* **89** 035132
- [9] Hearne G R, Sibanda W N, Carleschi E, Pischedda V and Attfield J P 2012 *Phys. Rev. B* **86** 195134
- [10] Klencsar Z, MossWinn 4.0Pre Manual, 2013.
- [11] Tjon J A and Blume M 1968 *Phys. Rev.* **165** 456
- [12] Herber R H and Eckert H 1985 *Phys. Rev. B* **31** 34
- [13] Litterst F J and Amthauer G 1984 *Phys Chem Minerals* **10** 250

Effects of catalyst:Ba molar ratio on the structure and luminescence properties of BaCO_3 :1% Eu^{3+} , 2% Dy^{3+} phosphors synthesized using sol-gel process.

AS Tebele*, SV Motloun and FB Dejene

Department of Physics, University of the Free State (Qwaqwa Campus), Private Bag X13, Phuthaditjhaba, 9866, South Africa

*Corresponding author.e-mail address: tebeleas@qwa.ufs.ac.za

Abstract. BaCO_3 :1% Eu^{3+} , 2% Dy^{3+} powders were synthesized by sol-gel process at a relatively low temperature $\sim 80^\circ\text{C}$. Metal nitrates were used as the source of metal ions and citric acid as a chelating agent. The catalyst:Ba molar ratio in the solution were varied from 1.0 - 2.5 during synthesis. The annealed powder samples were characterized by X-ray diffraction (XRD), scanning electron microscopy (SEM) and photoluminescence (PL) spectroscopy. The XRD data revealed that the annealed samples consist of orthorhombic BaCO_3 with extra peaks of BaO impurities at the lower catalyst content. Morphology of the phosphor was influence by the catalyst:Ba molar ratio. PL spectra indicated that the emission peaks appears at different wavelength positions 376, 531, 589 and 616 nm with the most intense peak at 616 nm. The emission at 376 nm is attributed to the changes in crystal field surrounding the activators. Emission peak at 531 nm is attributed to the host emission due to defects, while the emission peak at 589 nm is assigned to magnetic dipole $^5\text{D}_0$ - $^7\text{F}_1$ transitions in Eu^{3+} ion. The emission peak at 616 nm is attributed to hypersensitively force electric dipole $^5\text{D}_0$ - $^7\text{F}_2$ transitions in Eu^{3+} ion.

Keywords: Luminescence; BaCO_3 :1% Eu^{3+} , 2% Dy^{3+} ; catalyst:Ba; Sol-gel

1. Introduction

Metal carbonates are abundant minerals in nature and are of important interest in material research due to their promising applications such as optical materials, effective catalyst, separation technology, drug-delivery and pigment [1-5]. Barium carbonate (BaCO_3) also known as witherite is one of the well-known metal carbonates used in magnetic materials, television kinescope Glass-Shell and water cleaning [6]. Metal carbonates have shown to be good host lattice for various dopants or foreign atoms. Introduction of foreign atoms into the host lattice is known to affect the luminescence properties of the material [7-10]. It is well known rare earth ions have been playing an important role in modern lighting and display field due to abundant emission colors based on their transitions [11]. Furthermore, BaCO_3 is expected to show excellent luminescence properties by doping with rare-earth ions [12]. Eu^{3+} doped CaCO_3 phosphor with red emission prepared by microwave synthesis has been reported by Kang et al. [13]. Their results [13] shows that the red luminescence can be attributed to the transitions from the $^5\text{D}_0$ excited level to the $^7\text{F}_{J=0,4}$ levels of Eu^{3+} ions with the mainly electric dipole transition $^5\text{D}_0 \rightarrow ^7\text{F}_2$ (614 and 620 nm). SrCO_3 : Eu^{3+} microneedles through a large-scale and facile hydrothermal method without any template and further annealing treatment have been reported by Yang et al. [14]. In their results [14] they found a strong red

emission corresponding to the $^5D_0 \rightarrow ^7F_{J=1-4}$ transitions of Eu^{3+} ions. The luminescent properties of $\text{CaCO}_3:\text{Eu}^{3+}$, $\text{BaCO}_3:\text{Eu}^{3+}$ and $\text{SrCO}_3:\text{Eu}^{3+}$ prepared by simple and convenient synthetic route containing H_2O and EG (ethylene glycol) with various compositions has been reported by Shi et al. [15]. Furthermore, their results [15] show three red emission spectra which seems to be equal in shape except for the intensity differences between them. Therefore, it is expected that $\text{BaCO}_3:\text{Eu}^{3+}, \text{Dy}^{3+}$ can be used as an FED (Field Emission Display) phosphor for practical application. So it is very important to synthesize $\text{BaCO}_3:\text{Eu}^{3+}, \text{Dy}^{3+}$ and understand the luminescent mechanism. BaCO_3 crystallizes in the orthorhombic system [16]; hence, there is continued interest in synthesizing of BaCO_3 with controlled morphologies [17]. Sreedhar et al. [24] report the synthesis and characterization of BaCO_3 nanocrystallite using natural polymer, gum acacia and they obtain pure orthorhombic structure in XRD analysis. Wang et al. [6] report a new route to the synthesis of BaCO_3 crystals by the induction of bacillus pasteurri and they found sphere-like and rod-like morphology with a diameter of 7 and 1 μm , respectively. At present, numerous methods have been reported on the preparation of BaCO_3 nanostructures including precipitation-electrodeposition reaction [19], reversed micelle method [20], polyvinylpyrrolidone-assisted method [21], high-gravity method [22] and sol-gel method [23]. The sol-gel technique offers the advantages such as excellent control over the stoichiometry, homogeneous particle distribution, good reactivity between components, nano-size particles and allows lower processing temperature [16]. In this work, the effect of varying the catalyst:Ba molar ratio in $\text{BaCO}_3:1\% \text{Eu}^{3+}, 2\% \text{Dy}^{3+}$ nanophosphor synthesized by the sol-gel technique is investigated. The main aim is to develop red light emitting phosphor for applications in pigment.

2. Experimental

The powders of $\text{BaCO}_3:1\% \text{Eu}^{3+}, 2\% \text{Dy}^{3+}$ phosphors with different catalyst:Ba molar ratio ranging from 1.0 – 2.0 were prepared by sol-gel process. Note that the Ba concentration was kept constant at 1 mol. The nitrates; $\text{Ba}(\text{NO}_3)_2$ (99%), $\text{Al}(\text{NO}_3)_3 \cdot 9\text{H}_2\text{O}$ (98.5%), $\text{Eu}(\text{NO}_3)_3 \cdot 5\text{H}_2\text{O}$ (99.9%), $\text{Dy}(\text{NO}_3)_3 \cdot 5\text{H}_2\text{O}$ (99%) were used as the starting materials and $\text{C}_6\text{H}_8\text{O}_7 \cdot \text{H}_2\text{O}$ (citric acid) (99%) was used as a catalyst. These nitrates and the catalyst were dissolved in deionized water and stirred constantly on magnetic stirrer at a temperature $\sim 80^\circ\text{C}$ until gelation. The prepared white gels were dried at 130°C in an oven for 4 hour. The dried gel was ground and subsequently annealed at 800°C for 2 hours. The crystal structure of the samples was characterized by powder X-ray diffraction (Bruker AXS Discover diffractometer) with $\text{CuK}\alpha$ (1.5418\AA) radiation. The surface morphology of the phosphor powder was established using a Shimadzu Superscan ZU SSX-550 electron microscope (SEM). PL excitation and emission spectra were obtained by using Xenon lamp (Hitach F-7000 fluorescence spectrophotometer).

3. Results and discussion

3.1. Structure

For all samples the (111) diffraction peak were deconvoluted as shown in figure 1 (a) to distinguish between (111) and (102). The average d-spacing for the (111) peak was estimated to be 12.165\AA and the lattice parameters were calculated to be $a = 6.415\text{\AA}$, $b = 5.302\text{\AA}$ and $c = 8.870\text{\AA}$, which are very close to those of JCPDS card no. 45-1471 where $d = 12.202\text{\AA}$, and the lattice parameters are $a = 6.433$, $b = 5.314$ and $c = 8.903\text{\AA}$ [24]. Figure 1 (b) and (c) shows the diffractions patterns of the annealed $\text{BaCO}_3:1\% \text{Eu}^{3+}, 2\% \text{Dy}^{3+}$ powders for the co-doped and un-doped 1.0 catalyst:Ba molar ratio and for co-doped samples at varying catalyst:Ba molar ratios, respectively. The XRD patterns of all prepared samples matches with the peaks of BaCO_3 JCPDS card no. 45-1471. However, the XRD detected an impurity (shown by asterisk (*)) in figure 1 (b) and (c)), which can be attributed to the BaO [25]. If the catalyst content is too much as shown in figure 1 (c) (e.g. 2.5 molar ratio), the results reveal that there are no BaO impurity which suggest the single phase formation at the higher molar ratio. Furthermore, the single phase formation at higher molar ratio also suggests that the foreign atoms (e.g. Eu^{3+} and Dy^{3+}) were fully incorporated into the host matrix [26]. Therefore, catalyst:Ba molar ratio at 2.5 does not affect the crystal structure of the phosphor, which implies that catalyst at this stage helps BaCO_3 to fully incorporate with these foreign atoms (Eu^{3+} and Dy^{3+}). From Feilong et al. report [25], it is known that $\text{Ba}(\text{NO}_3)_2$ does not easily dissolve

in deionized water, and therefore, from this results we infer that more catalyst content might be necessary to fully catalyse the dissolving of $\text{Ba}(\text{NO}_3)_2$ in deionized water. Thus, this might serve as a good reason why the single phase structure is obtained at higher mole fraction as shown in figure 1 (c). The grain sizes of the powders with various Catalyst:Ba molar ratio for 1.0 (un-doped), 1.0 (co-doped), 1.3, 1.5, 1.8 and 2.5 were estimated from the main diffraction peak (111) by using the Scherrer formula [27] and they were found to be 85, 82, 84, 85, 69 and 70 nm, respectively. It is therefore, concluded that varying catalyst:Ba molar ratio influences the crystallites. Figure 1 (d) shows the analysis of the most intense peak (111). As it can be seen that there is slight shift to the higher diffraction angle in the (111) peak as the catalyst:Ba molar ratio is varied, and that is attributed to the change in the lattice spacing [26]. Shifting to the higher angle is also attributed to the decrease in d-spacing as the foreign atoms are incorporated into the BaCO_3 matrix [26] as shown in figure 1 (e). This is due to the substitution of bigger atom Ba^{2+} (atomic radius (1.61 Å)) [27] with smaller atoms Eu^{3+} (0.95 Å) [28] and Dy^{3+} (1.08 Å) [29]. It is therefore, proposed that the catalyst:Ba molar ratio influences the d-spacing. The d-spacing as a function of catalyst:Ba molar ratio shows the parabolic behaviour as depicted in figure 1 (e) for the (111) diffraction peak. The decrease in d-spacing is due to the increase in catalyst:Ba molar ratio, that is, due to the substitution of Ba^{2+} with Eu^{3+} and Dy^{3+} as mentioned above. Thus, the mole ratio is facilitating the substitutions or knocking-off the atoms in the host matrix. However, as the catalyst:Ba molar ratio is increased further, the decrease in d-spacing reaches a critical point at 1.9 (catalyst:Ba) molar ratio and then increases. Similar behaviour to our results has been observed by Motloung et al. [26] and the ref therein.

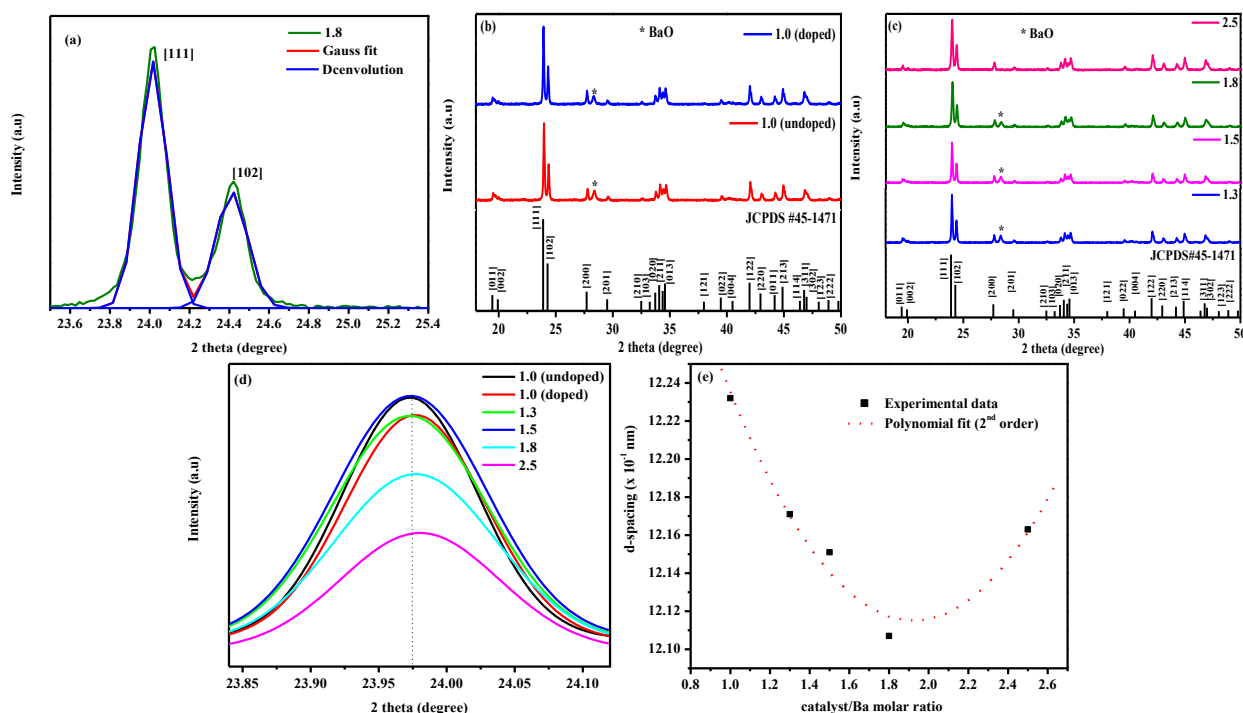


Figure 1 (a) Deconvolution of 1.8 catalyst:Ba molar ratio X-ray patterns of the (b) 1.0 (co-doped) and 1.0 (un-doped) (c) BaCO_3 :1% Eu^{3+} , 2% Dy^{3+} at varying catalyst:Ba molar ratio (d) Analysis of (111) Peak (e) d-spacing as a function of catalyst:Ba molar ratio.

3.2. Phosphor morphology

The surface morphologies of the annealed samples were observed by SEM micrographs shown in figure 2 (a) – (f). Fig. 2 (a) shows the dispersal of small irregular particles at 1.0 (catalyst:Ba) molar ratio for the

un-doped BaCO_3 which consists of voids morphology. At 1.0 molar ratio for the co-doped BaCO_3 the micrograph shows the dispersal of small irregular particles on the surface of the large agglomerated particle (see figure 2 (b)), which is similar to figure 2 (c) – (f). It is interesting to observe that as the catalyst:Ba molar ratio is increased to 2.5, as shown in figure 2 (f), the particles become more agglomerated with the porous morphology on the surface. These results showed that varying catalyst:Ba molar ratio does not affect morphology of the phosphor which agrees with XRD analysis. It is suggested that porous/voids are formed by degassing during annealing process [30], that is, pores originate from gas entrapment due to improper adhesion at certain places [31].

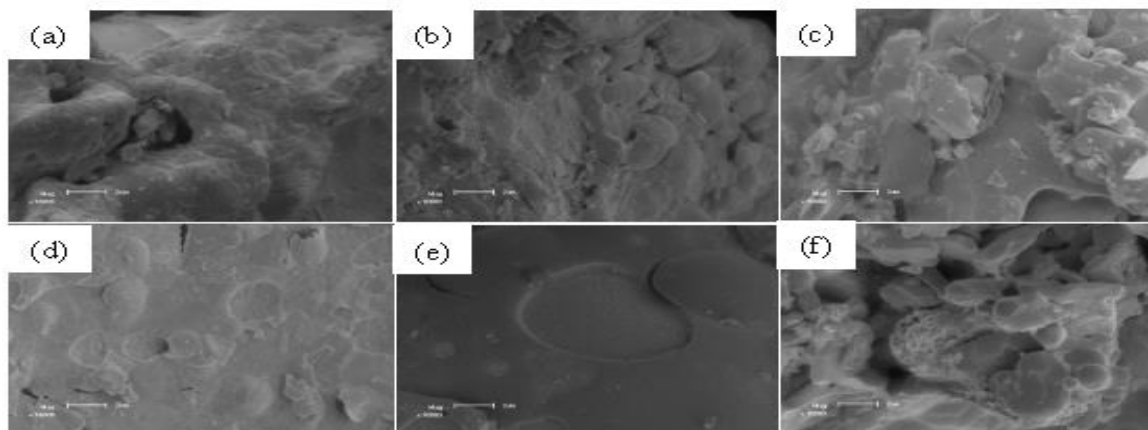


Figure 2 SEM images of the $\text{BaCO}_3:1\% \text{Eu}^{3+}, 2\% \text{Dy}^{3+}$ for various catalyst:Ba molar ratio for (a) 1.0 (un-doped), (b) 1.0 (co-doped), (c) 1.3, (d) 1.5, (e) 1.8 and (f) 2.5.

3.3 Photoluminescence characteristics

Excitation photoluminescence spectra of the $\text{BaCO}_3:1\% \text{Eu}^{3+}, 2\% \text{Dy}^{3+}$ are shown in figure 3 (a). It was observed that the main peaks are at 260 and 303 nm. The absorption band with maximum at 260 nm is attributed to the charge transfer (CT) state of $\text{Eu}^{3+} - \text{O}^{2-}$ [32]. The absorption at 303 nm corresponds to the host absorption [33]. The emission spectra of the $\text{BaCO}_3:1\% \text{Eu}^{3+}, 2\% \text{Dy}^{3+}$ powders are shown in figure 3 (b). The PL emission reveals that un-doped powder has peaks at 400 and 531 nm. These emission peaks are therefore attributed to originate from the host material. Both emission bands are assumed to be due to the band-to-band and defects level in the host. The co-doped powders have the emission peaks at different positions 376, 531, 589 and 616 nm with the most intense peak at 616 nm. The emission at 376 nm is attributed to the changes in crystal field surrounding the activators [34]. It can be seen that the emission peak at 531 nm arises from the un-doped 1.0 (catalyst:Ba) molar ratio and therefore is attributed to the defects emission in the host material (as mentioned above). The emission peak at 589 nm is assigned to magnetic dipole transitions $^5\text{D}_0-^7\text{F}_1$ of Eu^{3+} [35]. The emission peak at 616 nm is attributed to hypersensitively force electric dipole transitions $^5\text{D}_0-^7\text{F}_2$ of Eu^{3+} [35]. Figure 3 (c) shows the deconvolution of 2.5 (catalyst:Ba) molar ratio as the same as in figure 3 (b) when excited at the same excitation energy and it shows that there are two emission peaks at 350 and 376 nm. The emission peak at 350 nm is attributed to $^6\text{H}_{15/2}-^4\text{M}_{15/2}$, $^6\text{P}_{7/2}$ from Dy^{3+} ions [33] and 376 nm is attributed as mentioned above. This emission spectra show some changes in peak shape, especially at high energy (<400 nm) [33]. This behavior is observed at the higher molar ratio (e.g. 1.8 and 2.5) and it is assigned to changes in crystal field surrounding the activators [34], which agrees very well with the XRD findings. Moreover, this PL results clearly shows that varying catalyst:Ba molar ratio changes the dopants environment in the

host material. Thus, the variation of catalyst:Ba molar ratio affected the luminescent properties of $\text{BaCO}_3:1\% \text{Eu}^{3+}, 2\% \text{Dy}^{3+}$ phosphors.

Figure 3 (d) shows the afterglow curve of $\text{BaCO}_3:1\% \text{Eu}^{3+}, 2\% \text{Dy}^{3+}$ phosphor at room temperature after exciting by 260 nm when monitoring the red emission at 616 nm. The decay characteristics of this phosphor indicates that the shape of afterglow decay curves of the samples at the varying catalyst:Ba molar ratio are similar, which indicates that all the samples have the same afterglow decay mechanism. The afterglow intensity gradually becomes zero after 2750 ms. Therefore, when the excitation source is switched off, the relaxation of these secondary ions from deep traps is very slow; this leads to the long persistence of phosphor [18]. When the mole fraction is 1.5 it demonstrate a significant improvement in initial luminescent brightness and long afterglow. The results are consistence with the measurements of PL emission in figure 3 (b), where 1.5 catalyst:Ba molar ratio is observed with the highest emission intensity.

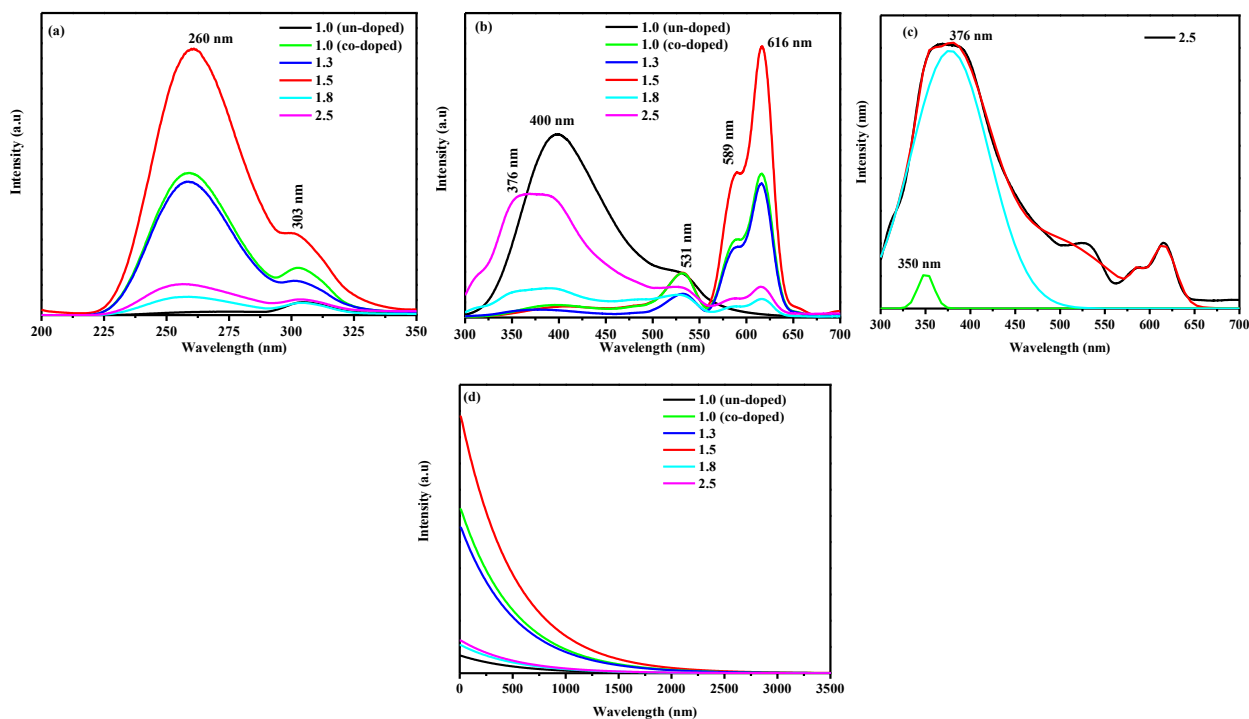


Figure 3 PL spectra of the annealed $\text{BaCO}_3:1\% \text{Eu}^{3+}, 2\% \text{Dy}^{3+}$ (a) excitation spectra ($\lambda_{\text{em}} = 616 \text{ nm}$) (b) emission spectra ($\lambda_{\text{ex}} = 260 \text{ nm}$) (c) 2.5 molar ratio deconvolution (d) Afterglow decay curve ($\lambda_{\text{ex}} = 260 \text{ nm}$).

4. Conclusion

$\text{BaCO}_3:1\% \text{Eu}^{3+}, 2\% \text{Dy}^{3+}$ phosphors powders were successfully synthesized using sol-gel process. XRD results show that when the powders are annealed at 800°C , the crystalline structure shows orthorhombic BaCO_3 with some impurities such as BaO. If the catalyst content is too much, the foreign atoms (e.g. Eu^{3+} and Dy^{3+}) distribute homogeneously in the host material to form pure orthorhombic structure. SEM images showed irregular particles morphology. The excitation peaks were found to be 260 and 303 nm, which are attributed to the charge transfer (CT) state of $\text{Eu}^{3+} - \text{O}^{2-}$ and the host absorption. The most intense emission peak at 616 nm is assigned to hypersensitively force electric dipole transitions $^5\text{D}_0 - ^7\text{F}_2$ transitions of Eu^{3+} . The amount of citric acid affected the luminescent properties of $\text{BaCO}_3:1\% \text{Eu}^{3+}, 2\% \text{Dy}^{3+}$ phosphors. The decay curve show that all samples have the same afterglow behaviour and it also shows the improvement in initial luminescent brightness and long afterglow.

Reference

- [1] Jacob D S, Joseph A, Mallenahalli S P, Shanmugam S, Makhluif S, Calderon-Moreno J, Kolytyn Y, Gedanken A 2005 *Angew Chem Int Ed* **44** 6560–6563.
- [2] Ma H C, Bai X T, Zheng L Q 2011 *CrystEngComm* **13** 3788–3793.
- [3] Wei W, Ma G H, Hu G, Di Y, Mcleish T, Su Z G, Shen Z Y 2008 *J. Am. Chem. Soc* **130** 15808–15810.
- [4] Macketta J J. 1977 *Encyclopedia of Chemicals Processing and Design* (Marcel Dekker: New York).
- [5] Dalas E, Klepetsanis P, Koutsoukos P G 1999 *Langmuir* **15** 8322.
- [6] Wang L N, Huo J C, Liu S X, Lei Y L 2011 *Chinese. J. Struct. Chem* **30** 738-742.
- [7] Ci Z, Wang Y 2009, *J. Electrochem. Soc.* **156** J267-J272.
- [8] Li P L, Wang Z J, Yang Z P, Guo Q L 2011 *Chin. Phys. Lett.* **1** 017801.
- [9] Chen R, Wang Y, Hu Y, Hu Z, Liu C 2008 *J. Lumin.* **128** 1180-1184.
- [10] Hwang K S, Kang B A, Kim S D, Hwangbo S, Kim J T 2011 *Bull. Mater. Sci* **34** 10591062.
- [11] Xue Y, Ren X, Yu M 2012 *solid state sciences* **14** 1086-1091.
- [12] Ci Z, Wang Y 2009, *J. Electrochem. Soc.* **156** J267-J272.
- [13] Kang M, Liu J, Yin G and Sun R 2009 rare metals **28** 439-444.
- [14] Yang J, Liu X, Li C Quan Z, Kang D and Lin J 2007 *J. Crystal Growth* **303** 480-486.
- [15] Bao SP, Chen XY, Li Z, Yang BJ and Wu Y C 2010 *CrystEngComm* **13** 2511-2520.
- [16] Zelati A, Amirabadizadeh A, Kompany A 2011 *IJCEA* **02**.
- [17] Sumy J, Sarala U, Kamathi V P 2009 *J. Chem. Sci.* **121** 685–691.
- [18] Sreedhar B, Vani C S, Devi D K, Rao M V B, Rambabu C 2012 *J. Mat. Sci.* **2** 5-13.
- [19] Bei C, Xiaoxiao Y, Jiagu Y, Xiujian Z 2006 *Rare Metals* **25** 382.
- [20] Karagiozov C, Momchilova D 2005 *Chem. Eng. Process.* **44** 115.
- [21] Lv S, Li P, Sheng J, Sun W 2007 *Mat. Lett.* **61** 4250.
- [22] Chen L, Shen Y, Xie A, Zhu J, Wu Z, Yang L 2007 *Cryst. Res. Technol.* **42** 886.
- [23] Clifford Y T, Chia-te T, Hwai-shen L 2006 *Chem. Eng. Sci.* **61** 7479.
- [24] Zelati A, Amirabadizadeh A, Kompany A, Larimi Z M 2011 *ICCCP* **10** 146-150.
- [25] Feilong S. Junwu Z 2011 *Journal of rare earth* **29** 326-329.
- [26] Motloung S V, Dejene B F, Swart H, Ntwaeaborwa O M 2014 *J. Sol-gel Sci. technol.* **70** 422-427.
- [27] Ohta R, Onda Y, Kojima S 2011 *J Korean phys.soc.* **59** 2471-2474.
- [28] Cheng B, Qu S, Zhou H, Wang Z 2006 *Nanoteq.* **17** 2982-2987.
- [29] Li Y, Chang Y, Lin Y, Chang Y, Lin Y 2007 *J. alloys compd.* **439** 367-375.
- [30] Granados-Correa F, Jime ´nez-Reyes M 2011 *Sep. Sci. Technol.* **46** 2360–2366.
- [31] Wang Z, Kulkani A, Deshpande S, Nakamura T, Herman H 2003 *Acta Mater.* **51** 5319-5334.
- [32] Ohta R, Onda Y, Kojima S 2011 *J Korean phys.soc.* **59** 2471-2474.
- [33] Li Y, Chang Y, Lin Y, Chang Y, Lin Y 2007 *J. alloys compd.* **439** 367-375.
- [34] Zhou L, Choy W C H, Shi J., Gong M, Liang H 2006, *Mat. Chem. Phys.* **100** 372-347.
- [35] Ntwaeaborwa O M, Nsimama P. D, Pitale S, Nagpure I M, Kumar V, Coetsee E, Terblans J J, Swart H C, Sechogela P T 2010 *J. Vac. Sci. Technol. A* **28**.

Synthesis and characterisation of ZnO-CNFs hybrid nanostructures for hydrogen storage applications

CT Thethwayo¹, SE Mavundla², TG Nyawo¹, PN Mbuyisa¹ and OM Ndwandwe¹

¹ Department of physics and engineering, University of Zululand, Private bag X1001, KwaDlangezwa 3886, South Africa

² Department of Chemistry, University of Zululand, Private bag X1001, KwaDlangezwa 3886, South Africa

Email: zhumane@gmail.com

Abstract. Zinc oxide carbon nanofibers hybrids (ZnO-CNFs) have been studied for hydrogen storage applications. Zinc Oxide nanoparticles (ZnO NPs) have been deposited using DC magnetron sputtering. We have then grown aligned ZnO nanorods (ZnO NRs) on the ZnO NPs, finally ZnO-CNFs synthesis have been carried out by chemical vapour deposition (CVD), using acetylene (C₂H₂) as a source of carbon. Elastic Recoil Detection Analysis (ERDA) results are presented and show that ZnO-CNFs are promising candidates for hydrogen storage. Results show that the temperature has an effect on the amount of hydrogen absorbed; at lower temperatures hydrogen is detected mostly on the surface and at higher temperatures hydrogen is detected on the surface and on the bulk which yield a higher hydrogen absorption/adsorption. SEM images show that ZnO-CNFs morphology depends on the growth temperature and is correlated to the ZnO NRs morphology.

Keywords: Hydrogen absorption/adsorption, ZnO-CNFs, ERDA, SEM, and EDS

1. Introduction

The world is facing a big challenge on finding the energy resources that are pollution free. Environmental concerns regarding the use of fossil fuels and their predicted exhaustion are globally important issues. The researchers have focus their studies on the energy resource that can replace fossil fuels, these energy resources should be environmental friendly, free of pollution and have high energy efficiency. Hydrogen is a leading candidate to eliminate petroleum dependence, associated air pollutants and greenhouse gases [1, 2]. Hydrogen can substitute fossil fuels in automobile applications, with additional benefit of potentially allowing the production of zero emission vehicles [2]. Hydrogen is the most abundant element in the universe constituting of about 75% of all baryonic mass and it can be found in different kind of materials [3]. Hydrogen can also be produce by splitting water, these process need energy. If hydrogen has to be used as the main source of energy in the world, it needs to be produced and be stored in high quantity. To store hydrogen we need a material that is light weight, safe, low cost and that will store high quantities of hydrogen in a little space. The goal is to pack

hydrogen as closely as possible in order to achieve the highest volumetric density by using as little additional material as possible.

Most researchers have shown a higher interest in carbon nanostructured materials for hydrogen application, due to their porous structure. Carbon nanofibers (CNFs) are easily synthesised with different methods; CVD, electro spinning, polymer blend techniques and more other methods [3, 4, 7]. Hydrogen can be absorbed/adsorbed in CNFs in two ways; physisorption and chemisorption. In physisorption, hydrogen is absorbed and stored in the pores of CNFs by weak Van der Waal forces. In chemisorption, hydrogen is chemically bonded into carbon forming the weak covalent bonding. The morphology of CNFs is controlled by the substrate/catalyst which was grown onto. In this research work, vertically aligned ZnO NRs were used as the substrate to grow ZnO-CNFs, which could be used for hydrogen storage application.

There are three available onboard automobile hydrogen storage methods: (a) gas compressed to high pressures [1, 6] (b) the use of a cryogenic liquid at temperatures near to its boiling point (20.3 K) [1, 6] and (c) reversible metal and chemical hydrides [1, 6]. Although each hydrogen storage method has desirable attributes, no approach satisfies all of the efficiency, size, weight, cost and safety requirements for personal transportation vehicles [5, 6]. Hydrogen stored as metal and chemical hydrides adds significant weight, cost, and thermal complexity to onboard storage systems [6]. The presence of ZnO NRs in CNFs will alter the properties of CNFs and we believe it will enhance the hydrogen absorption. The method of hydrogen profiling (ERDA) shows that our ZnO-CNFs can absorb/adsorb hydrogen since hydrogen is detected in both bulk and the surface of the sample.

2. Experimental Procedure

2.1. Synthesis of ZnO nanoparticles

ZnO NPs were deposited on the silicon (111) wafer substrate using dc reactive magnetron sputtering (AJA Orion 5 sputtering system). Substrates were chemically washed. The deposition chamber was evacuated to a base pressure of 10^{-5} Torr by oil-diffusion-pump and turbo-pump combination. High purity (99.995%) argon and oxygen were used as sputtering and reactive gases. Zinc metal of 99.999% purity was used as sputtering target. The flow ratio of argon to oxygen was controlled at 2:1. The sputtering pressure, which is a total pressure of argon and oxygen, was set to 3×10^{-3} Torr. The sputtering power was set to 70 W. The sputtering time was varied from 15 minutes to 60 minutes. After each deposition the samples were annealed at a substrate temperature of 300 °C for further oxidation. Annealing time and pressure was 60 minutes and 3×10^{-3} Torr. Zinc target was pre-sputtered in argon-oxygen atmosphere for 5 minutes to remove the surface oxide layer.

2.2. Synthesis of ZnO NRs

ZnO NRs were grown on ZnO NPs using the hydrothermal process [7]. The hydrothermal solution was prepared by dissolving ZnCl_2 on distilled water and hydrolysed with NH_3 (25%). Solution was transferred to autoclave bottles with silicon wafer vertically aligned. The autoclave bottles with substrates were put on the incubator, the temperature was set to 90 °C. The time for growing ZnO NRs was fixed at 2 hours. The incubator was pre-heated for 30 minutes to stabilise the temperature. The samples were removed suddenly after the required time to stop the process, then were washed with distilled water thoroughly to remove impurities on the surface of ZnO NRs and dried on the incubator for 5 minutes.

2.3. Synthesis of ZnO-CNFs

Chemical vapour deposition (CVD) was used for the nucleation of ZnO-CNFs using a dc reactive magnetron sputtering. High purity (99.995 %) of acetylene (C_2H_2) was used as the source of carbon. The flow rate of C_2H_2 was controlled at 10 sccm. Substrate temperature was varied from 500 °C to 800 °C to investigate the effect of temperature on the nucleation of ZnO-CNFs. ZnO NRs were used as a catalyst aligned and to control the morphology of ZnO-CNFs. CVD pressure was kept at $3 \times$

10^{-3} Torr and time was kept at 40 minutes for all the samples.

3. Results and Discussions

3.1. ZnO NPs analysis

Figure 1a-c shows SEM images of ZnO NPs deposited at 15 minutes, 30 minutes, and 60 minutes respectively. The average diameter of the particles sizes increases from 6.73 – 37.60 nm, at 15 minutes the average size was 6.73 nm, at 30 minutes it was 16.71 nm, and at 60 minutes the average diameter was measured to be 37.60 nm. It was observed that, as the deposition time changes, the colour of the film change. The change in the colour of the sample can allow us to use these ZnO NPs for other application i.e as a decorating coatings but on this work are used as seeding for ZnO NRs.

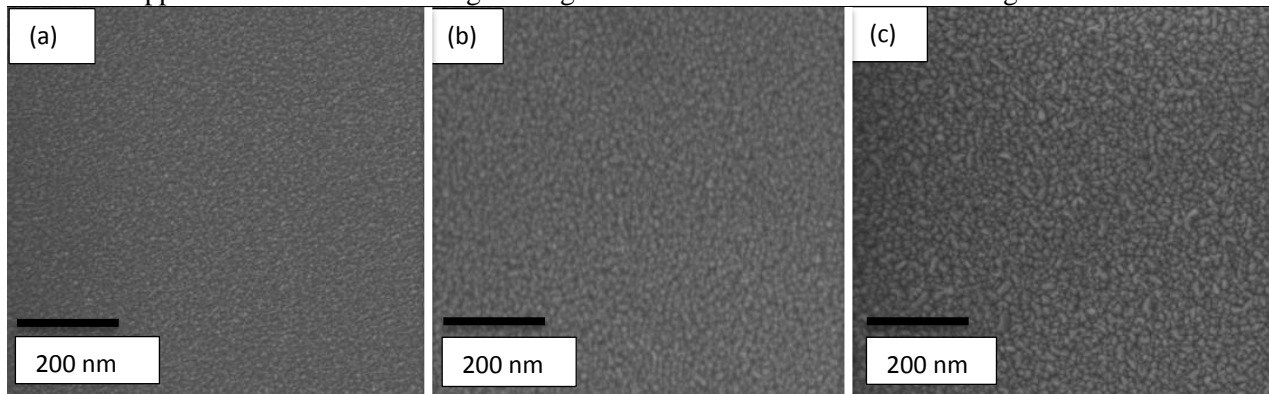


Figure 1: SEM images of ZnO NPs grown at different deposition time.

3.2. ZnO NRs analysis

Since ZnO NPs were used as a seeding for the growth of ZnO NRs, the effect of ZnO NPs on the synthesis of ZnO NRs was investigated.

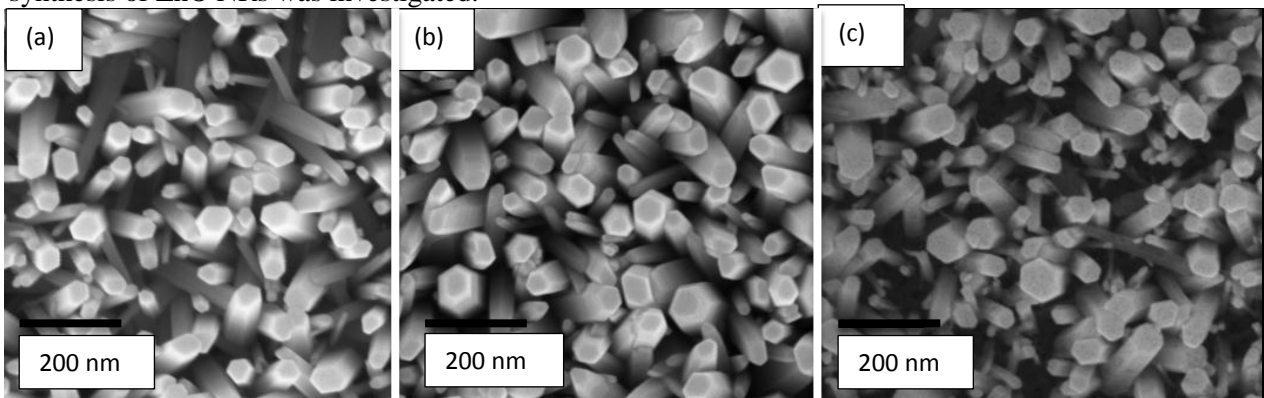


Figure 2: SEM images showing the effect of ZnO NPs grown at different deposition times on the growth of ZnO NRs. (a) 15 min, (b) 30 min and (c) 60 min.

Figure 2 shows the SEM images of ZnO NRs, that were synthesised at ZnO NPs sputtered at different deposition times. ZnO NRs were all synthesised for 2 hours at different seeding of NPs. Figure 2a shows SEM image of NRs grown on 15 minutes sputtered ZnO NPs, their average diameter and the length of these NRs were found to be 60.34 nm and 129.72 nm respectively. Figure 2b shows the NRs of 30 minutes sputtered NPs and their average diameter and length was found to be 75.40 nm and 105.90 nm respectively. Figure 2c show ZnO NRs of 60 minutes sputtered NPs, their average diameter and length was measured to be 90.31 nm and 114.05 nm respectively. The result show that the diameter of NRs increases with the increase in particles size of NPs.

3.3. Analysis of ZnO-CNFs

The effect of temperature on the nucleation of ZnO-CNFs was investigated using SEM imaging and EDS to study the chemical composition. Figure 3 show SEM images of ZnO-CNFs deposited at 500 °C, 600 °C, 700 °C, and 800 °C. Figure 3a shows that ZnO-CNFs take the morphology of ZnO NRs; only small amount of carbon is formed on the surface of NRs, this is confirmed by EDS result as shown in table 1.

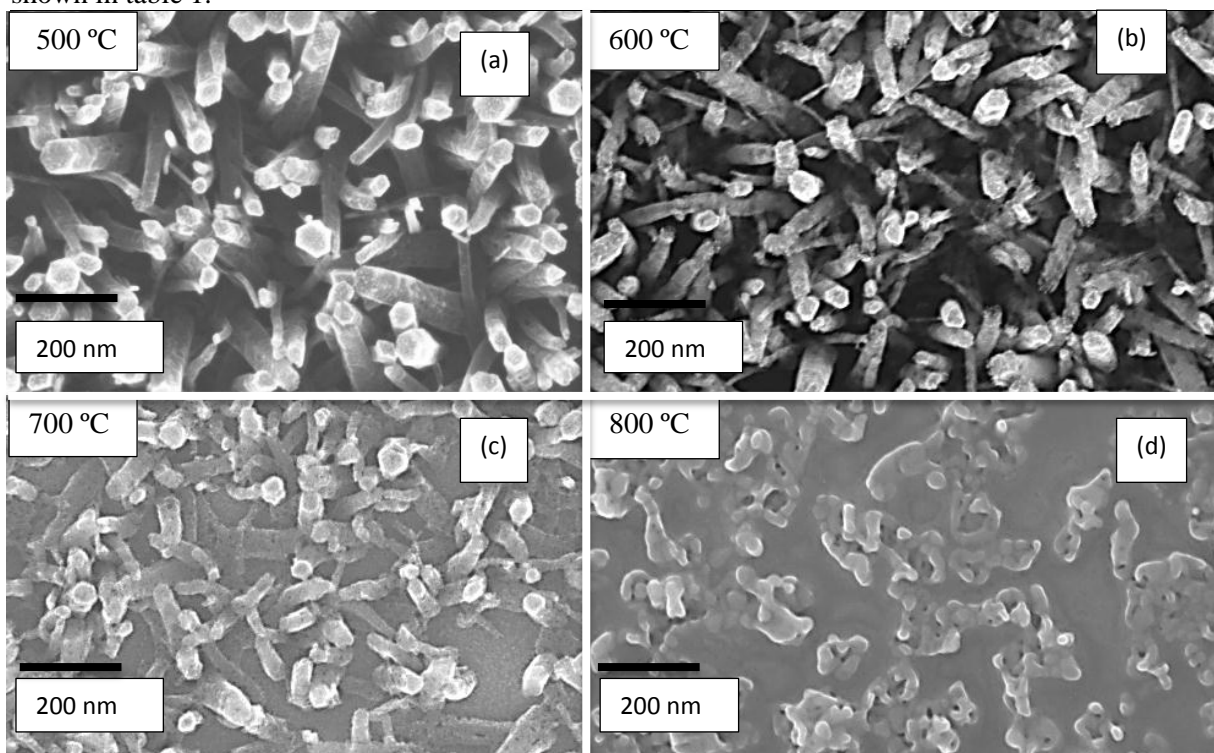


Figure 3: SEM images of ZnO-CNFs synthesised at different temperatures.

Figure 3b & 3c shows that when the temperature is increased (600 °C & 700 °C respectively), the hexagonal structure of ZnO NRs is lost and ZnO-CNFs take the shape of NRs. Figure 3d shows that at high temperatures ZnO-CNFs meltdown and loses the aligned morphology. EDS result show that carbon concentration increase with the increase in temperature (500 °C-700 °C) see table 1. At high temperatures (800 °C+) the content of carbon drops and Zinc and oxygen content increases. The process of CVD involve braking down of C_2H_2 into C^* and H^* . H^* will be attracted to oxygen of ZnO NRs, to form $H_2O_{(g)}$ causing vacancies on the NRs and $H_2O_{(g)}$ is pumped out the system. C^* will grow on that vacancies forming a weak covalent bonding with Zn (temporal bonding), as the concentration of C increases on the system C-C bonds are formed which are very strong covalent bonding. Zinc melts at 419.5 °C; hence when it loses it bonds (with oxygen, and with carbon) it will start to melt, living carbon building in the NRs taking the morphology of ZnO nanorods. As the temperature increases the reaction rate also increases, up until the process will be faster; hence Zn melts before C-C bond is form. CNFs need a catalyst to grow, thus the morphology of CNFs is controlled by a catalyst used. On figure 3d, ZnO-CNFs do not have a perfect morphology because the catalyst was melting down and it followed that structure.

Table 1: EDS table of Result, showing chemical composition in atomic %.

Elements	ZnO NP	ZnO NR	ZnO-CNFs 500 °C	ZnO-CNFs 600 °C	ZnO-CNFs 700 °C	ZnO-CNFs 800 °C
	At. %	At. %	At. %	At. %	At. %	At. %
C			31.29	75.16	84.21	35.11
O	59.74	66.76	45.84	20.74	14.75	63.23
Zn	40.26	33.24	22.87	4.10	1.04	1.66

ERDA method was used to study hydrogen adsorption/absorption on the samples. Samples were named as following; S9 deposited at 700 °C, Ze deposited at 600 °C, S8 deposited at 500 °C, and silicon wafer which is the substrate.

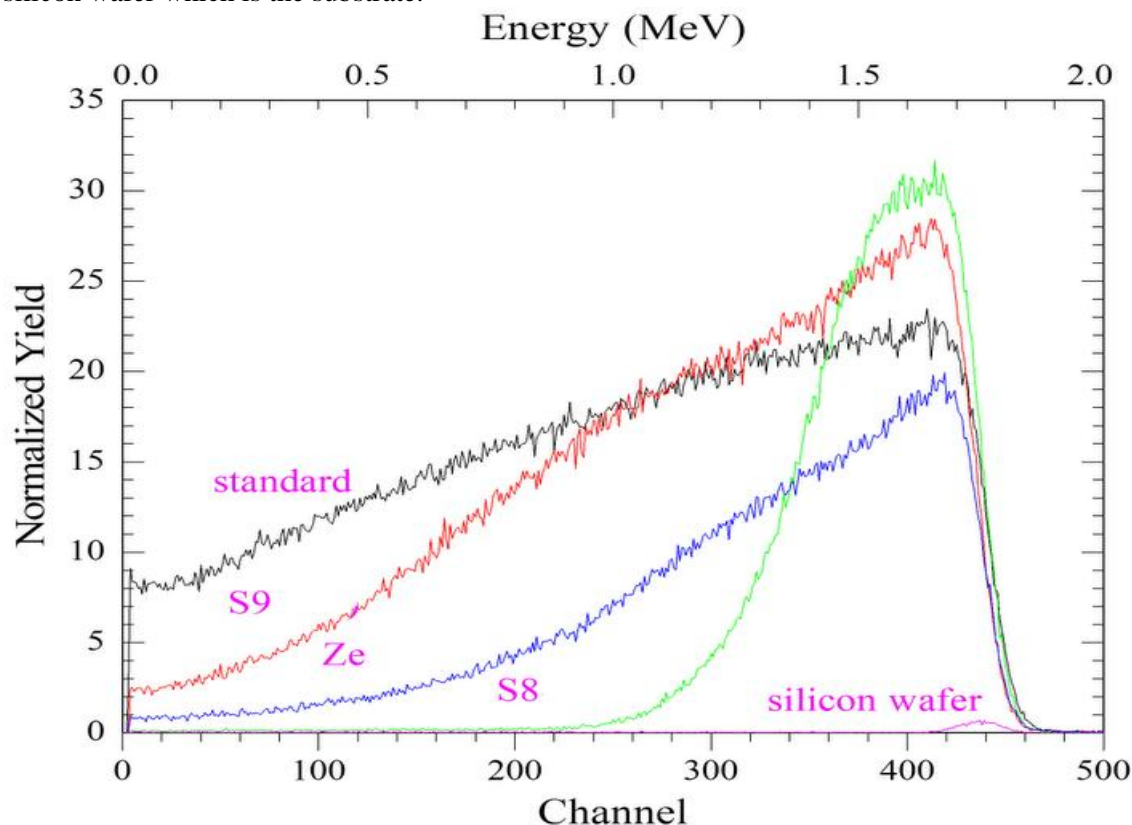


Figure 4: ERDA result, showing hydrogen absorption/adsorption.

The samples were compared to a standard material (kapton), a stable polyimide which has a very high concentration of hydrogen. In kapton hydrogen was detected in the bulk and also on the surface, on silicon wafer it was only detected on the surface. In the sample prepared at lower temperature (500 °C), hydrogen was detected mostly at the surface figure 4 (sample S8). The reason hydrogen is detected on the surface is because carbon was only formed on the surface of the ZnO NR see figure 3a and table 1. As the temperature increases, we started to detect hydrogen on the surface and in the bulk (figure 4 sample Ze deposited at 600 °C), since we still have more content on ZnO we still detect less amount of hydrogen. At 700 °C we start to detect more hydrogen on the surface and in the bulk sample S9 figure 4, these is due to the increase of carbon content on the sample, since we have 84.21 At.% of carbon. These tell us that by increasing the content of carbon on the sample the content of hydrogen increases, since it was absorbed /adsorbed by carbon only not by ZnO NRs.

4. Conclusion

ZnO NRs were grown ZnO NPs sputtered at different deposition time on silicon wafer substrate. Using SEM, it was shown that the size of ZnO nanoparticles increases with the increase of the deposition time and ZnO NRs growth depends on particles size of ZnO NPs. ZnO NPs also shows an interesting properties of changing colour with the change in time, which can be used to use them as a decorating coatings. SEM result showed that ZnO-CNFs morphology was controlled by ZnO NRs. Using ERDA it was determine that ZnO-CNFs have high absorption/adsorption of hydrogen at temperatures around 700 °C than lower temperature around 500 °C.

Acknowledgements

The financial assistance of the National Research Foundation (NRF) and UniZulu Research

committee towards this research is hereby acknowledged. We would like to extend our special thanks to; Ms A Adams and NW Paradza for running MANUS/MATSC program the way they do. We would like to thank all staff of the physics and engineering department at the UniZulu for their support.

References

- [1] Vicente Jimenez, Ana Ramirez-Lucas, Paula Sanchez, Jose Luis Valverde, Amaya Romero, *Hydrogen storage in different carbon materials: Influence of the porosity development by chemical activation*, Applied surface Science 258 (2012) 2498-2509.
- [2] James M Blackman, John W Patrick , Ana Arenillas, Wei Shi, Colin E. Snape, *Activation of carbon nanofibres for hydrogen storage*, Carbon 44 (2006) 1376-1385.
- [3] F Suarez-Garcia, E Vilaplana-Ortego, M. Kunowsky, M. Kimura, A. Oya, A. Linares-Solano, *Activation of polymer blend carbon nanofibres by alkaline hydroxides and their hydrogen storage performances*, INTERNATIONAL JOURNAL OF HYDROGEN STORAGE 34 (2009) 9141-9150.
- [4] Pierre Be´nard and Richard Chahine, *Storage of hydrogen by physisorption on carbon and nanostructured materials*, Scripta Materialia 56 (2007) 803-808.
- [5] Jun Yeon Hwang, Sang Ho Lee, Hyhu Sung Sim, Jong Won Kim, *Synthesis and Hydrogen storage of carbon nanofibers*, synthetic Metals 126 (2002) 81-83.
- [6] Vicente Jimenez, Ana Ramirez-Lucas, Paula Sanchez, Jose Luis Valverde, Amaya Romero, *Improving hydrogen storage in modified carbon materials*, INTERNATIONAL JOURNAL OF HYDROGEN STORAGE 37 (2012) 4144-4160.
- [7] RN Gayen, AK Pal, *Growth of carbon nanofibers on aligned zinc oxide nanorods and their field emission properties*, Applied Surface Science 256 (2010) 617-6178.

Characterization of cerium doped yttrium gadolinium aluminate garnet (Y-Gd)₃Al₅O₁₂:Ce³⁺ phosphor thin films fabricated by pulsed laser deposition

A H Wako^{1*}, F B Dejene¹, H C Swart²

¹Department of Physics, University of the Free State, QwaQwa Campus, Private Bag X13, Phuthaditjhaba 9866, South Africa

²Department of Physics, University of the Free State, P O Box 339, Bloemfontein, ZA-9300, South Africa

E-mail: wakoah@qwa.ufs.ac.za

Abstract. Thin films of cerium doped yttrium gadolinium aluminate garnet (Y-Gd)₃Al₅O₁₂:Ce³⁺ (YGAG:Ce) were grown on Si(100) substrates by a pulsed laser deposition (PLD) technique using a 266 nm Nd:YAG pulsed laser under varying deposition conditions, namely; substrate temperature, substrate – target distance, number of laser pulses and the working atmosphere during the film deposition process. The effect of annealing temperatures on the structure and luminescence properties of the as-deposited (YGAG:Ce) thin films were analysed. Photoluminescence (PL) data were collected in air at room temperature using an F-7000 FL Spectrophotometer. A slight shift in the wavelength of the PL spectra was observed from the thin films when compared to the PL spectra of the phosphor in powder form, which is probably due to a change in the crystal field. The PL intensity of the samples increased as the annealing temperature was increased from 400 °C to around 700 °C and then decreased with continued increase in the annealing temperature.

1. Introduction

Yttrium aluminium garnet (YAG) has proved to be an excellent host material compatible with lanthanides and most trivalent rare earth dopants. YAG are mostly applied in solid state lasers especially the white light emitting diodes (LEDs) by incorporating them with suitable rare earth dopants. YAG activated by trivalent Cerium (Ce³⁺) has excellent chemical and thermal stability and is also well known for good scintillators [1]. The optical band gap energy for YAG bulk powder is in the order of 6.6 eV comprising of filled O₂ 2p orbitals in the valence band and empty 4d orbitals [2].

Yttrium-gadolinium aluminium gallium garnet (Y-Gd)₃Al₅O₁₂:Ce³⁺ is a modified form of Y₃Al₅O₁₂:Ce³⁺ where Ce³⁺ and Gd³⁺ are incorporated into the Y₃Al₅O₁₂ lattice such that Gd³⁺ substitutes Y³⁺ in the system (Y_x-Gd_{1-x})₃Al₅O₁₂:Ce³⁺. Since the ionic radius of Gd³⁺ is larger than that of Y³⁺ replacing Y³⁺ with Gd³⁺ results in lattice expansion. This way depending on the concentration of Gd³⁺ the host matrix structure can be modified in order to blue-shift or red-shift the emission from the Ce³⁺. (Y-Gd)₃Al₅O₁₂:Ce³⁺ can efficiently absorb a wide range of excitation wavelengths from ultra-violet (UV) to visible light and it can give out a broad band emission in the spectral range of 510 nm to over 600 nm [3] which makes it suitable for white LED applications.

This colour shift can be attributed to the fact that the characteristic broad yellow emission band from the $5d^1 - 4f^1$ transitions of Ce^{3+} is highly sensitive to changes in the chemical environment of the host which induces crystal-field effect in the lowest $5d^1$ orbital [4].

Several methods have been used to grow thin films such as Physical Vapour Deposition (PVD) and/or Chemical Vapour Deposition (CVD) [5], Chemical Bath Deposition [6], rf magnetron sputtering [7] pulsed laser deposition (PLD) [8] and epoxide-catalyzed sol-gel methods [9]. However PLD has several advantages over other techniques such as the homogeneous stoichiometric evaporation of the ablated material on to the substrate [5] and good quality thin films can be obtained at low temperatures [10] by varying parameters such as substrate distance, atmosphere and substrate temperature among others. These parameters affect the thickness, roughness and overall quality of the obtained films which in turn influence the structure and photoluminescence (PL) properties of $(Y-Gd)_3Al_5O_{12}:Ce^{3+}$ phosphor.

In this study thin films of $(Y-Gd)_3Al_5O_{12}:Ce^{3+}$ were prepared using PLD. The effect of different annealing temperatures on the structure and photoluminescent (PL) properties of the $(Y-Gd)_3Al_5O_{12}:Ce^{3+}$ thin films were investigated.

2. Experimental

A Commercial yellow $(Y-Gd)_3Al_5O_{12}:Ce^{3+}$ powder obtained from Phosphor Technology (UK) was pressed into a pellet and used as a target for laser deposition after initial characterization. The pellet was then heated at 200 °C for 2hrs in a furnace to expel moisture that may be present and make it hard and strong for use as a target for laser ablation. The target was then loaded on to a rotatable carousel inside a PLD chamber. The carousel raster and rotates the target to prevent constant ablation by the laser beam from the same spot on the target.

Silicon (100) wafers of approximately 1 cm² were used as substrates. The wafers were first cleaned in an ultrasonic bath of acetone, ethanol and distilled water in that sequence then blow-dried with clean dry N₂ gas and placed on a fixed substrate holder in the PLD chamber perpendicular to the target at a fixed distance of 4.5 cm from the target. The chamber was then pumped down to vacuum of about 3.0x10⁻⁶ mbar. The target was ablated using a 266 nm Nd:YAG nanosecond (ns) pulsed laser with energy and frequency fixed at 45mJ/pulse (fluence of 0.2 Jcm⁻²) and 10Hz respectively.

The ablation of the $(Y-Gd)_3Al_5O_{12}:Ce^{3+}$ target was carried out in the PLD chamber with a vacuum pressure of 10⁻⁶ mbar at constant substrate temperature of 200 °C. To investigate the effect of annealing, the as-deposited films were first characterized and then post-annealed in air for 60 minutes at 400 °C, 800 °C and 1000 °C in an XD-1200 NT tube furnace. The room temperature PL excitation and emission fluorescence spectra were recorded using an F-7000 FL Spectrophotometer at scan speed of 60 nm/min at 400 PMT voltage and excitation and emission monochromator slit widths of 5 nm.

X-Ray Diffraction (XRD) continuous scans were recorded using a Bruker-AXS D8 Advance X-ray diffractometer operating at 40 kV and 4 mA using Cu Ka = 1.5406 nm from 15° to 65° (2θ), with a scan rate of 0.39° (2θ)/min and step scans with a step size of 0.02°(2θ).

3. Results and discussions

3.1. Structure

XRD patterns of the $(Y-Gd)_3Al_5O_{12}:Ce^{3+}$ phosphor powder and thin films that were post-calcined at 400 °C, 800 °C and 1000 °C are shown in figure 1(a). Except for impurity peaks the rest fitted well with the powder and the JCPDS card number 072-1315 for the cubic Y₃Al₅O₁₂ of space group Ia-3d(230), cell ratios a/b=1.0000 b/c=1.0000 c/a=1.0000 and cell parameters a=12.0160Å. FWHM is an indication of peak broadening [11]. From figure 1(a) it was observed that XRD peak broadening occurred with an increase in the annealing temperature up to around 700 °C (estimated from the smoothed curve) and then narrowed down as the temperature was increased to 1000 °C. This is also

displayed on figure 2(a) which shows the comparison of FWHM with annealing temperature. The average crystallite sizes calculated for the (420) direction were obtained using Scherrer's equation $D = K\lambda/(\beta \cos\theta)$ where D is the mean particle size, K is a geometric factor, λ is the X-ray wavelength and β is Full Width at Half Maximum (FWHM).

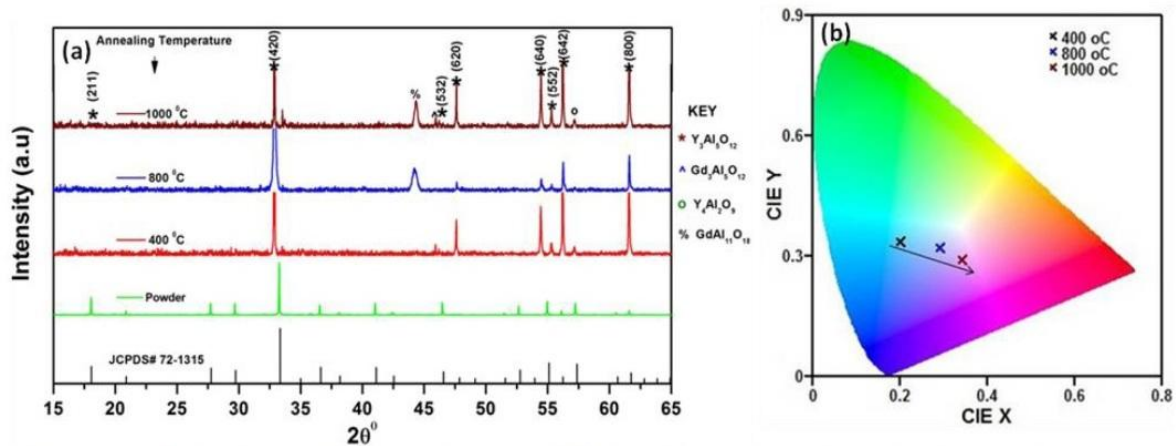


Figure 1. (a) XRD patterns of the (Y-Gd)₃Al₅O₁₂:Ce³⁺ phosphor powder and thin films that were deposited in vacuum at 300 °C for 10 minutes and annealed at 400 °C, 800 °C and 1000 °C respectively and (b): CIE colour chromaticity as a function of annealing temperature.

Figure 2 (a) and (b) also shows how the crystallite size varied with annealing temperature. The crystallite size reduced as the temperature increased up to around 700 °C suggesting improved crystallinity and growth [11] but as annealing temperature increased above 700 °C, the sample was gradually losing crystallinity which could be attributed to phase change that may have introduced impurities. As was also reported by Y Deng et al that Y₃Al₅O₁₂: Gd thin films deposited at 700 °C were amorphous, supports this fact [7].

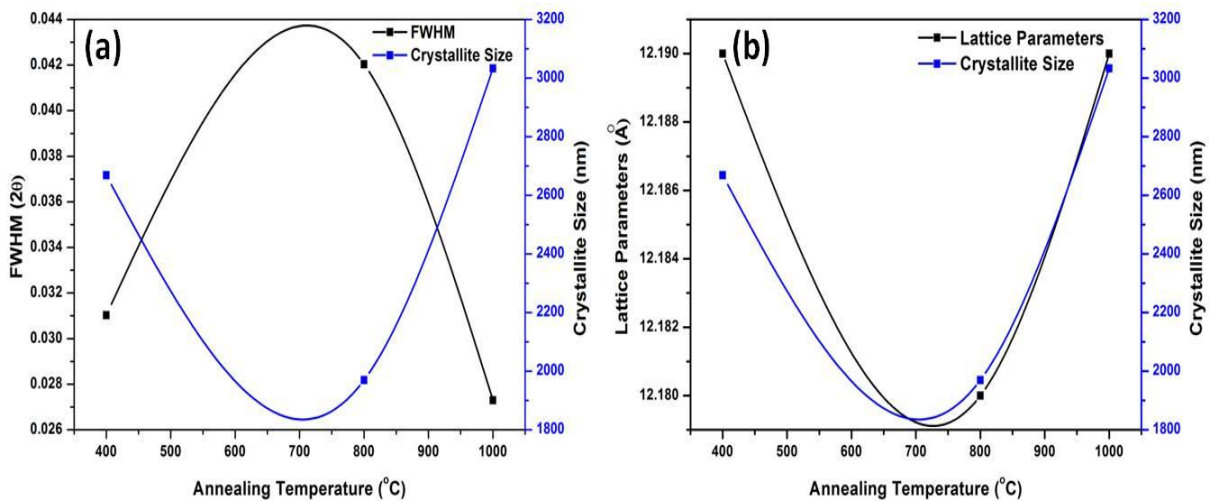


Figure 2(a) and (b). FWHM, Crystallite size and Lattice parameter as a function of annealing temperature.

It was also reported elsewhere that the phases; Y₃Al₅O₁₂ (cubic) and YAlO₃ (perovskite) were stable between room temperature and their melting points of 1970 and 1870 °C respectively but Y₄Al₂O₉ was

unstable below 1000 °C [12]. Yousif et al showed that new phases were formed after annealing the $\text{Y}_3(\text{Al,Ga})_5\text{O}_{12}:\text{Tb}^{3+}$ films at temperatures greater than 1073 K [8].

From figure 2(b) we observed that the lattice parameters are also influenced by the annealing temperature in a similar manner. It was observed to decrease with increasing temperature up to around 700 °C.

It was also observed that diffraction peaks in the directions; (420), (552), (640) and (642) shift toward smaller angles compared to those of the $(\text{Y-Gd})_3\text{Al}_5\text{O}_{12}:\text{Ce}^{3+}$ commercial powder and the JCPDS card #072-1315 which could be due to the impurities phases; $\text{Gd}_3\text{Al}_5\text{O}_{12}$, $\text{Y}_4\text{Al}_2\text{O}_9$ and $\text{GdAl}_{11}\text{O}_{18}$ that are formed as temperature was increased. It has been reported that the phases YAM ($\text{Y}_4\text{Al}_2\text{O}_9$), YAP (YAlO_3) and YAG ($\text{Y}_3\text{Al}_5\text{O}_{12}$) are often present in the $\text{Y}_2\text{O}_3\text{-Al}_2\text{O}_3$ system even if YAG is prepared stoichiometrically [13]. Therefore, this shift could be due to lattice expansion caused by differences in ionic radius when Y ($r=0.104$ nm) is substituted by Gd ($r=0.108$ nm) from the impurities phases [14] formed during the ablation process.

The crystallization temperature of the samples prepared by solid-state reaction (>1600 °C) is quite higher than the highest annealing temperature of 1000 °C in our experiment. But, during laser ablation, the temperature within the irradiated target volume rises rapidly exceeding the melting/boiling points of the target composition leading to explosive boiling because the laser heating rate significantly exceeds that of thermal diffusion and radiative losses [15].

3.2. Photoluminescence

Figures 3(a) and (b) show the PL excitation and emission spectra recorded from the $(\text{Y-Gd})_3\text{Al}_5\text{O}_{12}:\text{Ce}^{3+}$ as-prepared powder and thin films annealed at 400 °C, 800 °C and 1000 °C. The PL emission spectra from the powder sample with a maximum at 546 nm is a broad band ranging from 450 nm to 650 nm when excited with 337 nm UV due to the 4f-5d electronic transition of Ce^{3+} attributed to the de-localization of electrons from the lowest 5d level to the crystal field split 4f ($^2\text{F}_{5/2} / ^2\text{F}_{7/2}$) levels of Ce^{3+} .

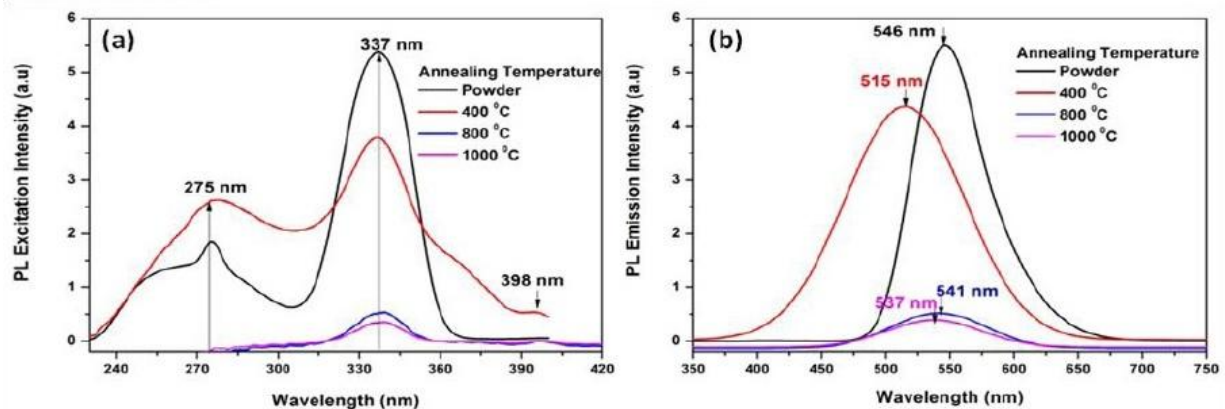


Figure 3 (a) and (b). Excitation and emission spectra of the $(\text{Y-Gd})_3\text{Al}_5\text{O}_{12}:\text{Ce}^{3+}$ phosphor powder and thin films that were deposited in vacuum at 200 °C for 10 minutes and annealed at 400 °C, 800 °C and 1000 °C respectively.

The normalized PL spectra for the annealed samples reveal changes in the peak positions and intensities relative to that of the powder sample. The PL intensity of the samples increased as the annealing temperature was increased from 400 °C to around 700 °C and then decreased with continued increase in the annealing temperature as shown in figure 4(a). Such observation was reported by Yousif et al who observed that intensity ratio of $I_{490\text{nm}}/I_{544\text{nm}}$ and $I_{625\text{nm}}/I_{544\text{nm}}$ for $\text{Y}_3(\text{Al,Ga})_5\text{O}_{12}:\text{Tb}^{3+}$ phosphor, gradually decreased when the annealing temperature was increased from 1073 K (800 °C) to 1473 K (1200 °C) [8]. This can be attributed to stress and cracking that led to the diffusion of Si into

the film at higher temperatures [16]. Also it can be due to the impurity phases, $\text{Gd}_3\text{Al}_5\text{O}_{12}$ and $\text{GdAl}_{11}\text{O}_{18}$ that are higher in Gd concentration since it is well known that for a phosphor to exhibit the required properties, purity is very important. The PL intensities of the films are generally low in all cases compared to those of the powder of the same chemical composition due to total internal reflection in thin films [17]. However thin films have advantages over the powders in that they possess better thermal stability, better adhesion to the substrate, higher lateral resolution and less out gassing in device applications [8,18]. As shown in figure 3(b) the centers of the emission bands red-shift from 515 to 541 nm as the annealing temperature increased. This red-shift can be explained by an increase in the crystal field effect when the Y^{3+} vacancies are occupied by the larger Gd^{3+} . Since the ionic size of Gd^{3+} is larger than that of Y^{3+} it results in lattice expansion confirmed by the shifting of XRD peaks towards the smaller angles, which correlates with the red-shift of Ce^{3+} emission. The variation of the lowest crystal-field energy level of the $5d^1$ state is strongly influenced by the electrostatic field in the surrounding host lattice which in turn depends on the crystalline structure of the host $(\text{Y-Gd})_3\text{Al}_5\text{O}_{12}$ lattice [3].

The CIE chromaticity co-ordinates of PL spectra of all the samples compared with the powder are shown in figure 1(b). It is clear that as the annealing temperature increases the co-ordinates shift to red as pointed out by the arrow. This is also supported by the PL spectra wavelength shifts in figure 3(b).

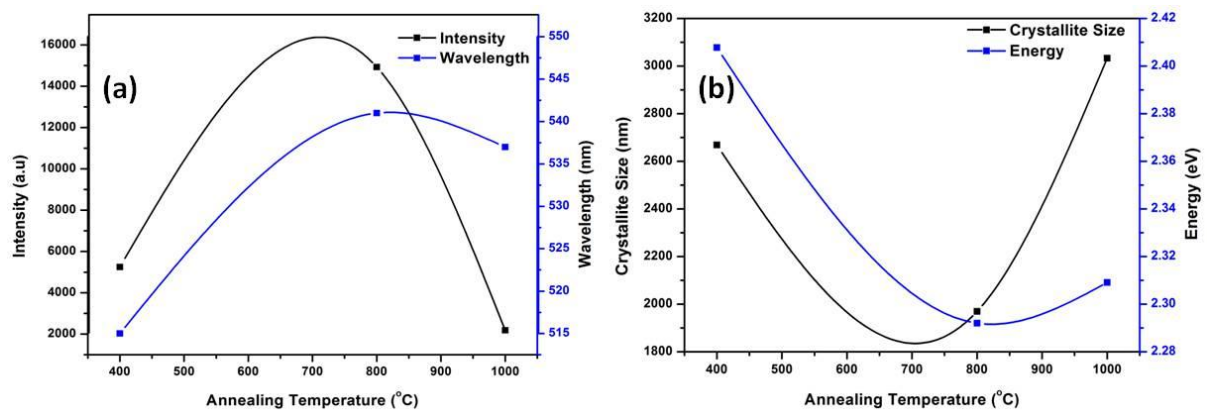


Figure 4: (a) PL intensity (a.u), Peak wavelength (nm) and (b) Peak Wavelength and band gap energy as a function of annealing temperature.

Figure 4(a) also shows how the PL wavelength peak positions vary as a function of annealing temperature. The wavelength peaks slightly shifted to higher wavelengths (red-shift) from 515 to 541 nm as the annealing temperature was increased from 400 – 800 °C and then blue-shifts to lower wavelengths at temperatures above 800 °C. The optical band gap calculated with respect to the PL peak positions shifted to a lower energy from 2.41 - 2.29 eV with increasing annealing temperature from 400 - 800 °C and then to higher energy 2.29 - 2.31 eV above 800 °C as shown in figure 4(b). These changes have been attributed to the morphological dependent properties of the energy band gap [19]. 515 to 541 nm. The band gap of the thin film nanophosphors was calculated to be ~2.34 eV, which is generally red shifted in comparison to the band gap (~6.6 eV) of its bulk counterpart. The blue shift in band gap is a clear signature of nanostructure formation and may be attributed to the quantum size effect [20].

4. Conclusion

In this study thin films of $(\text{Y-Gd})_3\text{Al}_5\text{O}_{12}:\text{Ce}^{3+}$ were prepared using the PLD technique. The effect of varied annealing temperature on the structure and PL properties of the $(\text{Y-Gd})_3\text{Al}_5\text{O}_{12}:\text{Ce}^{3+}$ thin films were investigated. It was found that as the annealing temperature increase broadening of the XRD peaks occurs. Since the ionic size of Gd^{3+} is larger than that of Y^{3+} substitution of Y^{3+} with Gd^{3+} leads

to lattice expansion confirmed by the shifting of the XRD peaks towards the lower angles which in turn cause slight shifts in the wavelength of the PL spectra toward red with respect to the annealing temperature which is probably due to a change in the crystal field caused by lattice expansion.

Acknowledgments

The authors are immensely grateful to the Prestige Cluster Bursary Program of the University of the Free State for financial support towards this project. The authors also thank the University of the Free State Physics Department for assistance with the research technique systems used to characterize materials for this study.

References

- [1] Potdevin A, Chadeyron G, Briois V and Mahiou R 2011 *Mater. Chem. Phys.* **130** 500
- [2] Dlamini S T S, Swart H C, Terblans J J and Ntwaeaborwa O M 2013 *Solid State Sci.* **23** 6
- [3] Pan Y X, Wang W, Liu G K, Skanthakumar S, Rosenberg R, Guo X Z and Li K K 2009 *J. Alloys Compd.* **488** 638
- [4] Yousif A, Swart H C and Ntwaeaborwa O M 2013 *J. Lumin.* **143** 201
- [5] Hirata G A, Lopez O A, Shea L E, Yi J Y, Cheeks T and Mckittrick J 1996 *J. Vac. Sci. Technol.* **14** 1694
- [6] Koao L F, Dejene F B and Swart H C 2014 *Mater. Sci. Semicond. Process.* **27** 33
- [7] Deng Y, Fowlkes J D, Rack P D and Fitz-Gerald J.M 2006 *Opt. Mater. (Amst).* **29** 183
- [8] Yousif A, Swart H C, Ntwaeaborwa O M and Coetsee E 2013 *Appl. Surf. Sci.* **270** 331
- [9] Murai S, Verschuuren M A, Lozano G, Pirruccio G, Koenderink A F and Rivas J G 2012 *Opt. Soc. Am.* **2** 1707
- [10] Sun X W and Kwok H S 1999 *Appl. Phys. A* **69** S39
- [11] Chun-jia L I U, Rui-min Y U, Zhi-wei X U, Jing C A I, Xing-huang Y A N, Xue-tao L U O, Science M and Engineering C 2007 *Trans. Nonferrous Met. Soc. China* **17** 1093
- [12] Cockayne, B. 1985 *J. Less-Comm.Met.* **114** 199
- [13] Huh Y, Cho Y and Do Y R 2002 *Bull. Korean Chem. Soc* **23** 1435
- [14] Uysal S S, Ege A, Ayvacikli M, Khatab A, Ekdal E, Popovici E J, Henini M and Can N 2012 *Opt. Mater. (Amst).* **34** 1921
- [15] Ashfold M N R, Claeysens F, Fuge G M and Henley S J 2004 *Chem. Soc. Rev.* **33** 23
- [16] Yousif A, Swart H C and Ntwaeaborwa O M 2013 *Proc. Int. Conf. of South Africa Institute of Physics (Zululand)* p 1
- [17] Choe J 2002 *Mater. Res Innov.* **6** 238
- [18] Hillie K T, Ntwaeaborwa O M and Swart H C 2004 *Phys. Status Solidi* **1** 2360
- [19] Ezema F I, Ekwealor B C and Osuji R U 2006 *Turk J Phys* **30** 157
- [20] Ankush V, Sanjeev G, Ravi K, Amit K C, Ramesh C, Nafa S, Keun H C 2012 *J. Alloys Compd.* **527** 1

Division B – Nuclear, Particle and Radiation Physics

Single top quark production in association with the Higgs: a feasibility study

Claire Antel¹

Department of Physics, University of Cape Town, Cape Town, 7701, South Africa

E-mail: claire.antel@cern.ch

Abstract. Due to the interference between top and W loops in the $H \rightarrow \gamma\gamma$ decay, there is a degeneracy in the minima of the Higgs coupling fits to fermions and bosons reported by ATLAS and CMS. Although one minimum is consistent with the Standard Model, the second minimum lies in a tantalizing region that is open to new physics.

This anomalous case is studied further by making use of a fermionic-bosonic interference that takes place in the single top production in association with a Higgs, in which the cross-section is significantly enhanced for anomalous coupling values. A truth level feasibility study of the tH process with $H \rightarrow b\bar{b}$ is conducted at 8 TeV and 21 fb⁻¹ of data. A sensitivity of 1.6 is found for the enhanced non-SM case which is insufficient for a detection. A 3σ value is reached at 100 fb⁻¹. However, this value falls to below 1σ when including a 10% uncertainty in background. This indicates that due to the worsening systematic effects at increased luminosities, the signal will likely remain undetectable for the upcoming 14 TeV run at the LHC.

1. Introduction

The year 2012 marks the discovery of the final particle predicted by the Standard Model. The CMS and ATLAS experiments at CERN concurrently announced the observation of the Higgs Boson at 125 GeV [1, 2]. Having found the Higgs, focus is now shifted to measuring its properties in order to determine if the new particle behaves like the SM Higgs or if properties are revealed to be anomalous, leading us into physics Beyond the Standard Model (BSM).

One of the properties of the Higgs is the strength of its coupling to other particles. Its coupling to fermions are known as Yukawa interactions. In this paper, the strength of the top Yukawa coupling is investigated by looking at the single top channel in association with Higgs production. This specific channel is of interest as it is sensitive to the sign of the top Yukawa coupling scale factor. Furthermore, there exist certain constructive interferences that lead to an enhanced signal production rate for an anomalous coupling value [3]. Despite the boost in signal, there is considerable background for this channel. This paper explores a truth-level feasibility study assuming a dataset of 21.3 fb⁻¹ and a centre of mass energy of 8 TeV using the ATLAS detector at the LHC.

2. Theory Background

Both CMS and ATLAS have published their best fits to the coupling strength of the Higgs to other SM particles. Figure 1 displays the coupling fits done by ATLAS. The vector boson and

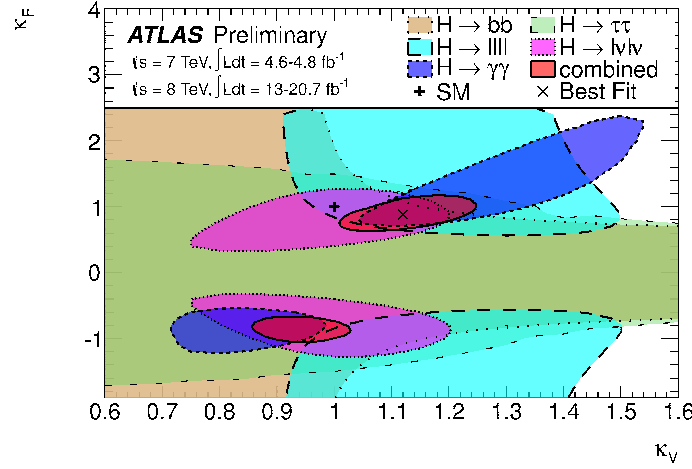


Figure 1. κ_V vs κ_F coupling fits for every Higgs decay channel and their combination. The dashed lines mark the 68% CL contours. The best fit and the SM prediction (1:1) are marked by crosses. Although the best fit agrees within 68% with SM, there remains an area of positive correlation at negative κ_F values which is significant at a $\sim 1\sigma$ level [4].

fermion coupling strengths are given in terms of κ_V and κ_F respectively. These are defined using a simplified parametrisation as

$$\kappa_V \equiv g_{hWW}/g_{hWW}^{SM} \quad \text{and} \quad \kappa_F \equiv g_{h\bar{t}t}/g_{h\bar{t}t}^{SM} \quad (1)$$

They essentially quantify the deviation from the SM prediction, g^{SM} .

In this fit, κ_V is constrained to > 0 ; nothing is lost by setting a positive constraint on one of the two factors because only the *relative* signs between the two coupling scale factors are physical. The 68% CL contour lines are traced out for each fit for a decay channel of the Higgs, as well as the combined fit. It is found that two minima result - one at (1:1), and the other at (1:-1). An important feature of the plot is the symmetry of the areas for every decay channel except for the $H \rightarrow \gamma\gamma$ decay. The symmetry in $\pm\kappa_F$ values is due to these channels being *insensitive* to the sign of the fermion coupling value. The $\gamma\gamma$ channel in contrast is dependant on the relative sign between κ_F and κ_V . This results from the dependence of the event yield on the coupling scale factors which goes like

$$\mathcal{N} \propto \frac{\kappa_x^2 \kappa_y^2}{\kappa_H} \quad (2)$$

where index x is from the production process and index y the decay channel. Because only the square of the κ values appear, it is only the absolute values of these that can be determined. However, the decay of Higgs to two photons is loop induced - the decay happens through a loop of heavy virtual particles. The major loop contributions are from the top, bottom and W boson. It is thus that κ_γ is a function of the more fundamental coupling scale factors, κ_F and κ_V , scaling as $|\alpha\kappa_V + \beta\kappa_F|^2$. This channel is therefore the only channel that can discriminate between signs. It however has the lowest branching ratio ($\sim 0.2\%$) and so statistics are not high enough to lift the degeneracy in minima evident in figure 1. For this reason, other channels that are also sensitive to the relative sign are needed. Several works (e.g. [3], [5]) have suggested the single top channel. Similar to the $H \rightarrow \gamma\gamma$, it too contains an interference term. The two main processes are presented in figure 2, where the Higgs is either radiated from a top or a W. The interference between the fermionic and bosonic process leads to an enhancement of the

number of signal events if the coupling factors are opposite in sign. It in fact leads to a boost a magnitude higher for a k_F value of -1. Different Higgs decay channels can subsequently be considered. Here we focus on the $H \rightarrow b\bar{b}$ decay which has the largest branching ratio ($\sim 58\%$), and so the highest statistics, but it however also suffers from significant background. It is investigated if optimised cuts may exclude enough background to be able to make a detection feasible.

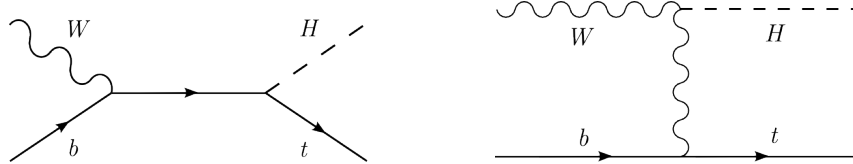


Figure 2. Feynman diagrams of the two major contributions (s channel, left, and t channel, right) to the core process $Wb \rightarrow tH$.

3. Monte Carlo Samples

The top is required to decay leptonically, i.e. $t \rightarrow W + b \rightarrow \ell + \nu + b$. The final signal is then $p + p \rightarrow t + H + j \rightarrow \ell + 3 \text{ b-jets} + j + MET$. The extra jet (j) is a by-product of the initial collision and a characteristic feature of tH production. It is generally produced at large rapidities and thus referred to as a forward jet. MET is missing transverse energy which originates from the neutrino that will shoot through the detector undetected.

Samples were generated at a centre-of-mass energy of 8 TeV. Two samples were generated for the signal: an SM-case, $\kappa_F = 1$ and a non-SM case, $\kappa_F = -1$.

The following backgrounds are considered:

- $t\bar{b}b\bar{j}$: an irreducible background
- tZj , where $Z \rightarrow b\bar{b}$: an irreducible background
- $t\bar{t}$, where $t \rightarrow bc\bar{s}$: a reducible background, where the c or \bar{s} are mistagged as a b-jet
- $t\bar{t}j$, where $t \rightarrow bc\bar{s}$: a reducible background, where the c or \bar{s} are mistagged as a b-jet and the third jet is missed.

The charge conjugate processes of the last two backgrounds are also included.

Only samples with at least one top decaying leptonically are used. The cross sections of the signal samples are 1.66 (SM) and 27.81 fb $^{-1}$ (non-SM). The cross sections for the background samples are 11.28, 6.27 and 210.85×10^3 fb $^{-1}$ for $t\bar{b}b\bar{j}$, tZj and $t\bar{t}$ (includes $t\bar{t}j$), respectively. The $t\bar{t}$ sample is significantly larger than the rest (by 4 to 5 magnitudes), but it is also reducible.

4. Event Selection

Although a lot of cuts were considered only cuts that improved the sensitivity are included in this paper.

Basic acceptance cuts are made from the outset, motivated by suppression of low-momentum pileup and underlying event particles and keeping the acceptance area to regions where the detector is of optimum efficiency. These cuts are a $p_T > 25$ GeV and $|\eta| < 2.5$ cut on all jets (save for the forward jet) and leptons.

4.1. lepton cut

Events containing exactly one electron or one muon are selected (τ leptons are ignored in this truth study due to a low efficiency rate). This cut serves as a *1 lepton* trigger. The lepton number distribution (τ s excluded) for each sample will result in the one lepton requirement cutting out $\sim 33\%$ of events in the signal and tZj cases (due to the elimination of τ lepton events), $\sim 40\%$ of events in the $t\bar{t}$ case and leaves the $t\bar{t}b\bar{b}j$ events untouched (due to the τ filter already been applied during sample file generation).

4.2. η distributions

An important distinction between the signal and $t\bar{t}$ samples lies in the η distributions. The distributions for the signal and $t\bar{t}$ samples are displayed in figure 3 (the distributions for the $k_F = 1$, $t\bar{t}b\bar{b}j$ and tZj samples look similar to the $k_F = -1$ distribution).

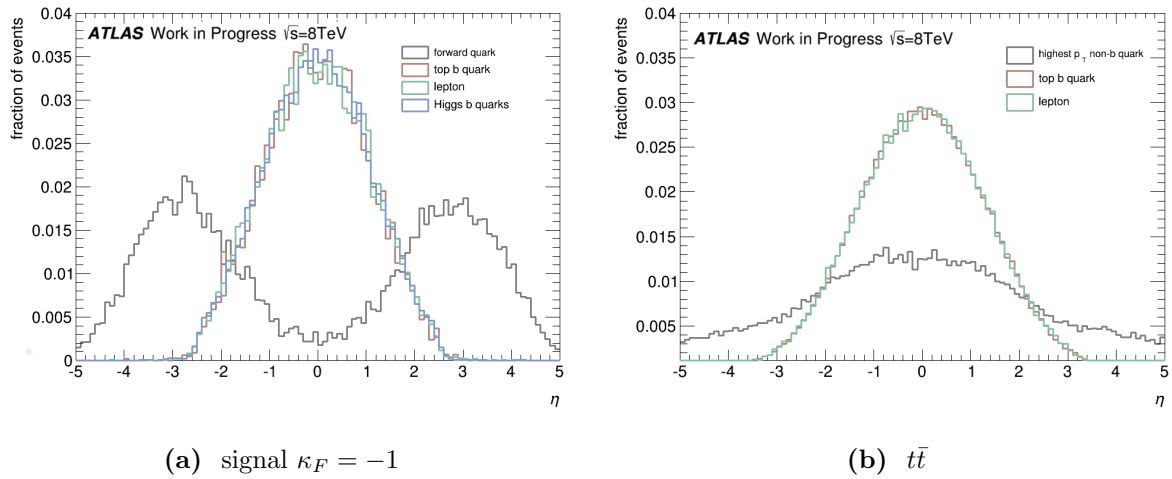


Figure 3. η comparisons: the distributions of MC truth η of light non-b quarks (grey) - in the case of signal events this is the forward quark, Higgs candidate b quarks (blue), the top b quark (red) and leptons (green) are shown for each sample.

For the signal events, two peaks flank the central region at high $|\eta|$ resulting from the forward jet. This feature is absent for $t\bar{t}$. Events are thus selected by requiring a high- p_T jet at $|\eta| > 2.5$.

4.3. b-jet number

The b-jet cut is the most effective cut that can be used. This is clear when looking at the b-jet number distribution for each sample. The $t\bar{t}$ background can be reduced by as much as 99% as the majority of events only include two b-tagged jets. Subsequently, keeping all events with 3 b-jets or more leads to a $\sim 34\%$ cut in signal events, $\sim 45\%$ in tZj events, $\sim 59\%$ cut in $t\bar{t}b\bar{b}j$ events and a fine $\sim 99\%$ cut in $t\bar{t}$ events. The 25 GeV cut is evidently more effective in the $t\bar{t}b\bar{b}j$ and tZj events, as a smaller fraction of 3 b-jet events remains.

4.4. Higgs mass reconstruction

The salient difference between the signal and the backgrounds is the presence of the Higgs boson. This means that the invariant mass of two of the b-jets in a signal event combined should reconstruct to a value that is close to the Higgs mass.

All samples include a W boson and a top. Thus, after having found the b-tagged jet that most likely originates from the top, one can expect the remaining 2 b-tagged jets (in 3 b-tagged jets

events) to reconstruct well to a Higgs mass for the signal samples and poorly for the background samples where no Higgs exists. A difficulty arises when reconstructing the W boson however, as neutrinos cannot be detected. Instead, the observable is the missing transverse energy assumed to be the neutrino. Only the transverse component of MET is reconstructed directly, the z-component needs to be estimated using additional information.

After permuting through all 3 or 4 b-tagged jets in the event, the b-tagged jet that together with the reconstructed W boson vector adds up to the invariant top mass closest, is selected. Finally, the invariant mass of the remaining b-tagged jets is computed. Normalized histograms of the $m_{b\bar{b}}$ distributions are shown in figure 4. In the case of 4 b-tagged jets events, the pair closest to the Higgs mass is chosen. The tZj sample especially shows a deviation from a Higgs-like peak: it leans more to the left, with a mean at 17 GeV above the Z boson mass. The RMS values of the distributions for $t\bar{b}b\bar{j}$ and $t\bar{t}$ are 10 – 17 GeV wider than the RMS values for the signal samples. Following this, the constraint $90 \text{ GeV} < m_{b\bar{b}} < 140 \text{ GeV}$ is applied to events. In reality, these distributions will be widened by experimental effects.

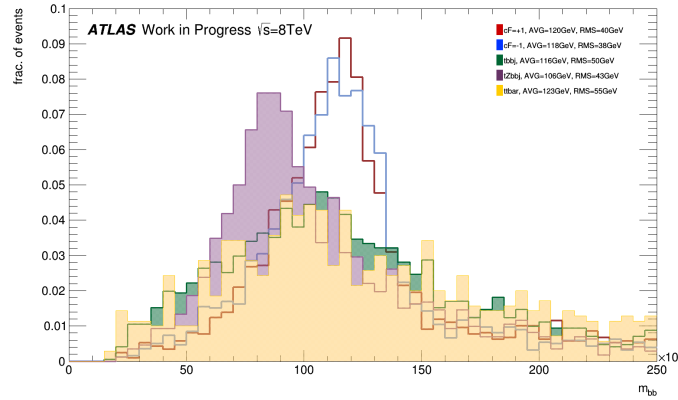


Figure 4. The mass distribution of the Higgs candidate for signal and background (normalised to 1): The b-tagged jet that is used for the top reconstruction is eliminated and the remaining pair of jets (or the pair that reconstructs closest to the Higgs mass in the case of a 4 b-tagged jets event) are used. Basic cuts, 3/4 b-jets, 1 lepton and 1 fwd jet cuts have been applied.

5. Discussion on Sensitivity

The cut flow for an integrated luminosity of 21 fb^{-1} is laid out in table 1. The significance is estimated as S/\sqrt{B} , which is valid for a large background and assuming poisson fluctuations. No systematic errors are included. The greatest boost to the significance is the 3 b-jet cut. The requirement of a forward jet is also useful and the constraints on $m_{b\bar{b}}$ marginally so.

The final significance, S/\sqrt{B} , is ~ 0.1 in the $\kappa_F = +1$ case and ~ 1.7 in the $\kappa_F = -1$ case. It is clear that for the Standard Model scenario the signal is basically indistinguishable from background noise. For the non-SM scenario the situation is improved due to its enhanced cross-section, however a significance value of at least 3 is needed to resolve the ambiguity in sign. Adding to that, the systematic uncertainties involved in measuring real data will obscure this value further. Data at 8 TeV and with 21 fb^{-1} integrated luminosity is likely to be insufficient to determine the sign of the top Yukawa coupling scale factor. In table 2, the significance estimates are summarized. Included are also the significance values assuming higher integrated luminosities

Table 1. Cut flow

cuts	$\kappa_F = +1$	$\kappa_F = -1$	$t\bar{b}\bar{b}j$	tZj	$t\bar{t}$	$\Sigma(\text{bg})$	$\frac{S}{\sqrt{B}}$ (SM)	$\frac{S}{\sqrt{B}}$ (non-SM)
initial	35	592	240	134	4491105	4492071	0.02	0.28
1 lepton	18	325	94	66	2218853	2219013	0.01	0.22
b jet no.	11	204	70	35	22984	23089	0.07	1.34
1 fwd jet	4	86	24	12	2813	2849	0.08	1.61
$m_{b\bar{b}}$	2	48	8	4	737	749	0.07	1.75

of 50 and 100 fb^{-1} , and an estimation of the significance if a 10% statistical uncertainty were assumed.

Table 2. Summary of significance calculations in the SM and non-SM case. Included are calculations for different integrated luminosities and assuming a systematic uncertainty of 10%.

	L (fb^{-1})	S/\sqrt{B}	$S/\sqrt{B + (0.1B)^2}$
$\kappa_F = +1 :$	21	0.07	0.03
	50	0.11	0.03
	100	0.16	0.03
$\kappa_F = -1 :$	21	1.75	0.60
	50	2.71	0.62
	100	3.83	0.63

For the next run of the LHC protons will be collided at 14 TeV and an integrated luminosity of 75-100 fb^{-1} is aimed for. With higher energies the cross-sections are increased which will boost the significance. However, the effects of the systematic uncertainties also increase with rising luminosity. The numbers at higher luminosities and 8 TeV cross-sections in table 2 fall well below 3σ if one includes a reasonable systematic uncertainty of 10%, indicating that even with the increased amount of data the signal events will remain undetectable for the upcoming run.

6. Conclusion

Degeneracies remain in the measurement of the fermionic and bosonic Yukawa couplings. It was studied whether the single top channel, being sensitive to the relative signs of these couplings, may be able to resolve these. A look at Monte Carlo events at truth level already suggests that a detection will not be possible. Projecting the sensitivity to higher integrated luminosities and assuming a 10% systematic uncertainty suggests that the signal yield for the upcoming run at the LHC, at higher centre-of-mass and an integrated luminosity 3-4 times greater, will be insufficient to lift the ambiguity in the sign of the couplings.

References

- [1] Collaboration A 2012 *Physics Letters B* **716** 1 – 29 ISSN 0370-2693 URL <http://www.sciencedirect.com/science/article/pii/S037026931200857X>
- [2] Collaboration C 2012 *Physics Letters B* **716** 30 – 61 ISSN 0370-2693 URL <http://www.sciencedirect.com/science/article/pii/S0370269312008581>
- [3] Farina M, Grojean C, Maltoni F, Salvioni E and Thamm A 2013 *JHEP* **1305** 022 (*Preprint* 1211.3736)
- [4] 2013 Combined coupling measurements of the Higgs-like boson with the ATLAS detector using up to 25 fb^{-1} of proton-proton collision data Tech. Rep. ATLAS-CONF-2013-034 CERN Geneva
- [5] Biswas S, Gabrielli E and Mele B 2013 *Journal of High Energy Physics* **2013** 1–16 URL <http://dx.doi.org/10.1007/JHEP01%282013%29088>

Representation of the few-group homogenized cross sections of a MOX fuel assembly

S Chifamba¹, D Botes², P M Bokov² and A Muronga¹

¹University of Johannesburg, P.O. Box 524, Auckland Park, 2006 Johannesburg, RSA

²The South African Nuclear Energy Corporation, P.O. Box 582, 0001 Pretoria, RSA

E-mail: saymorechifamba@gmail.com

Abstract. Nodal diffusion methods are often used to calculate the distribution of neutrons in a nuclear reactor core. They require few-group homogenized neutron cross sections for every heterogeneous sub-region of the core. The homogenized cross sections are pre-calculated at various reactor states and represented in a way that facilitates the reconstruction of cross sections at other possible states. In this study a number of such representations were built for the cross sections of a MOX (mixed oxide) fuel assembly via hierarchical Lagrange interpolation on Clenshaw-Curtis sparse grids. Traditionally, nodal reactor core simulators have employed cross sections with two energy groups, but there is evidence that more energy groups are needed to simulate reactor cores that contain MOX. Representations were therefore constructed for both the traditional two energy groups and a six energy group structure. Both the rate at which the representation accuracy improves with the number of samples and the complexity of the cross section dependence on individual state parameters were examined. The anisotropy feature of the representation procedure, which allows more samples to be taken for state parameters that are known to be more important to the representation accuracy than others, was applied throughout. The results show that the representation method allows both two-group and six-group cross sections to be represented in a computationally efficient manner to an industrially acceptable level of accuracy, despite additional complexity in the dependence of six-group cross sections on the state parameters.

1. Introduction

Plutonium is produced in reactors that use low-enriched uranium oxide (UOX) fuel through the transmutation of uranium. The fissile isotope $^{239}_{94}\text{Pu}$ can be recovered from spent UOX fuel and included during the fabrication of new fuel assemblies, which are then known as mixed oxide (MOX) fuel. Many pressurized water reactors (PWRs) load at least a part of the reactor core with MOX fuel, but this affects the neutron distribution in space and energy as well as the response of a reactor.

Computer models are used to calculate operational and safety parameters of a reactor core, both before start-up and during operations. These calculations must often be performed in a limited period of time, and computational efficiency is therefore an important consideration. One approach to achieve this efficiency is to use so-called *nodal methods* to numerically solve the diffusion approximation of the neutron transport equation [1]. Nodal methods use average material properties over relatively large regions in the core, and often over broad ranges of incident neutron energy. These material properties are described by cross

sections that are averaged by flux-volume weighting, which results in so-called *few-group homogenized cross sections*.

The homogenized cross sections depend on various thermo-hydraulic and material conditions that may exist in the reactor core. These are known as state parameters. The state parameters that are typically used in PWRs are: burnup, soluble boron concentration, fuel temperature, moderator temperature and moderator density [2].

Two types of homogenized cross sections are used in the full core calculations, namely microscopic and macroscopic cross sections. In this context, a macroscopic cross section is an effective cross section for several isotopes that are lumped together, which also takes into account the number density of each of these isotopes. A microscopic cross section is the probability that a given isotope will undergo a certain reaction, such as absorbing a neutron, averaged over the volume of the homogenized region and weighted by the neutron flux. The older generation of the full core simulators generally utilized macroscopic cross sections. However, there are several advantages of using microscopic homogenized cross sections for some neutronic important isotopes [3] and therefore this model is becoming more prevalent.

The few-group homogenized cross sections are pre-calculated for a limited number of reactor states and represented by mathematical functions in a way that would facilitate their reconstruction at other reactor states during the full core simulation. In the traditional methods, the homogenized cross sections are stored in multi-dimensional tables with either a linear interpolation rule, or a polynomial approximation which is usually limited to second order polynomials [4].

Nodal methods may be less accurate when applied to cores that contain MOX fuel than for cores that are loaded exclusively with UOX fuel [5]. Four ways to increase the accuracy with which both MOX and UOX fuelled cores are modelled are listed in [5], one of which is to increase the number of broad energy groups. A change in the number of broad energy groups may also impact the way in which the cross sections depend on the state parameters in each energy group and therefore the accuracy of a given representation method. The energy group structure that was used in this study is similar to the six-group structure proposed in [6] for modelling some of the important neutron-nuclear interactions for MOX fuelled cores.

The focus of this study is the representation of the few-group homogenized cross sections of a MOX fuel assembly using a sparse grid interpolation method [3, 7, 8], with the aim of limiting the cross section reconstruction errors during full core nodal calculations. This method, which uses a combination of sparse grid sampling and hierarchical polynomial interpolation with Lagrange basis functions, has been successfully applied for the few-group cross section representation of some light water reactors [3, 7, 8].

2. Problem description and methodology

This study addressed the problem of representing a few selected cross sections of a MOX fuel assembly whose specifications are contained in a Nuclear Energy Agency benchmark [9]. A model of this MOX fuel assembly was created in the HEADE (heterogeneous assembly depletion) code [10] of the OSCAR-4 (overall system for calculation of reactors, generation 4) system [11]. Using this model, several sets of the fission, nu-fission, transport and absorption homogenized cross sections were calculated at the sparse grid points in the state parameter space of the MOX fuel.

The homogenized cross sections were calculated for sixteen individual, neutronic important isotopes (namely ^{234}U , ^{235}U , ^{236}U , ^{237}U , ^{238}U , ^{237}Np , ^{239}Np , ^{238}Pu , ^{239}Pu , ^{240}Pu , ^{241}Pu , ^{242}Pu , ^{241}Am , ^{135}I , ^{135}Xe and ^{10}B) and two macroscopic cross sections. One of the macroscopic cross sections accounts for the effect of several isotopes lumped together, whilst another accounts for all the remaining isotopes from materials such as cladding and coolant. The homogenized cross sections were calculated using both two and six neutron energy group structures. For the six group cross sections, the group boundaries were set at: $1.10 \cdot 10^{-4}$ eV, $1.40 \cdot 10^{-1}$ eV, $6.25 \cdot 10^{-1}$ eV, 4.00 eV, $5.53 \cdot 10^3$ eV, $8.21 \cdot 10^5$ eV and $1.96 \cdot 10^7$ eV. The boundaries for the two group cross sections were set at $1.10 \cdot 10^{-4}$ eV, $6.25 \cdot 10^{-1}$ eV and $1.96 \cdot 10^7$ eV. The boundaries of the state parameter domain were chosen such that the cross section representation is applicable to both day to day reactor calculations and to transient analyses [3]. These intervals are listed in

Table 1. State parameter nominal conditions and boundaries.

State parameter	Nominal	Min	Max
Burnup [MWd/tU]	—	0	60 000
Moderator density [g/cm ³]	0.713	0.313	1.013
Moderator temperature [K]	579.4	279.4	979.4
Fuel temperature [K]	951.4	291.4	1 651.4
Boron concentration [ppm]	600	1	1 600

table 1.

All the cross sections were represented using hierarchical, multi-dimensional Lagrange interpolating polynomials as described in [3, 7, 8]. The interpolation was performed on several sparse grids with an increasing number of sample points in order to investigate how the representation accuracy improves with the number of samples. The representation accuracy can be characterized by the maximum relative error δ_{\max} and the average relative error δ_{mean} . For practical applications, it is required that these errors do not exceed a specified value that is considered acceptable. In this work, the upper limits for δ_{\max} and δ_{mean} were set as 0.2 % and 0.05 %, respectively. Although the representation method possesses a built-in way of accuracy assessment, in order to improve the quality of the error characterization, we estimated δ_{\max} and δ_{mean} using 4 096 independent and uniformly distributed test samples, calculated at the Sobol' quasi-random points [12].

Homogenized cross sections exhibit more complex dependence on some state parameters especially burnup [2]. Therefore in this study, both isotropic and anisotropic sparse grids were used [3, 7, 8]. An isotropic sparse grid contains the same number of samples in all dimensions. Anisotropic sparse grids are formed when any dimension in which the cross section has a prominent dependence is allowed to have more samples in order to improve the accuracy of the representation. The anisotropy of a sparse grid is described by an anisotropy vector $\alpha \in \mathbb{N}^d$, where d is the number of state parameters. The larger the value of the component α_i for a state parameter p_i ($i = 1, \dots, d$), the smaller the number of points that are sampled for that state parameter. We carried out a study to determine an optimal anisotropic vector that is suitable for the whole library. As a result of the study, the anisotropy vector $\alpha = (1, 2, 2, 2, 2)$ was chosen, where the order of the components corresponds to the order of the state parameters in table 1. This choice of α allows one to have more samples from burnup than from other state parameters.

3. Analysis of the obtained results

Constructing representations of every cross section, reaction type and energy group for all the microscopic and macroscopic materials would result in a library with a total of 152 two-group and 456 six-group cross sections. In this work we constructed representations for a subset of homogenized cross sections, which we considered to be either representative, or playing a prominent role in MOX fuel, or to be the most challenging. For the sake of brevity, the analysis in this paper is limited to three materials: a macroscopic one, which accounts for the effect of several isotopes lumped together, and two microscopic cross sections: $^{239}_{94}\text{Pu}$ and $^{238}_{92}\text{U}$. These materials are important for the operation of a reactor core that contains MOX fuel. For example, the cross sections for the macroscopic material describe the average effect of all the neutron-nuclear interactions that may occur in a reactor core. $^{238}_{92}\text{U}$ is an important resonance absorber which also has the highest number density of all the isotopes in the fuel. $^{239}_{94}\text{Pu}$ has the highest number density of all the fissile isotopes in MOX fuel.

Out of all the cross sections represented in our work, we are going to discuss the absorption cross sections. Neutron absorption is an important neutron-nuclear interaction which controls the rate of fission in a thermal reactor. The results will be presented for the thermal neutron energy groups, i.e. groups 5 and

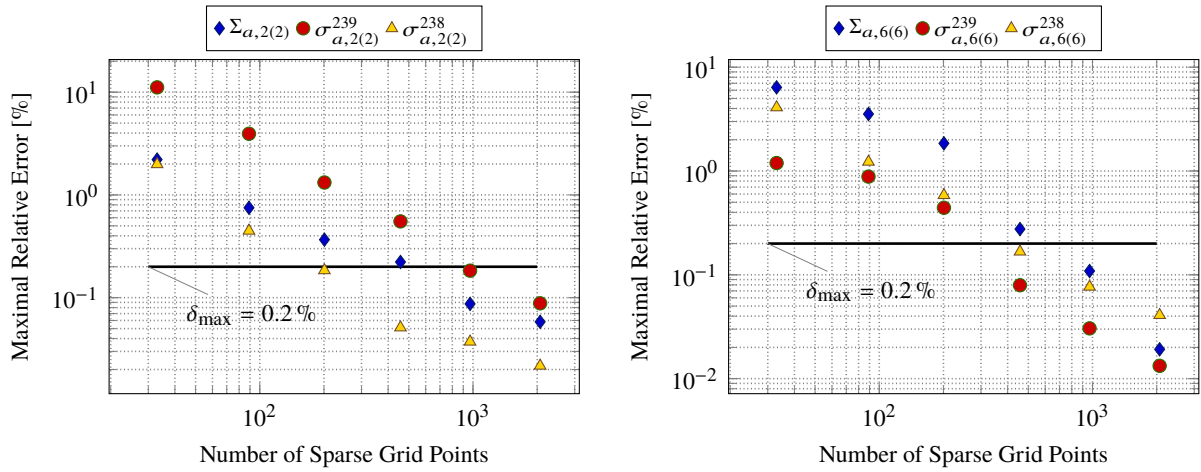


Figure 1. Dependence of the accuracy of the cross section representation on the number of sparse grid points.

6 for the six-group MOX fuel assembly model and group 2 for the two-group model. The boundaries for these groups are within the range in which nuclear fission reactions take place inside a thermal reactor core. We will denote the absorption cross sections for ^{239}Pu , ^{238}U and the macroscopic material by $\sigma_{a,g(G)}^{239}$, $\sigma_{a,g(G)}^{238}$ and $\Sigma_{a,g(G)}$ respectively, where $g(G)$ represents the energy group g for the G -group model, e.g. $\sigma_{a,5(6)}^{239}$ will be used for the group 5 microscopic cross section for the ^{239}Pu in six-group model.

Two and six-group absorption cross sections for ^{239}Pu , ^{238}U and macroscopic material were interpolated using sparse grids that were constructed hierarchically, using 33, 89, 201, 457, 969 and 2 065 points. The results for the rate at which the maximal relative error, δ_{\max} , decays with the number of sparse grid points¹ are shown in figure 1. As expected, the representation accuracy increases with the number of samples, which means that the method can be used to provide a range of representations, that are optimal in terms of balance between the accuracy and the number of samples.

Table 2 shows the number of samples that are required in order to obtain the target accuracy for each of the three cross sections for all the energy groups in both two- and six-group models. One can see from table 2 that the target accuracy is achieved with less than 1 000 samples in all the cases, which corresponds to a number of transport code runs acceptable in practice. Moreover, one can see that the thermal cross sections require more samples than the fast ones in order to achieve a target accuracy. The number of samples in a sparse grid is connected to the number of basis functions and it reflects the order of the polynomial of the representation function. The high polynomial order indicates the higher complexity of the approximation in terms of its deviation from a constant or linear shape. Therefore, numbers in table 2 demonstrate that the complexity of cross section dependencies increases from fast to thermal groups.

The complexity encountered in the construction of a representation for thermal cross sections is influenced by the degree to which the cross sections vary across the domain. Table 3 shows the magnitude of the percentage variation of the cross sections estimated with 4 096 independent, uniformly distributed test samples, normalized to their median values. In the all cases, one can see that the variation of thermal neutron cross sections is always greater than for the fast neutron cross sections.

As can be seen from figure 1, the representation accuracy for ^{238}U is smaller when the thermal homogenized cross sections are calculated using a six-group model. This pattern of results also applies to the absorption cross sections for all the neutronically important uranium isotopes. Thus, for uranium isotopes, the two-group cross sections can be approximated with lower order polynomials. This explains

¹ We report and analyze the maximal relative error only because, as our previous studies have demonstrated, it corresponds to the target accuracy, which is more difficult to meet.

Table 2. The number of sparse grid points required to obtain the target representation accuracy.

		Six-group						Two-group	
Energy group		1	2	3	4	5	6	1	2
Cross section	σ_a^{238}	89	33	89	33	457	457	89	201
	σ_a^{239}	89	89	201	89	969	457	89	969
	Σ_a	33	89	89	201	457	969	33	969

Table 3. Variation of the cross sections around the median value, in percent.

		Six-group						Two-group	
Energy group		1	2	3	4	5	6	1	2
Cross section	σ_a^{238}	11.8	1.6	45.0	5.0	21.5	27.2	12.8	30.4
	σ_a^{239}	4.8	0.4	31.6	19.5	56.7	11.8	13.4	120.3
	Σ_a	21.7	10.9	38.5	67.8	61.7	53.2	20.5	72.7

why the traditional methods were successful in representing two-group cross sections for UOX fuel.

One way to explain this shift in complexity is by examining the cross section dependencies on the individual state parameters. This may be achieved by constructing scatter plots from samples (the sparse grid or the independent test set) as a function of each of the five state parameters. Our analysis of such graphs proves that the variation of $\sigma_{a,6(6)}^{238}$ is dominated by the non-linear dependence on the burnup and the moderator temperature, while $\sigma_{a,2(2)}^{238}$ shows a strong dependence on fuel temperature and moderator density. This observation can be understood by considering that the calculation of the homogenized cross sections involves neutron flux weighting. However, the neutron flux in a reactor core is dependent on the state parameters. Changes in the values of the state parameters affect the neutron distribution in different energy groups differently. These variations result in different values of the neutron flux that are used for the weighting process.

Fewer samples are required to reach the target accuracy for $\sigma_{a,6(6)}^{239}$ than for $\sigma_{a,2(2)}^{239}$. However, the complexity of representing σ_a^{239} for the other thermal energy group, i.e. group 5, is comparable with that of $\sigma_{a,2(2)}^{239}$. This can be understood by considering that the cross section σ_a^{239} has a large resonance in the energy range $1.40 \cdot 10^{-1}$ eV to $6.25 \cdot 10^{-1}$ eV which coincides with group 5 for the six-group model. As a result, $\sigma_{a,5(6)}^{239}$ has a higher percentage variation across the state parameter space as compared to $\sigma_{a,6(6)}^{239}$. Therefore, for the six-group model, the representation error is isolated into a smaller energy range.

The analysis of thermal reactors that are loaded with MOX fuel can be done with better accuracy if the cross sections are calculated using an energy group structure that removes the complexity of the cross section dependencies away from the energy range in which fission reactions are most likely to take place in plutonium. This is done by increasing the number of energy groups. Whilst an increase in the number of energy groups yields better results for plutonium, this does not hold true for other isotopes. However, the sparse grid method is still capable of representing these cross sections to an acceptable accuracy.

4. Conclusion

The paper addressed the problem of representing a few selected few-group homogenized neutron cross sections of a MOX fuel assembly using the sparse grid method. This method utilizes a combination of

sparse grid sampling and hierarchical polynomial interpolation with multivariate Lagrange basis functions.

In order to improve the efficiency of the cross section representation, a mini-study with a goal to find an optimal (i.e. which can provide the best accuracy for a given number of samples for all cross sections in the library) sparse grid anisotropy vector has been performed. A candidate vector with an equally reduced number of samples for all the state parameters, except the burnup, was identified and used through the rest of this work. Further studies has confirmed that, though different cross sections can exhibit a behavior dominated by different state parameters, the burnup is a state parameter, importance of which is shared by all the homogenized cross sections.

The application of the method to the representation of several two- and six-group cross sections has produced excellent results. Results for the microscopic absorption cross sections for $^{239}_{94}\text{Pu}$ and $^{238}_{92}\text{U}$ and for the macroscopic absorption cross sections are reported and analyzed in this paper. The desired representation accuracy, $\delta_{\max} = 0.2\%$, was achieved with less than 1 000 samples which make the method suitable for practical application. These results are impressive if we consider that the cross sections used in this work were calculated using intervals of state parameters that are suitable for transient analysis. This means that the cross sections we interpolated have a more complex dependence on the state parameters than the cross sections that are usually used for the normal day-to-day reactor operations.

The primary fissile isotope in MOX fuel, $^{239}_{94}\text{Pu}$, poses a problem when the two-group model is used: the accuracy of the representation of its thermal group cross section is exceptionally low. Nevertheless, using the six-group model allows one to palliate the problem by isolating the complexity of the dependence (and the corresponding representation error) in group 5, thus improving the representation accuracy of the sixth group, where the nuclear fission reactions are most likely to take place. This does not address all the limitations of nodal diffusion methods when simulating MOX-fuelled PWR cores, which is an area of study that spans much wider than just cross section representation.

Acknowledgments

The financial assistance of the South African National Research Foundation (NRF) towards this study through an Innovation Master's Scholarship that was awarded to the first author (Grant 89093) and the through Incentive Funding, awarded to the third author (Grant 70730), is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the NRF.

References

- [1] Lawrence R D 1986 *Progress in Nuclear Energy* **17** 271–301
- [2] Zimin V G and Semenov A A 2005 *Annals of Nuclear Energy* **32** 119–136
- [3] Botes D 2012 *Few-group Cross Section Representation Based on Sparse Grid Methods* Master's thesis Potchefstroom Campus of the North-West University, South Africa
- [4] Dufek J 2011 *Annals of Nuclear Energy* **38** 1569–1577
- [5] Downar T, Lee C H and Jiang G 2000 An Assessment of Advanced Nodal Methods for MOX Fuel Analysis in Light Water Reactors *PHYSOR'2000* (Pittsburgh, PA, USA)
- [6] Müller E Z, Joubert W R and Reitsma F 1994 An Energy Group Structure Study for PWR and PWR/MOX Cores Tech. Rep. PIN 1334 (B/R) Radiation and Reactor Theory, Atomic Energy Corporation, South Africa
- [7] Bokov P M, Botes D and Zimin V G 2012 Pseudo-spectral Chebyshev Representation of Few-group Cross Sections on Sparse Grids *PHYSOR 2012* (Knoxville, Tennessee, USA)
- [8] Botes D and Bokov P M 2014 *Annals of Nuclear Energy* **64** 156
- [9] Kozlowski T and Downar T J 2006 PWR MOX/UO₂ Core Transient Benchmark: Final Report Tech. Rep. NEA/NSC/DOC(2006)20 Nuclear Energy Agency
- [10] Reitsma F and Joubert W R 1999 A Calculational System to Aid Economical Use of MTRs *International Conference on Research Reactor Fuel Management (RRFM'99)* (Brugges, Belgium)
- [11] Müller E Z, Joubert W R, Schutte H C, Stoker C C and Reitsma F 1994 Development of a Core Follow Calculational System for Research Reactors *9th Pacific Basin Nuclear Conference* (Sydney, Australia)
- [12] Sobol' I M 1967 *Zhurnal Vychislitelnoi Matematiki i Matematicheskoi Fiziki* **7** 784–802

Towards a crystal undulator

S H Connell¹, J H"artwig², A Masvaure¹, D Mavunda^{1,3} and T N Tran Thi¹

¹ University of Johannesburg, Johannesburg, South Africa

² European Synchrotron Radiation Facility (ESRF), Grenoble, France

³ Nuclear Energy Corporation of South Africa (Necsa), South Africa

E-mail: shconnell@uj.ac.za

Abstract. The CUTE FP7 project proposes to produce ultimately a MeV range gamma ray laser by the FEL principal in a crystal undulator. The GeV range electron beam would need to be captured in a low index crystallographic channel of a crystal superlattice, in such a way that the varying electrostatic crystalline field would resemble a Tesla range periodically varying magnetic field with a few micron pitch, when viewed in the reference frame of the undulating electron or positron. We have investigated a prototype diamond superlattice using x-ray diffraction topography. The undulator fabrication principle involved CVD growth of diamond on a diamond substrate while varying the concentration of boron in the gas phase during growth. This should lead to the periodic variation of the lattice dilatation by the varying concentration of the single substitutional boron impurity atom. The validation via x-ray diffraction topography proved non-trivial but was eventually promising.

1. Introduction

A crystal undulator is similar to a normal undulator as typically found at a synchrotron for the production of extremely brilliant X-ray beams. The difference is the magnetic lattice is realized by the periodic electrostatic potential of a crystal lattice seen from the reference frame of the GeV range electron or positron beam. The extremely relativistic incident particle beam would need to be captured in a high index crystallographic channel of a crystal superlattice. The particle beam will then see a many Tesla range periodically varying magnetic field with a micron scale pitch. This method could theoretically lead to an MeV range gamma ray laser by the FEL principal.

We have investigated a prototype diamond superlattice using x-ray diffraction topography at the ESRF. The undulator fabrication principle involved CVD growth of boron doped diamond on a diamond substrate while varying the concentration of boron in the gas phase during growth. This should lead to the periodic variation of the lattice dilatation corresponding to the varying concentration of the single substitutional boron impurity atom. This is the realisation of the diamond superlattice, which has periodic layers of graded composition, coherently registered on each other, with a smoothly varying periodic lattice constant.

2. Parameters for the crystal undulator and its realisation

The theoretical feasibility study to produce powerful monochromatic undulator radiation in the gamma ray region by means of a crystal undulator (periodically bent crystal) and

ultra relativistic light lepton beams impinging on the lattice with aligned incidence to a crystallographic axis is presented in several references [1, 2, 3] and references therein.

Figure 1 indicates the trajectory of a charged particle captured in the channeling condition in a low index crystallographic direction. The channelled trajectory has an oscillation within the channel, leading to the emission of lower energy channeling radiation. In addition the particle follows the periodic undulations of the lattice, and as such, it will also emit MeV range undulator radiation.

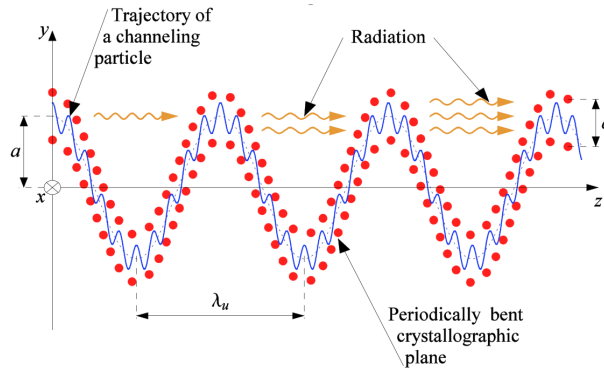


Figure 1. A charged particle captured in a undulating crystal channel, indicated by the periodically bent lattice of dots.

The parameters of the crystal undulator are typically in the range shown in table 1, where an important condition is $\lambda_u \gg a \gg d$.

Table 1. Typical physical parameters for a crystal undulator.

Undulator wavelength = λ_u	≈ 0.1 mm
Undulator amplitude = a	≈ 50 Å
Interplanar distance = d	1 – 2 Å
Crystal thickness = t	1 – 4 mm
Number of oscillations = $N_u = t/\lambda_u$	10

These parameters are obtained considering the energy and particle type dependent dechanneling length and the beam and crystal physical properties. Diamond suggests itself as a premier material for a crystal undulator for several reasons. The lattice is extremely radiation hard (it sustains 1×10^{15} mips/cm²). This consideration is of special importance for a material that could conceivably withstand the beam bunch intensities for the SASE (Self Amplified Spontaneous Emission) version of the crystalline undulator. The high Debye temperature results in diminished latticed vibrations, which increases the dechanneling length. This also enhances the coherence length for phenomena which are associated with radiation emission. The low atomic number leads to a lower channelling potential but this is offset somewhat by the very high atomic packing density. The $\langle 110 \rangle$ channelling direction in diamond is particularly favourable for the channelling of positive particles, as both the core spatial distribution and the electron spatial distribution contribute to a deep potential well which is well separated from sources of hard scattering.

Previous attempts to fabricate a diamond crystal undulator deployed the permanent periodic deformation of the diamond by inducing a periodic surface strain using a laser scoring procedure

[4]. As the surface strain relaxes towards the bulk, this method produced a distribution of undulator amplitudes. The monochromaticity of the energy of the undulator radiation depends on all transmitted channeled particles experiencing the same undulator amplitude, and so the surface deformation technique is not suitable [5].

This has led to the need for the development of a graded superlattice, as an alternative method to produce a crystal undulator.

3. The diamond graded superlattice

Doping diamond with single substitutional atoms of either nitrogen or boron dilates the lattice. Both of these atoms are soluble in diamond, and can be introduced during the growth of synthetic diamond in the CVD process in a regulated way. It is in fact the case that one must go to considerable lengths to exclude these dopants should one desire intrinsic diamond (considered as having impurity concentrations in the low ppb range). The lattice dilatation for nitrogen [6] (single substitutional) is

$$\frac{\Delta a}{a_0} = 0.125 \pm 0.006 \times C_N,$$

while that for boron [7, 8, 9] (single substitutional) is

$$\frac{\Delta a}{a_0} = 0.144 \times C_B.$$

C_N and C_B are the atomic concentrations of nitrogen and boron respectively as fractions. For the purpose of deliberate lattice dilation under conditions of minimising extended defects and more complex point defects (dislocations, stacking faults, interface coherence, aggregation) boron is favoured. Experience has shown that one may obtain even heavily doped boron concentrations of $C_B < 1.5$ at% while still maintaining excellent local lattice quality. In exceptional cases, the interface definition can be on the nanometric, near atomic, scale [10].

The undulator amplitude to be achieved by a dilatation at each lattice site over the undulator period (see table 1) requires

$$n\Delta a = \left(\frac{d}{a_0}\right) \Delta a = 50\text{\AA}.$$

Using the dilatation formula for boron incorporation (above) and considering a 45° channeling trajectory with respect to layers of the superlattice, we find a maximum doping level of $C_B = 1000$ ppm is required. Accordingly, Element Six Technologies prepared the samples as indicated in figure 2 below, as a pilot study of the graded diamond super lattice principle. The substrates for epitaxial CVD overgrowth are High Pressure High Temperature (HPHT) synthetic diamond type Ib. It is well known that the substrate lattice quality affects the lattice quality of the CVD layer [11]. In this case no special attention was made to select high quality substrates, as this was a first proof of principle investigation. This could be seen in this study as the typical clusters of dislocations threading through the CVD layer, originating at imperfections in the substrate. A future study would select for substrates which were clear of extended defects, as was done in the reference [11].

4. Characterisation with X-ray diffraction techniques

The graded boron doped layer was studied with various X-ray diffraction techniques at the ESRF in Grenoble, beam line BM05. The X-ray photon energy was 20 keV and the reflection studied was the (400) reflection with a Bragg angle of $\theta_B = 14.23^\circ$. This study was rather difficult, as the thickness of the layer falls between two techniques for investigation of this type of local lattice parameter changes. On the smaller scale, there are interferometric techniques, and on

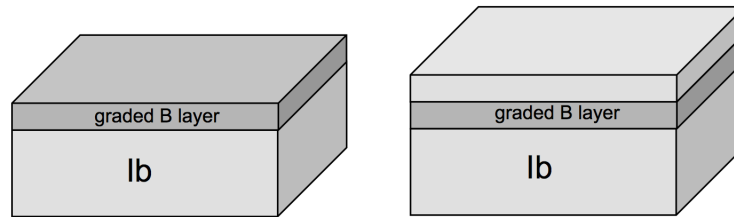


Figure 2. 50 μm thick graded B-doped layer (CVD) on a lb diamond substrate (side 4 mm). The layer profile is triangular, extending to 3000 ppm. The sample on the right is capped with an additional "intrinsic" layer of diamond ($C_B < 50$ ppm).

the larger scale there are differential sectioning techniques. Considerable effort has therefore been expended to study these graded B-doped layers.

We report here on the investigations using plane wave monochromatic diffraction. This means the incident beam was highly collimated and monochromatic, essentially with an effective energy width which is larger but comparable to the acceptance of the actual sample crystal and reflection used. The theoretical reflectivity curve of the Si-(111) monochromator is about 4" and that for the C^{*}-(400) diamond sample is about 0.75". Furthermore, the configuration was in the dispersive mode, which arises because the monochromating crystal and reflection were different to that of the sample crystal and reflection, and neither crystal was bent. This means that the Bragg condition is typically satisfied for only a given spatial range of the sample crystal, where this spatial range is a narrow band perpendicular to the incident direction of the beam. The image selected in figure 3 below is an integral of an angular rocking curve scan, so that a full rocking curve is integrated for each pixel on the sample surface.

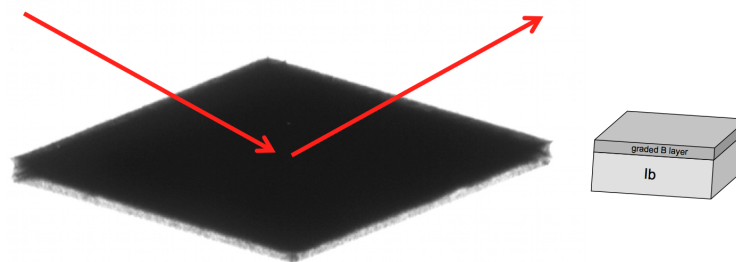


Figure 3. An integral of a rocking curve scan for the (400) reflection in the Bragg case in plane wave monochromatic diffraction.

The diffraction vector is parallel to the maximal change of the lattice constant in the graded doped layer. The geometry chosen presents a perspective view of the diamond crystal. The penetration depth of the x-ray beam is a complex mix of effects related to the angular and energy dispersion of the beam, the quality of the crystal, and the angle at which the rocking curve is taken. As discussed, in this case, the image presented is the integral of the full rocking curves for each pixel of the sample surface. The reflectivity of the graded doped layer appears significant over the whole crystal surface. It is interesting to first evaluate that the graded layer has a spread of lattice parameters (due to the range of dilatation) that corresponds to a spread of effective local diffracting angles, $\Delta d/d \approx 10^{-4}$ corresponding to $\Delta\theta \approx 21 \gg \theta_D$, the theoretical reflectivity curve for diamond, the Darwin width. This is the condition for multi-layer mirror

behaviour. There is a continuous variation of the lattice constant over a relatively shallow depth in such a way that the acceptance of the crystal in the graded doped layer is much wider than the reflectivity curve of the interrogating beam. This means that the graded doped layer can yield an integral reflectivity which is much larger than the substrate of higher lattice quality diamond. As we can see in the perspective view, where the edge is not screened by the surface layer, one can separate the graded layer and the substrate. Within the substrate, one still has the usual view of the defects within. This is because the defects once again create a local strain which also broadens the local reflectivity curve. What is important here is that we see clear evidence for successful implementation of the graded B-doped layer.

In the next figure 4, a plane wave monochromatic dispersive case topograph at just a single point on the flank of the rocking curve is presented. The geometry used optimises the view of the sample edge. It was taken with a shield over the diamond main surface to reduce the Laue case component of the diffraction and emphasise the more surface sensitive Bragg case diffraction. A component of the strain due to the dilatation should still be visible, and this seems to be the case. The analysis of the image shows a variation of the local Bragg angle across the edge where the graded B-doped layer is visible.

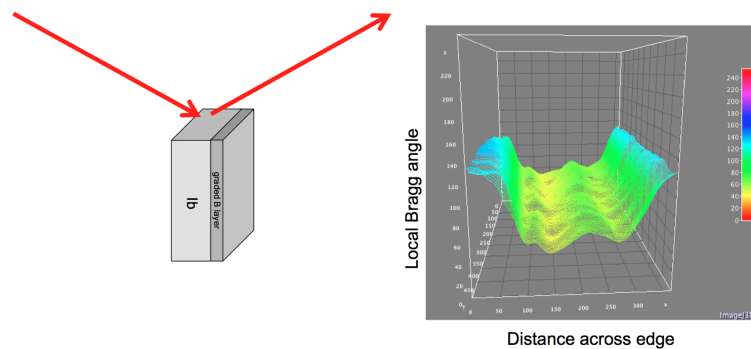


Figure 4. A plane wave monochromatic dispersive case topograph at just a single point on the flank of the rocking curve is presented with a view of the sample edge.

5. Conclusion

This paper has reported on a pilot study to produce and characterise graded B-doped superlayers. This has proved to be a non-trivial task due to the length scale of the layers. The pitch of the layer is much larger than the nano-scale, so the interference effects from Bragg scattering normal to the layer are not visible. The pitch of the layer is nonetheless somewhat small to use collimation in the entrance and exit channel to section the diffracting volume within the crystal. A large data set of both Bragg and Laue condition Rocking Curve Imaging measurements and Plane Wave Topography measurements indicated that indeed the lattice parameter variation for the superlattice could be detected. Two examples have been presented in this paper. Indeed, the lattice parameter evolves over the depth of the graded layer. The graded B-doped layer acts as an effective multi-layer mirror. The studies are continuing. It would be preferable to develop the graded layer on the best possible (lattice) quality substrates. This is the start of the iterative process to produce a diamond crystal undulator, for the first time, and the results are promising.

References

- [1] Tabrizi M, Korol A V, Solovyov A V and Greiner W 2007 *Physical Review Letters* **98** 164801
- [2] Korol A V, Solovyov A V and Greiner W 1998 *Journal of Physics G* **24** L45

- [3] Korol A V, Solovyov A V and Greiner W 1999 *International Journal of Modern Physics* **E8** 49
- [4] Balling P, Esberg J, Kirsebom K, Le D Q S, Uggerhoj U I, Connell S, Haertwig J, Masiello F and Rommeveaux A 2009 *Nuclear Instruments and Methods in Physics Research B* **267** 29522957
- [5] Kostyuk A, Korol A, Solovyov A and Greiner W 2009 *Nuclear Instruments and Methods in Physics Research B* **266** 972987
- [6] Bragg W and Bragg W 1913 *Proceedings of the Royal Society (London)* **89A** 277–290
- [7] Brunet F, Germin P, Pernet M, Deneuve A, Gheeraert E, Laugier F, Burdin M and Rolland G 1998 *Diamond and Related Materials* **7** 869–873
- [8] Brazhkin V V, Ekimov E A, Lyapin A G, Popova S V, Rakhmanina A V, Stishov S M, Lebedev V M, Katayama Y and Kato K 2006 *Physical Review B* **74** 140502
- [9] Wojewoda T, Achatz P, Ortéga L, Omnés F, Marcenat C, Bourgeois E, Blase X, Jomard F and Bustarret E 2008 *Diamond and Related Materials* **17** 1302–1306
- [10] El-Hajj H, Denisenko A, Bergmaier A, Dollinger G, Kubovic M and Kohn E 2008 *Diamond and Related Materials* **17** 409–414
- [11] Martineau P M, Gaukroger M P, Guy K B, Lawson S C, Twitchen D J, Friel I, Hansen J O, Summerton G C, Addison T P G and Burns R 2009 *Journal of Physics: Condens. Matter* **21** 364205

The Development of a General Purpose Processing Unit for the Upgraded Electronics of the ATLAS Tile Calorimeter

Mitchell A. Cox and Bruce Mellado

School of Physics, University of the Witwatersrand. 1 Jan Smuts Avenue, Braamfontein, Johannesburg, South Africa, 2000

E-mail: mitchell.cox@students.wits.ac.za

Abstract. The Large Hadron Collider at CERN generates enormous amounts of raw data which presents a serious computing challenge. After planned upgrades in 2022, the data output from the ATLAS Tile Calorimeter will increase by 200 times to over 40 Tb/s! System on Chips (SoC) such as ARM and Intel Atom are common in mobile devices due to their low cost, low energy consumption and high performance. It is proposed that a cost-effective, high data throughput Processing Unit (PU) can be developed by using several consumer SoCs in a cluster configuration to allow aggregated processing performance and data throughput while maintaining minimal software design difficulty for the end-user. This PU could be used for a variety of high-level functions on the high-throughput raw data such as spectral analysis and histograms to detect possible issues in the detector at a low level. High-throughput I/O interfaces are not typical in consumer SoCs but high data throughput capabilities of greater than 20 Gb/s per PU is feasible via the novel use of PCI-Express as the I/O interface to the SoCs. An overview of the PU is given and the results of throughput testing of Freescale Semiconductor i.MX6 quad-core ARM Cortex-A9 processors are presented.

1. Introduction

Projects such as the Large Hadron Collider (LHC) generate enormous amounts of raw data which presents a serious computing challenge. After planned upgrades in 2022, the data output from the ATLAS Tile Calorimeter (TileCal) will increase by 200 times to over 40 Tb/s (Terabits/s) [1]. It is infeasible to store this data for offline computation.

A paradigm shift is necessary to deal with these future workloads and the cost, energy efficiency, processing performance and I/O throughput of the computing system to achieve this task are vitally important to the success of future big science projects. Current x86-based microprocessors such as those commonly found in personal computers and servers are biased towards processing performance and not I/O throughput and are therefore less-suitable for high data throughput applications otherwise known as Data Stream Computing [2].

ARM System on Chips (SoCs) are found in almost all mobile devices due to their low energy consumption, high performance and low cost [3]. One of the first steps to a true Data Stream Computing system is a high data throughput Processing Unit (PU). The author is developing an ARM-based PU for use by ATLAS TileCal as a high throughput, general purpose co-processor to the read-out system Super Read Out Driver (sROD) which will be used to combat the issue of

pile-up. A general purpose co-processor is able to run more sophisticated and memory intensive algorithms than the FPGA-based sROD, although the latency is inferior which is why the sROD is used in the data path.

A brief discussion of the ATLAS TileCal read out architecture and the sROD is given in Section 2. An overview of the energy reconstruction methodology for Photo Multiplier Tubes (PMTs) is provided in Section 3. Possible issues with the current methodology is also given. The Processing Unit (PU) is described in Section 4 along with how it could be used to help solve the issues with the PMT energy reconstruction algorithms. Section 5 concludes with a brief discussion of future work.

2. TileCal Read Out Architecture

The TileCal read out architecture is required to digitize the analog signals produced by the PMTs located on the Tile Calorimeter. In the existing hardware, shown in Fig. 1, analog circuitry is used extensively. At the time of design and construction, digital electronics were not fast enough to satisfy the requirements. In the upgraded system, shown in Fig. 2, the digital electronics have been superseded by the sROD which is based on high-end Field Programmable Gate Arrays (FPGAs). There is also an upgraded front-end which also contains FGPA's for digitising the analog signals.

The sROD is located in the back-end, off the detector to avoid the requirement for expensive radiation-hard electronics. A photo of the prototype sROD circuit board before assembly is shown in Fig. 3. The sROD will be located in an industry standard AdvancedTCA (ATCA) chassis (seen in Fig. 3) which enables comprehensive redundancy and monitoring to ensure maximum uptime.

In both the existing and the upgraded systems, a pipeline is used to store events until the level one trigger provides an accept signal. This short delay is required while the level one trigger performs computations. In the upgraded system the sROD is able to perform some calculations before sending data to the rest of the triggering and data acquisition system. Some of these sROD calculations will be explained in Section 3. A general purpose Processing Unit can be used to enhance this functionality.

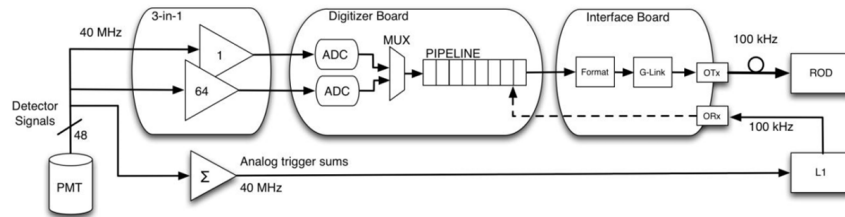


Figure 1: ATLAS TileCal current read out architecture [1].

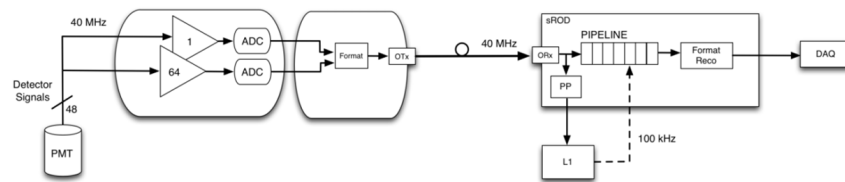


Figure 2: ATLAS TileCal upgraded read out architecture showing the sROD [1].

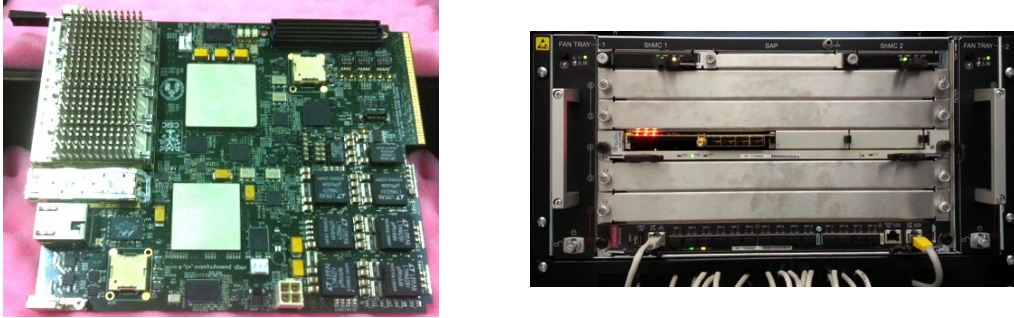


Figure 3: Photo of the sROD (left) and AdvancedTCA chassis (right).

3. TileCal Energy Reconstruction

Bunches Crossings (BC) in the ATLAS detector happen at a rate of 40 MHz. In reality there can be over 20 separate collisions in a single bunch crossing which leads to pile-up or Out Of Time (OOT) signals in the detector. The Tile Calorimeter is made up by layers of scintillator and steel. When a particle interacts with the scintillator, a pulse of light is produced which is converted to an electrical signal by PMTs.

This electrical signal is conditioned and spread into a pulse with a length of 150 ns and full width at half maximum of 50 ns. An Analog to Digital Converter (ADC) samples this pulse every 25 ns resulting in seven samples per pulse. A reference pulse showing example sampling points is visible in Fig. 4 with three main parameters of the pulse also illustrated.

Optimal Filtering (OF) or a Matched Filter (MF) are two methods by which the amplitude, A , phase, τ , and base-line pedestal, p , parameters can be calculated from the seven ADC samples of a pulse [4]. The pulse shape and therefore the energy can be reconstructed, when required, from these three parameters [5].

For both the OF and MF algorithms, each parameter (A , τ or p) can be found by multiplying the ADC samples by a specific set of weights which are calculated ahead of time. Both algorithms work well in low luminosity operation where the background noise of the PMT signals is gaussian and uncorrelated. This assumption fails for high luminosity operation (above about $\sqrt{s} = 8$ TeV) where the background noise is no longer uncorrelated due to pile-up [4]. This effect is visible in Fig. 4 where the algorithm energy calculations are not correlated with multiple OOT signals as well as the negative energies produced by the OF algorithm.

3.1. Energy Reconstruction Supervision

As discussed in Section 3, the energy reconstruction algorithms fail to work as expected under high pile-up conditions caused by higher luminosity operation of the LHC. A general purpose PU, as described in Section 4, can be used to alleviate this issue through the use of higher level programming of more sophisticated algorithms and the availability of more memory.

The PU is unlikely to operate within the latency constraints of the level one trigger and sROD and therefore cannot be in the critical data path. It can, however, be used to plot a histogram of the raw data and run its own analysis to determine if the algorithms on the sROD are functioning correctly. If they are not, the operators or engineers will be notified. This analysis is very difficult without the use of a PU. It is feasible that the PU can also be used to recompute the weights used by the OF or MF algorithms when an issue is detected which would greatly enhance the precision and reliability of the TileCal.

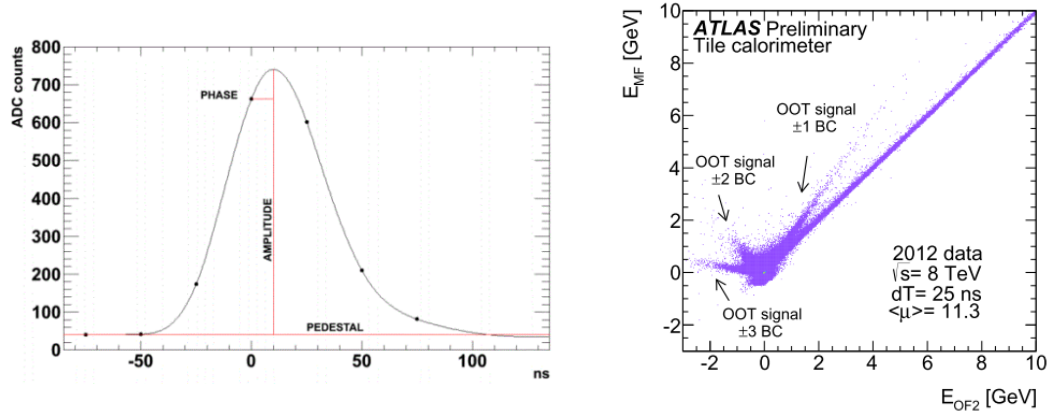


Figure 4: TileCal reference pulse from the PMT showing 7 sample points (left) and the energy reconstruction mismatch between the OF and MF algorithms under higher luminosity (right) [4].

4. Processing Unit

ARM System on Chips (SoCs) are low cost, energy efficient and high performance which has led to their extensive use in mobile devices. ARM performance and energy efficiency results have been published previously [2]. An ARM-based Processing Unit (PU) is under development to complement the sROD with higher-level computational tasks such as those discussed in Section 3.1.

The completed PU will be located in the ATCA chassis on an Advanced Mezzanine Card next to the sROD or as a separate board connected to the back-plane. The PU will be able to process up to 40 Gb/s raw data, fed via the ATCA from the sROD. A PCI-Express I/O interface will be used to link the FPGA on the sROD to a cluster of ARM SoCs on the PU.

PCI-Express throughput tests have been performed on a pair of Freescale i.MX6 quad-core ARM Cortex-A9 SoCs clocked at 1 GHz, located on Wandboard development boards [6]. The results are presented in Tab. 1 and a photo of the custom test setup designed by the author is in Fig. 5. Three tests were run to ascertain the maximum data throughput that can be obtained from the i.MX6 SoC: a simple CPU based memcpy command and two Direct Memory Access (DMA) transfers, initiated by the Endpoint (EP) or slave and the Root Complex (RC) which is the host.

Table 1: PCI-Express throughput results of a i.MX6 pair.

	CPU memcpy	DMA (EP)	DMA (RC)
Read (MB/s)	$94.8 \pm 1.1\%$	$174.1 \pm 0.3\%$	$236.4 \pm 0.2\%$
Write (MB/s)	$283.3 \pm 0.3\%$	$352.2 \pm 0.3\%$	$357.9 \pm 0.4\%$

The theoretical maximum throughput for the PCI-Express Gen 2 x1 link that was used is 500 MB/s. The best result is using DMA initiated by the RC but it is only 72% of the theoretical maximum. The RC-mode drivers are more optimized than the EP-mode drivers due to limited manufacturer support for EP-mode. The read results are lower than write because of overheads to initiate the read. The PU architecture will take these differences into account and use a data push rather than a pull based approach.

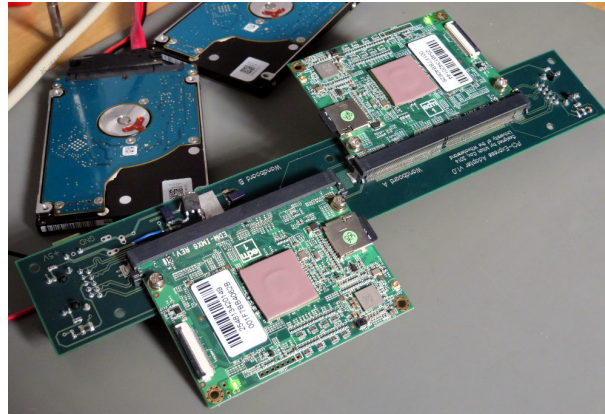


Figure 5: PCI-Express test setup for a pair of i.MX6 SoCs.

5. Discussion, Conclusions and Future Work

High data throughput computing, or Data Stream Computing, is required for projects such as the LHC which produce enormous amounts of raw data. A general purpose ARM System on Chip based processing unit is being developed which will be used as a co-processor to the sROD to help mitigate the energy reconstruction issues caused by pile-up under higher luminosity (over $\sqrt{s} = 8$ TeV) operation of the LHC.

A PCI-Express interface will be used for the raw data transfer between the sROD and the PU. Initial throughput measurements presented for a pair of Freescale i.MX6 quad-core Cortex-A9 SoCs are 72% of the theoretical maximum 500 MB/s for the available x1 link. Twelve of these SoCs will therefore be connected in parallel to provide the required 40 Gb/s throughput.

The next stage of research by the author is to measure the data throughput through a PCI-Express switch with up to eight i.MX6 SoCs connected. This will enable a close prototype of the final PU design to be tested.

Acknowledgements

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. We would also like to acknowledge the School of Physics, the Faculty of Science and the Research Office at the University of the Witwatersrand, Johannesburg.

References

- [1] F Carrió *et al.*, *The sROD Module for the ATLAS Tile Calorimeter Phase-II Upgrade Demonstrator*, Meyrin, Switzerland, 2013. [Online]. Available: <http://cds.cern.ch/record/1628753?ln=en>.
- [2] M. A. Cox, R. Reed, T. Wrigley, G. Harmsen, and B. Mellado, "Performance Characterisation of ARM Cortex-A7, A9 and A15 System on Chips for Data Stream Computing," *In Review*, 2014.
- [3] T. Krazit, *ARMed for the living room*, 2006. [Online]. Available: http://news.cnet.com/ARMed-for-the-living-room/2100-1006_3-6056729.html.
- [4] B. S. Peralva, "The Tilecal Energy Reconstruction for Collision Data Using the Matched Filter," *ATLAS TileCal*, 2013. [Online]. Available: <https://cds.cern.ch/record/1629575/files/ATL-TILECAL-PROC-2013-023.pdf>.
- [5] E. Fullana *et al.*, "Digital Signal Reconstruction in the ATLAS Hadronic Tile Calorimeter," *IEEE Transactions on Nuclear Science*, vol. 53, no. 4, pp. 2139–2143, Aug. 2006, ISSN: 0018-9499. DOI: 10.1109/TNS.2006.877267. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1684077>.
- [6] Wandboard.org, *Wandboard - Freescale i.MX6 ARM Cortex-A9 Opensource Community Development Board*, 2012. [Online]. Available: <http://www.wandboard.org/>.

Affordable and power efficient computing for high energy physics: CPU and FFT benchmarks of ARM processors

Mitchell A Cox, Robert Reed and Bruce Mellado

School of Physics, University of the Witwatersrand. 1 Jan Smuts Avenue, Braamfontein, Johannesburg, South Africa, 2000

E-mail: `mitchell.cox@students.wits.ac.za`

Abstract. Projects such as the Large Hadron Collider at CERN generate enormous amounts of raw data which presents a serious computing challenge. After planned upgrades in 2022, the data output from the ATLAS Tile Calorimeter will increase by 200 times to over 40 Tb/s. ARM System on Chips are common in mobile devices due to their low cost, low energy consumption and high performance and may be an affordable alternative to standard x86 based servers where massive parallelism is required. High Performance Linpack and CoreMark benchmark applications are used to test ARM Cortex-A7, A9 and A15 System on Chips CPU performance while their power consumption is measured. In addition to synthetic benchmarking, the FFTW library is used to test the single precision Fast Fourier Transform (FFT) performance of the ARM processors and the results obtained are converted to theoretical data throughputs for a range of FFT lengths. These results can be used to assist with specifying ARM rather than x86-based compute farms for upgrades and upcoming scientific projects.

1. Introduction

Projects such as the Large Hadron Collider (LHC) generate enormous amounts of raw data which presents a serious computing challenge. After planned upgrades in 2022, the data output from the ATLAS Tile Calorimeter will increase by 200 times to over 41 Tb/s (Terabits/s) [1]. It is not feasible to store this data for offline computation.

A paradigm shift is necessary to deal with these future workloads and the cost, energy efficiency, processing performance and I/O throughput of the computing system to achieve this task are vitally important to the success of future big science projects. Current x86-based microprocessors such as those commonly found in personal computers and servers are biased towards processing performance and not I/O throughput and are therefore less-suitable for high data throughput applications otherwise known as Data Stream Computing [2].

ARM System on Chips (SoCs) are found in almost all mobile devices due to their low energy consumption, high performance and low cost [3]. The modern ARM Cortex-A range of processors have 32- and 64-bit cores and clock speeds of up to 2.5 GHz, making them potential alternatives to common x86 CPUs for scientific computing. This paper presents benchmarks of three common ARM Cortex CPUs, namely the Cortex-A7, A9 and A15 with more information on these platforms in Table 1. The benchmark results are useful for specifying an ARM-based system in new scientific projects such as ATLAS read-out and trigger system upgrades.

A brief discussion of the ATLAS Triggering and Data Acquisition System (TDAQ) and where ARM processors may potentially be used is presented in Section 2. CPU benchmark results are given in Section 3. Fast Fourier Transform (FFT) benchmarks are in Section 4. Section 5 concludes with a brief discussion of future work.

Table 1: Specifications and other details of the ARM platforms used.

	Cortex-A7	Cortex-A9	Cortex-A15
Platform	Cubieboard A20	Wandboard Quad	ODROID-XU+E
SoC	Allwinner A20	Freescale i.MX6Q	Samsung 5410
Cores	2	4	4 (+ 4 Cortex-A7)
Max. CPU Clock (MHz)	1008	996	1600
L2 Cache (kB)	256	1024	2048
Floating Point Unit	VFPv4 + NEONv2	VFPv3 + NEON	VFPv4 + NEONv2
RAM (MB)	1024	2048	2048
RAM Type	432 MHz 32 bit	528 MHz 64 bit	800 MHz 64 bit
2014 Retail (USD)	65	129	169
Linux Kernel	3.4.61	3.10.17	3.4.5
GCC	4.7.1	4.7.3	4.7.3

2. ATLAS Triggering and Data Acquisition System

The ATLAS experiment is composed of several sub-detectors, each of which has separate data processing requirements. The massive amount of raw data is reduced by a process called triggering. In the Tile Calorimeter, there are currently three main levels of triggering, shown in Figure 1. The read-out system is based on FPGAs (Field Programmable Gate Arrays) and DSPs (Digital Signal Processors) to form the level one trigger which serves to reduce the data rate (event rate) from 40 MHz to about 100 kHz. Each ATLAS event consists of about 1.5 MB data, of which a portion is made up from TileCal data. Some components of the read-out subsystems will be replaced by the Super Read Out Driver (superROD or sROD) in the 2022 upgrade [1].

An ARM-based Processing Unit (PU) is in development at the University of the Witwatersrand, Johannesburg, to complement the sROD with higher level processing tasks on the raw data, before it has been reduced by the triggering system.

The level two and three triggers (LVL2 and Event Filter) are implemented with compute clusters. Data from the level one trigger system is fed to the Read Out System (ROS) which passes the data to the Event Builder and LVL2 trigger at a rate of about 120 GB/s. [4]

2.1. Level Two Trigger, Event Builder and Event Filter

The LVL2 filter only works on a portion of the data to determine whether it is interesting - if the portion is interesting then all of the associated data is let through to the Event Builder. The LVL2 cluster is built from approximately 500 machines, each of which has two 2.5 GHz quad-core Intel Harpertown CPUs, two 1 Gb/s Ethernet interfaces and 16 GB RAM [5]. This CPU achieves about 59 000 CoreMarks.

The Event Builder consists of about 100 rack-mounted servers, each of which has two 2.6 GHz AMD Opteron 252 CPUs with 1 Gb/s Ethernet. According to the CoreMark online database, the AMD Opteron 254 CPU - which is a dual core variant of the Opteron 252 used in the Event Builder - achieves about 13700 CoreMarks [6].

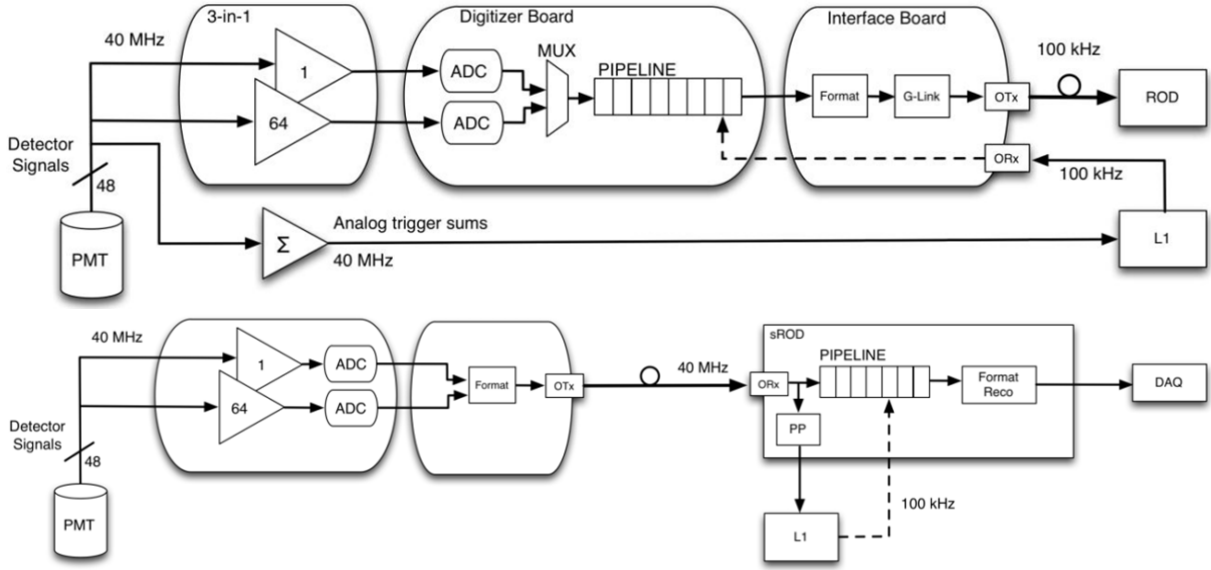


Figure 1: ATLAS TileCal Trigger and Data Acquisition System flow diagram before (above) and after the 2022 upgrade (below) [1].

The Event Filter is a much larger cluster of about 1900 machines which is the last level of triggering. The same machines as for the LVL2 filter are used. Data is output from the Event Filter at about 300 MB/s for storage [5].

3. General CPU Benchmarks

High Performance Linpack (HPL) is an industry standard benchmark application for measuring floating point performance [7]. HPL is the primary benchmark used when rating supercomputers on the Green500 and Top500 lists [8]. Both single and double precision floating point were tested.

CoreMark is another industry standard benchmark for measuring general CPU performance with a balance of integer, floating point and common algorithms [6]. ARM has recommended that CoreMark be used to replace the older Dhrystone benchmark for integer MIPS performance [9].

All three CPUs were manually forced to 1 GHz for a fair comparison. The peak power consumption of the test platforms were measured. The HPL and CoreMark results as well as the power consumption is presented in Table 2.

A cluster of eight Wandboards has been built at The University of the Witwatersrand in order to test scientific algorithms on ARM. In total, 32 1 GHz Cortex-A9 cores are available, with 16 GB RAM and interconnected with 1 Gb/s Ethernet. A photo of the cluster mounted in a rack is shown in Figure 2. No in depth testing is available at present.

4. Fast Fourier Transform Benchmarks

FFTs have numerous uses in science, mathematics and engineering. FFTs are computationally intensive and are an $O(n \log(n))$ algorithm which stresses CPU and memory subsystems.

FFTW is an open source, high performance FFT library which has a benchmark facility which reports the time, t , (in microseconds) and the estimated MFLOPS of a run [10]. The length, N , and type of FFT is specified for each run. One dimensional, single-precision complex FFTs (8 bytes per point) were tested. The CPUs were manually set to their maximum frequencies.

Figure 3 a) shows the MFLOPS results for a wide range of FFT lengths. The maximum results of multi-core and summed multi-process runs are reported. This methodology was used

Table 2: General CPU benchmarking results and power consumption.

	Cortex-A7	Cortex-A9	Cortex-A15
CPU Clock (MHz)	1008	996	1000
CPU Cores	2	4	4
HPL (Single Precision GFLOPS)	1.76	5.12	10.56
HPL (Double Precision GFLOPS)	0.70	2.40	6.04
CoreMark	4858	11327	14994
Peak Power (W)	2.85	5.03	7.48
Double Precision GFLOPS/Watt	0.25	0.48	0.81
Single Precision GFLOPS/Watt	0.62	1.02	1.41

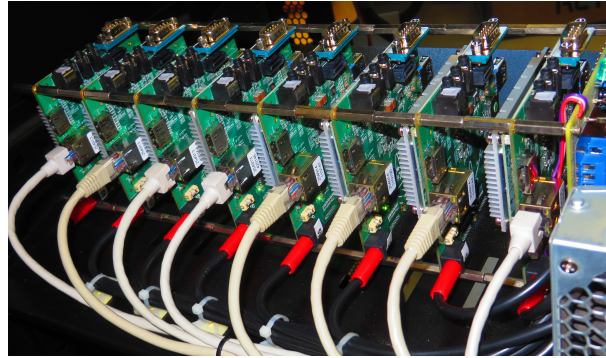


Figure 2: Photo of the 32 core Cortex-A9 cluster at The University of the Witwatersrand, Johannesburg.

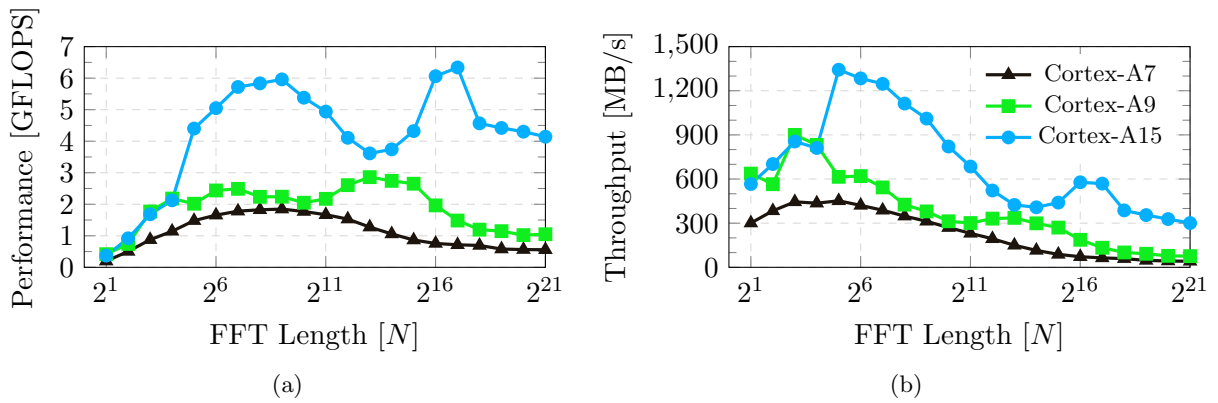


Figure 3: a) The maximum score of multi-core and multi-process runs to always utilise all processors. b) Theoretical maximum calculated FFT throughput.

to ensure 100% utilisation of the CPU for the tests. Figure 3 b) shows the calculated theoretical FFT throughput based on the FFT size in Bytes and the run time: $\text{Throughput} = (8N)/t$ [MB/s].

5. Discussion, Conclusions and Future Work

The processing performance of the Cortex-A7, A9 and A15 CPUs is comparable to an older x86 CPU, but the power efficiency is excellent. The Cortex-A15 achieves 0.81 GFLOPS/W with double precision floating point operations. The ARM processors that have been tested are optimised for single precision floating point and so are likely to be used in applications where double precision floating point is not necessary. The power efficiency is above 1 GFLOPS/W for the Cortex-A9 and A15 and is 0.62 GFLOPS/W for the Cortex-A7.

The compute performance of the Cortex-A15 is significantly higher when it is run at its maximum clock speed of 1.6 GHz but this would be true for any processor at a higher clock speed. The CoreMark results confirm that the Cortex-A15 is significantly higher performance than the Cortex-A7 and A9.

It should be noted that the Cortex-A9 and A15 SoCs that were tested could be used as a substitute based on CoreMark results and specifications for the AMD CPU used in the Event Builder section of the ATLAS TDAQ system. If ARM SoCs were used, power consumption would decrease by an order of magnitude with the ARM SoC consuming approximately 5 W and the current AMD-based system consuming over 68 W [11].

The FFT performance also indicates that the Cortex-A15 is superior to the Cortex-A9 and A7. The theoretical FFT throughputs are over 300 MB/s for most FFT lengths with the Cortex-A15 sustaining over 300 MB/s up to the largest FFT tested at 2097152 points.

Based on the results presented in this paper, specifically the low cost, high performance and good power efficiency, it is clear that ARM SoCs should be considered for future upgrades and new computing systems in big science projects.

Acknowledgements

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. We would also like to acknowledge the School of Physics, the Faculty of Science and the Research Office at the University of the Witwatersrand, Johannesburg.

References

- [1] Carrió F *et al.* 2014 *Journal of Instrumentation* **9** C02019–C02019 ISSN 1748-0221 URL <http://stacks.iop.org/1748-0221/9/i=02/a=C02019>
- [2] Cox M A, Reed R, Wrigley T, Harmsen G and Mellado B 2014 *In Review*
- [3] Krazit T 2006 ARMed for the living room URL http://news.cnet.com/ARMed-for-the-living-room/2100-1006_3-6056729.html
- [4] Beck H P *et al.* 2008 *IEEE Transactions on Nuclear Science* **55** 176–181
- [5] Winklmeier F 2009 The ATLAS High Level Trigger infrastructure, performance and future developments *2009 16th IEEE-NPSS Real Time Conference* (IEEE) pp 183–188 ISBN 978-1-4244-4454-0 URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5321918>
- [6] EEMBC 2014 CoreMark Online Database URL <http://www.eembc.org/coremark/>
- [7] Petitet A, Whaley R C, Dongarra J and Cleary A 2008 HPL - A Portable Implementation of the High-Performance Linpack Benchmark for Distributed-Memory Computers URL <http://www.netlib.org/benchmark/hpl/>
- [8] Feng W c and Cameron K 2007 *Computer* **40** 50–55 ISSN 0018-9162 URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4404810>
- [9] ARM 2011 Application Note 273: Dhrystone Benchmarking for ARM Cortex Processors URL http://infocenter.arm.com/help/topic/com.arm.doc.dai0273a/DAI0273A_dhrystone_benchmarking.pdf
- [10] Frigo M and Johnson S 2005 *Proceedings of the IEEE* **93** 216–231 URL <http://www.fftw.org>
- [11] CPU World 2014 AMD Opteron 252 Specifications URL <http://www.cpu-world.com/CPUs/K8/AMD-Opteron252-OSP252FAA5BL.html>

Cluster model analysis of exotic decay in actinide nuclei

EJ du Toit

Department of Physics, University of Stellenbosch, Post Office Box 1529, Stellenbosch, 7599, South Africa

E-mail: ejdutoit@sun.ac.za

Abstract. A simple method is described to determine the possible cluster structures forming within a nucleus. This is used to describe the multiple cluster decays seen from certain nuclei and reproduce exotic decay data for a range of actinide nuclei, using the Binary Cluster Model, along with a phenomenological potential. Further, energy spectra and $B(E2)$ values are calculated in order to establish whether these cluster structures could exist within the nucleus as a bound state.

1. Introduction

The emission of a heavy ion was first observed experimentally by Rose and Jones [1], when they detected the emission of ^{14}C from ^{223}Ra . In the following years, decays via the emission of clusters heavier than Carbon and via the emission of multiple clusters were observed [2, 3, 4, 5], while several theoretical models were developed in order to understand this process of decay by reproducing experimental data and predicting other possible cluster decays [6, 7].

One of the more simple theoretical models is that of the Binary Cluster Model (BCM) [8], where a cluster consisting of (Z_2, N_2) protons and neutrons orbits an inert core of (Z_1, N_1) protons and neutrons as a description of the nucleus consisting of (Z_T, N_T) protons and neutrons, where $Z_T = Z_1 + Z_2$ and $N_T = N_1 + N_2$. Exotic decay is then easily exhibited, by simply choosing the cluster as the emitted ion, and identifying the remaining nucleons as the core.

By utilizing the work of Gurvitz and Kalbermann [9, 10], the exotic decay half-life can be calculated with $T_{1/2} = \frac{\hbar \ln 2}{\Gamma}$ where the decay width Γ is given by

$$\Gamma = \frac{\hbar^2}{2\mu} \frac{\exp \left[-2 \int_{r_2}^{r_3} |p(r)| dr \right]}{\int_{r_1}^{r_2} \frac{1}{p(r)} dr} \quad (1)$$

with r_1 , r_2 and r_3 corresponding to the classical turning points; the intersection of Q and $V(r)$, where Q refers to the Q -value of the decay. The reduced mass is given by $\mu = \frac{A_1 A_2}{A_1 + A_2}$, and the classical momentum $p(r) = \sqrt{\frac{2\mu}{\hbar^2}(Q - V(r))}$.

Buck *et al.* successfully reproduced the exotic decay half-lives of several nuclei by using [11, 12], along with a Woods-Saxon and cubic Woods-Saxon form of the nuclear potential [13], given by

$$V_N(r) = -V_0 \left(\frac{x}{1 + \exp[(r - R_0)/a]} + \frac{1 - x}{(1 + \exp[(r - R_0)/3a])^3} \right) \quad (2)$$

where the potential depth is given by V_0 , the diffuseness by a , the radius by R_0 , and the relative contributions made by the Woods-Saxon and cubic Woods-Saxon terms are indicated by x .

Experiments have shown, however, that it is possible for nuclei to decay via the emission of multiple clusters, which poses certain problems in describing such decays theoretically. Further, calculating the exotic decay half-life for nuclei for which no experimental decay data exists also poses a problem in selecting the core and cluster. To this end, a method has been developed which enables us to consider the nucleus as a mixture of up to four different core-cluster decompositions, each with an associated preformation probability [14, 15].

These core-cluster structures can account for multiple cluster emissions, and the decay half-lives can be computed, by weighting the decay width with the preformation probability, and comparing it to the experimental values, where applicable. Further confirmation of these cluster structures can be found by calculating energy spectra and $B(E2)$ values within the BCM.

Section 2 provides a brief summary of the Binary Cluster Model, while Section 3 describes the method of determining the core and cluster sizes. Section 4 contains the results where this method has been applied on a range of nuclei.

2. Binary Cluster Model

The Binary Cluster Model (BCM) assumes that a cluster (A_2, Z_2) orbits an inert core (A_1, Z_1) . In order to satisfy the Pauli Exclusion Principle, the nucleons contained in the cluster have to be placed above the Fermi surface of the core nucleons. The global quantum number is introduced for this purpose, and defined as,

$$G = 2n + L = \sum_{i=1}^{n_c} (2n_i + l_i) \quad (3)$$

where n_c is the number of nucleons in the cluster. The values of n_i and l_i correspond to the filling of shell orbitals above the closed core, satisfying the Pauli principle by hand [8]. The quantum numbers L and n refer to the orbital angular momentum and the number of nodes contained in the wavefunction of the core-cluster relative motion, respectively.

For actinide nuclei, the value of G can be estimated through the simple scaling formula $G = 5A_2$, first introduced by Buck *et al.* [16]. This gives rise to a natural band of states, corresponding to the energies E_L , where $L = 0, 2, 4, \dots$, which can be calculated by using the Bohr-Sommerfeld (BS) relation [17],

$$\int_{r_1}^{r_2} \sqrt{\frac{2\mu}{\hbar^2} (E_L - V(r))} dr = (G - L + 1) \frac{\pi}{2} \quad (4)$$

where r_1 and r_2 represents the innermost classical turning points. The potential $V(r)$ consists of three parts; a coulomb term $V_c(r)$ [18]; a Langer-modified centrifugal potential

$V_L(r) = \frac{(L + \frac{1}{2})^{1/2} \hbar^2}{2\mu r^2}$ [19]; and a nuclear potential given by the phenomenological Woods-Saxon and cubic Woods-Saxon form, given by (2). By fitting experimental decay data to the cluster emission of several thorium isotopes, the best-fit parameters were found to be,

$$a = 0.75 \quad x = 0.36 \quad V_0 = 54.7A_2 \quad G = 5A_2 \quad (5)$$

while the radius R_0 is obtained by fitting its value to the energy of the ground state, such that (4) is satisfied. These parameters were used to investigate the cluster emission and -structure of nuclei known to undergo exotic decay.

3. Cluster Formation Probability

The advantage of the Binary Cluster Model (BCM) is that the emission of a heavy ion is very easily exhibited, although a consistent method of determining the core and cluster sizes are required. Experimental data serves as a good guide, but these data are mostly lacking. To this end, we investigate the hypothesis of utilizing the binding energy of core and cluster as a guide in determining these two components [14, 15].

From known exotic decay data, the most favourable combination seems to be if either the core or cluster is doubly, or at least singly, magic. The most tightly bound nuclei are then found by determining the partition of a fixed charge and mass (Z_T, A_T) into various possible pairs of nuclides with charge and mass values (Z_1, A_1) and (Z_2, A_2) which maximize the difference:

$$D(Z_1, A_1, Z_2, A_2) = [B_A(Z_1, A_1) - B_L(Z_1, A_1)] + [B_A(Z_2, A_2) - B_L(Z_2, A_2)] \quad (6)$$

between the actual binding energy B_A and some smoothly varying average B_L obtained from a liquid-drop mass formula.

Since even-even nuclei are more tightly bound than other nuclei, only even-even (Z_T, A_T) , (Z_1, A_1) and (Z_2, A_2) are considered. The current version of the liquid-drop formula, or Semi-Empirical Mass Formula (SEMF), for even-even nuclei is used [20]

$$B_L = a_v A - a_s A^{2/3} - a_c \frac{Z(Z-1)}{A^{1/3}} - a_a \frac{(A-2Z)^2}{A} + \frac{a_p}{A^{1/2}} \quad (7)$$

where

$$a_v = 15.56 \text{ MeV} \quad a_s = 17.23 \text{ MeV} \quad a_c = 0.697 \text{ MeV} \quad a_a = 23.285 \text{ MeV} \quad a_p = 12 \text{ MeV} \quad (8)$$

The idea is to establish a consistent method where the only information required is the total charge and mass, (Z_T, A_T) , along with readily available mass tables and (6). Additional input stems from the experimental observation that electric dipole transitions in heavy nuclei are very weak, i.e. $B(E1; 1^- \rightarrow 0^+) \approx 0$. Within the BCM, this requires, to a good approximation [15]

$$\frac{Z_1}{A_1} = \frac{Z_2}{A_2} = \frac{Z_T}{A_T} \quad ; \quad \frac{N_1}{A_1} = \frac{N_2}{A_2} = \frac{N_T}{A_T} \quad (9)$$

where N represents the number of neutrons.

In order to satisfy (9), the nucleus is considered to consist of a mixture of up to four core-cluster decompositions, with associated probabilities. The clusters are neighbouring isotopes and isotones, to ensure the satisfaction of these equations. This allows us to select an arbitrary cluster charge \bar{Z}_2 as plotting parameter, compute the corresponding mean neutron number \bar{N}_2 according to

$$\bar{N}_2 = \frac{N_T \bar{Z}_2}{Z_T} \quad (10)$$

and determine the number of protons and neutrons contained in the cluster in such a manner that they bracket the mean values, i.e.

$$Z_2 \geq \bar{Z}_2 \geq Z_2 - 2 \quad ; \quad N_2 \geq \bar{N}_2 \geq N_2 - 2 \quad (11)$$

with weights

$$\begin{aligned} p(Z_2) &= (1/2)[\bar{Z}_2 - (Z_2 - 2)] \quad ; \quad p(Z_2 - 2) = (1/2)[Z_2 - \bar{Z}_2] \\ p(N_2) &= (1/2)[\bar{N}_2 - (N_2 - 2)] \quad ; \quad p(N_2 - 2) = (1/2)[N_2 - \bar{N}_2] \end{aligned} \quad (12)$$

while the weighted average $\bar{D}(\bar{Z}_2, \bar{N}_2)$ is calculated as

$$\bar{D}(\bar{Z}_2, \bar{N}_2) = \sum_{Z_2, N_2} p(Z_2)p(N_2)D(Z_2, N_2). \quad (13)$$

This method allows us to select mean cluster charges across a range of values, and calculate the corresponding weighted average $\bar{D}(\bar{Z}_2, \bar{N}_2)$. These results can be plotted, with the maximum corresponding to the most likely mean cluster charge. Decomposing this mean cluster charge into four core-cluster decompositions will then reveal the most likely core-cluster structure of the parent nucleus, with its associated probability of formation.

4. Results

In order to establish the possible core-cluster structures of a nucleus, it is necessary to determine the mean cluster charge \bar{Z}_2 for which $\bar{D}(\bar{Z}_2, \bar{N}_2)$ is a maximum. Figure (1) shows the plot of $\bar{D}(\bar{Z}_2, \bar{N}_2)$ for ^{226}Th as a function of cluster charge. Since the full calculation produces two peaks, corresponding to $Z_2 = 7$ and $Z_2 = 9$, it is necessary to introduce some smoothing of the curve, in order to obtain a single maximum.

A previous method based on a Fourier-based filtering process has been used [14] to remove the irregularities, but a similar result can be obtained by simply averaging over neighbouring values. At a specific mean cluster charge \bar{Z}_2 , an average value of $\bar{D}(\bar{Z}_2, \bar{N}_2)$ is then obtained by averaging the value of $\bar{D}(\bar{Z}_2, \bar{N}_2)$ from a set number of \bar{Z}_2 values surrounding the mean cluster charge of interest.

This method eliminates irregularities, by producing a single maximum, and takes into account relative contributions from multiple peaks and regions next to peaks. For the example of ^{226}Th , the two peaks corresponding to $Z_2 = 7$ and $Z_2 = 9$ is reduced to a single peak at $Z_2 = 8$, a cluster charge which is contained in both peaks of the full calculation, since only even-even clusters are considered, and should therefore be the most likely cluster charge.

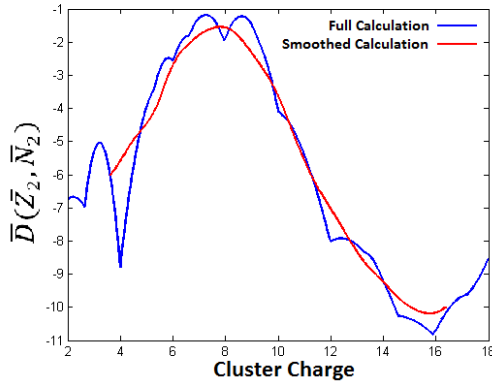


Figure 1. Calculations of $\bar{D}(\bar{Z}_2, \bar{N}_2)$ as a function of cluster charge for ^{226}Th .

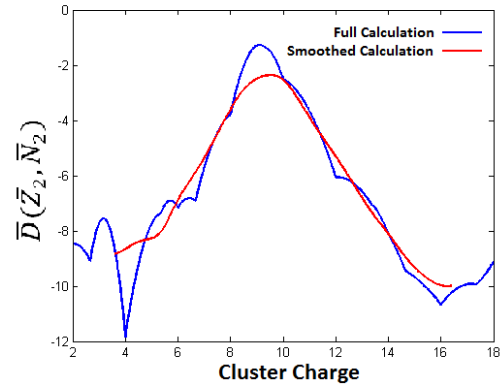


Figure 2. Calculations of $\bar{D}(\bar{Z}_2, \bar{N}_2)$ as a function of cluster charge for ^{230}U .

The maximum of $\bar{D}(\bar{Z}_2, \bar{N}_2)$ for ^{226}Th is obtained for a mean cluster charge of $\bar{Z}_2 = 9.26$, while for ^{230}U $\bar{Z}_2 = 9.52$. These mean cluster charges can be broken down into the four different core-cluster decompositions in order to reveal the possible cluster structures, along with their preformation probabilities. The exotic decay half-life can then be calculated for each cluster, and compared to experimental data, where available.

It was found that cluster decays which are not observed experimentally tend to have a small preformation probability and a long half-life, while the experimentally observed cluster decays tend to have greater preformation probabilities and shorter half-lives. Should the half-life for a

cluster decay be calculated to be orders of magnitude greater than another (observed) cluster decay, it will be nearly impossible to detect the decay of such a cluster, due to the competition from α -decay and spontaneous fission that occurs on a much shorter time-scale.

A summary of the experimentally observed cluster decays seen in the actinide nuclei considered are shown in table 1. Of all the nuclei investigated, only one experimentally observed cluster was not predicted in the model: the ^{28}Mg decay of ^{238}Pu , although this shortcoming can be fixed by extending the model such that the nucleus consists of eight core-cluster structures. As can be seen from table 1, more than half of the calculated half-lives are in reasonable agreement with experimentally measured values.

The choice of parameters failed to reproduce the cluster decay of carbon and oxygen. While previous work shows [11, 12] that the BCM is not successful in accurately describing the cluster emission of oxygen, the inaccuracy in the carbon decay of ^{224}Ra puts an additional bound on the mass and charge region of where the particular choice of parameters are valid. The potential parameters were chosen to reproduce the cluster decay of thorium isotopes, and applied to the calculations of radium, uranium and plutonium isotopes. A more consistent method of determining the potential parameters is therefore needed to accurately calculate exotic decay data for all nuclei considered, such as calculating parameter values from a microscopic potential, as was done for alpha clusters [21].

The calculated data show that multiple cluster decays can be described theoretically within this model, and that, where experiments are unable to determine the mass of an emitted cluster, this theoretical approach can be used to deduce the most likely emitted cluster. Such an example is the cluster emission of neon isotopes from ^{232}Th and ^{234}U [22].

Table 1. Summary of Exotic Decay Half-Lives of all experimentally observed decays. All considered actinide nuclei are shown with the emitted cluster and whether or not the model predicted this cluster. If the model predicts the cluster, the preformation probability P is also shown. Should the calculated half-life be within a factor 3 of the experimentally measured half-life, it is considered as a good approximation.

Nucleus	Cluster	P	Observed	Factor ~ 3	$T_{1/2}^{\text{calc}}[\text{y}]$	$T_{1/2}^{\text{expt}}[\text{y}]$
^{222}Ra	^{14}C	0.573	✓	✓	6590	4017 ± 528
^{224}Ra	^{14}C	0.115	✓		22.7×10^8	$(24.9 \pm 1.6) \times 10^7$
^{226}Th	^{18}O	0.099	✓		21.9×10^{10}	1.82×10^9
^{228}Th	^{20}O	0.38	✓		24.0×10^{13}	$(16.9 \pm 1.4) \times 10^{12}$
^{230}Th	^{24}Ne	0.504	✓	✓	13.3×10^{16}	$(13.0 \pm 0.52) \times 10^{16}$
^{232}Th	^{24}Ne	0.332	✓		6.54×10^{22}	5.06×10^{21}
^{232}Th	^{26}Ne	0.478	✓	✓	4.87×10^{21}	5.06×10^{21}
^{230}U	^{24}Ne	0.654	✓		20.6×10^{13}	$(11.9 \pm 2.8) \times 10^{11}$
^{232}U	^{24}Ne	0.432	✓	✓	13.7×10^{12}	$(77.4 \pm 4.5) \times 10^{11}$
^{234}U	^{24}Ne	0.025	✓		48.7×10^{18}	$(27.3 \pm 6.7) \times 10^{17}$
^{234}U	^{26}Ne	0.815	✓	✓	43.0×10^{17}	$(27.3 \pm 6.7) \times 10^{17}$
^{234}U	^{28}Mg	0.155	✓	✓	15.8×10^{17}	$(17.5 \pm 4.3) \times 10^{17}$
^{236}Pu	^{28}Mg	0.129	✓	✓	1.14×10^{14}	1.43×10^{14}
^{238}Pu	^{28}Mg					1.46×10^{18}
^{238}Pu	^{30}Mg	0.470	✓	✓	2.51×10^{18}	1.46×10^{18}
^{238}Pu	^{32}Si	0.070	✓		4.46×10^{18}	6.27×10^{17}
^{240}Pu	^{34}Si	0.454	✓	✓	14.1×10^{18}	$(50.5 \pm 5.4) \times 10^{17}$

In order to determine whether the proposed cluster structures could exist within the nucleus as a bound state, structure observables such as energy spectra and $B(E2)$ values were computed. The $B(E2)$ values can be computed according to [23, 24]

$$B(E\ell; L_i \rightarrow L_f) = \frac{2\ell + 1}{4\pi} \left[Z_1 \left(-\frac{A_2}{A} \right)^\ell + Z_2 \left(\frac{A_1}{A} \right)^\ell \right]^2 |\langle L_i 0 \ell 0 | L_f 0 \rangle|^2 |\langle \psi_{L_f} | r^\ell | \psi_{L_i} \rangle|^2 \quad (14)$$

where ψ_L is the wavefunction, and a small effective charge is introduced by replacing $Z_i \rightarrow Z_i + \epsilon A_i$, with $\epsilon = 0.17$.

Since the nucleus is considered to be a mixture of four core-cluster decompositions, the observables are calculated for each core-cluster decomposition, and a weighted average is calculated, according to the preformation probabilities. This weighted average value is then compared to the experimental value.

A summary of the calculated $B(E2; 2^+ \rightarrow 0^+)$ and the highest known energy state are shown in table 2. Since the radius R_0 is fitted to the Q -value of the ground state, the ground state energy is inherently correct, and should the highest known energy state correspond to the experimental value, all energy levels in between should be in good agreement with the experimental values.

From table 2 it can be deduced that the proposed cluster structures can exist as bound states, since all calculated $B(E2)$ values, except for plutonium isotopes, are in agreement with the experimental values, as well as the energy levels. The inaccuracy in reproducing data of the plutonium isotopes places an additional bound on the region where the choice of parameters are valid.

Table 2. The calculated $B(E2; 2^+ \rightarrow 0^+)$ values for all actinide nuclei considered are compared to known experimental values. The highest known energy levels are also calculated and compared to experimental values.

Nucleus	$B(E2)^{expt}$ [W.u.]	$B(E2)^{calc}$ [W.u.]	E_L^*	E_L^{expt} [MeV \pm eV]	E_L^{calc} [MeV]
^{222}Ra	111 ± 9	80	E_{20}^*	3.29 ± 11	3.36
^{224}Ra	97 ± 4	96	E_{12}^*	1.41 ± 4	1.36
^{226}Th	164 ± 10	145	E_{20}^*	3.10 ± 8	2.84
^{228}Th	167 ± 6	172	E_{18}^*	2.41 ± 7	2.29
^{230}Th	196 ± 8	206	E_{24}^*	3.82	2.59
^{232}Th	198 ± 11	219	E_{30}^*	5.16 ± 3	5.07
^{230}U	222 ± 27	213	E_{22}^*	3.24 ± 4	3.01
^{232}U	241 ± 21	231	E_{20}^*	2.66 ± 9	2.51
^{234}U	236 ± 10	251	E_{30}^*	4.81	4.95
^{236}Pu	-	314	E_{16}^*	1.79 ± 5	1.51
^{238}Pu	285 ± 5	347	E_{26}^*	4.27 ± 9	3.44
^{240}Pu	287 ± 11	393	E_{32}^*	5.82 ± 8	4.76

In conclusion, the proposed method of determining possible core-cluster decompositions within actinide nuclei and using the binary cluster model to calculate decay half-lives, electromagnetic transition probabilities and energy spectra, has shown that multiple cluster decays can be described theoretically, and that these core-cluster structures can possibly exist within heavy nuclei as bound states.

A fixed choice of potential parameters places a bound on the region where it can be applied, but within this region, the calculated decay values are in good agreement with experimental values. As such, it is possible to deduce the mass of emitted clusters where experiments are

unable to. Calculations of $B(E2)$ values and energy spectra also indicate that these cluster structures can exist as bound states, although a more careful approach to obtaining the nuclear potential parameters might be needed.

References

- [1] Rose H J and Jones G A 1984 *Nature* **307** p 245
- [2] Barwick S W, Price P B and Stevenson J D 1985 *Phys. Rev.* **C31** p 1984
- [3] Schicheng Wang, Price P B, Barwick S W, Moody K J and Hulet E K 1987 *Phys. Rev.* **C36** p 2717
- [4] Price P B 1989 *Annu. Rev. Nucl. Part. Sci.* **39** p 19
- [5] Tretyakova S P and Mikheev V L 1997 *Il Nuovo Cimento* **110A** p 1043
- [6] Greiner W 2007 *Rom. Rep. Phys.* **59** p 193
- [7] Bonetti R and Guglielmetti A 2007 *Rom. Rep. Phys.* **59** p 301
- [8] Buck B, Dover C B and Vary J P 1975 *Phys. Rev.* **C11** p 1803
- [9] Gurvitz S A and Kalbermann G 1987 *Phys. Rev. Lett.* **59** p 262
- [10] Gurvitz S A 1988 *Phys. Rev.* **A38** p 1747
- [11] Buck B, Merchant A C and Perez S M 1996 *Phys. Rev. Lett.* **76** p 380
- [12] Buck B, Merchant A C, Perez S M and Seals ME 2005 *J. Phys.* **G31** p 1499
- [13] Buck B, Merchant A C and Perez S M 1995 *Phys. Rev.* **C51** p 559
- [14] Buck B, Merchant A C, Horner M J and Perez S M 2000 *Phys. Rev.* **C61** p 024314
- [15] Buck B, Merchant A C and Perez S M 2000 *Few-Body Systems* **29** p 53
- [16] Buck B, Merchant A C, Perez S M and Tripe P 1994 *J. Phys.* **G20** p 351
- [17] Zettili N 2001 *Quantum Mechanics: Concepts and Applications* (John Wiley and Sons)
- [18] Buck B, Merchant A C and Perez S M 1992 *Phys. Rev.* **C45** p 2247
- [19] Langer R E 1937 *Phys. Rev.* **51** p 669
- [20] Martin B.R 2006 *Nuclear and Particle Physics: An Introduction* (John Wiley and Sons)
- [21] Ibrahim T T, Perez S M and Wyngaardt S M 2010 *Phys. Rev.* **C82** p 034302
- [22] du Toit E J, Wyngaardt S M and Perez S M 2015 *J. Phys. G* **42** p 015103
- [23] Wong S M 1990 *introductory Nuclear Physics* (Prentice Hall)
- [24] Ibrahim T T 2009 Ph.D. Thesis, Stellenbosch University (unpublished)

Monte Carlo simulation of secondary gamma production during proton therapy for dose verification purposes

Jeyasingam Jeyasugiththan^{1,2} and Stephen Peterson¹

¹Department of Physics, University of Cape Town, Rondebosch, 7701, South Africa

²Department of Clinical Oncology, Teaching Hospital, Jaffna, Sri Lanka

E-mail: jeyasugiththan@yahoo.com

Abstract. High energy protons can kill both cancerous as well as normal tissue, so any range uncertainty during a proton radiotherapy treatment will strongly affect the success of the overall dose delivery. In recent years, detection of prompt gammas produced by inelastic nuclear reactions between protons and nuclei within human tissue has been proposed for online treatment verification. The aim of this work is to simulate these discrete prompt gammas using the Geant4 (v9.6.02) Monte Carlo toolkit that provides several models for low energy proton inelastic nuclear reactions: binary cascade (BIC), Precompound (PRECO) and Intra nuclear cascade (INCLXX). The selection of an appropriate physics model would increase the accuracy of the prompt gamma simulations. The suitability of these models for discrete gamma emission from excited states of ^{16}O , ^{12}C and ^{14}N nuclei were tested by comparing simulated inelastic gamma production cross section data against available experimental data in the energy range 0 to 200 MeV. Among the Geant4 physics models, the Precompound model was found to be most suitable for producing reasonable prompt gamma spectra. Moreover, Fermi break up was activated below 20 MeV for complete ^{16}O simulation. A combination of different physics models in different energy regions was tested to fit a model for prompt gamma emission. Finally, a water phantom was simulated with 200 MeV proton passive beam and the prompt gamma energy spectrum was acquired by a LaBr₃ detector actively shielded by a BGO detector. Time-of-flight (TOF) techniques were used to eliminate scattered gammas from the beam line elements and secondary neutrons from the target. The Geant4 simulations confirmed the ability of our physics model to produce reasonable prompt gamma spectra that will be used in further studies for comparison to measured spectra.

1. Introduction

The use of x-ray photons in cancer treatment has a long history and is a well-defined method of treatment. Recent developments in radiation therapy, attacking the tumour with accelerated protons, presents unprecedented control over the dose deposition and greatly reduces the side-effects of treatment as compared to traditional x-ray radiotherapy. The proton range in a patient is uncertain for many reasons and can cause either *undershoot* to the tumour or *overshoot* to the normal tissue [1]. Although uncertainties during the calculation of the dose and patient preparation for treatment could be minimized, any range uncertainty presented during the treatment will affect the overall success of tumour control. Typically, an additional safety margin around the treatment area is applied to ensure tumour coverage [2]. This additional

margin can over-dose the normal tissue surrounding the tumour and cause serious damage to any important organs situated very close to the tumour [3], [1]. The use of a PET scanner for range verification was proposed for dose verification either immediately following treatment (on-time) or sometime after the treatment [4] [5]. Unfortunately, the treatment verification by PET/CT is seen to be insufficient at this time [5], [6]. Yet, it so far is the only available method that can be used as an *in-vivo* proton range verification device until a more precise method for online verification is developed.

A technique using prompt gamma emission from excited nuclei induced by inelastic nuclear reactions during proton radiotherapy has been attracting the attention of scientists engaged in developing a device for on-line range verification [7], [8], [9]. However, the measurement of this prompt gamma radiation in the clinical environment is still difficult due to the presence of secondary neutrons and scattered photons. In the passive beam delivery mode, additional secondary radiation activated in the beamline elements and patient final collimator is another issue. In this paper we investigate the feasibility of Geant4 Monte Carlo simulations of secondary gamma production during passive proton therapy for dose verification purposes.

2. Method

2.1. Geant4 simulations

Geant4 is an object oriented toolkit and implemented in the C++ environment [10]. It is being used to simulate the interactions of particles as they traverse through matter and is applicable in the fields of particle physics, nuclear physics, astro physics, accelerator design and medical physics. The first version of the Geant4 was released in 1998, and the following year (1999) the Geant4 collaboration was established for further development, maintenance and user support. The present work is based on the Geant4.9.6.p02 released in 2013.

Geant4 provides pre-build physics models for low energy hadronic interactions: binary cascade, precompound and intra nuclear cascade. Because of its transparency, physics models can be modified to meet user requirements by fitting of their experiment's data. In this study, the recommended reference physics list for proton radiotherapy (QGSP_BIC_EMY [11]) was used to validate the prompt gamma emissions from the elements (^{16}O , ^{12}C and ^{14}N) mostly abundant in human tissue. By default, the selected physics list provides the binary cascade model for inelastic nuclear reactions in which, the interaction is described as a two-particle binary inelastic collision between the incident proton and nucleons in the target nucleus. Further interactions between the remaining nucleons in the target nucleus and any resulting secondaries are allowed to create an intra nuclear cascade. To check the resulting secondaries, the Fermi exclusion principle is applied. If the momentum of a secondary particle falls below the Fermi level (momentum), the interaction is suppressed. Therefore the original primary particle is taken to the next interaction. On the other hand, if an interaction occurs, the secondaries are treated like primary particles. Any particle propagation into the nuclear field is determined by solving the equation of motion numerically. The cascade is terminated if the secondaries have not reached the threshold energy required for the interaction. After each interaction, particle-hole states or excitons will be added to the target nucleus and at the end of cascade, the remaining residual nuclear system with the exciton state is treated by the precompound and de-excitation models.

With respect to validations, the binary cascade model was found to underestimate the peak of the 6.13 MeV prompt gamma-ray from ^{16}O and therefore replaced by the precompound model with a modified initial exciton number of 2. The initial precompound nucleus is described by the atomic mass and charge of the residual nucleus, its four-momentum vector, its nuclear excitation energy and the number of excitons which is calculated by summing particles above the holes below Fermi level and vacant under the Fermi level of the compound nucleus. Geant4 then loops through the transitions until an equilibrium condition is reached for de-excitation which uses only the excitation energy and not the number of excitons. Several nuclear transitions are

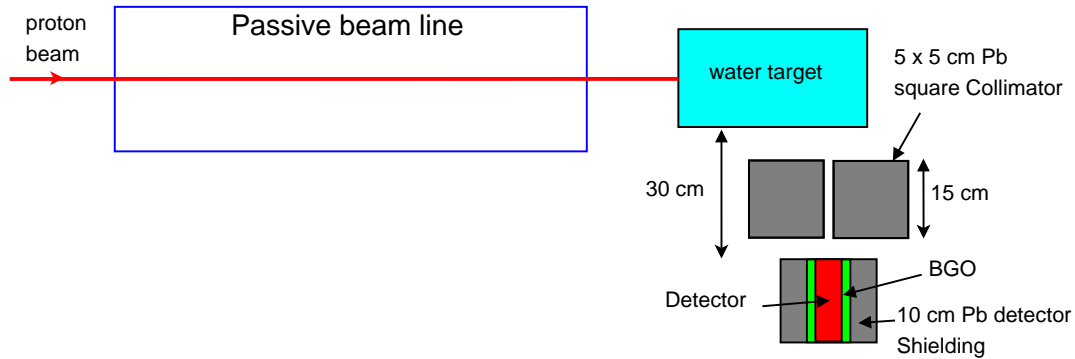


Figure 1: Simulation geometry setup with LaBr₃ detector.

possible with $\Delta n = \pm 2, 0$ associated with their transition probability depending on the exciton number and the excitation energy. If there is particle emission (neutrons, protons, deuterons, tritium and helium nuclei) before equilibrium is reached, the above steps are repeated with the new nuclear fragment. The particle emission is also associated with their emission probability which is a function of the exciton number and excitation energy. At statistical equilibrium, the simulation will be handled by the equilibrium model for emission of photons, nucleons and light fragments from the residual state [12]. The evaporation model is considered for the emission of nuclear fragments or gamma-rays from the excited nucleus through five different channels handled by G4ExcitationHandler: Evaporation as the main de-excitation, Fission for heavy nuclei, Fermi Break-up for light ion, Photon evaporation as competitive channel in evaporation and multifragmentation for very excited nuclei [13].

2.2. Geant4 simulation set-up

The simulation set-up is shown in the figure 1. The target is a cylindrically shaped water phantom of radius 20 cm and length 30 cm. A cylindrical LaBr₃ (2 inch x 2 inch) detector was surrounded by both 18 mm thick BGO ($Bi_4Ge_3O_{12}$) active shielding and 10 cm thick lead shielding, and modelled perpendicular to the beam axis with a distance between detector front face and beam axis of 30 cm. A 15 cm thick lead collimator having a 5 cm x 5 cm square hole was used to collimate the prompt gammas between the target and the detector. The previously validated Monte Carlo passive-scatter beam line model (for more details see [14]) was used to generate a 200 MeV (191 MeV at the iso-centre) proton beam of circular field with diameter of 10 cm. The modelled detector (also previously validated) was used to reproduce the detector response[14].

The purpose of the BGO active shielding was to reduce the continuous Compton background from the incomplete Compton-scattered gamma energy deposition in the LaBr₃ crystal using an anti-coincidence method. Also the BGO reduces the background gamma-rays produced by the neutrons hitting the lead detector shielding. The uncollimated gamma rays scattered from objects surrounding the detector were also attenuated at the lead shielding. A time-of-flight (TOF) approach was introduced in the simulation in order to reduce secondary neutrons impinging the detector from the target and scattered gamma background rays, a significant problem in passive beam radiotherapy. Inelastic nuclear reactions were simulated by using the precompound model, in which the Fermi break-up was activated for proton energy below 20 MeV.

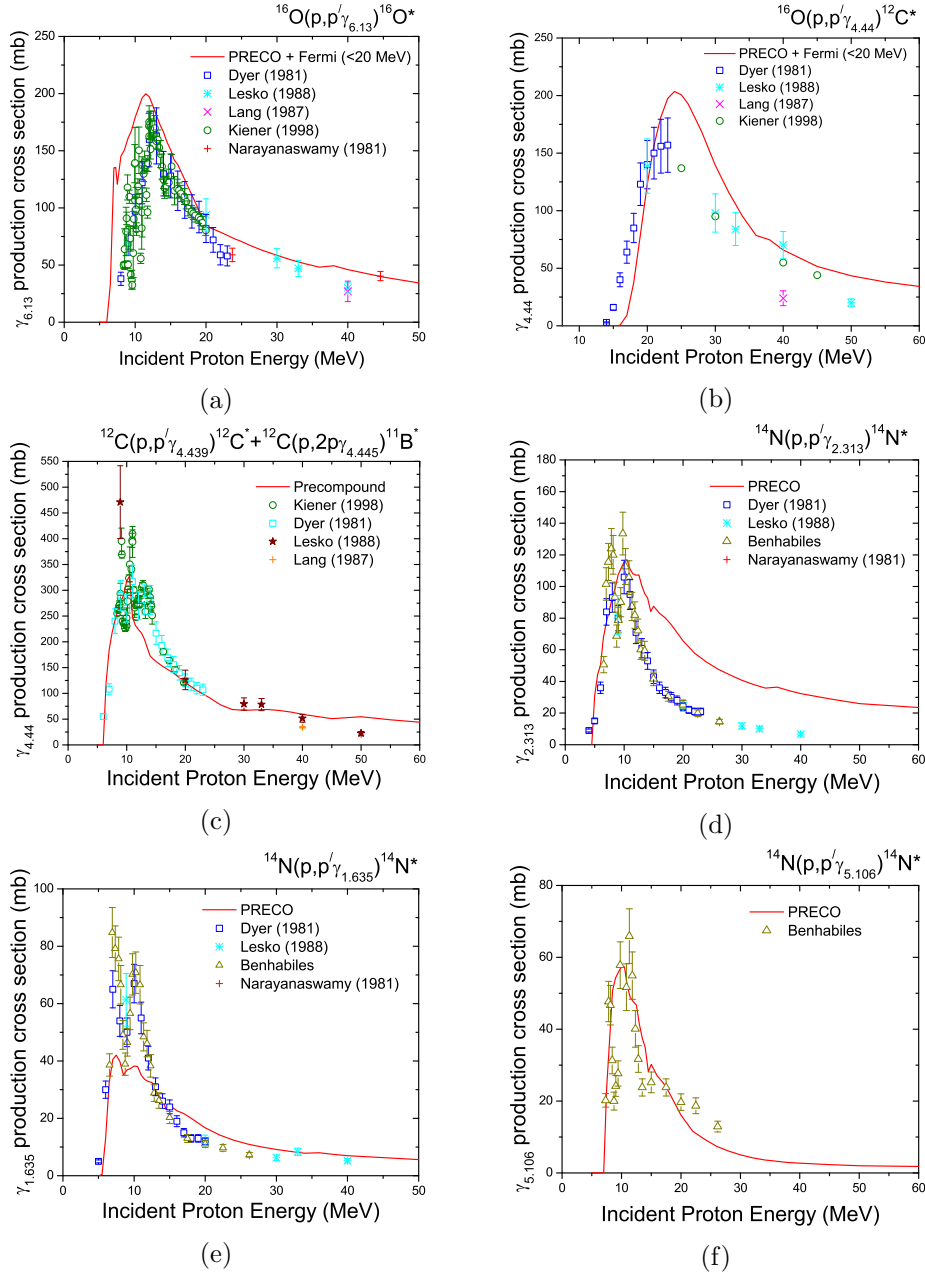


Figure 2: Comparisons between simulation and experimental cross section data for six important gamma lines produced by proton inelastic reactions on ^{16}O , ^{12}C and ^{14}N .

3. Results and Discussions

3.1. Geant4 physics validation

Figure 2 shows the comparisons between simulation and experimental cross section data currently available [15], [16], [17], [18], [19], [20] for six important gamma lines produced by proton inelastic reactions on ^{16}O , ^{12}C and ^{14}N which are the most abundant elements in human tissue. These simulations were performed using the precompound model. There were discrepancies in the gamma production performed with the default total inelastic cross sections of Willish and Axen(1996) for ^{12}C and ^{14}N . Therefore the cross sections of Tripathi et al. light ions [21] were used alternatively. The Fermi break-up below 20 MeV was required for the ^{16}O

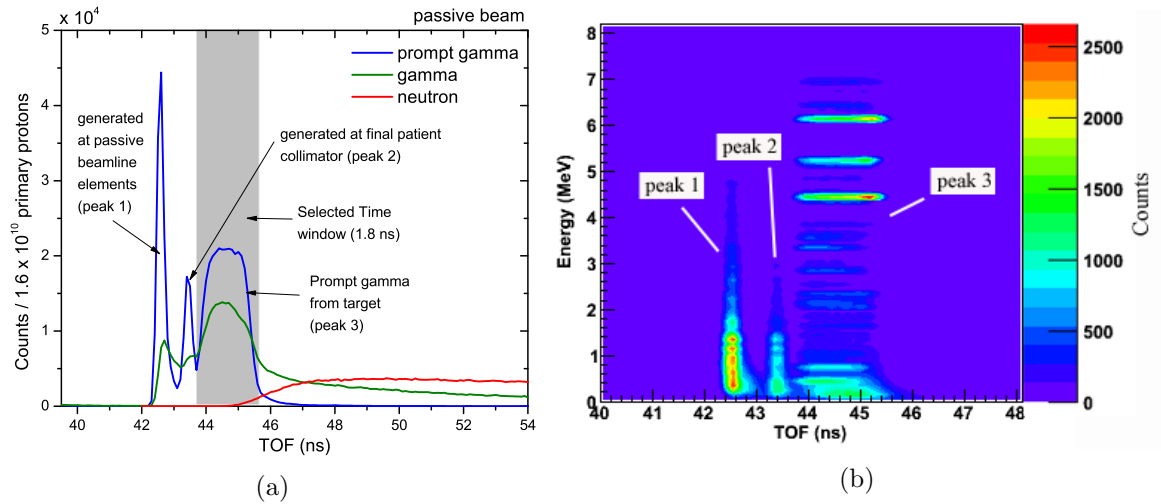


Figure 3: (a) TOF spectra of particles and (b) Time-energy spectrum of prompt gamma impinging on the LaBr_3 detector.

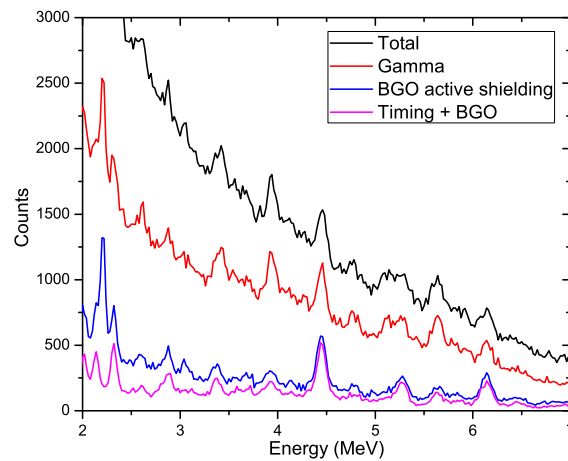


Figure 4: Comparison of energy spectra with and without background reduction methods.

simulation. Simulations of the 6.13 MeV and 4.44 MeV lines from ^{16}O , the 4.44 MeV line from ^{12}C and the 5.11 MeV line from ^{14}N all agreed well with the experimental data, unlike the other gamma-rays emitted from ^{14}N .

3.2. Geant4 simulation for prompt gamma detection

Monte Carlo time of flight (TOF) spectra are shown in figure 3a. The time of flight for the prompt gamma rays (peak 3) is shorter than slower neutrons, but greater than the scattered secondary gamma rays generated in the passive beam line elements (peak 1) and in the final patient collimator (peak 2). Therefore, a TOF window of 1.8 ns can filter out the late-arriving neutrons (about 99%) from target and as well as early-arriving background gamma rays. Moreover, the TOF window can also be coupled with an energy selection (see figure 3b). Figure 4 shows a comparison of the energy spectra simulated with and without background reduction methods applied. The simulation was carried out with 2.5×10^{10} incident protons entered into the passive beam line. Among them, only 8% reached the target and the remaining protons were stopped in the beamline elements and in the final collimator. The results indicate that the use of both

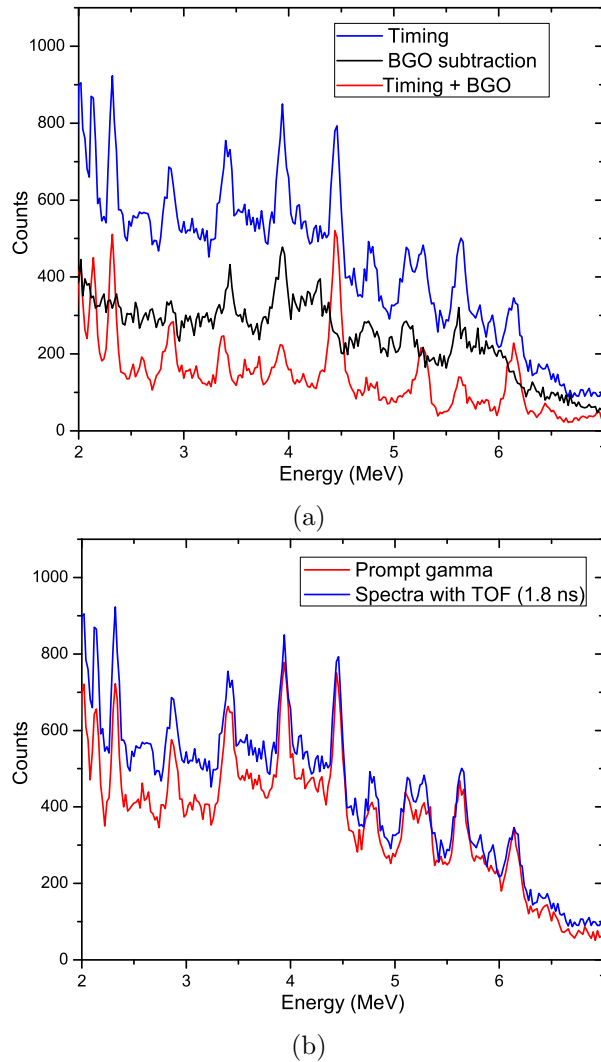


Figure 5: Comparison (a) of the energy spectra using a 1.8 ns TOF window (with and without BGO active shielding) and (b) of the energy spectra of the prompt gammas produced within the target to the TOF spectra from the LaBr₃ detector.

timing and active shielding is able to remove 90% of the background radiation which includes a 10% reduction due to BGO subtraction as shown in figure 5a. Peaks at 6.13, 4.44 and 5.24 MeV from $^{16}\text{O}^*$, $^{12}\text{C}^*$ and $^{15}\text{O}^*$ respectively are clearly identified when both timing and BGO subtraction are used. Subtraction of Compton-scattered background is required to resolve the 5.24 MeV peak from the second escape peak of 6.13 MeV. Comparison between the prompt gamma spectra detected from the target and the spectra using the TOF window shown in figure 5b is promising, showing a prompt gamma detection with an efficiency of 72.4%.

4. Conclusion

Our study confirmed the feasibility of Geant4 simulation for prompt gamma-rays in the passive-scatter proton beam mode. The background reduction using a TOF window was excellent. Also BGO subtraction helped to resolve the prompt gamma peaks from Compton background. The physics model is able to reproduce the prompt gamma spectra that will be used in our further studies for comparison to measured data.

Acknowledgments

The authors would like to thank the Center for High Performance Computing (CHPC), CSIR Campus, 15 Lower Hope St., Rosebank, Cape Town, in South Africa for providing them access to their computing resources.

References

- [1] Lu H M 2008 *Physics in Medicine and Biology* **53** N415
- [2] Paganetti H 2002 *Physics in Medicine and Biology* **47** 747
- [3] Trofimov A, Nguyen P L, Coen J J, Doppke K P, Schneider R J, Adams J A, Bortfeld T R, Zietman A L, DeLaney T F and Shipley W U 2007 *International Journal of Radiation Oncology*Biophysics* **69** 444 – 453 ISSN 0360-3016
- [4] Maccabee H D, Madhvanath U and Raju M R 1969 *Phys. Med. Biol.* **14** 213–24
- [5] Parodi K, Bortfeld T and Haberer T 2008 *International Journal of Radiation Oncology*Biophysics* **71** 945 – 956 ISSN 0360-3016
- [6] Knopf A, Parodi K, Bortfeld T, Shih H A and Paganetti H 2009 *Physics in Medicine and Biology* **54** 4477
- [7] Min C, Kim C H, Youn M and Kim J 2006 *Appl. Phys. Lett.* **89** 183517
- [8] Polf J C, Peterson S, Ciangaru G, Gillin M and Beddar S 2009a *Phys. Med. Biol.* **54** 731
- [9] Verburg J M, Riley K, Bortfeld T and Seco J 2013 *Physics in Medicine and Biology* **58** L37
- [10] Agostinelli S, Allison J, Amako K and Apostolakis J 2003 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** 250 – 303 ISSN 0168-9002
- [11] Cirrone G A P, Cuttone G, Mazzaglia S E, Romano F, Sardina D, Agodi C, Attili A, Blancato A A, Napoli M D, Rosa F D, Kaitaniemi P, Marchetti F, Petrovic I, Ristic-Fira A, Shin J, Tarnavsky N, Tropea S and Zacharatou C 2011 *Progress in nuclear science and technology* **2** 207–212
- [12] Jarlskog C Z and Paganetti H 2008 *IEEE Trans. Nucl. Sci.* **55** 1018–1025
- [13] Lara V and Wellisch J P 2000 *9th Int. Conf. Calorimetry in High Energy Physics* 449–452
- [14] Jeyasugiththan J, Peterson S, Camero J N and Symons J 2014 *Proceedings of the 58th Annual Conference of the South African Institute of Physics, edited by Roelf Botha and Thulani Jili pp. 260-267* ISBN: 978-0-620-62819-8. Available online at <http://events.saip.org.za>
- [15] Dyer P, Bodansky D, Seamster A G, Norman E B and Maxson D R 1981 *Phys. Rev. C* **23** 1865–82
- [16] Lesko K T, Norman E B, Larimer R M, Kuhn S, Meekhof D M, Crane S G and Bussell H G 1988 *Phys. Rev. C* **37**(5) 1808–1817
- [17] Lang F L, Werntz C W, Crannell C J, Trombka J I and Chang C C 1987 *Phys. Rev. C* **35** 1214–27
- [18] Kiener J, Berheide M, Achouri N L, Boughrara A, Coc A, Lefebvre A, de Oliveira Santos F and Vieu C 1998 *Phys. Rev. C* **58**(4) 2174–2179
- [19] Narayanaswamy J, Dyer P, Faber S R and Austin S M 1981 *Phys. Rev. C* **24** 2727–30
- [20] Benhabiles-Mezhoud H, Kiener J, Thibaud J P, Tatischeff V, Deloncle I, Coc A, Duprat J, Hamadache C, Lefebvre-Schuhl A, Dalouzy J C, de Grancey F, de Oliveira F, Dayras F, de Séréville N, Pellegriti M G, Lamia L and Ouichaoui S 2011 *Phys. Rev. C* **83**(2) 024603
- [21] Tripathi R K, Cucinotta F A and Wilson J W 1999 *Technical Report TP-1999-209726 National Aeronautics and Space Administration (NASA)*

Radiation hardness of plastic scintillators for the Tile Calorimeter of the ATLAS detector

H Jivan^{1,2}, R Erasmus¹, B Mellado¹, G Peters¹, K Sekonya³ and E Sideras-Haddad^{1,2}

¹ University of the Witwatersrand, 1 Jan Smuts Avenue, Braamfontein 2000, Johannesburg.

² DST-NRF Centre of Excellence in Strong Materials.

³ iThemba LABS. Gauteng.

Email: harshna.jivan@gmail.com

Abstract The Tile Calorimeter of the ATLAS detector employs plastic scintillators to aid in the detection of hadrons, taus, and jets of quarks and gluons that arise from proton-proton collisions within the Large Hadron Collider of CERN. During the first data taking period, plastic scintillators in the GAP region of the detector were exposed to a radiation environment of up to 10 kGy/year. With the LHC set to run proton collisions at almost double the previous centre of mass energy and higher luminosity, the radiation environment is expected to become much harsher. In order to ensure that the radiation damage sustained does not compromise the detector performance, these scintillators will be replaced with more radiation hard plastics. To aid in the choice of scintillator, this preliminary comparative investigation looks into the radiation damage undergone by three polyvinyl toluene (PVT) based plastic scintillators (EJ200, EJ208 and EJ260) provided by ELJEN technology. Samples of each scintillator grade were subject to 6 MeV proton irradiation with doses of the order $O(10^3)$, $O(10^5)$ and $O(10^6)$ Grays using the tandem accelerator of iThemba LABS, Gauteng. Light transmission spectroscopy analysis indicated that a loss of scintillator transparency occurs with increasing irradiation dose, with EJ208 exhibiting the least transmission loss. Raman spectroscopy showed that losses to the benzene ring and vinyl bonding structure occurred for the Mega Gray irradiated samples. This could be correlated with transmission light loss since hydrogen degassing from the breaking of these bonds would result in the formation of free radicals which affect the way in which light is absorbed by the scintillators.

1. Introduction

The Tile Calorimeter of the ATLAS detector, is a hadronic calorimeter responsible for recording the trajectory and energy of hadrons, taus as well as jets of quarks and gluons that result from the proton-proton collisions within the Large Hadron Collider of CERN. Plastic scintillators form an integral component [1] of this calorimeter and are specifically chosen for their properties of high optical transmission and fast rise and decay times. These enable efficient data capture since fast signal pulses can be generated [2]. The main problem encountered by plastic scintillators however, is radiation damage incurred due to their interaction with the ionizing particles to be detected. This damage causes a significant decrease in the light yield of the scintillator and introduces an error into the time-of flight data acquired.

As part of the phase two upgrade of the Tile Calorimeter planned for 2018, a comparative study was conducted into the radiation hardness of several plastic scintillators. The results would aid in choosing the best grade for replacing the current crack scintillators within the Gap region of the Tile Calorimeter. In this paper, we present the results for radiation damage in three polyvinyl toluene based scintillator samples; EJ200, EJ208 and EJ260 obtained from ELJEN technologies [3].

2. Materials and Methodology

The three plastic scintillators under study are composed of a polyvinyl toluene base and 3% of added organic fluors. Differences in the organic fluor dopants give rise to the different characteristic properties observed between the three samples. Some of the characteristic properties are listed in Table. 1 below.

Table 1: Properties of the plastic scintillators as provided from the manufacturers [3].

	EJ200	EJ208	EJ260
Light Output (% Anthracene)	64	60	60
Wavelength of maximum emission (nm)	425	435	490
Rise time (ns)	0.9	1.0	Not available
Decay time (ns)	2.1	3.3	9.2
Refractive index	1.58	1.58	1.58

The tandem accelerator of iThemba LABS in Gauteng was used to irradiate samples with 6 MeV protons. To ensure that the study simulated a similar type of particle-scintillator interaction as observed in the Tile Calorimeter, protons would be required to pass through the samples whilst imparting energy primarily through ionization losses. The stopping range of 6 MeV protons within the scintillator material was determined using the SRIM software package which includes a preset PVT scintillator layer in its materials database [4]. The layer was adjusted to accommodate the CH ratio and density of the plastics as indicated by the manufacturer [3]. The stopping range was found to occur at $\sim 472 \mu\text{m}$. Several samples of each grade were then cut and polished to dimensions of 1 cm by 1 cm, with thickness of $250 \mu\text{m}$. A polishing procedure based on standard metallographic techniques was employed.

Three samples of each scintillator grade were subjected to irradiation with beam intensities of approximately 1 nA and 10 pA with exposure times of 60 min and 6 min in order to achieve radiation doses ranging of the order $O(10^3)$ Grays to $O(10^6)$ Grays. Samples were mounted in a hexagonal sample carousel and housed within the microprobe chamber. The beam current was determined by measuring the charge across a metal plate situated on the side opposite to the sample on the carousel. The beam was scanned in the x and y plane using a uniform raster-scan pattern to achieve an irradiated area of approximately 4 mm by 4 mm. After irradiation, samples were contained in vials that were wrapped in aluminium foil. This was done in order to prevent the recovery of damage through photo-bleaching effects that result from exposure to visible light.

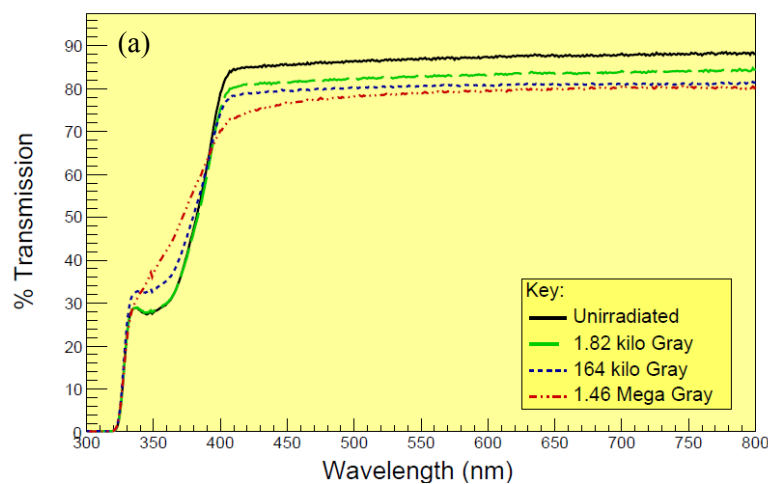
Transmission spectroscopy was conducted using the Varian Carry 500 spectrophotometer. Light transmission was measured relative to transmission in air over a range of 300-800 nm. Thereafter, the Raman spectra for un-irradiated and the highest dose irradiated samples were obtained using the Horiba Jobin-Yvon Raman spectrograph. An Argon laser was used to provide a 515 nm excitation wavelength. The Raman analysis was used to provide information on changes to the bonding structure induced by radiation damage.

3. Light Transmission Results and Analysis

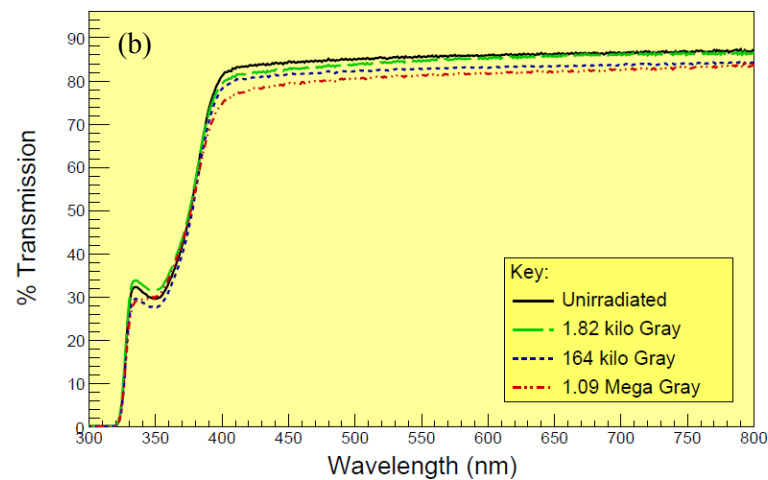
The results of light transmission spectroscopy for each grade relative to light transmission in air are shown in figure 1. The absorptive edge completely falls away by 320 nm in both EJ200 and EJ208, whilst this occurs at 340 nm for EJ260. It is observed that the overall transmission decreases as dose exposure increases. An increase in transmission within the region of 340 nm to 380 nm is observed prominently within EJ200 and could indicate damage to the fluorescent dyes. Furthermore, an absorptive tint is observed to form along the absorptive edge in EJ260 which shifts towards higher wavelengths as dose increases. This could either indicate the formation of free radicals or the development of a competing absorptive process between the primary and secondary fluor dopants. This additional absorption component may reduce the attenuation length of the scintillation light as described [5].

We consider the transmission loss at a wavelength of approximately 450 nm as this corresponds to the peak absorption wavelength of the fiber that these scintillators are coupled to within the Tile Calorimeter. For dose exposure in the range of Mega Grays, 9.8% light transmission loss is observed in EJ200, whilst EJ260 shows an 8% loss and EJ208 shows a 5.5% loss. At this dose, visible yellowing of the samples occur within the irradiated region. A systematic error of 3% in the transmission loss arose due to slight changes of sample position on the sample mount within the spectrophotometer upon repetition of experiments.

Transmission vs Wavelength For EJ200 at different Exposure Doses



Transmission vs Wavelength For EJ208 at different Exposure Doses



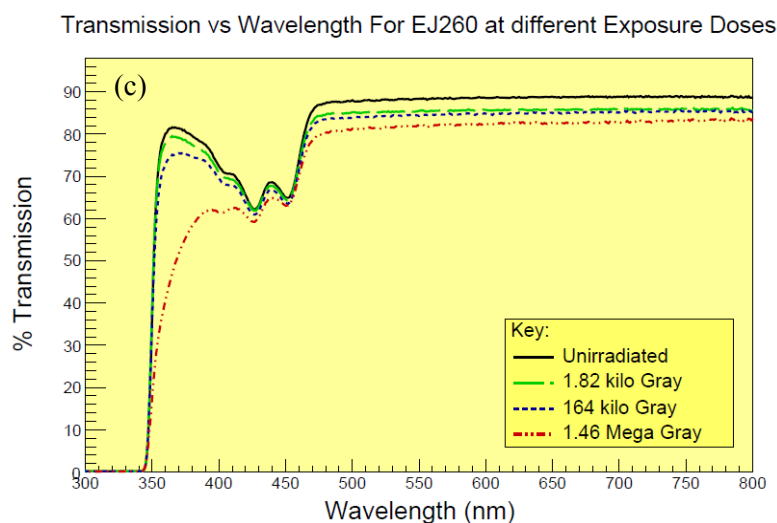


Figure 1: Transmission spectrum for un-irradiated and irradiated samples over several doses relative to air for (a) EJ200, (b) EJ208 and (c) EJ260.

4. Raman Spectroscopy Results and Analysis

Raman spectra were obtained for un-irradiated and Mega Gray dose irradiated samples of each grade. An increased amount of background fluorescence could be observed in the radiation damaged samples which may arise from free radicals interacting with the excitation light. It should be noted that transmission spectra also indicate a loss in scintillator transmission at this excitation wavelength of 515 nm. EJ260 had the most drastic occurrence of background fluorescence as it is a green emitting scintillator and absorbs light at higher wavelengths as compared to EJ200 and EJ208. Plots of the Raman spectra after subtraction of the fluorescent background are shown in figure 2. The Raman peaks observed were then allocated to the characteristic functional groups or vibrational groups using the Raman Peak assignment datasheet [6]. The peak assignment is summarized in table 2. In order to assess change to the bonding structure and minimize errors caused from fluctuations in laser intensity and background subtraction, the ratio of the intensity of each peak to the intensity of a control peak on each of their respective spectra were plotted as shown in figure 3. Comparing changes to these ratios then gave an indication to the relative increase or decrease to that particular bonding structure.

Table 2: Allocation of peaks to characteristic functional groups or vibrational groups.

Functional Group/Vibration	Peak Assignment for EJ200 and EJ208	Peak assignment for EJ260
$\delta(\text{C-C})$ aliphatic	1-2	1
$\nu(\text{C-C})$ alicyclic or aliphatic chain vibrations	3-5, 7-9	2-4, 6-7
$\nu(\text{C-C})$ aromatic ring chain vibrations	6	5
$\delta(\text{CH}_3)$	10	/
$\delta(\text{CH}_2)$ or $\delta(\text{CH}_3)$ asymmetric	11	8
$\nu(\text{C=C})$	12	9
$\nu(\text{C-H})$	14	11-12
$\nu(=\text{C-H})$	15	13-14
Control peak	13	10

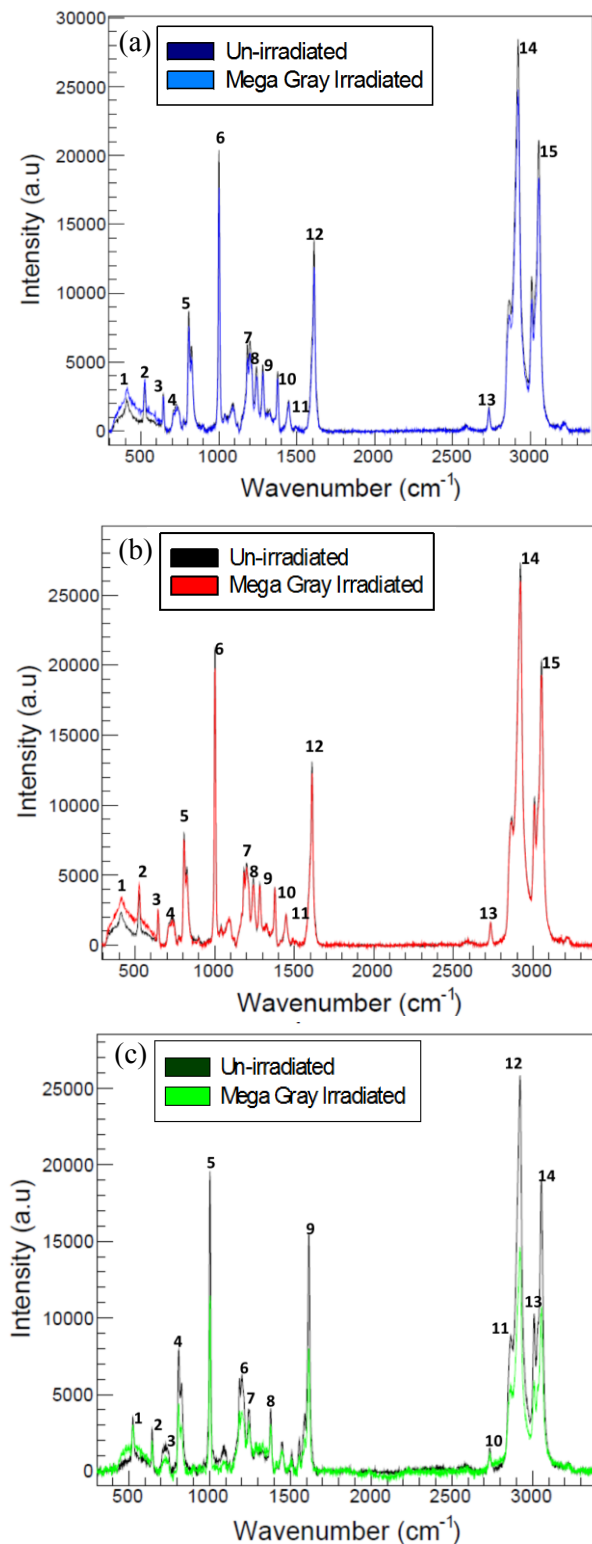


Figure 2: Background subtracted Raman spectra for (a) EJ200, (b) EJ208 and (c) EJ260.

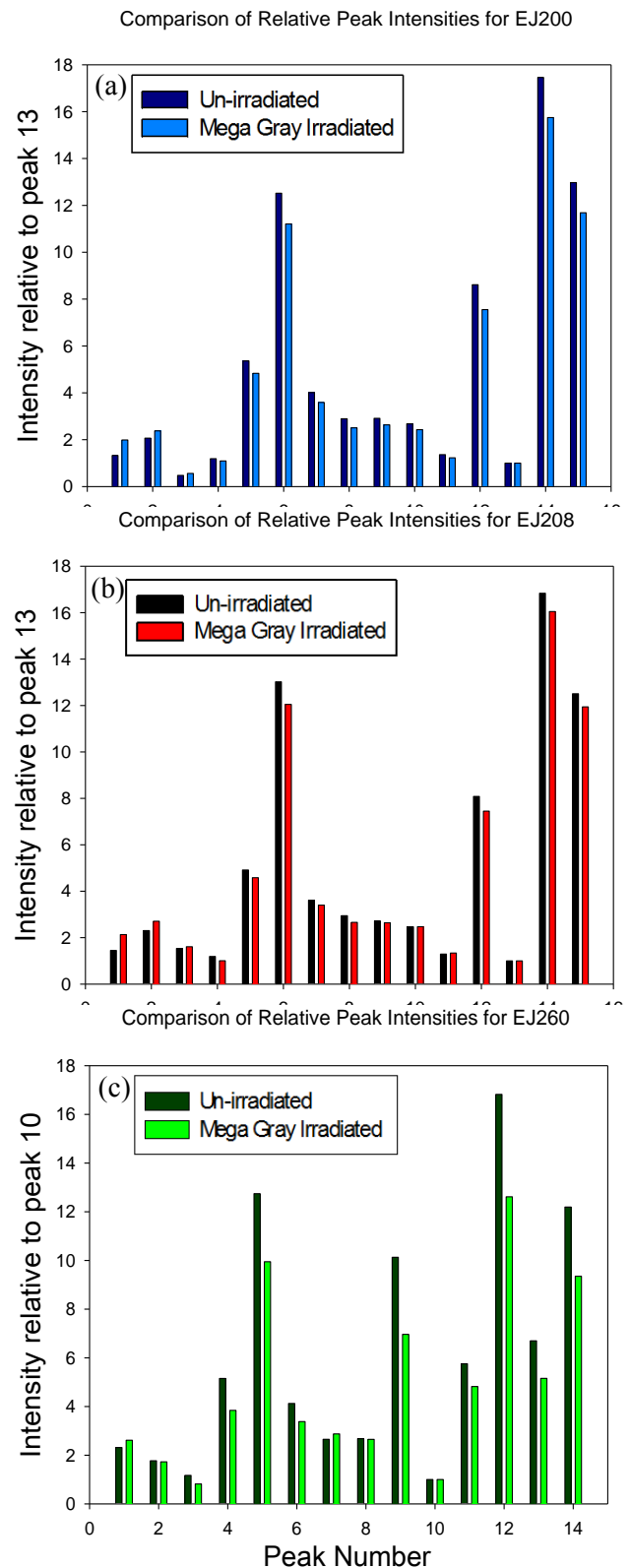


Figure 3: Plots of intensity ratios between peaks and control peak for (a) EJ200, (b) EJ208 and (c) EJ260.

The spectra for irradiated and un-irradiated samples are fairly similar, with a few subtle changes. No broadening of peaks or the formation of new peaks were observed. All three samples show an increase in the C-C aliphatic or alicyclic vibrations indicated by peaks 1 and 2 in EJ200 and EJ208, and peaks 1 and 7 in EJ260. These bonds occur in chains not belonging to the aromatic benzene ring structure. All the other peaks show a decrease, with the most prominent occurring for peaks 6, 12, 14 and 15 in EJ200 and EJ208 and peaks 4, 5, 9 and 11-14 in EJ260. These typically represent the C-C aromatic ring bonds, as well as C=C, C-H and $\text{C}=\text{C}-\text{H}$ which could belong to either the benzene ring or the vinyl backbone. Damage to the benzene ring structure directly affects the scintillation mechanism which would lead to a degradation in scintillation light yield. The Mega Gray irradiated doses of each scintillator grade contain more carbon type species and less hydrogen type species than the un-irradiated samples. This indicates that hydrogen degassing occurs when the C-H bonds are broken.

5. Conclusions

According to the results found during this investigation, exposure to proton irradiation leads to a decrease in light transmission as radiation dose exposure is increased. Overall, EJ200 exhibited the most loss in transparency at a wavelength of 450 nm when irradiated with a Mega Gray range dose, and EJ208 showed the least transmission loss. Although EJ260 showed less transmission loss than EJ200 at 450 nm, the development of an absorptive tint shift could indicate additional damage effects such as free radical formation. Other features observed on the spectra such as the formation of additional transmission peaks and dips indicate that radiation damage could be affecting the light transfer mechanism. This would greatly affect the scintillator light conversion efficiency and hence light yield studies will need to be conducted.

Surface scratches and structural damage caused during the sample preparation procedure could have caused minor additional transmission loss and therefore provide an experimental error. The extent of this random error could not be gauged accurately as each sample was polished separately and most of this damage was not directly visible to the naked eye. A more gentle polishing procedure incrementing over more polishing stages and finishing with a finer polishing stage, as well as using a larger sample group could reduce this effect in future studies conducted. It would also be imperative that transmission tests are conducted on individual samples before and after irradiation as opposed to comparing transmission of irradiated samples to that of an un-irradiated control sample.

Raman analysis indicates that structural changes occur within the samples. The species of C-C aromatic, C=C, C-H and $\text{C}=\text{C}-\text{H}$ bonds in the benzene ring structure and vinyl backbone decrease whilst a small amount of additional C-C aliphatic bonds are formed when the scintillators are exposed to a dose in the Mega Gray range. This indicates that hydrogen degassing occurs and the damaged scintillators become more carbon rich. The observation of an increased background fluorescence in the Raman spectra of the Mega Gray irradiated samples of each grade is further indicative of free radical formation. This is particularly emphasized in EJ260.

From this preliminary study, EJ208 appears as a good potential candidate as it suffers the least transmission loss and structural differences in its Raman spectra are less apparent than in EJ200 and EJ260, however further investigation is required. This study has provided a good baseline for the experimental techniques required for investigating the radiation damage of plastic scintillators and improvements will be implemented for further studies to come.

6. Upcoming work

Several scintillator samples obtained from Bicron, the Joint Institute for Nuclear Research (Dubna) as well as samples manufactured particularly for the Tile Calorimeter in Protvino, will be added to the study. New samples have been polished to dimensions of 5 mm x 5 mm x 0.35 mm to prevent inhomogeneity resulting from machining stress. An analysis of the changes in light response toward beta electrons from a strontium source is underway. Light transmission and light yield testing after

irradiation will be conducted at several times over the course of a year in order to establish long term changes to the optical properties due to the chemical reactions which may arise from free radical interactions with the environment. A larger sample group will be looked at with several samples being irradiated per dose in order to ensure repeatability in the effects observed. Electron paramagnetic resonance studies conducted by C. Pelwan *et al.* [7] are also underway in order to determine types of free radical species that may form as a result of radiation damage. An investigation into the damage undergone by previous scintillator tiles used within the Tile Calorimeter during the previous data taking period at ATLAS is also ongoing by L. Maphanga, *et al.* [8].

7. References

- [1] ATLAS Collaboration 2008 *The ATLAS Experiment at the CERN Large Hadron Collider* (IOP Publishing and SISSA)
- [2] Knoll G F 1999 *Radiation Detection and Measurement, Third Edition* (Michigan: John Wiley & Sons Inc.) chapter 8 pp 220-222
- [3] ELJEN Technology 2013 *Products: Plastic Scintillators* [Online]. Available: <http://www.eljentechnology.com/index.php/products/plastic-scintillators>.
- [4] Ziegler J 2013 *SRIM and TRIM : Particle interactions with matter* [Online]. Available: <http://www.srim.org>.
- [5] Zorn C 1993 A pedestrian's guide to radiation damage in plastic scintillators *Radiat. Phys. Chem.* **41** 37-43
- [6] G. Socrates 2004 *Infrared and Raman Characteristic Group Frequencies Tables and Charts, Third Edition*, (London: John Wiley and Sons) pp 35-47
- [7] Pelwan C, *et al.* 2014 *Electron Paramagnetic Resonance Analysis of Plastic Scintillators for the Tile Calorimeter of the ATLAS detector* (submitted to the SAIP 2014 Conference)
- [8] Maphanga L, *et al.* 2014 *Characterization of Damage in in-situ Radiated Plastic Scintillators at the Tile Calorimeter of ATLAS* (submitted to the SAIP 2014 Conference)

Coupled-channel studies of nucleon scattering from oxygen isotopes

S Karataglidis^{1,2}, J P Svenne³, K Amos^{2,1}, L Canton⁴, P R Fraser^{4,2},
and D van der Knijff²

¹ Department of Physics, University of Johannesburg, P.O. Box 524 Auckland Park, 2006, South Africa

² School of Physics, University of Melbourne, Victoria 3010, Australia

³ Department of Physics and Astronomy, University of Manitoba, and Winnipeg Institute for Theoretical Physics, Winnipeg, Manitoba, Canada R3T 2N2

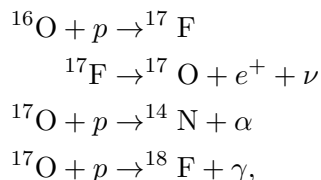
⁴ Istituto Nazionale di Fisica Nucleare, Sezione di Padova, Padova I-35131, Italia

E-mail: ¹stevenka@uj.ac.za

Abstract. The Multi-Channel Algebraic Scattering (MCAS) method for the description of nucleon-nucleus scattering has been used with a (collective) rotational model of structure describing the target. The success of that model, when incorporating the Pauli Principle in the interactions describing the scattering and the formation of the compound systems, has been quite good. We extend that method to include the vibrational model in describing the target states, and apply the method to the scattering of low-energy nucleons from oxygen isotopes. Preliminary results for neutron scattering from ¹⁶O, leading to states in ¹⁷O will be reported.

1. Introduction

The structures of ¹⁷O and ¹⁷F play an important role in the synthesis, in the stellar environment, of elements beyond carbon. Once carbon burning begins, with successive proton capture reactions [1] and β -decays, elements to nitrogen and oxygen are formed, as well as fluorine. Of particular note within the CNOF cycle are the reactions [2]



the last reaction being the breakout from the CNO cycle. Proton capture on ¹⁶O can proceed via either the ground state of ¹⁷F, capture to the $d_{5/2}$ level, which has a Q value of 600 keV, or via the first excited state, capture to the $s_{1/2}$ level, with a Q value of 105 keV. It is likely to proceed via the first excited state, as that state is a proton halo [3].

Yet the two nuclei are of interest for their own sake. They are each a single nucleon outside an ¹⁶O core, and are mirror nuclei, with the first three positive-parity states reflecting the single particle energies of the $0d_{5/2}$, $1s_{1/2}$, and $0d_{3/2}$ levels in the sd -shell model. But the model for each nucleus is not so simple: in a $(0 + 2)\hbar\omega$ prescription, there is significant admixing of $2\hbar\omega$

components, $\sim 25\%$, in the ground states. This largely stems from 2p-2h components giving rise to additional nucleons in the sd shell. With this in mind, it is instructive to compare the extreme shell model picture, with one particle in the sd shell, or the more general $(0 + 2)\hbar\omega$ model, to the collective model description contained in the Multi-Channel Algebraic Scattering theory (MCAS) [4], which describes low-energy nucleon-nucleus scattering, and the resonances (both bound-state and scattering) in the compound nucleus. For the mass-17 nuclei, the model would entail coupling a single nucleon to the ^{16}O core. This is, however, not a trivial exercise, as it is well-known that the description of the spectrum of ^{16}O requires a $4\hbar\omega$ shell model at the minimum [5, 6, 7].

2. MCAS

MCAS has been described in detail elsewhere [4], and so we present a brief description for the purposes of highlighting those aspects which are of relevance to the present calculation. The method is a means of solving the coupled-channel Lippmann-Schwinger (LS) equations, describing a coupled two-cluster system, in momentum space. For the most part, it is assumed that the two clusters are a nucleon and an even-even nucleus. It is also assumed that the target (core) is described by a collective model, which allows for the specification of the matrix of interaction potentials defining the coupled-channel problem. Nothing is assumed of the spectrum of the target states. The matrix of potentials is expanded in terms of sturmian functions, and a finite set of ~ 30 are used to ensure convergence. Pauli-exclusion is handled by the use of orthogonalising pseudo-potentials (OPP) with weights of 10^6 MeV to guarantee that the sturmians are orthogonal to any states corresponding to a nucleon coupling to a filled orbit in the target [8, 9].

Once the sturmians and OPPs are set, the matrix of potentials is re-expressed as a sum of separable potentials in momentum space as input to the LS equations. As the equations are expressed in momentum space, solutions of the LS equations may be found for both bound (negative energy) and scattering (positive energy) states in the compound system. Energies and widths are found for the states in the compound system, assuming that the channels specified are only those of a nucleon coupled to a target nucleus, in which case the widths are partial widths corresponding only to nucleon emission from the compound nuclear states.

3. Shell model considerations

While the spectrum of ^{16}O requires a full $4\hbar\omega$ shell model for description [5], the ground state found by Brown and Green is dominated by $0\hbar\omega$ and $2\hbar\omega$ components which correspond approximately to those found from a pure $(0 + 2)\hbar\omega$ shell model [7]. In that respect, we may calculate the spectra of ^{17}O and ^{17}F in a $(0 + 2)\hbar\omega$ model space, for the positive parity states, and a $(1 + 3)\hbar\omega$ model space for the negative parity states. In both sets of calculations all shells from the $0s$ to the $0f1p$ are used, with all particles active. We calculate the spectrum using OXBASH [10] with the WBP interaction of Warburton and Brown [11]. Any calculations of transition rates between states in ^{16}O to indicate the strength of the couplings, however, will require a full $4\hbar\omega$ model space calculation to ensure a proper description of the ^{16}O spectrum; that is work in progress. The resultant spectrum, together with the known spectra for ^{17}O and ^{17}F [12], is shown in Fig. 1.

It is clear that the spectrum obtained from the shell model compares well with both spectra. Discrepancies between the model spectrum and the known spectra may be due to limitations in the model space and/or the underlying limitation on the ground state of ^{16}O . Nevertheless, this result serves to illustrate that the extreme single-particle picture of the mass-17 system is too simplistic. It points to the need for a coupled-channel description of the nuclei, with as many states in the target (^{16}O) spectrum included in the coupling as possible.

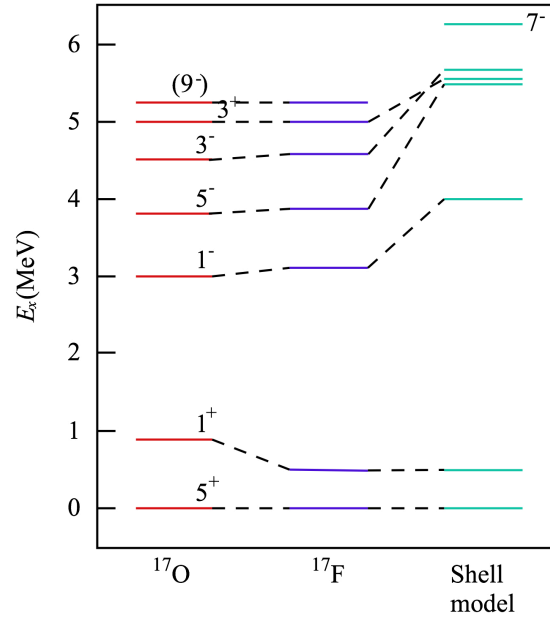


Figure 1. Spectra for ^{17}O and ^{17}F [12], with zero energy corresponding to the ground states of each. The state labels denote $2J^\pi$.

4. MCAS considerations

The MCAS has been applied to $n + ^{16}\text{O}$, leading to compound states in ^{17}O , using the vibrational model and 5 target states in ^{16}O , namely the ground state, the 0_2^+ state (6.049 MeV), the 3_1^- state (6.130 MeV), the 2_1^+ state (6.917 MeV), and the 1_1^- state (7.117 MeV). The parameters for the Woods-Saxon functions used in the calculation are $R_0 = 3.15$ fm, and $a = 0.65$ fm. The potential parameters (positive and negative parity) are, in units of MeV,

$$\begin{aligned} V_0^- &= -47.50; V_0^+ = -50.50 \\ V_u^- &= 2.55; V_u^+ = 0.00 \\ V_{ls}^- &= 6.90; V_{ls}^+ = 7.20 \\ V_{ss}^- &= 2.50; V_{ss}^+ = -2.0, \end{aligned} \quad (1)$$

with two deformation parameters, $\beta_2 = 0.21$ and $\beta_3 = 0.42$. To obtain the spectrum of ^{17}F , we add a Coulomb potential. Also, while Pauli blocking of the $0s$ and $0p_{3/2}$ orbits have been incorporated by the addition of the OPP, Paul hindrance [9] has also been included for the higher orbits. In particular, this includes the $0p_{1/2}$ orbit, to account for the ground state of ^{16}O being $4\hbar\omega$ in character. It should be noted that the results presented herein from these calculations are preliminary.

Fig. 2 shows the spectra of ^{17}O and ^{17}F as compared to the result obtained from MCAS. Agreement with the known spectra is quite good, with the low-lying states well-reproduced. Comparison to Fig 1 shows the results from MCAS agree well also with the results from the shell model. Above the nucleon thresholds, the density of states make identification of states difficult, though the trends in the groupings of states in the known spectra are reproduced. Changes to the Coulomb potential parameters have little influence on the spectrum of ^{17}F .

The low-energy neutron scattering cross section from ^{16}O is shown in Fig. 3. It is clear from Fig. 3 that the cross section to 300 keV obtained from MCAS is in agreement with the

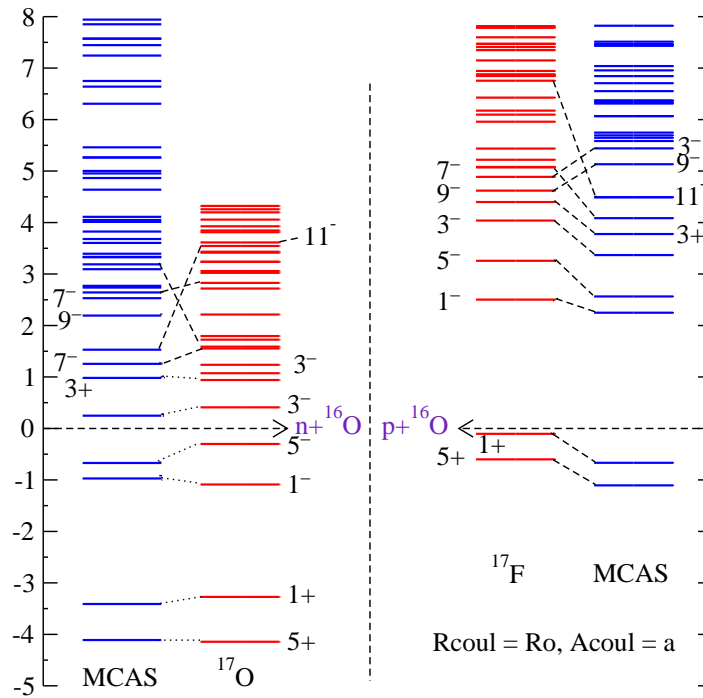


Figure 2. Spectra of ^{17}O and ^{17}F obtained from MCAS, as compared to the known spectra. Notation is as for Fig. 1, and the zero energy corresponds to the nucleon scattering threshold.

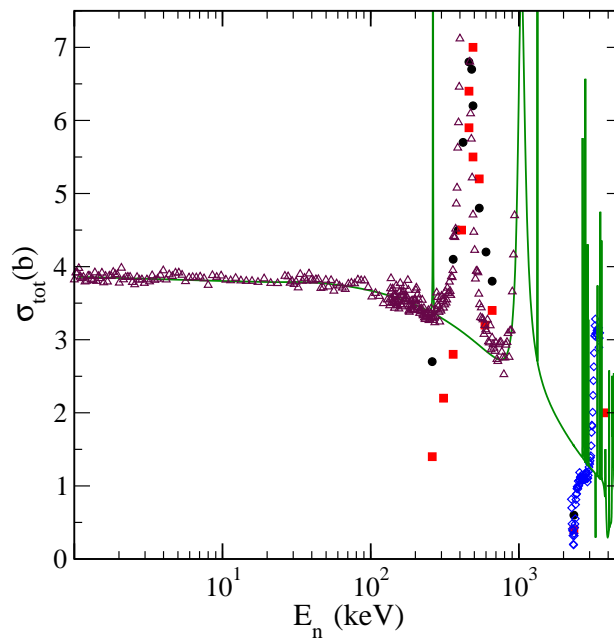


Figure 3. Low-energy neutron scattering cross section from ^{16}O , showing states in ^{17}O . Note that the energy is on a log-scale.

data. However, the resonance observed at 450 keV is not in the correct position; the MCAS result places it at 1 MeV. That resonance corresponds to the $3/2^-$ state in ^{17}O . Further work is continuing in order to understand this discrepancy. Work is also continuing on the low-energy proton scattering cross sections with regards to states in ^{17}F .

5. Conclusions

We have calculated the spectra of ^{17}O and ^{17}F from the shell model and also from MCAS. Results of both models agree well with the known spectra. However, preliminary results for the neutron scattering cross section from ^{16}O , leading to states in the compound ^{17}O does not reproduce the resonances observed. Further work is being pursued to understand this discrepancy. This has to be resolved before investigating the proton scattering cross sections, leading to states in ^{17}F .

SK acknowledges support of the National Research Foundation. JPS acknowledges supports from the Natural Sciences and Engineering Council of Canada (NSERC). PRF acknowledges funds from the Dipartimento di Fisica e Astronomia dell'Università di Padova and the PRIN research project 2009TWL3MX.

References

- [1] Rolfs C 1973 *Nucl. Phys.* **A217** 29
- [2] Arnett D 1996 *Supernovae and nucleosynthesis* (New Jersey: Princeton University Press)
- [3] Karataglidis S and Bennhold C 1998 *Phys. Rev. Lett.* **80** 1614
- [4] Amos K, Canton L, Pisent G, Svenne J P and van der Knijff D 2003 *Nucl. Phys.* **A728** 65
- [5] Brown G E and Green A M 1966 *Nucl. Phys.* **75** 401
- [6] Haxton W C and Johnson C 1990 *Phys. Rev. Lett.* **65** 1325
- [7] Karataglidis S, Dortmans P J, Amos K and de Swiniarski R 1996 *Phys. Rev. C* **53** 838
- [8] Canton L, Pisent G, Svenne J P, Amos K and Karataglidis S 2006 *Phys. Rev. Lett.* **96** 072502
- [9] Amos K, Canton L, Fraser P R, Karataglidis S, Svenne J P and van der Knijff D 2013 *Nucl. Phys.* **A912** 7
- [10] OXBASH-MSU (the Oxford-Buenos-Aries-Michigan State University shell model code). A. Etchegoyen, W.D.M. Rae, and N.S. Godwin (MSU version by B.A. Brown, 1986); B.A. Brown, A. Etchegoyen, and W.D.M. Rae, MSUCL Report Number 524 (1986)
- [11] Warburton E K and Brown B A 1992 *Phys. Rev. C* **46** 923
- [12] TUNL Nuclear Data Evaluation, <http://www.tunl.duke.edu/nucldata/index.shtml>

Firmware development for the upgrade of the Tile Calorimeter of the ATLAS detector

C O Kureba and B Mellado

School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africa

E-mail: Chamunorwa.Kureba@wits.ac.za

Abstract. The Large Hadron Collider is scheduled to be upgraded in the year 2022, in order to increase its instantaneous luminosity. The High Luminosity LHC, also referred to as upgrade Phase-II, means an inevitable complete re-design of the read-out electronics in the Tile Calorimeter in which the completed new read-out architecture is expected to have the front-end electronics transmit fully digitized information of the full detector to the back-end electronics system. The back-end system will provide digital calibrated information with greater precision and granularity to the first level trigger, thereby resulting in improved trigger efficiencies. In Phase II, the current Mobile Drawer Integrity CheckKing (MobiDICK4) test-bench will be replaced by the next generation test-bench for the TileCal super-drawers, the new Prometeo (A Portable ReadOut Module for Tilecal ElectrOnics). The Prometeo's prototype is being assembled by the University of the Witwatersrand and installed at CERN for further developing, tuning and tests. This note presents some details on the design of the Prometeo, with particular emphasis on the firmware development for the Xilinx Virtex-7 XC7VX485T Field Programmable Gate Array (FPGA), which is the backbone of its VC707 motherboard. Some aspects are outlined on the VHDL hardware description language and the Xilinx Integrated Software Environment design suite which are employed in the firmware development.

1. Introduction

At CERN, the European Organization for Nuclear Research, the fundamental structure of matter is being probed by scientists and engineers. The Large Hadron Collider (LHC) accelerates and collides protons, and also heavy ions. The A Toroidal LHC Apparatus (ATLAS) [1] is one of two general purpose detectors used for detecting the sub-atomic particles produced during these high-energy collisions. The Tile Calorimeter (TileCal) [2] is the central hadronic calorimeter of the ATLAS detector. The year 2022 has been scheduled to see an upgrade of the LHC in order to increase its instantaneous luminosity. The High Luminosity LHC, also referred to as upgrade Phase-II, means an inevitable complete re-design of the read-out electronics in the TileCal [3]. The completed new read-out architecture is expected to have the front-end electronics transmit fully digitized information of the full detector to the back-end electronics system. The back-end system will provide digital calibrated information with greater precision and granularity to the first level trigger, thereby resulting in improved trigger efficiencies and background rejection. The upper section of Figure 1 shows a schematic of the current front-end electronics while the lower section depicts the equivalent electronics set-up for Phase-II upgrade. In the current set-up, pipeline memories are used to store digitized data samples before they can be trigger-selected. The read-out electronics systems in each superdrawer are daisy-chained, resulting in

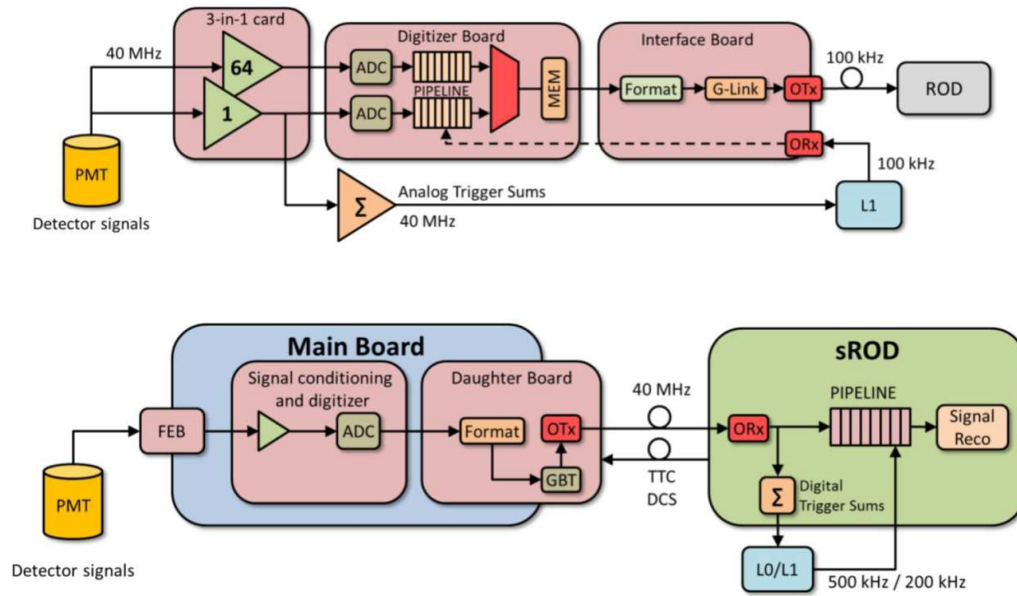


Figure 1. *Top:* Current front-end electronics; *Bottom:* Equivalent electronics for Phase-II upgrade.

the sharing of a data connection to the Read-Out Driver (ROD). In the future electronics design, each superdrawer has been divided into 4 minidrawers. One minidrawer hosts 12 PMTs and 1 daughter board. Each daughter board has one link to the super Read-Out Driver (sROD), thereby reducing failure rate by 25% with respect to ATLAS. There are up to 48 PMTs in one superdrawer, grouped in groups of 6 in a digitizer board. There are 16 digitizers for one superdrawer, interfaced by 1 interface card. It is envisaged that the future electronics will have:

- a super Read-Out Driver (sROD), capable of receiving data at 40 MHz as opposed to the current Read-Out Driver (ROD) which only handles 100 kHz,
- an increase in the number of point-to-point links with the front-end electronics,
- improved radiation tolerance,
- a higher read-out bandwidth due to the need to read-out all sampled data to avoid corruption in the front-end pipeline memories.

An evaluation of this new proposed architecture is currently being carried out in the demonstrator project, where a small fraction of the detector ($1/256$) will be evaluated in test beams and inserted into ATLAS at the next shutdown of the LHC. In Phase II, the current Mobile Drawer Integrity CheckKing (MobiDICK4) system [4, 5] test-bench will be replaced by the next generation test-bench for the TileCal superdrawers, the new Prometeo (A Portable Read-Out Module for Tilecal ElectrOnics) [6]. The MobiDICK system faces challenges against aging and new technologies [7]. The Prometeo is designed to certificate the TileCal front-end electronics by performing multiple tests. The Prometeo's prototype is being assembled by the University of the Witwatersrand and installed at CERN for further developing, tuning and tests.

2. Prometeo

Figure 2 shows a schematic layout of the Prometeo. Prometeo has been designed to have the ability to: read-out all channels at the LHC bunch crossing frequency, assess the quality of data in real-time, diagnose malfunctions in each minidrawer, be self-contained and portable

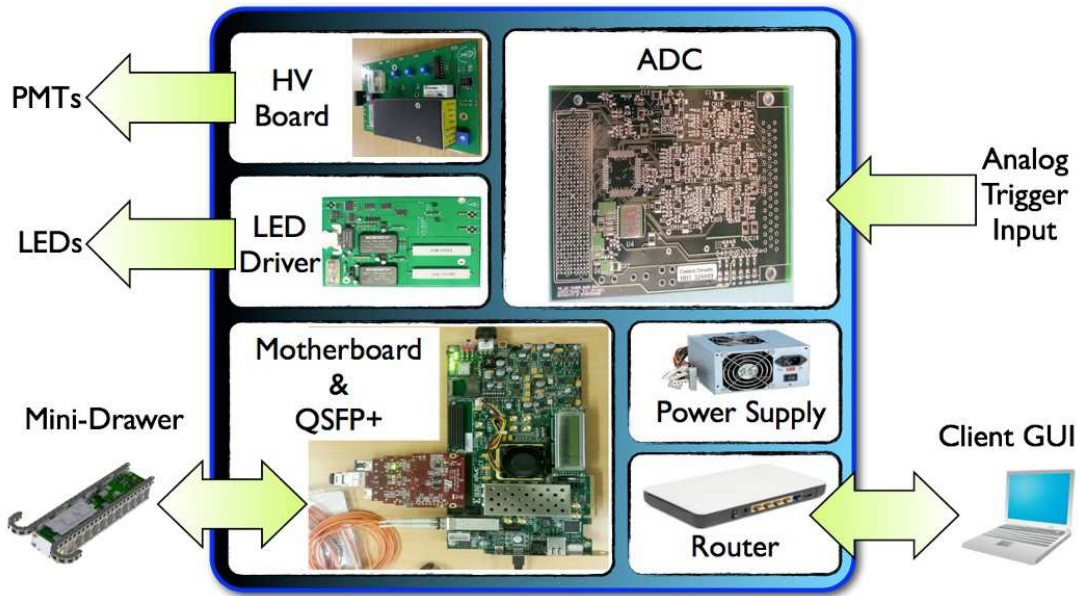


Figure 2. Schematic of the Prometeo.

for maintenance inside the detector, and be low-cost and scalable for network usage. Custom firmware is being developed for the FPGA for it to control and manage the data read-out. The Prometeo consists of several parts as follows:

- Main board: Xilinx VC707, Virtex-7 FPGA chip, 1 GB DDR3 RAM, two FMC connectors, received through SFP on the mainboard,
- Dual QSFP FMC card for digital communication with 2 minidrawers,
- Analogue to Digital Converter (ADC) FMC card: for the digitization of data in the hybrid demonstrator,
- High Voltage (HV) board: to turn on/off the HV and provide the -830 V voltage to power on the photo-multipliers (PMTs),
- Light Emitting Diode (LED) board: to illuminate the PMTs,
- Commercial ATX power supply,
- Ethernet router,
- Aluminum enclosure: whose dimensions are 50 cm \times 35 cm \times 20 cm in terms of length, width and depth, thickness of 3 mm and weight of 8 kg.

The design inherits features from the presently-used MobiDICK4 demonstrator. However, the latter makes use of a Xilinx ML507 evaluation board, with a virtex-5 FPGA. Figure 3 shows the upper and lower sides of the ADC board for Prometeo. This board was recently manufactured in South Africa. The LED and HV boards were also manufactured in South Africa.

3. Firmware development

The Xilinx Integrated Software Environment (ISE) design suite [9] is being used to develop firmware for FPGAs, where the coding language being used is VHDL. VHDL is the VHSIC Hardware Description Language, where VHSIC, in turn, stands for Very High Speed Integrated Circuit, which was a joint program between the US Department of Defense and IEEE in the mid-1980s to research on high-performance Integrated Circuit (IC) technology. The ISE design flow is illustrated in Figure 4. It comprises the following steps: design entry, design synthesis,

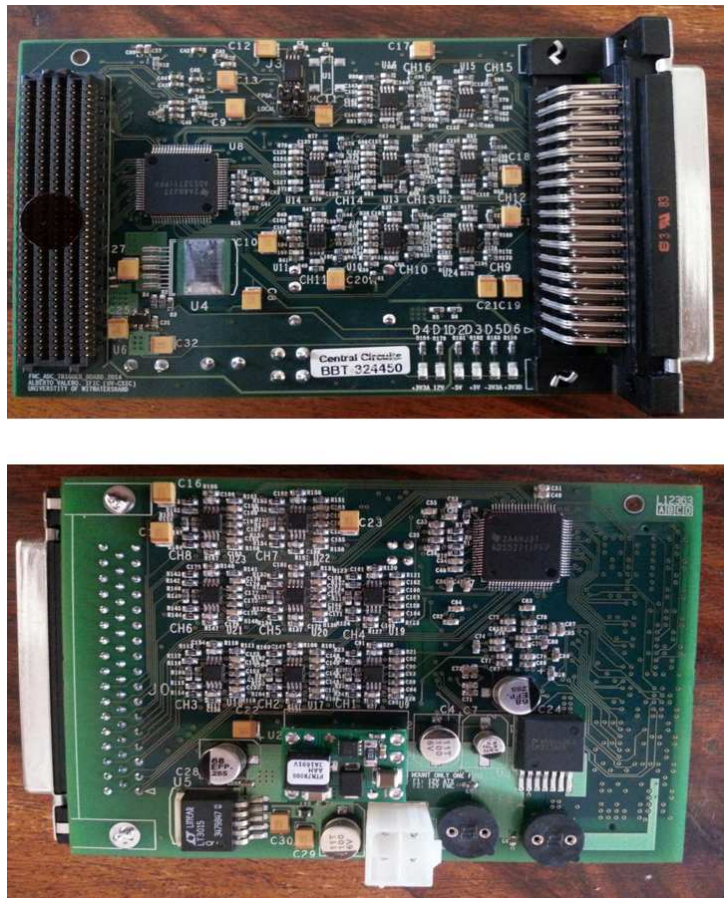


Figure 3. ADC board for Prometeo.

design implementation and Xilinx device programming. The design verification, which includes both functional verification and timing verification, takes place at different points during the design flow.

- (i) **Design Entry:** One creates an ISE project in which files are created and added to the project, including a User Constraints File (UCF). Any existing files are assigned to the project at this stage. Timing constraints, pin assignments, and area constraints are also assigned
- (ii) **Functional Verification:** One can verify the functionality of their design at different points in the design flow as follows: Behavioral simulation (also known as Register Transfer Level (RTL) simulation) is run before synthesis. Functional simulation (also known as gate-level simulation) is run after "Translate", using the SIMPRIM library. After device programming, one runs in-circuit verification. Behavioral simulation employs a high level of abstraction to model the design. A behavioral design might, for example, contain high-level operations, without containing specifics on how the design will be implemented. Synthesis tools then take these behavioral designs and infer the actual gate structures and connections to be used, generating a netlist description. Of the three simulation methods (behavioral, structural, and timing), behavioral simulation runs the fastest but provides the least design information. Behavioral simulation allows you to verify syntax and functionality without timing information. During design development, most verification is accomplished through behavioral simulation. Errors identified early in the design cycle are inexpensive

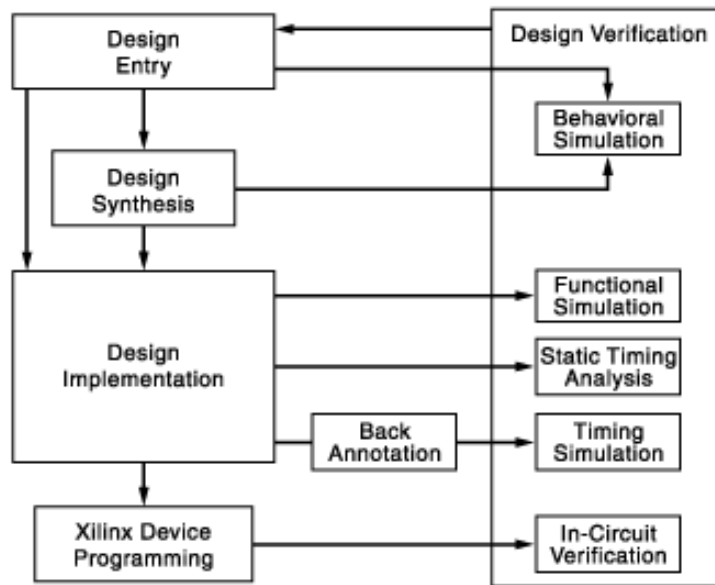


Figure 4. Integrated Software Environment (ISE) design flow.

to fix compared to functional errors identified during silicon debug. After the required functionality is achieved, structural and timing simulation methods can be implemented to obtain more detailed verification data.

- (iii) **Design Synthesis:** After design entry and optional simulation, one runs synthesis. The ISE software includes Xilinx Synthesis Technology (XST), which synthesizes VHDL, Verilog, or mixed language designs to create Xilinx-specific netlist files known as NGC files. NGC files contain both logical design data and constraints. XST places the NGC file in the project directory and the file is accepted as input to the Translate (NGDBuild) step of the Implement Design process.
- (iv) **Design Implementation:** After synthesis, one runs design implementation, which comprises the following steps:
 - Translate - merges the incoming netlists and constraints into a Xilinx design file,
 - Map - fits the design into the available resources on the target device,
 - Place & Route - places and routes the design to the timing constraints,
 - Generation of Programming File - creates a bitstream file that can be downloaded to the device.
- (v) **Timing Verification:** One can verify the timing of their design at different points in the design flow by running static timing analysis after Map and after Place & Route. Timing simulations are performed after Map (for a partial timing analysis of CLB and IOB delays) and after Place and Route (for full timing analysis of block and net delays)
- (vi) **Xilinx Device Programming:** This is done through creating a programming file (BIT) to program the target FPGA. The next step is to generate a PROM or ACE file for debugging or to download to one's device. Optionally, one may also create a JTAG file. Finally, one uses iMPACT to program the device with a programming cable.

The Prometeo firmware is being designed to perform both linearity and stability tests on each of the following:

- Charge Injection System (CIS)

- High Voltage (HV)
- Integrator
- Pedestal

These results are used for diagnosis of faulty PMTs, 3 in 1 cards, mainboards, daughter boards, optical links and analogue cards.

4. Conclusion

The Prometeo's prototype is being assembled by the University of the Witwatersrand and installed at CERN for further developing, tuning and tests. Firmware development for the Virtex-7 FPGA in Prometeo is ongoing, having been successfully completed for the Virtex-5 FPGA in the mobiDICK4 system and also for the Virtex-6 FPGA.

References

- [1] ATLAS Collaboration 2008 The ATLAS Experiment at CERN Large Hadron Collider *JINST* **3** S08003
- [2] Gangiobbe V and Calvet D 2010 Readiness of the ATLAS Tile Calorimeter for LHC collisions *Eur. Phys. J. C* **70**
- [3] Carrio F *et al* 2013 The sROD module for the ATLAS Tile Calorimeter Phase-II upgrade demonstrator *ATL-TILECAL-PROC-2013-020*
- [4] Calvet D and Gangiobbe V 2008 Performance of the TileCal super-drawers from a global analysis of the MobiDICK tests *ATL-TILECAL-PUB-2008-007*
- [5] Carrio F *et al* 2013 Design of an FPGA-based embedded system for the ATLAS Tile Calorimeter front-end electronics test-bench *ATL-TILECAL-PROC-2013-017*
- [6] Ruan X and Mellado B 2014 Prometeo: The next generation test-bench for the Tile Calorimeter of the ATLAS detector, submitted to the SAIP
- [7] Solans C *et al* 2014 Computing challenges in the certification of ATLAS Tile Calorimeter front-end electronics during maintenance periods *J. Phys.: Conf. Ser.* **513** 012035
- [8] Xilinx 2014 VC707 Evaluation Board for the Virtex-7 FPGA, User Guide UG885 (v1.4) May 12
- [9] Xilinx 2013 ISE design suite: System Edition, Version 14.7

Search for the Standard Model Higgs boson and its coupling to fermions with the ATLAS detector

Claire A. Lee

University of Johannesburg, South Africa

E-mail: claire.lee@cern.ch

Abstract. The Higgs boson discovered in 2012 by the ATLAS and CMS experiments at the LHC has been shown to couple to bosons through its discovery channels. The emphasis since the discovery has since shifted to measurements of its properties. The direct observation of the coupling of the new boson to fermions is of particular importance for determination whether it is indeed the Higgs boson predicted by the Standard Model. A review of the search for the Higgs boson through its decay into b-quark and muon pairs is given.

1. Introduction

The discovery of a particle with a mass of 125.5 GeV by the ATLAS [1] and CMS [2] experiments in 2012 is one of the great successes of the Large Hadron Collider program (LHC) at CERN. The observation was made through the bosonic decay channels $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ$ and $H \rightarrow WW$, indicating that the new particle was itself a boson. Subsequent measurements of its properties focused on the bosonic decay modes due to the better resolution in those channels, but an important question remained: whether or not the new particle also coupled to fermions.

The non-zero vacuum expectation value of the Higgs field spontaneously breaks the electroweak symmetry of the Standard Model (SM), imparting mass on the weak vector bosons and generating a massive scalar boson in the process. The single Higgs field predicted within the framework of the SM also bestows mass on the fermions through their Yukawa coupling, with a strength proportional to their masses. Evidence of these couplings and measurement of their strengths is therefore necessary for determining whether the newly-discovered particle is indeed the SM Higgs boson.

The dominant Higgs production mechanism at the LHC is gluon-gluon fusion (ggF), a process that proceeds through a top quark loop and is therefore an indirect probe of the coupling of the Higgs to the top quark. Additional production mechanisms are vector boson fusion (VBF) which has a signature of two well-separated jets; associated production with a vector boson (VH)¹ where one can trigger on the resulting signature from the V decay; and associated production with a pair of top quarks ($t\bar{t}H$) whose decays give b-tagged jets, leptons and missing transverse momentum in the detector. Values of the cross sections for a 125 GeV Higgs boson at 7 and 8 TeV centre of mass energy are shown in table 1 below:

¹ also called Higgs-strahlung

Process	$\sigma_{7 \text{ TeV}}$ (pb)	$\sigma_{8 \text{ TeV}}$ (pb)
ggF	15.13	19.27
VBF	1.22	1.58
WH	0.58	0.70
ZH	0.34	0.42
$t\bar{t}H$	0.09	0.13

Table 1: Production cross sections for 125 GeV Higgs boson at 7 and 8 TeV centre of mass energy at the LHC.

The final production mechanism, $t\bar{t}H$, allows us to directly probe the top Yukawa coupling via the various decay channels, though it has the smallest production cross section. Coupling of the Higgs to the down-type fermions, however, requires direct observation of their decays, though these channels are particularly challenging at hadron colliders, and require very high performance reconstruction of objects (electrons, muons, tau leptons, jets and the missing transverse energy). The ATLAS experiment has performed searches in the $H \rightarrow \tau\tau$ [3], $H \rightarrow b\bar{b}$ [4, 5], and $H \rightarrow \mu\mu$ [6] fermionic decay channels. The $H \rightarrow \tau\tau$ channel is covered in detail by K. Bristow in this same conference, and an overview of the $H \rightarrow b\bar{b}$ and $H \rightarrow \mu\mu$ analyses is given here.

2. The ATLAS Detector

The ATLAS detector [7] is a multipurpose system of particle detectors with a forward-backward symmetric cylindrical geometry and near 4π solid angle coverage. It is composed of 3 core systems arranged in a barrel-plus-endcaps format². The inner detector (ID), covering the pseudorapidity range $|\eta| < 2.5$ ³, consists of a silicon pixel detector, a silicon microstrip detector (SCT) and a transition radiation tracker (TRT) within the $|\eta| < 2$ range. The ID is surrounded by a thin superconducting solenoid providing a 2 T magnetic field. A high granularity lead/liquid argon (LAr) sampling electromagnetic calorimeter covers the region $|\eta| < 3.2$ while a steel/scintillator-tile calorimeter provides hadronic coverage within $|\eta| < 1.7$. A Copper-LAr hadronic calorimeter is used in the end-cap region $1.5 < |\eta| < 3.2$. In the forward region, $3.1 < |\eta| < 4.9$, a copper-LAr electromagnetic calorimeter and a copper/tungsten-LAr hadronic calorimeter for a full energy measurement. Finally, the muon spectrometer (MS) surrounds the calorimeters, consisting of three large air-core superconducting toroid magnets, precision tracking chambers providing accurate muon momentum measurement out to $|\eta| < 2.7$, and additional detectors for muon triggers in the region $|\eta| < 2.4$.

3. Search for $VH \rightarrow b\bar{b}$

At 57.7%, $H \rightarrow b\bar{b}$ has the highest branching ratio of all Higgs decays for a 125 GeV Higgs boson, but only associated production channels are possible at hadron colliders due to the large multijet background. At ATLAS we focus on the VH and $t\bar{t}H$ channels separately; here we discuss the $VH \rightarrow b\bar{b}$ results, with the $t\bar{t}H \rightarrow b\bar{b}$ results in section 4.

The $VH \rightarrow b\bar{b}$ analysis covers both 2011 and 2012 LHC data, with 4.7 fb^{-1} at 7 TeV and 20.3 fb^{-1} at 8 TeV. The analysis follows a cut-based approach, with a profile likelihood fit to the m_{bb}

² The barrel-region detectors are arranged in concentric cylinders parallel to the beam axis, while the endcap detectors are disks oriented perpendicular to the beam axis.

³ In ATLAS, the positive x -axis is defined as pointing from the interaction point to the center of the LHC ring, the positive y -axis is defined as pointing upwards, and the positive z -axis corresponds to protons running anticlockwise. The polar angle θ is measured from the beam axis (z -axis), the azimuthal angle ϕ is measured in the transverse (xy)-plane, and the pseudorapidity is defined by $\eta = -\log(\tan(\theta/2))$.

invariant mass, and events are categorised by the number of high p_T isolated leptons into the 0,1, or 2 lepton channel, depending on the vector boson decay channel ($Z \rightarrow \nu\nu$, $W \rightarrow \ell\nu$, and $Z \rightarrow \ell\ell$ respectively). All selected events for the analysis require exactly two b-tagged jets, with a p_T of 45 (20) GeV for the leading (sub-leading) jet. At most there may be one extra jet in the event.

In the 0 lepton channel, no leptons passing the "loose" identification criteria are allowed, with the $E_T^{\text{miss}} > 120$ GeV, track-based $p_T^{\text{miss}} > 30$ GeV, and a directional requirement matching the E_T^{miss} and p_T^{miss} .⁴ For the 1 lepton channel, the lepton is required to pass the "tight" identification criteria, a $E_T^{\text{miss}} > 25$ GeV cut is applied, and the transverse mass of the W boson is required to be between 40 and 120 GeV. In the 2 lepton case we require 1 medium and 1 loose leptons of same flavour but opposite sign, a dilepton invariant mass within 18 GeV of the Z boson mass, and $E_T^{\text{miss}} < 60$ GeV. Control regions are defined as events with only one b-tagged jet, or events with two different-flavour leptons (top control region).

Events are sub-categorised in bins of vector boson p_T and jet multiplicity, with topological requirements to reduce the backgrounds in the different bins individually and thus increase overall sensitivity. A global fit is performed on the $b\bar{b}$ invariant mass distributions in the signal regions and the event yields in the different control regions. Systematics contributing to the overall uncertainties originate from $t\bar{t}$ background modelling uncertainties, the two to three jet ratio, vector boson p_T , c-jet tagging efficiencies, multijet normalisations and signal acceptance, and amount to about 3% on the background and 12% on the signal after the fit. A separate diboson fit using the same configuration as the Higgs signal is performed to extract the signal strength parameter μ_{VZ} from $VZ \rightarrow b\bar{b}$ decay and validate the analysis method, with a result of $\mu_{VZ} = 0.9 \pm 0.2$, consistent with the SM.

Figure 1 shows the 95% CL upper limits [8, 9, 10] on the Higgs boson production cross section in the range 110 - 150 GeV for this channel. No significant excess is observed, with a best-fit signal strength of $\mu = 0.2 \pm 0.5(\text{stat.}) \pm 0.4(\text{syst.})$. The observed (expected) 95% CL upper limit on the $VH \rightarrow b\bar{b}$ signal strength for a 125 GeV Higgs boson is set as 1.4 (1.3) times the SM production cross section.

4. Search for $t\bar{t}H \rightarrow b\bar{b}$

In the $t\bar{t}H \rightarrow b\bar{b}$ channel, a multivariate analysis on 20.3 fb^{-1} of 8 TeV LHC data is performed to discriminate the signal from the overwhelming $t\bar{t} + \text{jets}$ background. Events are categorised by number of high p_T leptons (1 or 2), allowing for events where one of the W bosons decays either hadronically or leptonically (the other W boson is always required to decay leptonically). The events are further classified depending on their jet and b-tagged jet multiplicities; In the single lepton channel nine independent regions are considered, requiring at least 4 jets where at least 2 of which are b-tagged, while six independent regions cover the dilepton channel, where at least 2 jets with at least 2 b-tags are required.

Events are selected that contain either 1 or 2 leptons with p_T greater than 25 GeV (or 15 GeV for the sub-leading lepton in the dilepton channel). Additionally, in the dilepton channel further cuts are applied to suppress contributions from resonances such as the J/ψ , Υ and Z that decay to lepton pairs. No requirement on the E_T^{miss} is made, since after all the selections

⁴ The track-based transverse missing momentum, p_T^{miss} , is the transverse momentum calculated solely from tracks in the ATLAS Inner Detector originating from the hard scatter vertex. It is almost completely decoupled from the calorimetric E_T^{miss} measurement and independent of the effects of pileup, making it a good complement to the E_T^{miss} measurement and a useful tool for rejecting events with fake E_T^{miss} .

the sample is dominated by $t\bar{t}$ + jets in both channels.

Signal-rich regions, depending on their jet multiplicities, are defined according to their S/\sqrt{B} ratio, combining events with both e and μ flavours of leptons. The signal-rich regions are the ($\geq 5j$, $\geq 4b$), ($\geq 6j$, $3b$) and ($\geq 6j$, $\geq 4b$) regions for single lepton events, and ($\geq 4j$, $3b$) and ($\geq 4j$, $\geq 4b$) regions for dilepton events.⁵ In these regions, a neural network (NN) is used to discriminate signal from background, while in the other, signal-depleted regions, the scalar sum of the jet p_T H_T^{had} (jet and lepton p_T H_T in the dilepton channel) is used due to the lower sensitivity. The NNs are built using the NeuroBayes package [11], with 10 variables entering the discriminant to give a close-to-optimal signal-to-background separation. The variables used cover object kinematics, global event variables, event shape and object pair properties, and those used in one region are required to be well described by Monte Carlo simulations in other regions too.

A simultaneous binned likelihood fit to the discriminants of each of the analysis regions (NN output in the signal-rich regions, and $H_T^{(had)}$ in the signal-depleted regions) is performed under the signal-plus-background hypothesis, to extract the signal strength parameter μ (which is allowed to float freely in the fit but required to be equal in all fit regions). Dominant systematics are due to the normalisation of the $t\bar{t}$ + heavy flavour background, flavour tagging efficiencies, $t\bar{t}$ reweighting and the jet energy scale, and amount to around 7% on the signal and 20% on the background after the fit.

The fitted signal strength for a 125 GeV Higgs boson in the $t\bar{t}H \rightarrow b\bar{b}$ channel is found to be $\mu = 1.7 \pm 1.4$, and a signal 4.1 times larger than predicted by the SM is excluded at 95% CL (the expected exclusion limit is 2.6 larger). The post-fit event yields per bin ordered by $\log(S/B)$ are shown for all bins and both channels combined in figure 2, where the signal is normalised to the best-fit value of μ .

5. Search for $H \rightarrow \mu\mu$

Finally, a search spanning the ggF, VH and VBF production mechanisms is conducted on 24.8 fb⁻¹ of 7 and 8 TeV data in the $\mu\mu$ decay channel, the only channel currently available for measuring the Higgs coupling to second generation fermions. This channel provides a clean final state, but suffers from a very small branching ratio (0.0219%) and large SM backgrounds. Exactly two muons are selected in the final state, with $p_T > 25$ (15) GeV for the leading (sub-leading) muon, and events are required to have $E_T^{miss} < 80$ GeV to suppress $t\bar{t}$ and diboson backgrounds.

To maximise sensitivity to the Higgs signal, 7 mutually exclusive categories are defined. Events with two forward jets with an invariant mass greater than 500 GeV, $|\eta_{jet1} - \eta_{jet2}| > 3$, and $\eta_{jet1} \times \eta_{jet2} < 0$ are selected for the VBF category, based on the resulting dijet signature from the VBF production (events with jets identified as originating from a b-quark are vetoed). The events that are not selected for the VBF category are then classified using $p_T^{\mu\mu}$ into low ($p_{T\mu\mu} < 15$ GeV), medium ($15 < p_{T\mu\mu} < 50$ GeV) and high ($p_{T\mu\mu} > 50$ GeV) regions. Each category is then further divided into two categories - central and non-central (based on the $|\eta|$ distributions of the two selected muons) - to exploit the more precise muon p_T measurement in the central region.

A binned maximum likelihood fit is simultaneously performed on the $m_{\mu\mu}$ distribution in

⁵ (xj , yb) implies events with x jets, of which y are b-tagged.

each category, though separate distributions are used for the 7 and 8 TeV data samples. The background PDF for the $p_{T\mu\mu}$ categories is defined as the sum of a Breit-Wigner (BW) function convolved with a Gaussian, and an exponential function divided by $(m_{\mu_1\mu_2})^3$. For the VBF category, the product of a BW and an exponential function is used. For all categories, the mean and width of the BW peak are fixed to the world average values for those of the Z boson, of $m_{BW} = 91.2$ GeV and $\Gamma_{BW} = 2.49$ GeV. The signal PDF is obtained from Higgs MC samples with ggF, VBF and VH processes combined, and is parametrised as a sum of a Crystal Ball and Gaussian PDF.

Main systematic uncertainties for the number of expected signal events include the uncertainty on the $H \rightarrow \mu\mu$ branching ratio ($\pm 7\%$), QCD scale and PDF predictions ($\pm 8\%$ each for ggF) and the ggF uncertainty in VBF ($\pm 22\%$). Experimentally, the dominant sources of uncertainty come from the luminosity estimation, of $\pm 1.8\%$ for 7 TeV data, and $\pm 2.8\%$ for 7 TeV data, and the jet energy scale $\sim 4.5\%$ (VBF).

The invariant dimuon mass spectrum $m_{\mu\mu}$ is shown in figure 3 for both 7 and 8 TeV data combined. No evidence for a narrow peak in the $m_{\mu\mu}$ distribution is observed. A 95% CL upper limit on the signal strength for a 125.5 GeV Higgs boson is set at 7.0 times the SM prediction, where 7.2 is expected. Assuming a Higgs boson mass of 125.2 GeV and the SM production cross section, the 95% CL upper limit on the $H \rightarrow \mu\mu$ branching ratio is set at 1.5×10^{-3} .

6. Conclusions

A search for the Standard Model Higgs boson of mass 125 GeV through its fermionic decay channels has been performed using 7 and 8 TeV data from Run 1 at the LHC. In the $VH \rightarrow b\bar{b}$, $t\bar{t}H \rightarrow b\bar{b}$ and $H \rightarrow \mu\mu$ channels, no significant excesses were observed over the background, with observed (expected) limits of 1.4 (1.3), 4.1 (2.6) and 7.0 (7.2) \times SM respectively. Overall, the results show no significant deviations from the Standard Model, and we look forward to final updated results on the Run 1 analyses and to the higher energy regime that will be probed during the LHC Run 2.

References

- [1] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B 716 (2012) 1, arXiv:1207.7214 [hep-ex].
- [2] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B 716 (2012) 30, arXiv:1207.7235 [hep-ex].
- [3] ATLAS Collaboration, *Evidence for Higgs Boson Decays to the $\tau^+\tau^-$ Final State with the ATLAS Detector*, ATLAS-CONF-2013-108.
- [4] ATLAS Collaboration, *Search for the $b\bar{b}$ decay of the Standard Model Higgs boson in associated (W/Z)H production with the ATLAS detector*, ATLAS-CONF-2013-079.
- [5] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector at the LHC*, ATLAS-CONF-2014-011.
- [6] ATLAS Collaboration, *Search for the Standard Model Higgs boson decay to $\mu + \mu^-$ with the ATLAS detector*, arXiv:1406.7663v1 [hep-ex].
- [7] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 (2008) S08003.
- [8] T. Junk, *Condence level computation for combining searches with small statistics*, Nucl. Instr. Meth. A 434 (1999) 435, arXiv:9902006 [hep-ex].
- [9] A. L. Read, *Presentation of search results: the CLs technique*, J. Phys. G 28 (2002) 2693.
- [10] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C 71 (2011) 1554, arXiv:1007.1727 [physics.data-an].
- [11] Phi-T GmbH, *NeuroBayes package*, <http://neurobayes.phi-t.de/>

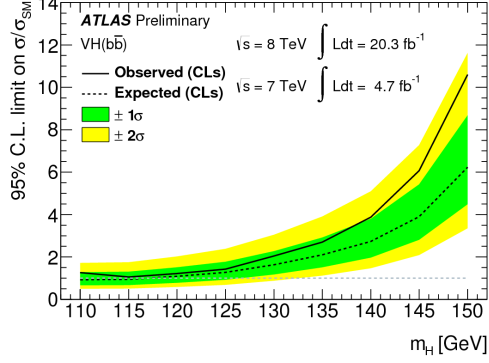


Figure 1: 95% CL upper limits on the Higgs boson production cross section in the range 110 - 150 GeV for the $VH \rightarrow b\bar{b}$ channel.

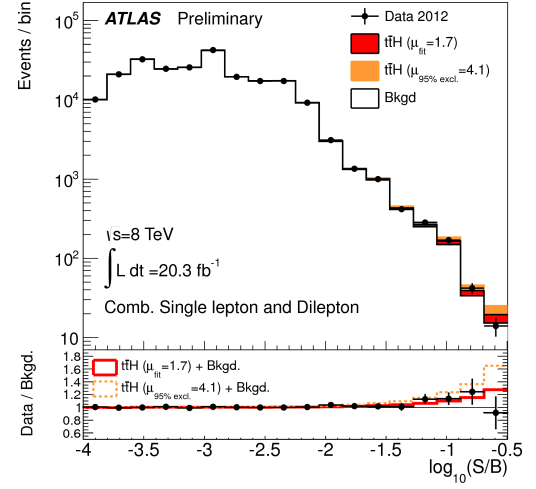


Figure 2: Post-fit $t\bar{t}H \rightarrow b\bar{b}$ event yields per bin ordered by $\log(S/B)$ for all bins and both channels combined. The signal is normalised to the best-fit value of μ .

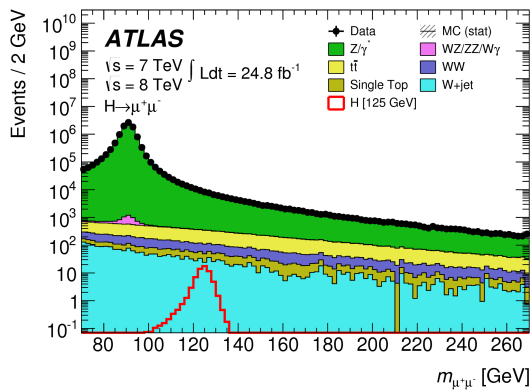


Figure 3: Invariant dimuon mass spectrum $m_{\mu\mu}$ for 7 and 8 TeV data combined.

Characterization of damage in in-situ radiated plastic scintillators at the Tile calorimeter of the ATLAS

L Maphanga¹, R Erasmus¹, H Jivan¹, B Mellado¹, G Peters¹ and E Sideras-Haddad¹

School of Physics, University of the Witwatersrand, 1 Jan Smuts Avenue,
Braamfontein, Johannesburg, South Africa, 2000

Linah.Maphanga@students.wits.ac.za

Abstract. Ukrainian Polystyrene-based plastic scintillator (UPS923A) samples manufactured by the Institute of Scintillating materials in Ukraine were investigated using Raman scattering techniques. Three irradiated samples of the same type (UPS923A) from the Minimum Bias Trigger scintillator (MBTS) taken from three different positions: Top (T1), Bottom 1 (B1) and Bottom 2 (B2) in the ATLAS detector were investigated. And compared to an un-irradiated sample (D1) of the same type and a 40 MegaGy UPS923A irradiated sample. The aim of this investigative project was to understand how the molecular structures of plastic scintillators are damaged due to high energy collisions at LHC using Raman scattering techniques. It was also to observe if there are any variations in the molecular damage of the plastic at different positions in the MBTS. It was found that the Raman spectra of the irradiated samples at three different positions in MBTS are similar in shape and have similar peaks, thus the extent of the molecular damage at this positions is not easily distinguishable between the samples. It was also observed that the intensities of the Raman spectra peaks of irradiated samples are smaller in magnitude than the intensities of the peaks in the un-irradiated sample, thus bond breaking occurred during radiation interactions to decrease the amount of specific species in the molecular structure of the plastic. The 40 MegaGy irradiated sample Raman spectrum shows extensive molecular damage. Using Raman scattering analysis, it was observed that the benzene rings in UPS923A molecular structure was damaged due to radiation.

1. Introduction

The Tile calorimeter of the ATLAS (A Toroidal LHC Apparatus) experiment at CERN, in Geneva uses plastic scintillators as one of its detection systems to detect high energy proton-to-proton collisions happening inside the Large Hadron Collider (LHC). The ATLAS experiment is also used to detect high energy particles such as the Higgs bosons and thus these experiments will help in the understanding of how these particles behave and to determine which role they have played in the formation of the universe. In July 2012 the ATLAS team announced their discovery of these particles.

The ATLAS detector has few different detector systems, some of them are: Minimum Bias Trigger Scintillators, Zero-Degree Calorimeter, Luminosity Cherenkov Integrating detector and the ALFA Roman pots. It thus uses two minimum bias trigger scintillators (MBTS). These MBTS detectors are installed in a forward calorimeter cryostats located approximately 3.6 m away from the collision point [1].

The MBTS detectors are designed in such a way that they contain 16 alternating 2cm polystyrene scintillators separated by metal plates each and connected to a nominal center consisting of two rings (inner and outer) [1].

It is very important to understand how MBTS detectors are damaged by radiation as this would affect the data collected. Therefore studies have been conducted in this subject by ATLAS teams in different countries around the world. New plastic scintillators have been proposed as a replacement of the old one since the ATLAS detector at LHC is undergoing upgrades. As a result a more radiation resistant plastic scintillator has to be chosen as a replacement to improve the quality of the collected data [1].

Polystyrene Based Plastic Scintillators (UPS923A) was manufactured by the Institute of Scintillating Materials in Ukraine. These plastic scintillators have a high radiation hardness compared to PVT (Polyvinyltholuene). And also they do not degrade naturally. They are made of polystyrene which contains 2% PTP (1, 4-diphenylbenzene) and 0.03% POPOP (1, 4-di-(5-phenyl-2-oxazolyl) benzene) [2].

This study was conducted using three irradiated UPS923A samples from the MBTS detector found on the Tile calorimeter and a 40 MegaGy irradiated UPS923A sample. These samples were irradiated in situ. These samples are then compared to an un-irradiated UPS923A sample. This study uses Raman scattering as a tool to understand how the molecular structure of the plastic scintillator have been destroyed. The three irradiated samples were taken from the top and bottom positions in the MBTS detector. This study will help in identifying any differences between the samples in terms of radiation damage in the molecular level of the plastic. Thus this study will help in the ATLAS upgrades as it outlines how UPS923A scintillator becomes damaged due to radiation at different levels in the MBTS detector. This investigation is more concern about the molecular damage of the plastic. Therefore it will help in determining how strong UPS923A scintillator is.

2. Experimental procedure

Raman spectroscopy is one of the light scattering methods used to investigate the molecular structures of materials without disturbing the material's molecular packaging [3].

A Raman microscope called Horiba Lab Ram manufactured by Horiba Scientific through Jobin Yyou technologies was used to get the data for Raman spectra of each sample. An Argon (green) laser with wavelength of 574nm was used to cause Raman-effect in all the samples except for the 40MegaGy sample. For this sample a 785 nm laser was used instead. A silicon sample was used as a reference to find the reference spectrum. The data collected from Horiba Lab Ram microscope was then used to plot the different Raman spectra for all the samples.

3. Raman spectra analysis for each sample

The Raman spectra for each sample were plotted in order to be able to observe the change in the molecular structure of each sample. This will help in the understanding of how UPS923A plastic scintillator responds to radiation. And also to investigate which bonds are broken during radiation and how these changes affect the scintillation mechanism of the plastic scintillator. Thus the Raman spectra for all the irradiated samples are compared to the un-irradiated sample Raman spectrum to investigate the changes that occurred. Also the peak intensities give a clue on how the plastic scintillator becomes damaged due to radiation.

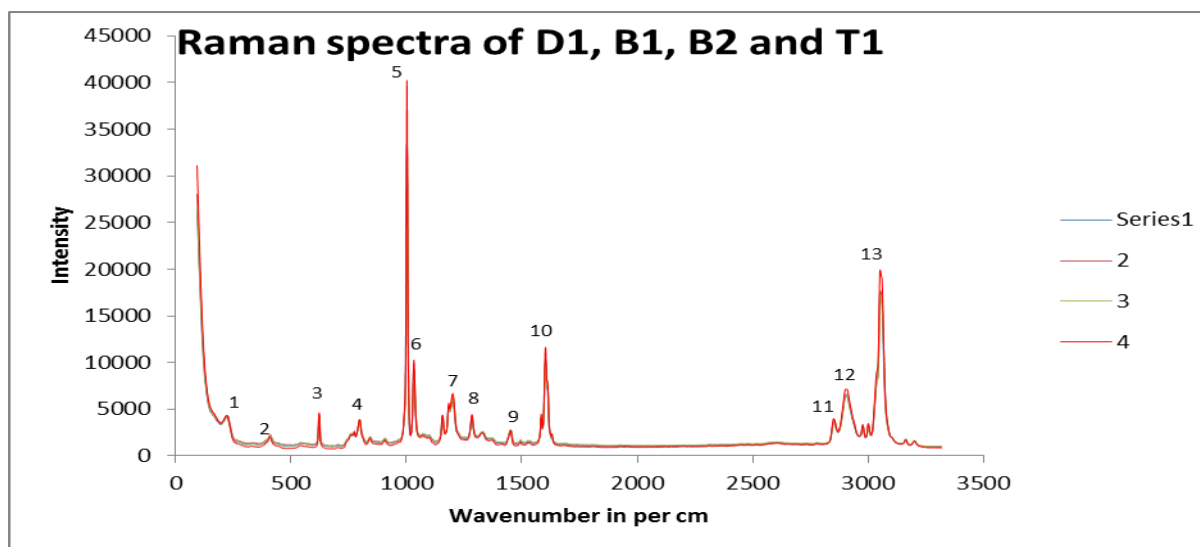


Figure 1. Raman spectra of all irradiated sample from the Tile calorimeter compared to an equivalent un-irradiated sample (see text).

Red represents D1 (un-irradiated sample), Brown represents B2 (2nd bottom irradiated sample), Green represents T1 (top irradiated sample) and Blue represent B1 (bottom irradiated sample). Figure 1 shows the Raman spectra of the irradiated samples and those spectra have numbered peaks intensities which are defined in Table 1.

Using Figure 1, it can be argued that the spectrum of each sample in the MBTS detector, at each position: 1st bottom (B1), 2nd bottom (B2) and top (T1) irradiated samples have similar Raman spectra shape as the Raman spectra of un-irradiated D1 sample. All these samples have 13 peaks in their Raman spectra. It is also noted that the peaks intensities for D1 un-irradiated sample Raman spectrum are larger than that of the irradiated samples, which suggest that some bonds in the plastic scintillator before irradiation were broken.

It can be noted that the molecular structure of irradiated samples and the un-irradiated sample compared is the same. Thus the peaks of D1, B1, B2 and T1 have the same wavenumber and therefore have the same functional group in each sample. But the intensities of the peaks for each irradiated samples are different from the intensities of the un-irradiated sample. Looking at Figure 3, it can be seen that peak 2 of D1 has smaller intensity compared to peak 2 of B1, B2 and T1, which suggest that the irradiated samples have more δ (CC) aliphatic chains compared to the un-irradiated sample.

Looking at peak 5, it can be seen that the intensity of this peak in D1 is very high compared to intensities of the same peak in the irradiated samples. These suggest that there is a depletion of ν (CC) aromatic ring chains after irradiation, thus bonds were broken in the benzene ring. But when B1, B2 and T1 are compared to each other, it is noted that radiation damage is not constant in MBTS detector since their peak intensities are not equal (see Figure 3).

Using Figure 2, it can be noted that the Raman spectrum of UPS923A 40 MegaGy sample is different from the other Raman spectra of the samples from MBTS detector. Again using Figure 2, it can be seen that all the peaks in the Raman spectra ranges from 700-900 cm^{-1} . This region corresponds to the ν (CC) alicyclic, aliphatic chains vibrations region when using Horiba Raman band analysis (refer to [4] and Table 1) for comparison. This suggests that the benzene ring was completely destroyed, thus that the material was extensively damaged by radiation. This also tells us that as the radiation dose is increased, radiation damage increases, thus more bonds are destroyed in the molecular structure of the plastic scintillator.

Table 1. Peak analysis using Horiba Raman Analysis Table [5] for all samples.

Assigned Peak number:	Functional group/vibration	Region In per cm	Peak Wavenumber in per cm	Raman
1,2,6,7,8	$\delta(\text{CC})$ aliphatic chains	250- 400	221.334	strong
3,4	$\nu(\text{CC})$ alicyclic, aliphatic chains vibrations	600- 1300	621.92	medium
5	$\nu(\text{CC})$ aromatic ring chain vibrations	1000	1003.73	Strong/medium
9	$\delta(\text{CH}_2)$ $\delta(\text{CH}_3)$ asym	1400- 1470	1449.52	medium
10	$\nu(\text{C}=\text{C})$	1500- 1900	1603.65	strong
11,12	$\nu(\text{C}-\text{H})$	2800- 3000	2851.61	strong
13	$\nu(=\text{C}-\text{H})$	3000- 3100	3053,72	strong

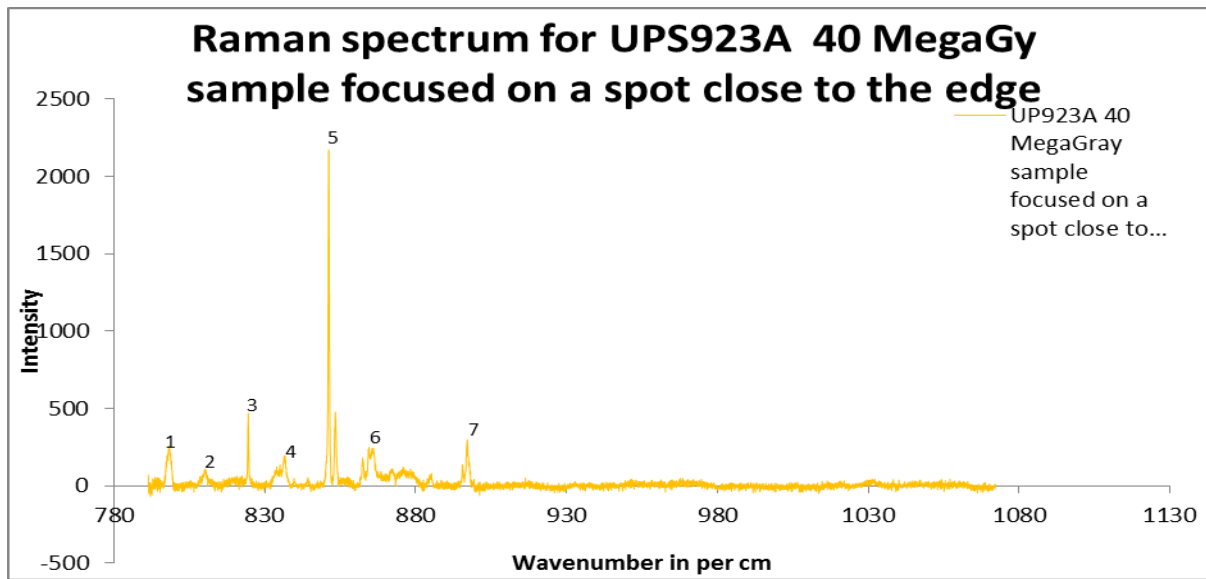


Figure 2. Raman spectrum for UPS923A 40 MegaGy sample.

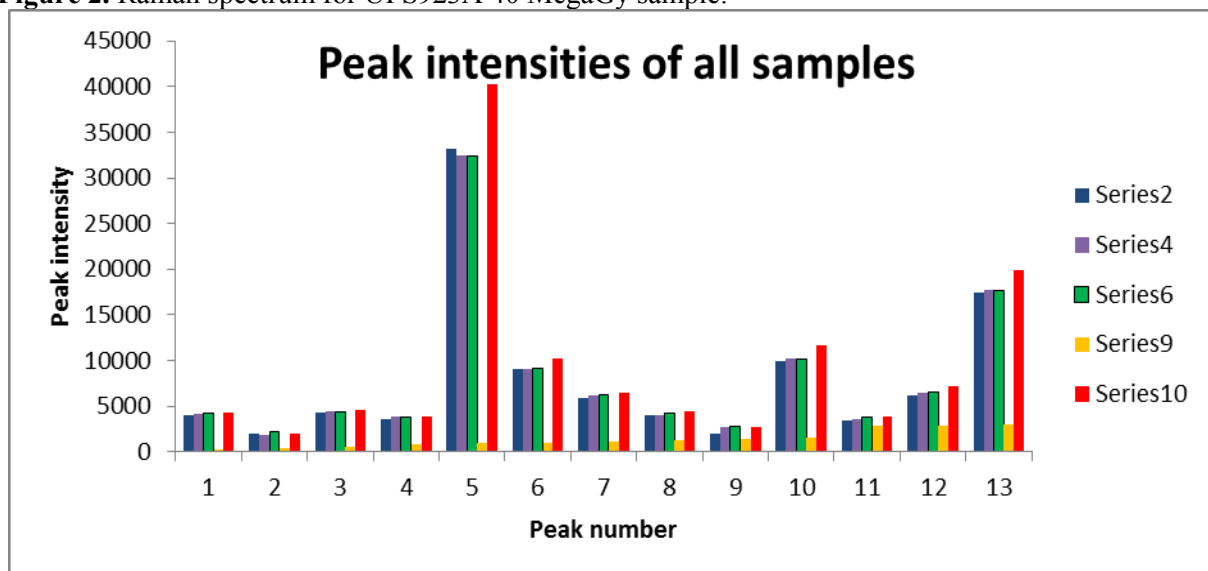


Figure 3. Compared peak intensities of all samples (see text).

Red represent D1 (un-irradiated sample), Blue represent B1 (1st bottom irradiated sample), Purple represent T1 (top irradiated sample) and Orange represent 40 MegaGy irradiated sample.

It can be noted that the peaks intensities of B1, B2 and T1 are not the same, which suggest that radiation damage in the MBTS detector is not constant; it depends on where the plastic scintillator is located. According to Figure 3, the bottom sample (B1 and B2) was damaged more than the top (T1) sample plastic scintillator from MBTS detector. It can be noted that the peak intensities for peak 1,2,3,4,6,7,8,9,10,11,12,13 of B1 are smaller than those of T1 and B2. These suggest that the radiation damage in B1 is greater than the radiation damage of the same type of plastic scintillator at position B2 in MBTS detector.

4. Summary

Radiation damage in the MBTS detector is not constant; it depends on the position of the plastic scintillator relative to action or collision point. This can be supported by the differences observed for the peaks intensities of B1, B2 and T1. Therefore, it has been observed that B1 (bottom irradiated sample) becomes damaged more than the other plastic scintillator samples taken from the top position. This suggests that radiation damage to plastic scintillators depends on the amount of radiation exposure, that is, the higher the radiation exposure the more the plastic scintillator becomes damaged.

By comparing the 40 MegaGy sample with the other samples from MBTS detector, it can be observed that radiation damage destroys the molecular structure of plastic scintillators and the extent of the damage depends on the amount of radiation concern. It is also observed that the benzene ring in plastic scintillator's molecular structures is one of the highly destroyed functional groups and this observation is consistent with theoretical literature. This can be supported by the high difference between the peak intensities of peak 5 for each sample from MBTS detector compared to an un-irradiated equivalent sample D1. This suggests that these samples molecular structures are destroyed due to radiation damage. For the 40 MegaGy irradiated sample, the benzene ring is entirely destroyed due to radiation.

UPS923A plastic scintillator is relatively stronger as it can be seen in the results, that the benzene rings are not completely destroyed. This means that the plastic scintillator can still undergo scintillation process. But this will depend on the amount of radiation exposed.

References

- [1] ATLAS Experiment,[Online].Available: <http://atlas.web.cern.ch/Atlas/Collaboration/>. [Accessed 18 March 2014].
- [2] Artikov A, Budagov A, Chirikov-Zorin I, Chokheli D, Lyablin M, Bellettini G, Menzione A, Tokar S, Giokaris N, Manousakis-Katsikakis A 2005 *Properties of the Ukrainian Polystyrene-Based Plastic Scintillator UPS923A* (Dzhelepov Laboratory of Nuclear Problems, Dubna) p.16
- [3] Long D 2002 Survey of Light-scattering Phenomena *The Raman Effect* (West Sussex, John Wiley & Sons Ltd) chapter 1 pp 3–71
- [4] LabRam Horiba, “Products: LabRam band Analysis”, [Online]. Available: <http://www.horiba.com/fileadmin/uploads/Scientific/Documents/Raman/bands.pdf> . [Accessed 18 March 2014].

Measurement of Higgs production in association with high p_T jets with the ATLAS detector

Bruce Mellado¹

University of the Witwatersrand, 1 Jan Smuts Avenue, Johannesburg 2000, South Africa

Abstract. The first measurement of the differential cross section of the Higgs boson, performed in the diphoton decay channel, is presented. The dataset used corresponds to 20.3 fb^{-1} of proton-proton collisions at the center of mass of 8 TeV, produced by the LHC and collected by the ATLAS detector in 2012. With its high signal selection efficiency the diphoton decay channel is well suited to probe the underlying kinematic properties of the signal production and decay. Measurements for several diphoton and jet distributions are made for isolated photons within the geometric acceptance of the detector and they are corrected for experimental acceptance and resolution. Results are compared to theoretical predictions at the particle level. Prospects for Run 2 are discussed.

1. Introduction

In the Standard Model, SM, of electro-weak, EW, and strong interactions, there are four types of gauge vector bosons (gluon, photon, W and Z) and twelve types of fermions (six quarks and six leptons) [1, 2, 3, 4]. These particles have been observed experimentally. At present, all the data obtained from the many experiments in particle physics are in agreement with the Standard Model. In the Standard Model, there is one particle, the Higgs boson field, that is responsible for giving masses to all the elementary particles [5, 6, 7, 8]. In this sense, the Higgs particle occupies a unique position.

In July 2012 the ATLAS and CMS experiments reported the discovery of a boson, a Higgs-like particle with a mass $m_H \approx 125 \text{ GeV}$ based on the data accumulated during 2011 and part of 2012 periods [9, 10]. It is also relevant to note that no additional Higgs bosons with couplings as in the Standard Model have been observed in the range of $m_H < 600 \text{ GeV}$.

The interest of the community has now shifted towards understanding the properties of the newly discovered boson. One of the relevant aspects is the production of the Higgs boson in association with high transverse momentum, p_T , hadronic jets. Together with the exploration of inclusive kinematics these measurements are a test of the SM. Some of these observables are sensitive to physics beyond the SM. Ultimately, the detailed measurement of the Higgs boson in association with jets will help improve the uncertainty of this important background to the extraction of the Vector Boson Fusion (VBF) mechanism. This production mechanism entails the scattering of two energetic quarks that exchange a weak boson. This weak boson radiates a Higgs boson. This production mechanism displays two high p_T jets at leading order and with a very particular topological configuration. Presence of admixtures of anomalous couplings in the

¹ E-mail: Bruce.Mellado@wits.ac.za

$HVV, V = Z, W$ vertex would lead to the distortion of the kinematic pattern, opening a window of opportunity to potentially observe physics beyond the SM [11, 12].

The results shown here are the first of their kind and were reported by the ATLAS collaboration in Ref. [13]. In section 2 the event selection, definition of fiducial cuts and a summary of the systematic errors are described. In section 3 the unfolding procedure is briefly outlined. Section 4 shows differential cross-section distributions for relevant observables. The note is concluded with a brief summary and conclusions in section 5.

2. Event Selection and Systematic Errors

The event and object selection used in this analysis follow closely the approach of earlier ATLAS analyses. This includes the generation of Monte Carlo (MC) events. Events are recorded using a diphoton trigger with an efficiency above 99% with respect to the offline selection, requiring two energetic clusters that match criteria according to expectations for photon-induced electromagnetic showers. One of the clusters must have transverse energy, E_T , greater than 35 GeV, and the other one must be greater than 25 GeV. They must lie within the fiducial acceptance of $|\eta| < 2.37$, where η is the pseudorapidity, excluding the transition region between the barrel and endcap calorimeters, $1.37 < |\eta| < 1.56$. Basic data quality requirements ensure that all necessary components of the detector are operational. Each photon of the pair must further satisfy isolation requirements for the inner detector and calorimeter. The total energy in a cone around the photon of $\Delta R < 0.4$ must be less than 6 GeV.

MC generation is an important element to the analysis. Particle level description of the final state is defined after the process of hadronization in the MC. The particle level fiducial definition is chosen to mirror the event selection in data. This minimizes the extrapolation from detector level to particle level quantities. The selection criteria are as follows: the two highest- E_T , isolated final state photons, within $|\eta| < 2.37$ and with $105 \text{ GeV} < m_{\gamma\gamma} < 160 \text{ GeV}$ are selected. Note that the transition region between the barrel and endcap calorimeters is not removed in this definition. After the pair is selected, the same cut on $E_T/m_{\gamma\gamma}$ is applied as in the event selection in data: $E_T/m_{\gamma\gamma} > 0.35(0.25)$ for the higher- E_T (lower- E_T) photon. The isolation criterion is defined as follows: the sum of the p_T of all stable particles excluding muons and neutrinos is required to be less than 14 GeV within $\Delta R < 0.4$ of the photon. This requirement was found to correspond approximately to the calorimetric isolation cut of 6 GeV at reconstruction level.

Jets are selected both at particle and reconstruction level, using the anti-kt algorithm with a distance parameter of $R = 0.4$. At reconstruction level, the inputs are clusters of energy in the electromagnetic and hadronic calorimeters. At particle level all stable particles excluding muons and neutrinos serve as input. The jet must have transverse momentum exceeding 30 GeV, and rapidity $|y| < 4.4$.

A comprehensive set of systematic uncertainties are considered. This is comprised of three groups:

- Uncertainties on the signal yield. These include systematics on the luminosity, trigger and isolation efficiencies.
- Uncertainties on photon energy resolution.
- Migration uncertainties. These include the jet energy resolution, jet energy scale, the jet vertex fraction ² and jets from pileup

A summary of the size of all uncertainties considered and those that are neglected can be found in Ref. [13].

² The jet vertex fraction is introduced as a measure to quantify the probability that a jet originates from the primary vertex and not from other collisions in the event.

3. Unfolding Procedure

The data yields extracted are corrected for detector effects using bin-by-bin factors. These are derived as the ratio of the yields from particle level to reconstruction level from simulated Higgs boson events, according to the SM expectation. In each bin,

$$c_i = n_i^{Particle\ level} / n_i^{Reconstructed} \quad (1)$$

is used to correct the extracted signal yield obtained in the data. This unfolding procedure corrects for all efficiencies, acceptances, and resolution effects. The correction factors range from 1.2 to 1.8, and include the extrapolation (about 20%, across all bins and observables) over the small region in rapidity (see section 2) excluded from reconstructed photon candidates. The method is formally unbiased provided that $c^{MC} = c^{Data}$. In practice, the requirement to use this method is that the purity of events reconstructed into the same bin in which they were generated should not be significantly below 50%.

The application of the unfolding coefficients defined in expression 1 assumes that the Higgs is produced as predicted by the SM. This feature of the measurement unavoidably results in the generation of certain biases. These biases need to be quantified and can be classified into sample composition and shape. Four groups of variations that are orthogonal to each other, are considered:

- Apply scale factors to the various production mechanisms according to measurements of their signal strengths at ATLAS. This check ensures that assuming the SM with regards to the relative rate of various production mechanisms does not generate significant biases.
- Impact of missing higher order perturbative corrections in gluon-gluon fusion production.
- Disable multi-parton interactions to test the impact of the modelling of the underlying event.
- Reweight the combined SM prediction in $p_T^{\gamma\gamma}$ and $|y^{\gamma\gamma}|$ or p_T^{j1} spectra (see section 4) to match the unfolded results in these variables. This check ensures that the shape of the differential distribution predicted by the SM does not bias the measurement.

A complete study of these biases is reported in Ref. [13]. Overall, the measurement is dominated by statistical uncertainties. At this point in time model dependencies do not play a critical role.

4. Results

Figure 1 displays the measured differential cross-sections for inclusive observables: the Higgs boson candidate transverse momentum, $p_T^{\gamma\gamma}$, its rapidity $|y^{\gamma\gamma}|$ and the helicity angle $|\cos(\theta^*)|$. The lower right plot displays the transverse momentum of the leading jet in the event, p_T^{j1} . Figure 2 shows the results of jet-related quantities: the exclusive rate of jets for events with up to three jets, the ratio of the exclusive rate to the inclusive rate, $\sigma_i/\sigma_{\geq i}$, the azimuthal angle difference between the two leading jets in the event, $\Delta\phi_{jj}$, and the transverse momentum of the di-photon and di-jet system in events with at least two jets, $p_T^{\gamma\gamma jj}$.

Results are reported for the region of the phase-space described in section 2. The results are compared to MC predictions with the SM expectation. All predictions were generated at $m_H = 126.8$ GeV, the central value of the measured Higgs boson mass in the diphoton channel by ATLAS. For the gluon-gluon fusion mechanism three different predictions are considered: Higgs boson production at NLO POWHEG [14], Higgs boson + 1 jet from MINLO [15] at NLO, and HRes at NNLO+NNLL [16]. The POWHEG prediction is shown for all 8 observables. HRes is presented for $p_T^{\gamma\gamma}$ and $|y^{\gamma\gamma}|$, and MINLO is used for the remaining 6 observables.

The measurements reported here are limited by statistical uncertainties. A statistical analysis to quantify the compatibility of the data with the SM prediction, including the Higgs boson p_T

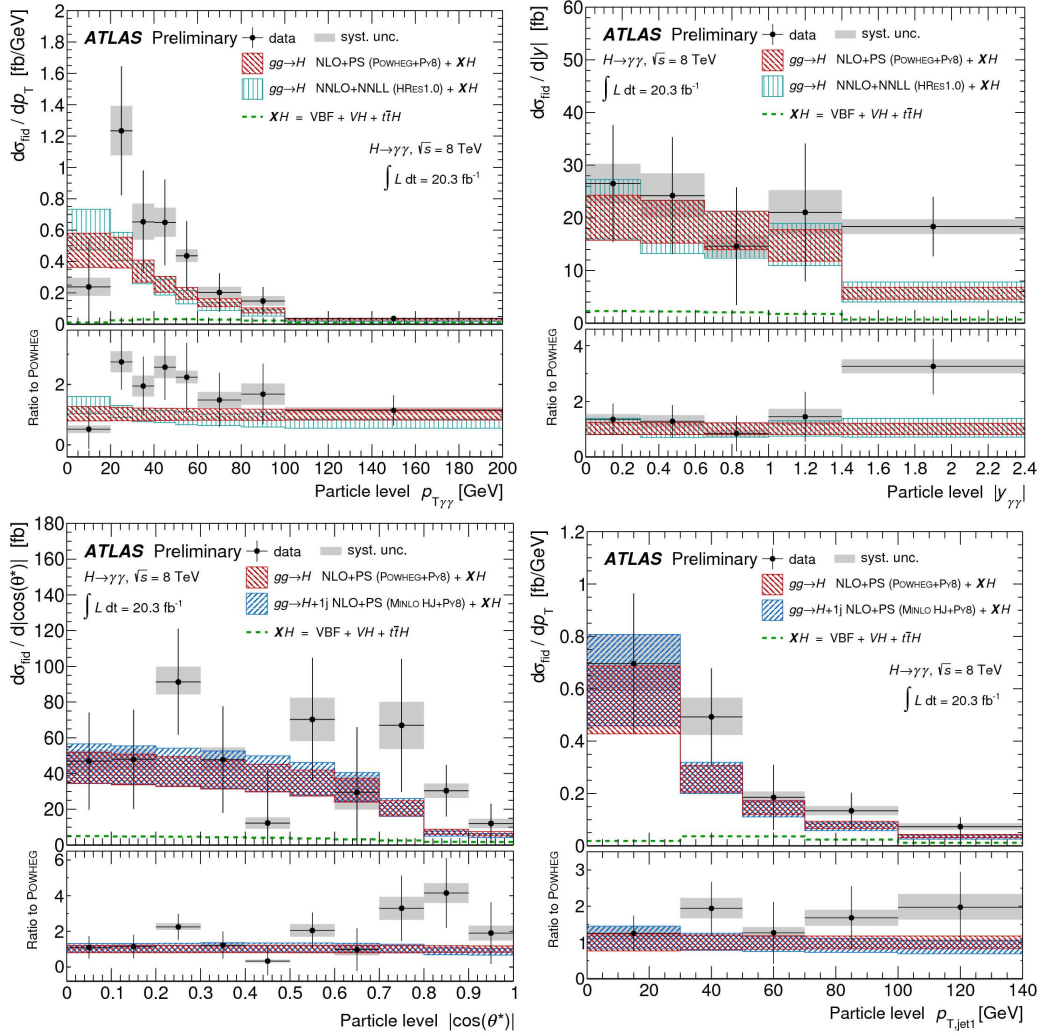


Figure 1. Observed differential cross sections of the Higgs bosons decaying into two isolated photons. Systematic uncertainties are presented in grey, and the black bars represent the quadratic sum of statistical and systematic errors. The hatched histograms present theoretical predictions for the Standard Model at $\sqrt{s} = 8$ TeV and $m_H = 126.8$ GeV. Their width represents the theory uncertainties from missing higher order corrections, the PDF set used, the simulation of the underlying event, and the branching fraction. The sum of VBF with WH , ZH , and ttH is denoted XH . These are added to the simulated ggH predictions from POWHEG, MINLO, and HRes (see text).

distribution. The compatibility of the data and MC is better than 10%. One can conclude that the reported measurements are consistent with the predictions by the SM within uncertainties.

5. Summary and Conclusions

Following the discovery of a new scalar boson, the measurement of properties becomes a focus of investigation by the experimentalists. The measurement of differential cross-sections of inclusive observables and the production in association with hadronic jets represents a test of the SM and a window of opportunity for searches of physics beyond.

The measurements reported here are the first of its kind. The comparison of these

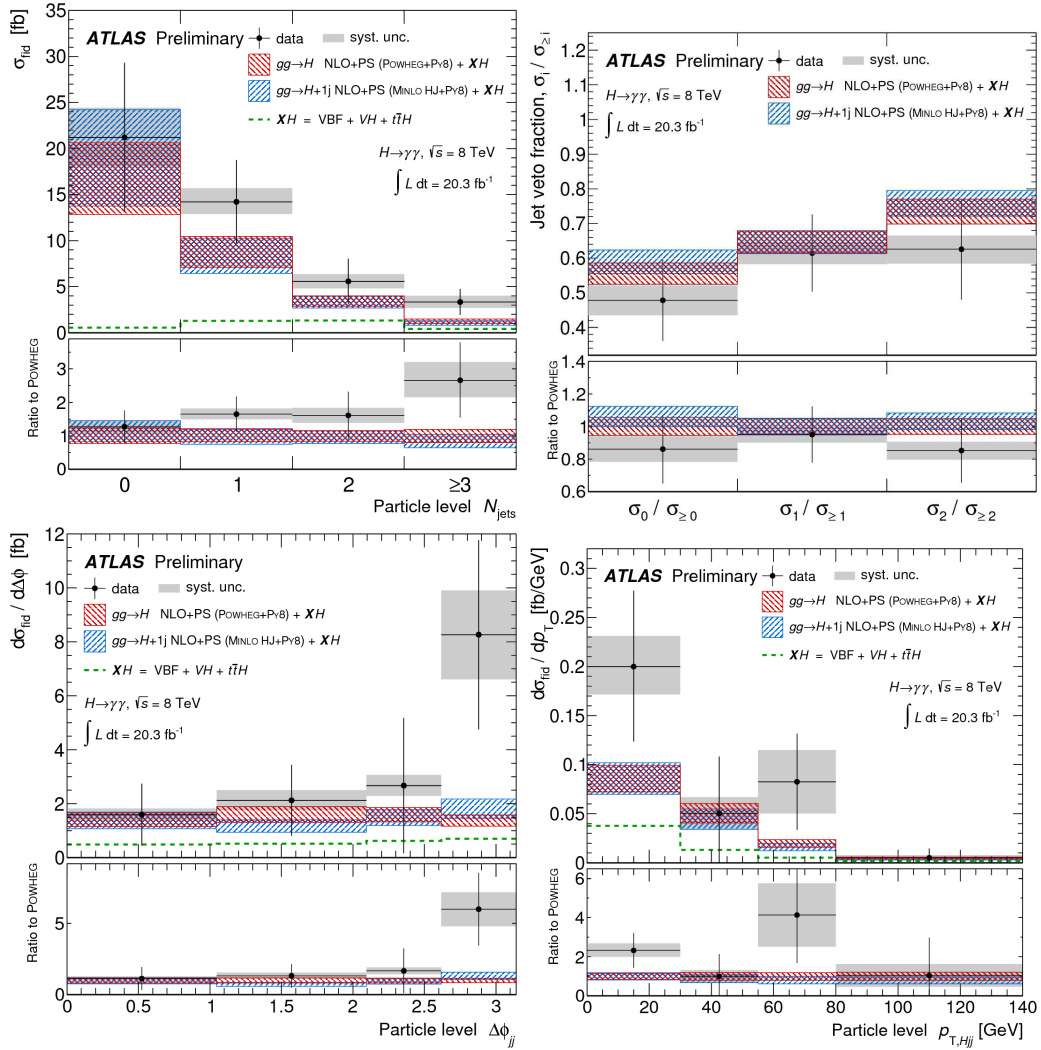


Figure 2. Same as Fig. 1 for multi-jet related variables (see text).

measurement with state-of-the-art MC tools and calculations indicate that Higgs kinematics and that of the associated jets are consistent with the SM within the errors quoted. This measurement needs to be taken as a first step towards the establishment of precision measurements that will lead to the reduction of the uncertainties in the understanding of the Higgs production via gluon-gluon fusion in association with jets. This will be essential in furthering the accuracy of the measurements of the Higgs production in VBF.

The measurements presented here are limited by the statistical uncertainty. The prospects of re-doing these measurements with Run 2 data are very exciting. With the increased proton-proton center of mass energy the Higgs production cross-section in certain corners of the phase-space will increase considerably. These effects, in addition to the collection of about 4-5 times the integrated luminosity, render Run 2 a very exciting period for the exploration of Higgs boson physics at the LHC.

References

- [1] Glashow S 1961 *Nucl.Phys.* **22** 579

- [2] Weinberg S 1967 *Phys.Rev.Lett.* **19** 1264
- [3] Salam A 1968 Proceedings to the eighth nobel symposium, may 1968, ed: N. svartholm 357
- [4] Glashow S , Iliopoulos J and Maiani L 1970 *Phys.Rev.* **D2**, 1285
- [5] Englert F and Brout R 1964 *Phys.Rev.Lett.* **13** 321
- [6] Higgs P 1964 *Phys.Lett.* **12** 132
- [7] Higgs P 1964 *Phys.Rev.Lett.* **13** 508
- [8] Guralnik G, Hagen C and Kibble T 1964 *Phys.Rev.Lett.* **13** 585
- [9] ATLAS Collaboration (G. Aad *et al.*) 2012 *Phys.Lett.* **B716** 1
- [10] CMS Collaboration (S. Chatrchyan *et al.*) 2012 *Phys.Lett.* **B716** 30
- [11] Englert C, Goncalves-Netto D, Mawatari K and Plehn T 2013 JHEP **1301** 148
- [12] Djouadi A, Godbole R , Mellado B and Mohan K 2013 Phys. Lett. B **723** 307
- [13] The ATLAS collaboration 2013 ATLAS-CONF-2013-072 (<http://cds.cern.ch/record/1562925>)
- [14] Nason P and Oleari C 2010 JHEP **1002** 037
- [15] Hamilton K, Nason P and Zanderighi G 2012 JHEP **1210** 155
- [16] de Florian D, Ferrera G, Grazzini M and Tommasini D 2012 JHEP **1206** 132

The status of the LHeC project and its impact on Higgs Physics

Bruce Mellado¹

University of the Witwatersrand, 1 Jan Smuts Avenue, Johannesburg 2000, South Africa

Abstract. The LHeC is envisioned to collide electrons and protons concurrently with collisions at the LHC. The overall status of the project is summarized. This comprises a review of the accelerator facility, the Energy Recovery Linac, and the detector design.

The ATLAS and CMS collaborations at the Large Hadron Collider have observed a new particle consistent with a scalar boson and with a mass of about 125 GeV. The prospects of studying this newly discovered boson at the LHeC are reviewed. This includes ability to isolate the $H \rightarrow b\bar{b}$ decay with a large signal-to-background ratio of better than $S/B = 2$ and the model independent exploration of the CP-properties of the HVV , $V = W, Z$ couplings. The latter is a unique capability of ep collisions. The prospects of other decay channels will also be discussed. An enhanced instantaneous luminosity scenario of $L = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ is considered. In this scenario the LHeC becomes a Higgs facility.

1. Introduction

With the discovery of a Higgs boson by the ATLAS and CMS collaborations [1, 2] a new era in particle physics has begun. While the Standard Model predicts the existence of a scalar field, the observation of this new particle opens a window of opportunity to search for new physics beyond. The exploration of the Higgs boson coupling to other particles in the Standard Model is sensitive to new interactions. The Large Hadron-electron Collider (LHeC) offers an excellent setup for precision measurement of these couplings.

Deep inelastic lepton-hadron scattering is the cleanest and most precise probe of parton dynamics in protons and nuclei. The LHeC is the only current proposal for TeV-scale lepton-hadron scattering and the only medium-term potential complement to the LHC pp , AA and pA programme at the maximum center of mass energy. As such, it has a rich and diverse physics programme of its own, as documented extensively in the recent conceptual design report (CDR) [3] and summarised in an initial submission by the LHeC Study Group to the European Strategy of Particle Physics (ESPP) discussion prior to the Cracow Symposium [4].

2. Accelerator Design

The LHeC envisions electron-proton (ep) and electron-ion (eA) collisions as a complement of the Large Hadron (pp and AA/p) collider. At the LHeC the lepton-quark interactions would reach the TeV scale. It is important to note that the LHeC is envisioned to run concurrently with LHC operations. The LHeC is designed not to disrupt operation of the LHC. In the CDR [3] two configurations were considered:

¹ E-mail: Bruce.Mellado@wits.ac.za

Table 1. LHeC ep and eA collider parameters (from reference [5]). The numbers give the default CDR values, with optimum values for maximum ep luminosity in parentheses and values for the ePb configuration separated by a comma.

Parameter [unit]	LHeC	
species	e^-	$p, {}^{208}\text{Pb}^{82+}$
beam energy (/nucleon) [GeV]	60	7000, 2760
bunch spacing [ns]	25, 100	25, 100
bunch intensity (nucleon) [10^{10}]	0.1 (0.2), 0.4	17 (22), 2.5
beam current [mA]	6.4 (12.8)	860 (1110), 6
rms bunch length [mm]	0.6	75.5
polarization [%]	90	none, none
normalized rms emittance [μm]	50	3.75 (2.0), 1.5
geometric rms emittance [nm]	0.43	0.50 (0.31)
IP beta function $\beta_{x,y}^*$ [m]	0.12 (0.032)	0.1 (0.05)
IP spot size [μm]	7.2 (3.7)	7.2 (3.7)
synchrotron tune Q_s	—	1.9×10^{-3}
hadron beam-beam parameter	0.0001 (0.0002)	
lepton disruption parameter D	6 (30)	
crossing angle	0 (detector-integrated dipole)	
hourglass reduction factor H_{hg}	0.91 (0.67)	
pinch enhancement factor H_D	1.35	
CM energy [TeV]	1300, 810	
luminosity / nucleon [$10^{33} \text{ cm}^{-2}\text{s}^{-1}$]	1 (10), 0.2	

- Ring-Ring (RR) option. A ring with a new electron beam would be mounted on top of the proton rings of the LHC.
- Linac-Ring (LR) option. An Energy Recovery Linac (ERL) in race track configuration providing electrons to collide head on with a proton beam of the LHC.

Careful consideration of a number of factors led to choosing the LR option as the default one (see reference [5] and references therein). Each linac accelerates the beam to 10 GeV, which leads to a 60 GeV electron energy at the interaction point after three passes through the opposite linear structures of 60 cavity-cryo modules each. The arc radius is about 1 km, mainly determined by the synchrotron radiation loss of the 60 GeV beam which is decelerated for recovering the beam power after having passed the IP. The default tunnel circumference is 1/3 that of the LHC. The tunnel is designed to be tangential to IP2 or IP8 (see Section 3). Detailed civil engineering considerations are described in the CDR.

The preference for the LR option has been recently further reiterated in [6]. A proposal for the construction of a ERL facility at CERN is materializing as an experimental facility for research and development in accelerator physics [6].

An important development is the recent progress to demonstrate the feasibility of reaching the instantaneous luminosity of $L = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. While remaining a challenge, confidence among accelerator physicists at CERN remains healthy [6]. The target parameters to reach this goal are summarized in Tab. 1.

3. Status of Detector Design

detector with backward-forward asymmetry, due to the beam configuration. It is a classical design with strong similarities to the ATLAS detector, except for the use of the return magnetid

field for the measurement of the momenta of escaping muons [3].

The time schedule of the LHeC project is determined by the evolution of the the LHC project and the upgrade plans. One can expect that the detector would need to be ready within about 10 – 12 years. A first installation study was made considering pre-mounting the detector at the surface, lowering and installing it at the pit. The LHC complex has eight caverns where detectors and infrastructure, such as the beam dump and the RF cavities are located. Based on these constraints Interaction Point, IP, number two² was considered in the CDR. Recently IP8³ has been considered as an alternative to IP2 [6]. Further studies are expected to make progress in the near term. In case of a possible installation in IP2 the detector should be small enough to fit into the L3 magnet structure of 11.2 m diameter, which is still resident in IP2 and would be available as mechanical support. Based on the design, as detailed in the CDR, it is estimated that the whole installation can be done in 30 months, which appears to be compliant with the operations currently foreseen in the LS3 shutdown in the early twenties.

A number of developments have taken place after the publishing of the CDR. These include improvements in the design of the beam-pipe and the tracking detectors. Discussions are ongoing in view of the strength of the Higgs boson physics at the LHeC to re-optimize the detector by de-emphasizing electromagnetic calorimetry in favor of hadronic calorimetry. The latter is essential for the reconstruction of the Higgs boson with the $b\bar{b}$ decay. The Wits group is contributing to the design of the hadronic calorimeter. Currently, a new generation of plastic scintillators are being tested using facilities in South Africa. This includes determination of radiation hardness, detailed understanding of the mechanisms of radiation damage and the optical properties. Results in this area are reported in separate proceedings [7, 8, 9].

4. Higgs Boson Production in High Energy ep Collisions

The leading production mechanism for the SM Higgs boson at the LHeC is

$$eq \rightarrow \nu_e H q' \quad \text{and} \quad eq \rightarrow e H q, \quad (1)$$

via the Vector Boson Fusion processes (VBF), as depicted in figure 1. It is remarkable that the Higgs boson production via VBF was first calculated for lepton-nucleon interactions (for a review of this question see [10] and references therein).

The production rate for the Charge Current, CC, process is larger than that of the Neutral Current, NC, process by about a factor of 4–6. This is mainly due to the accidentally suppressed NC coupling to the electrons. Here we have used the package MadGraph [11] for the full matrix element calculations at tree-level, and adopted the parton distribution functions CTEQ6L1 [12]. We choose the renormalization and factorization scales to be at the W -mass, which characterizes the typical momentum transfer for the signal processes.

In order to appreciate the unique kinematics of the VBF process it is most intuitive to express the cross section in a factorized form. Consider a fermion f of a c.m. energy E radiating a gauge boson V ($s \gg M_V^2$), the cross section of the scattering $fa \rightarrow f'X$ via V exchange can be expressed as:

$$\sigma(fa \rightarrow f'X) \approx \int dx \, dp_T^2 \, P_{V/f}(x, p_T^2) \, \sigma(Va \rightarrow X) \quad (2)$$

where $\sigma(Va \rightarrow X)$ is the cross-section of the $Va \rightarrow X$ scattering and $P_{V/f}$ can be viewed as the probability distribution for a weak boson V of energy xE and transverse momentum p_T . These expressions lead us to the following observations:

- 1 Unlike the QCD partons that scale like $1/p_T^2$ at the low transverse momentum, the final state quark f' typically has $p_T \sim \sqrt{1-x}M_V \leq M_W$.

² The ALICE detector is currently located in IP2.

³ The LHCb detector is currently located in IP8.

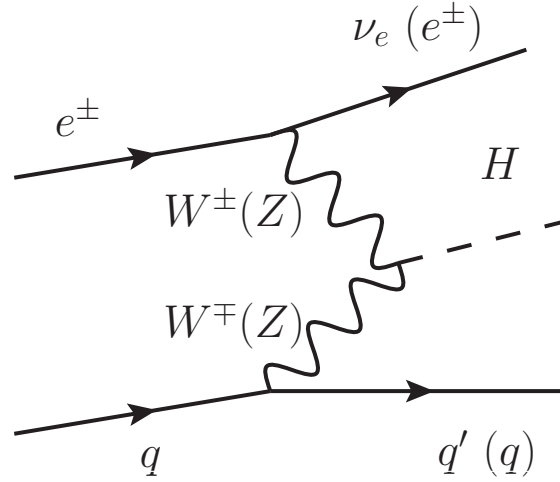


Figure 1. Leading order diagram for the production of a Standard Model Higgs boson in ep collisions for the charged current and neutral current processes.

- 2 Due to the $1/x$ behavior for the gauge boson distribution, the out-going parton energy $(1-x)E$ tends to be high. Consequently, it leads to an energetic forward jet with small, but finite, angle with respect to the beam.
- 3 At high p_T , $P_{V/f}^T \sim 1/p_T^2$ and $P_{V/f}^L \sim 1/p_T^4$, and thus the contribution from the longitudinally polarized gauge bosons is relatively suppressed at high p_T to that of the transversely polarized.

Items 1 and 3 clearly motivate a tagging for a forward jet to separate the QCD backgrounds [13, 14], while item 3 suggests a veto of central jets with high p_T to suppress the backgrounds initiated from the transversely polarized gauge bosons, and from other high p_T sources such as top quarks [15].

The mere identification as a Higgs boson is not enough, for it will leave open a host of other questions, such as whether this scalar is elementary or composite, CP-conserving or CP-violating. In general the tensor structure of the coupling to weak bosons needs to be investigated in order to assess whether the newly discovery boson is related to new physics. Collisions at the LHeC provide a unique opportunity to separate the HWW and HZZ vertices while allowing for the independent exploration of the azimuthal correlation of the scattered fermions [16]. This is a unique feature of ep collisions not present in pp and e^+e^- collisions. Deviations from the SM can be parametrized using two dimension-5 operators:

$$\Gamma_{\mu\nu} \propto [\lambda(p \cdot q g_{\mu\nu} - p_\nu q_\mu) + i\lambda' \epsilon_{\mu\nu\rho\sigma} p^\rho q^\sigma], \quad (3)$$

where p and q are four momenta of the weak bosons, and λ and λ' are effective coupling strengths for the anomalous CP-conserving and the CP-violating operators, respectively.

5. Results

Studies reported in the LHeC CDR [3] based on a fast simulation of signal and background using the CC reaction for the nominal 7 TeV LHC proton beam and electron beam energies of 60 and 150 GeV. Simple and robust cuts are identified and found to reject effectively e.g. the dominant single-top background, providing an excellent S/B ratio of about 1 at the LHeC, which may be further refined using sophisticated neural network techniques. At the default electron beam energy of 60 GeV, for 80% e^- polarisation and an integrated luminosity of 100 fb^{-1} , the Hbb coupling is estimated to be measurable with a statistical precision of about 4%, which is not far from the current theoretical uncertainty. It is important to note that the

Table 2. Cross sections and rates of Higgs production in ep scattering with the LHeC. The cross sections are obtained with MADGRAPH5 (v1.5.4) using the p_T of the scattered quark as scale, CTEQ6L1 partons and $m_H = 125$ GeV. The assumed branching ratios, Br , to different decays are given.

LHeC Higgs		CC (e^-p)	NC (e^-p)	CC (e^+p)
Polarisation		-0.8	-0.8	0
Luminosity [ab^{-1}]		1	1	0.1
Cross Section [fb]		196	25	58
Decay	$Br(H \rightarrow X)$	$N_{CC}^H e^-p$	$N_{NC}^H e^-p$	$N_{CC}^H e^+p$
$H \rightarrow b\bar{b}$	0.577	113 100	13 900	3 350
$H \rightarrow c\bar{c}$	0.029	5 700	700	170
$H \rightarrow \tau\tau$	0.063	12 350	1 600	370
$H \rightarrow \mu\mu$	0.00022	50	5	—
$H \rightarrow 4l$	0.00013	30	3	—
$H \rightarrow 2l2\nu$	0.0106	2 080	250	60
$H \rightarrow gg$	0.086	16 850	2 050	500
$H \rightarrow WW$	0.215	42 100	5 150	1 250
$H \rightarrow ZZ$	0.0264	5 200	600	150
$H \rightarrow \gamma\gamma$	0.00228	450	60	15
$H \rightarrow Z\gamma$	0.00154	300	40	10

instantaneous luminosity was assumed $L = 10^{33} \text{cm}^{-2} \text{s}^{-1}$. Given these promising results efforts are being made in order to consider a higher luminosity scenario, namely $L = 10^{34} \text{cm}^{-2} \text{s}^{-1}$. Various parameters assumed in Reference [3] have been re-assessed in order to achieve the desired instantaneous luminosity [5]. In this scenario the LHeC could be considered a Higgs facility on its own right. Table 2 gives estimates of Higgs boson cross-sections and yields for various production mechanisms and decay channels. With the high luminosity scenario the exploration of the $H \rightarrow b\bar{b}$ and the CP properties of the HWW coupling will enter into the realm of precision. Other decays, such as $H \rightarrow \tau\tau, VV, gg, c\bar{c}$ will become accessible with sizeable statistics.

The LHC is believed to display inferior sensitivity to couplings compared to that of a linear collider. Part of this statement comes from large uncertainties, which are related to the imperfect knowledge of the PDFs and theory parameters. The LHeC, with its high precision PDF and QCD programme, will render many of these uncertainties unimportant.

6. Conclusions

The progress of the LHeC project is summarized. In the area of accelerator studies significant progress has been made in the definition of the project for the Electron Recovery Linac. This project, while broad in nature, is essential to the Ring-Linac option that is considered to be default at the LHeC. In the area of detector development progress has been made with regards to more detailed designs of the beam pipe and the tracking detectors. With regards to the location of the interaction point, in addition to IP2, IP8 is also being considered. Currently, various activities are ongoing in facilities in South Africa to understand the performance and radiation resistance of a new generation of plastic scintillators for the LHeC hadronic calorimeter [7, 8, 9].

When considering the scenario of instantaneous luminosity $L = 10^{34} \text{cm}^{-2} \text{s}^{-1}$, the LHeC is rendered a Higgs facility. Here hundred of thousands of Higgs candidates could be reconstructed over a decade of data taking. This provides strong sensitivity to a number of couplings beyond the capabilities of the LHC and competitive with the ILC. The LHeC has the unique feature of

being able to separate the HWW and HZZ coupling measurements.

References

- [1] ATLAS Collaboration (Aad G *et al.*) 2012 *Phys.Lett.* **B716** 1
- [2] CMS Collaboration (Chatrchyan S *et al.*) 2012 *Phys.Lett.* **B716** 30
- [3] Abelleira Fernandez J L *et al.* [LHeC Study Group] 2012 *J.Phys.G.* **39** 075001
- [4] LHeC Study Group 2012 Contribution (No 147) to the Cracow meeting on the ESPP *A Large Hadron Electron Collider at CERN* LHeC-Note-2012-004 GEN CERN
- [5] Bruening O and Klein M 2013 *Mod. Phys. Lett. A* **28** no. 16 1330011
- [6] Meeting of the LHeC International Advisory Committee, CERN, Geneva, 26/06/14, <http://indico.cern.ch/event/315848/>
- [7] Jivan H *et al.* 2014 *Radiation Hardness of Plastic Scintillators for the Tile Calorimeter of the ATLAS Detector* submitted to the SAIP 2014 Conference
- [8] Maphanga L *et al.* 2014 *Characterization of Damage in in-situ Radiated Plastic Scintillators at the Tile Calorimeter of the ATLAS* submitted to the SAIP 2014 Conference
- [9] Pelwan C *et al.* 2014 *Electron Paramagnetic Resonance Analysis of Plastic Scintillators for the Tile Calorimeter of the ATLAS detector* submitted to the SAIP 2014 Conference
- [10] Han T and Mellado B 2010 *Phys. Rev. D* **82** 016009
- [11] Alwall J *et al.* 2007 *JHEP* **0709** 028
- [12] Pumplin J *et al.* 2002 *JHEP* **0207** 012
- [13] Kleiss R and Stirling W J 1988 *Phys. Lett. B* **200** 193
- [14] Barger V D, Han T and Phillips R J N 1988 *Phys. Rev. D* **37** 2005
- [15] Barger V D, Cheung K M, Han T and Phillips R J N 1990 *Phys. Rev. D* **42** 3052
- [16] Biswal S S, Godbole R M, Mellado B and Raychaudhuri S 2012 *Phys. Rev. Lett.* **109** 261801

Validation of the performance of Geant4 in the simulation of neutron induced reactions relevant to reactor studies.

Rotondwa Mudau¹, Azwinndini Muronga¹, Simon Connell¹

University of Johannesburg, Physics Department¹

Nithec, Stellenbosch²

E-mail: rotondwamda@hotmail.co.za

Abstract. Geant4 is a Monte Carlo simulation toolkit that is used for the simulation of particles through matter. It was developed by a world wide collaboration of about 100 scientists. It has applications in high energy physics, astrophysics, medical physics, nuclear physics experiments, accelerator and space science studies. The Evaluated Nuclear Data File, ENDF/B-VII database is included in the simulation toolkit, making it feasible to use the Geant4 toolkit for low energy neutron-physics simulations. The ENDF/B-VII database stores evaluated nuclear reaction data files from the major evaluated libraries. In this presentation Geant4 was used to perform simple single event neutron scattering simulations on materials that are typical for a nuclear reactor. Some of the materials used are found in the SAFARI 1 reactor at Necsa, Pelindaba. These include Aluminium, Beryllium. This was done in order to validate the Geant4 implementation of primary processes relevant to reactor studies within the ENDF/B-VII database of cross sections.

1. Introduction

Geant4 is a Monte Carlo based simulation toolkit that is used for the Geometry and Tracking of particles through matter. Although initially developed for high energy physics it has been extended and it has capabilities for applications in the low energy physics spectrum [1]. It is an open source modern object oriented code. It has proven success at a very high level of complexity for the geometry construction, materials specification, tracking algorithms and the physics lists in the high energy physics regime [2].

The energy loss of neutrons is mainly due to elastic scattering, absorption and nuclear collisions [3]. A typical nuclear reaction is written as [4]:

$$a + X \rightarrow Y + b \quad (1)$$

where a is the incident particle and X is the target and Y and b are the reaction products. If the outgoing and the incident particles are the same i.e Y and X are the same nucleus then the nuclear reaction is a scattering process. If Y and b are in their ground states it is an elastic scattering.

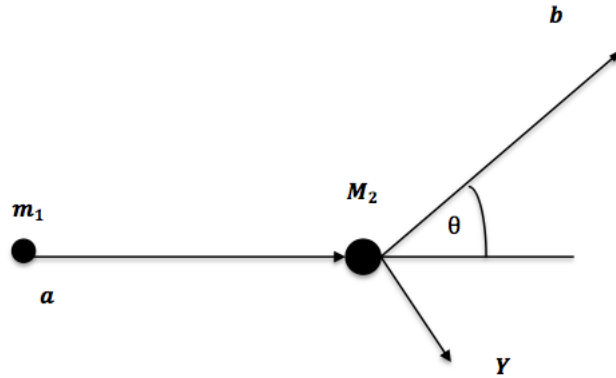


Figure 1. Kinematics of a nuclear reaction of the type $a + X \rightarrow Y + b$, where m_1 is the incident particle, M_2 is the target nucleus, θ is the scattering angle, Y and b are the reaction products.

Conservation of total relativistic energy in a basic reaction is:

$$m_X c^2 + E_X + m_a c^2 + E_a = M_Y c^2 + E_Y + m_b c^2 + E_b \quad (2)$$

where the E 's are the kinetic energies and the m 's are the rest masses [4]. The Q value which may be positive, negative or zero becomes:

$$Q = \{m_{initial} - m_{final}\}c^2 \quad (3)$$

This is the same as the access kinetic energy of the final products:

$$Q = E_{final} - E_{initial} = E_Y + E_b - E_X - E_a \quad (4)$$

Conservation of linear momentum (p) along and perpendicular to the beam direction is given by:

$$p_a = p_b \cos \theta + p_Y \cos \phi \quad (5)$$

and

$$0 = p_b \sin \theta - p_Y \sin \phi \quad (6)$$

where θ is the neutron scattering angle and ϕ is the nucleus recoil angle.

If particle Y is not observed which is usually the case, then ϕ and E_Y can be eliminated from the equations and a relationship between E_b and θ can be found.

$$E_b^{\frac{1}{2}} = \frac{(m_a m_b E_a)^{\frac{1}{2}} \cos \theta \pm \{m_a m_b E_a \cos^2 \theta + (M_Y + m_b)[M_Y Q + (M_Y - m_a)E_a]\}^{\frac{1}{2}}}{M_Y + m_b} \quad (7)$$

This result is then used to sort the Geant4 events as elastic or inelastic scattering by using the Q value.

2. Simulation

To validate the performance of Geant4 in the implementation of primary neutron nuclear processes, a few conventional neutron scattering experiments were configured and simulated and the results were compared with the ENDF/B-VII database of cross sections.

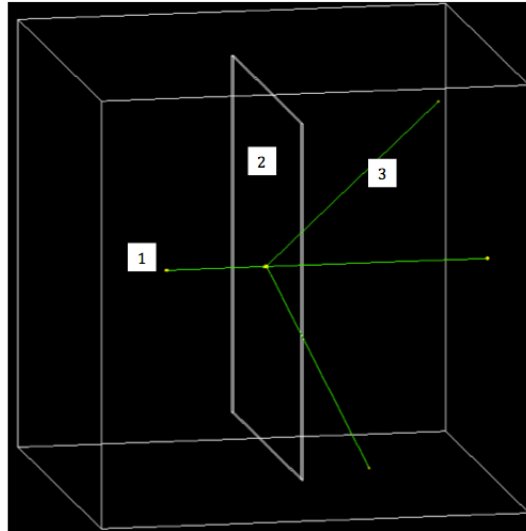


Figure 2. Illustration of the simulation of neutron scattering using Geant4. Point 1 represents a neutron generator; point 2 represents a target, which is one of the materials; point 3 represents neutron trajectories.

The simulation was a scattering experiment which consisted of a neutron beam that had randomly generated energies between 0 MeV and 10 MeV, and targets of three different well chosen materials with a thickness of 0.3 cm. The experiment was conducted in an empty box with the targets 10 cm away from the source. The target thickness was chosen in such a way as to avoid any secondary hard scattering events. The materials within which the neutron scattering was observed are Aluminium, Beryllium and Carbon. In order to make a quantitative analysis of the simulation, current information of the incident particle needs to be accessible to the user. This, Geant4 makes viable via one or many of the useraction classes in the Geant4 kernel [5].

As an object oriented toolkit, a standard Geant4 simulation consists of the main program which is the executable. This main program is executed via the G4RunManager class, which controls the flow of the program and manages the event loop/loops within a run, and the G4UImanager class which creates a pointer to the interface manager, and allows the user to issue commands to the program [6]. The G4RunManager is also responsible for managing initialization procedures, including methods in the user initialization classes. A simulation program consists of some mandatory classes that are defined by the user, and these fall under the user initialisation classes and the useraction classes. The mandatory user initialisation classes are derived from the Geant4 base classes, .e.g. G4VUserDetectorConstruction from which the experimental setup is designed and the G4VUserPhysicsList wherein one can specify the kind of physics that is to take place between the target and the incident particles of choice. The mandatory useraction class is derived from the G4VUserPrimaryGeneratorAction base class, in this class one gives dynamic information about the incident particle. Other useraction classes that can be added to the main include the SteppingAction, TrackingAction, EventAction classes etc. In these classes the user can extract information about the track or event [2].

Geant4.9.5 was used together with a reference physics list called QGSP-BIC-HP, Quark-Gluon String-Precompound Binary-Cascade High-Precision model. This reference physics list is a high precision neutron package, NeutronHP for the transport of neutrons below 20 MeV down to thermal energies and is therefore the recommended physics list for low energy neutron physics [2]. A step is a single discrete movement of a particle and a track is a series of sequential steps that make up the history of the particle. The G4Step class, stores the transient information of a step. This includes the two endpoints of the step, PreStepPoint and PostStepPoint, which contain the points' coordinates and the volumes containing the points. G4Step also stores the change in track properties between the two points. These properties, such as energy and momentum, are updated as the various active processes are invoked [2]. Therefore information about the track was accessed via the SteppingAction class. The histograms were stored via the EventAction class, which stores collective information of an event.

3. Results and discussion

The results below show a comparison of the ENDF/B-VII cross sections with the the reconstructed cross sections obtained from the Geant4 simulation. Since the experimental configuration is a scattering one and the kinematics are typical neutron nuclear reaction kinematics, we were able to extract the elastic cross section by selecting those events which had obeyed the elastic scattering kinematics. This works, because for elastic reactions, the Q value, or the access energy is zero. However, it becomes cumbersome to determine the Q values of nonelastic reactions. One can also get the cross sections that Geant4 uses for a reaction directly from Geant4, by just requesting a cross section for that particular reaction. However the purpose of this simulation was not only to see which cross sections Geant4 is using, but rather to also determine how Geant4 implements the physics of these reactions. To get these results, the Geant4 kernel was used. In addition to the transient information of a step, the G4Step class has various other Get methods of accessing more information about a step. For these results the GetProcess and GetProcessType methods were used to get information about the kind of processes that took place at each step. An algorithm that sorted between the different hadronic processes that took place at each event was written and the histograms were populated by the incident energy thereby giving the cross sections.

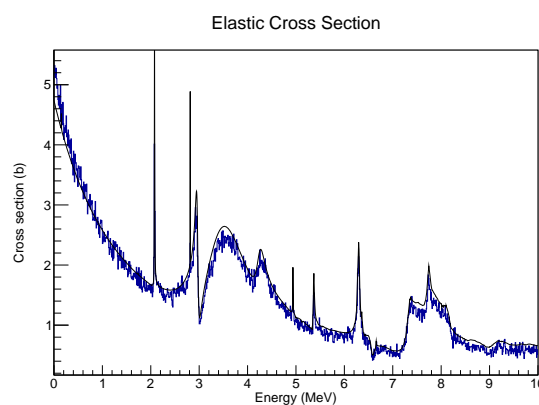


Figure 3. A comparison between ENDF/B-VII(Black) and Geant4(Blue) elastic cross sections of ^{12}C

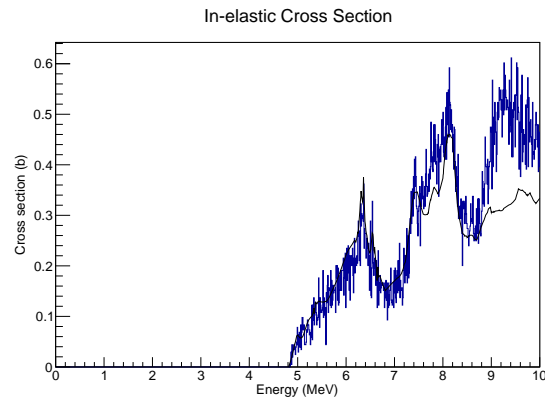


Figure 4. A comparison between ENDF/B-VII(Black) and Geant4(Blue) inelastic cross sections of ^{12}C

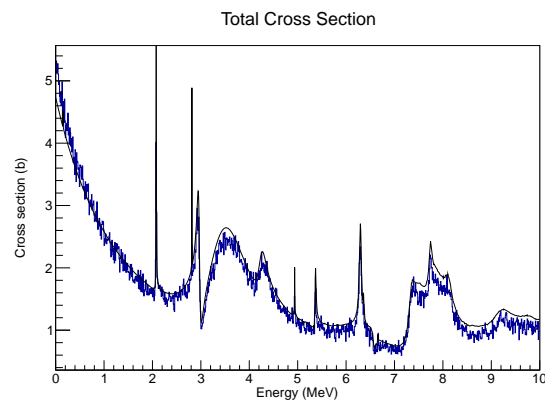


Figure 5. A comparison between ENDF/B-VII(Black) and Geant4(Blue) total cross sections of ^{12}C

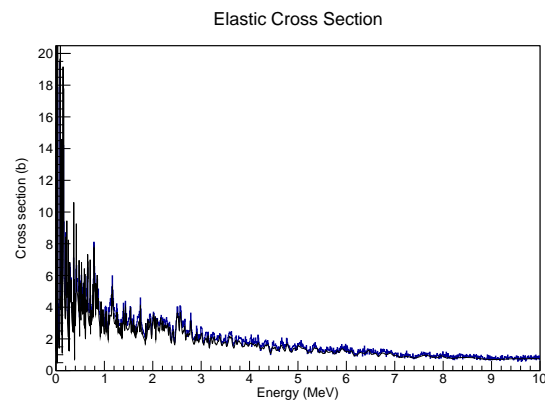


Figure 6. A comparison between ENDF/B-VII(Black) and Geant4(Blue) elastic cross sections of ^{27}Al

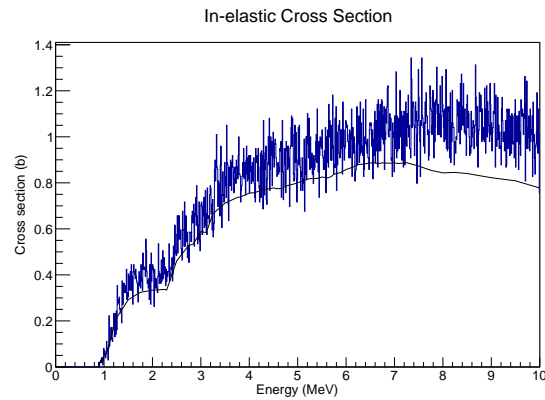


Figure 7. A comparison between ENDF/B-VII(Black) and Geant4(Blue) inelastic cross sections of ^{27}Al

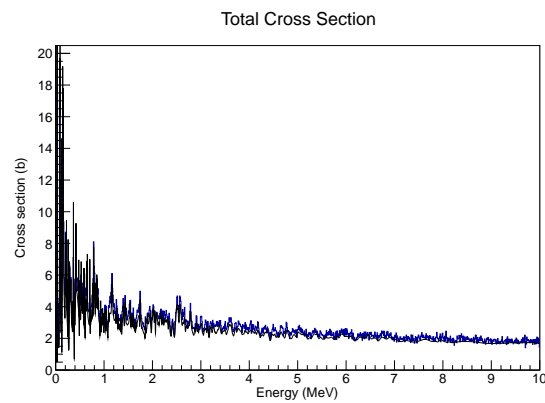


Figure 8. A comparison between ENDF/B-VII(Black) and Geant4(Blue) total cross sections of ^{27}Al

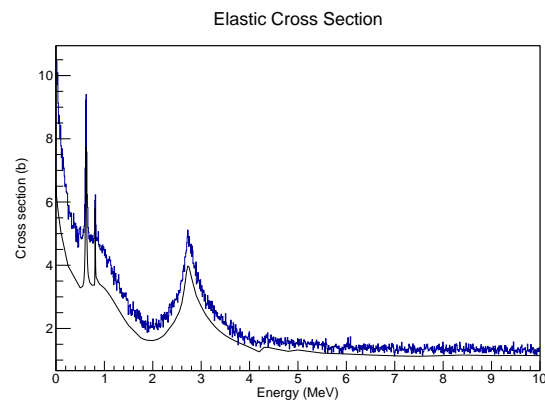


Figure 9. A comparison between ENDF/B-VII(Black) and Geant4(Blue) elastic cross sections of ^9Be

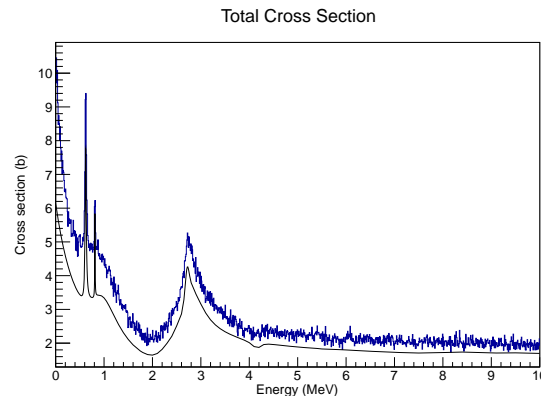


Figure 10. A comparison between ENDF/B-VII(Black) and Geant4(Blue) total cross sections of ^9Be

4. Conclusion

The histograms above show a fair correlation between the Geant4 cross section results and the ENDF/B-VII database cross sections. This partially validates the performance of Geant4 in the simulation of neutron induced reactions relevant to reactor studies. Although there are some inconsistencies in the absolute cross sections in terms of the scaling e.g Figure 10 and Figure 9 and some discrepancies at higher energies for the inelastic cross sections Figure 7 and Figure 4 the overall shape is fairly correlated. The discrepancies could be due to several issues. For example, there could be differences between the reference libraries and the libraries in this version of Geant4. There are also issues with the performance and configuration tuning of the neutron Physics Class [7]. These results are going to be a part of the continual investigation in the validation of Geant4 as a toolkit for reactor studies. There have been improvements [2] in the lower energy neutron physics in the latest version, Geant4 10.0 but those are yet to be tested.

5. References

- [1] Catsaros N, Gaveau B, Jaekel M, Maurel G, Savva P, Silva J, Varvayanni M and Zisis T 2009 *Annals of Nuclear Energy* **Vol.36**
- [2] <http://geant.cern.ch>
- [3] Leo W 1994 *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach* (New York: Springer) ISBN 9783540572800
- [4] Krane S 1987 *Introductory Nuclear Physics* (New York: John Wiley & Sons, Inc.)
- [5] Rujiwarodom R, Asavapibhop B and Shang P 2009 *Kasetsart Journal:Natural Science*. **Vol. 043**
- [6] Agostinelli S *et al.* 2003 *Nuclear Instruments and Methods A* **506-3** 250–303
- [7] Mendoza E and Cano-Ott D 30 November 2012 Recommendations for running the G4NeutronHP package Tech. rep. Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Spain.

Electron paramagnetic resonance analysis of plastic scintillators for the Tile Calorimeter of the ATLAS detector

Chad Pelwan

University of Witwatersrand, *School of Physics*, P242

E-mail: chad.pelwan@gmail.com

Abstract. In an attempt to understand the effects of ionizing radiation on various plastic scintillators, a number of studies are currently in progress with a hope that favourable properties of plastic scintillators, such as high light output and fast decay time, can be optimized. The Tile Calorimeter (TileCal) is a hadronic calorimeter able to detect hadrons, leptons and other energetic subatomic particles. In this investigation, polyvinyl toluene based plastic scintillators and a polystyrene based plastic scintillator, identical to those situated on the TileCal of the ATLAS detector at CERN, were irradiated and sent to the University of Witwatersrand where they were prepared for electron paramagnetic resonance (EPR) analysis. EPR spectroscopy allows for the study of unpaired electrons within these scintillators and offers a deeper insight into the organic or inorganic free radicals present. This technique was used to validate the assumption that dangling bonds in the plastics were as a result of ionizing radiation damage caused in the irradiation phase. This was done by detecting the existence of paramagnetic centres and, in addition, magnetic properties of these centres could be characterized. Three polyvinyl toluene Eljen scintillator plastics, EJ200, EJ208, and EJ260 were used in this investigation as well as one polystyrene Dubna sample. These samples were irradiated at the iThemba LABS, Gauteng. Experimental results thus far show a difference in the signal detected for irradiated samples compared to un-irradiated ones, and that a higher radiation dose produces varying EPR results for the Eljen and Dubna samples. It was also observed that other ions are detected after irradiation and that, over time, certain bonds would re-form within the plastics. Further investigation is required to understand this effect. Over all, the results from the EPR analysis form a small, yet vital, contribution into understanding the various effects of ionizing radiation in plastic scintillators.

Introduction

The Tie Calorimeter (TileCal) is a hadronic calorimeter designed to measure high energy positron jets, τ leptons and other subatomic particles like gluons and quarks that result from two proton collisions. The TileCal is situated within the ATLAS detector; a larger detector made of several sub-detectors. The ATLAS detector, shown in figure 1, is 44 meters in width and 22 meters in diameter. It is the largest of four detectors within the Large Hadron Collider (LHC) which is 27 kilometres in circumference and is 100 meters underground between the borders of Switzerland and France [1].

The TileCal consists of plastic scintillators that are able to relay information about the incident particle on them by emitting light. This light is then passed to photo-multiplier tubes (PMTs) which digitizes this signal. However, these plastic scintillators are bombarded with ionizing particles and γ -rays that cause radiation damage to the scintillators affecting their efficiency [2]. In order to understand how the radiation damage effects the plastics on a molecular scale, electron paramagnetic resonance (EPR) is used. EPR is a type of spectroscopy used to study molecules or atoms with unpaired electrons. The phase two upgrade in currently in progress for the TileCal and looks to replace the plastic scintillators situated in the gap region of the TileCal. Hopefully, EPR studies can shed light on which plastic scintillator is best suited for this. In the paper we present EPR results obtained from irradiated and un-irradiated polyvinyl toluene based scintillator samples: EJ200, EJ208 and EJ260 produced by Eljen technologies as well as the Dubna sample identical to that located in the TileCal produced by Dubna.

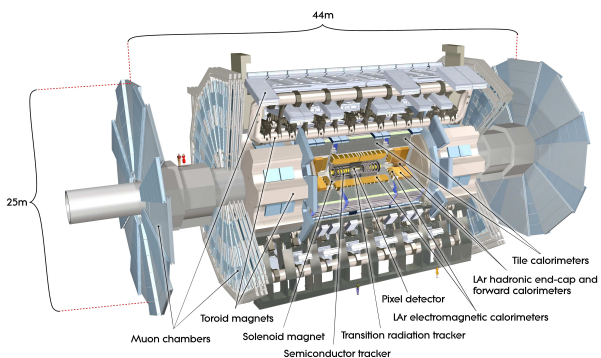


Figure 1. Schematic of the ATLAS detector

Materials

Since plastic scintillators in the TileCal mainly undergo ionization from the incident particles, there was a need to replicate this in order to study how radiation damage effects the plastic scintillators. The tandem accelerator at iThemba LABS, Gauteng was instrumental in the replication of this damage as it accelerated protons to an energy of 6 MeV which, in turn, were incident on the plastic scintillators. In order for the protons to impart energy to the plastic scintillators mainly through ionization, each plastic scintillator was cut and polished to a width less than the stopping range of the 6 MeV protons in both polystyrene and polyvinyl based plastics. SRIM (Stopping Range of Ions in Materials) is a computational method employed to find this range which was around $470\text{ }\mu\text{m}$. Thus, each plastic scintillator sample was machined to a width of $250\text{ }\mu\text{m}$ and volume of $250\text{ }\mu\text{m} \times 500\text{ }\mu\text{m} \times 500\text{ }\mu\text{m}$. The SRIM computations also found that the average energy of each proton was approximately 2.07 MeV. This result was used when calculating the dose each sample was irradiated to. Samples were irradiated to two doses: 0.164 MGy and 1.46 MGy, and an un-irradiated sample was used as a control.

Methodology

Electron paramagnetic resonance (EPR) measurements were performed on the plastic scintillators by placing them in a strong magnetic field while monochromatic electromagnetic (EM) radiation is applied; this is usually within the microwave region. If there are unpaired electrons in the plastic scintillators, they will absorb a small range of EM radiation which is detectable. We have made the assumption that ionizing radiation would cause bonds to break within the plastic scintillators forming free electrons and ions which are detectable by the EPR spectrometer. Experiments were performed at room temperature at the University of Witwatersrand in the NMR LAB using the Bruker ESP 380e spectrometer.

Results

Results in figure 2 show that paramagnetic centres are found in the polystyrene samples before irradiation (left image). A similar result was seen in the other Eljen samples. The intensity of

the absorption peak derivative also gives an indication as to how many paramagnetic centres can be found. In order to directly compare the samples, a mass normalisation is usually necessary, however there was no need for it as the samples all have the same volume and density. After irradiation (right image), the peak intensity of the absorption peak derivative decreased in the Eljen samples.

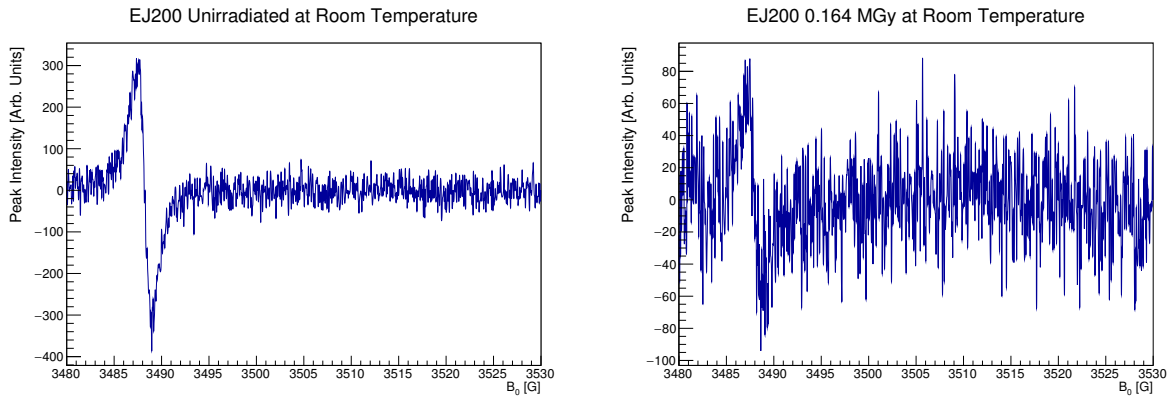


Figure 2. Comparison of an unirradiated EJ200 sample (left) and an irradiated sample at a dose of 0.164 MGy (right)

The left image in figure 3 also indicates that there are paramagnetic centres in the polystyrene based Dubna sample before irradiation. It was noted that, in general, the absorption peak intensity derivative was lower for the Dubna sample than the Eljen samples. However, the right image in figure 3 clearly shows an increase in the intensity after the samples were irradiated. Therefore more paramagnetic centres can be found the Dubna sample after irradiation.

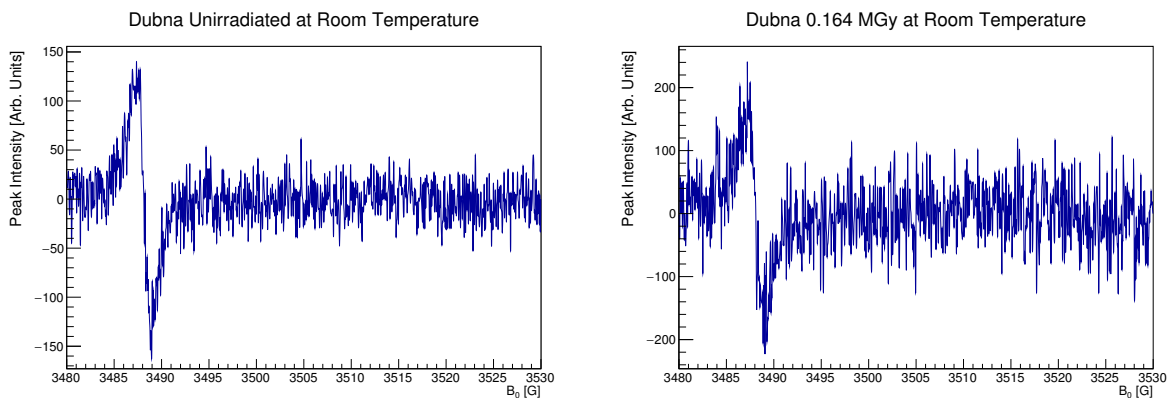


Figure 3. Comparison of an unirradiated Dubna sample (left) and an irradiated sample at a dose of 0.164 MGy (right)

In both cases, the signal to noise ratio increases with the increase of dose; measurements will be performed at cryogenic temperatures which would focus on the unpaired electrons and decrease the noise in the spectra. It has also been hypothesised that bonds within the samples could reform overtime when the unpaired electrons reform bonds or bond with ions in the material [3].

Conclusion

From this investigation it was found that exposure to proton irradiation causes structural damage to both the polystyrene and the polyvinyl toluene plastic scintillators studied. Further analysis of the data can now be done to investigate exactly which bonds break as a direct result of radiation damage. It was also observed that an increase in dose varied the number of unpaired electrons within the different samples but over time many of these bonds would reform.

References

- [1] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. August 2008.
- [2] V.M. Gorbachev. Plastic Scintillators with High Radiation Stability. *Atomnaya Energiya*, 6:427-429, June 1975.
- [3] K. Wick, U. Holm. Radiation Stability of Plastic Scintillators and Wave Length Shifters. *IEEE Tans. Nucl. Sci.*, 1:163, 1989

An ATCA framework for the upgraded ATLAS read out electronics at the LHC

Robert Reed

School of Physics, University of the Witwatersrand, Johannesburg, South Africa

E-mail: robert.reed@cern.ch

Abstract. The Large Hadron Collider at CERN is scheduled to undergo a major upgrade in 2022. The ATLAS collaboration will do major modifications to the detector to account for the increased luminosity. More specifically, a large proportion of the current front-end electronics, on the Tile Calorimeter sub-detector, will be upgraded and relocated to the back-end. A Demonstrator program has been established as a proof of principle. A new system will be required to house, manage and connect this new hardware. The proposed solution will be an Advanced Telecommunication Computing Architecture (ATCA) which will not only house but also allow advanced management features and control at a hardware level by integrating the ATCA chassis into the Detector Control System.

1. Introduction

The Large Hadron Collider (LHC) is situated between the borders of France and Switzerland. Inside the 27 km ring two beams of protons are collided at very high energies. This is achieved by guiding the beam using superconducting magnets inside an ultra-high vacuum beam pipe. The proton packets are crossed every 25 nano seconds with 20 or more collisions happening at each instant.

This results in over 600 million collisions every second. To detect and identify the fragments in each collision, the ATLAS detector is sub-divided into numerous sub-detectors, each one having a specific function. Figure 1 shows the sub detectors of ATLAS. It consists of the Inner Detector, Electromagnetic Calorimeters, Hadronic Calorimeters and Muon Detectors. The Tile Calorimeter (TileCAL), shown in the central region, is designed to measure energies and directions of hadrons, jets, τ leptons. The TileCAL consists of plastic scintillators which, when particles pass through, emit light to photo-multiplier tubes (PMTs) which is digitized and passed on to where the first layer of triggering occurs. This happens on the front-end electronics.

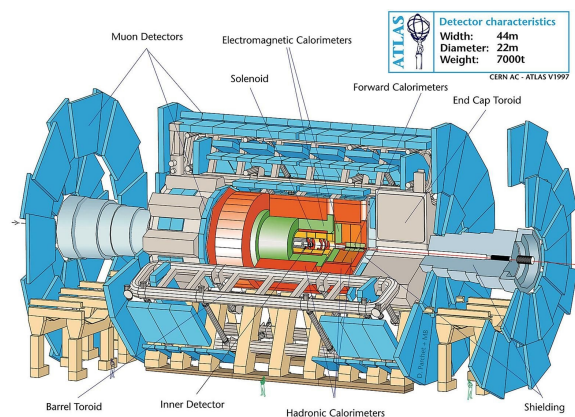


Figure 1. Representation of ATLAS detector showing structure of the sub-detector systems.

The front-end electronics are housed in superdrawers located on the outside of the TileCAL. If an event passes the triggering process, it is moved through to the Read Out Driver (ROD) to be sent to level two triggering. The ROD is located outside the detector in the back-end electronics inside a Versa Module Europa crate (VME crate). In order to meet the challenge of operating ATLAS at High Luminosity it is envisioned that the entire front-end electronics will be replaced [1].

1.1. Super Read Out Driver

The current front-end electronics consist of a long drawer which houses a 3-in-1 board, digitizer board and the interface board which packages the data for transmitting. Figure 2 shows the flow of data from the PMTs to the ROD. Data is generated at 40 MHz due to the number of bunch crossings per second and leaves the front-end electronics at 100 KHz to the back-end system.

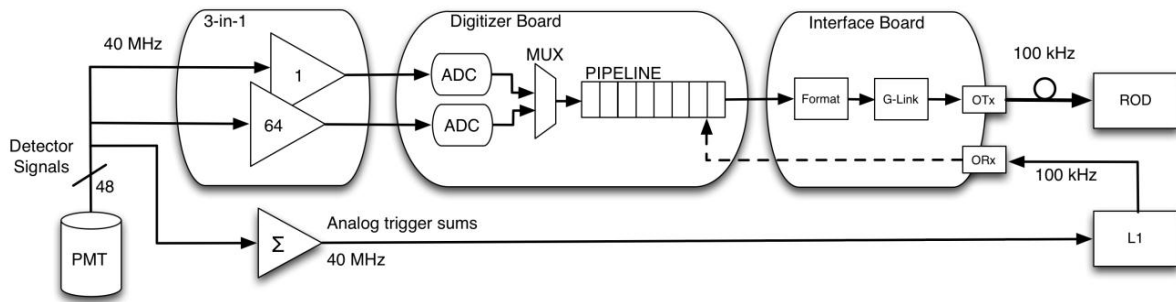


Figure 2. Current read out electronics.

During the 2022 upgrade the front-end electronics will be replaced with a new design. The triggering will become fully digital as well as being moved to the back-end but externally to the sROD. This will allow the full data rate of 40 MHz to exit the detector. Figure 3 shows the new data flow with the pipelines and triggering contained in the new super ROD (sROD) and accepting data at 40 MHz. Since the sROD is housed in the back-end it will be accessible during run time allowing easy maintenance and troubleshooting. Table 1 shows the expected bandwidth that will be provided by the new sROD and the TileCAL as a whole [2]. The sROD will be housed in an Advanced Telecommunications Computing Architecture (ATCA) chassis. The older VME crates will be phased out. The first prototype of the super Read Out Driver is currently in production. When it has been tested the designs for a revision board will be brought to South Africa where the PCB will be manufactured and all components mounted. The sROD will be inserted into the ATCA system at the University of the Witwatersrand and the integration with the WinCC software can be tested.

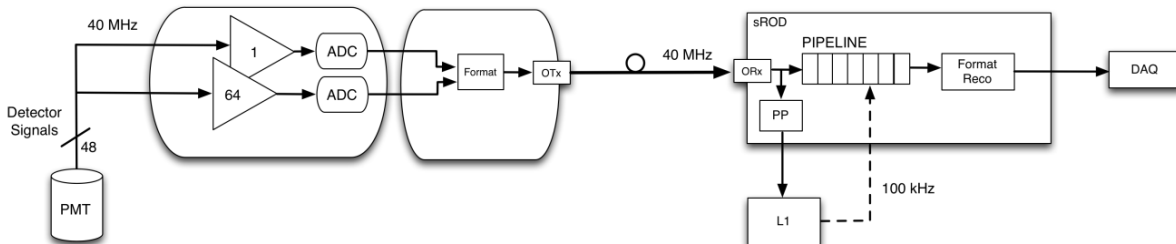


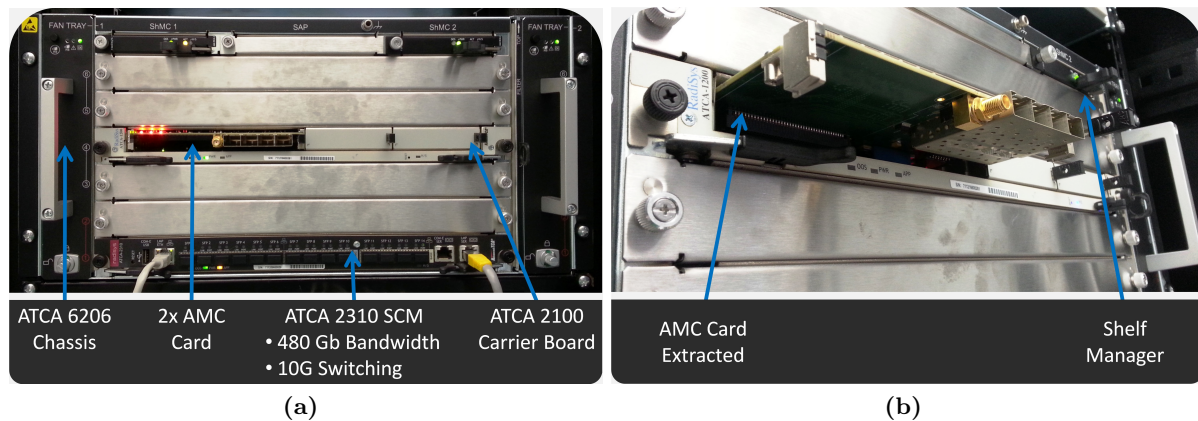
Figure 3. Upgraded read out electronics.

Table 1. Table Showing Bandwidth for Present and Future Design.

Phase	Present	Upgrade
Number of fibers	256	4096
Fiber Bandwidth	800 Mbps	10 Gbps
Total Bandwidth	205 Gbps	41 Tbps

2. ATCA Framework

The ATCA chassis is an intelligent shelf with high speed backplane connectivity. The chassis contains the backplane, power entry modules, fan trays and interfaces for individual blades to be installed. Figure 4 shows a front view of the chassis that is currently installed at the University of the Witwatersrand. This model has six slots with the first two connecting to all the others. This is called a dual star topology and offers the option of having a redundant switch installed. The connectivity is managed by a 10 Gbps switch module, providing routing and switching configuration. A carrier board provides power and monitoring to smaller Advanced Mezzanine Cards (AMC). The sROD is a doubled sized AMC card that can be inserted as shown in figure 4 (b). All boards can be "Hot Swapped" which means inserted or extracted while the chassis is still operational. The functionality is controlled by the Shelf Manager which provides various methods of accessing and controlling the devices such as the Simple Network Management Protocol (SNMP). The ATCA follows the PCI Industrial Computer Manufacturers Group (PICMG) standard. PICMG, as defined on their website, is "a consortium of companies who collaboratively develop open specifications for high performance telecommunications and industrial computing applications." [3].


Figure 4. a) Front of ATCA chassis and b) AMC extracted from chassis at Wits.

3. Integration into the Detector Control System

The ATLAS Detector Control System (DCS) is a complex system of hardware monitoring drivers arranged in tiers according to geographic, function or dependency groups. The software used at CERN is called WinCC Open Architecture designed by Siemens [4]. The suggested channel to insert the ATCA system can be seen in figure 5. The demonstrator setup will be installed in the Long Barrel A Partition of the TileCAL.

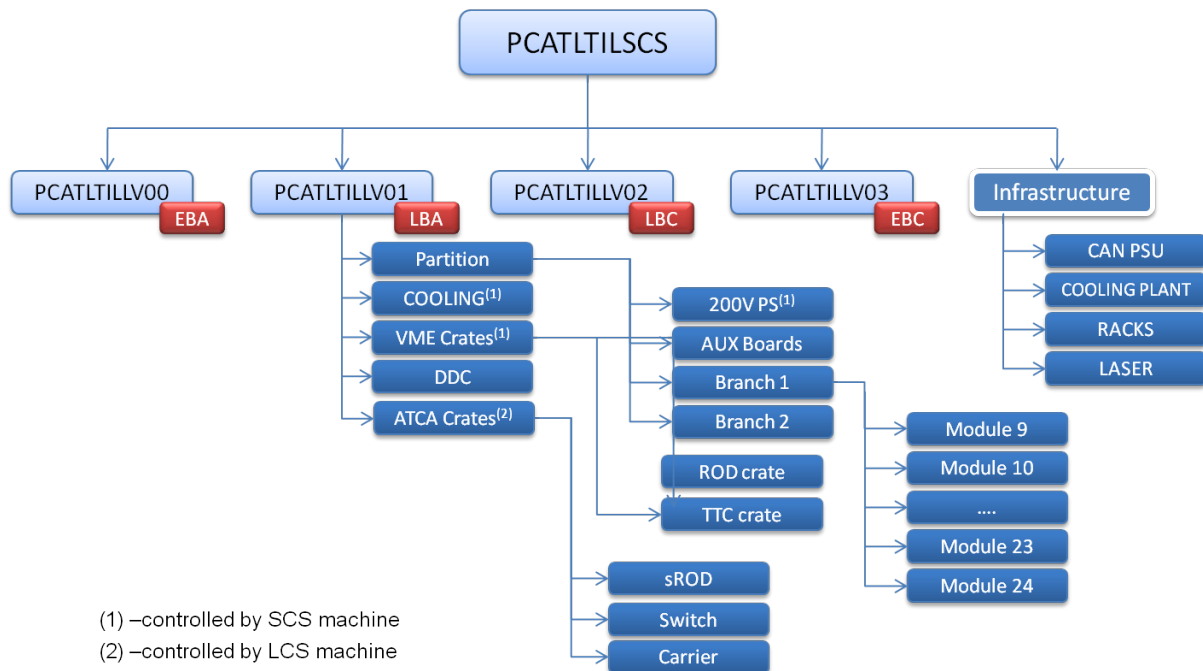
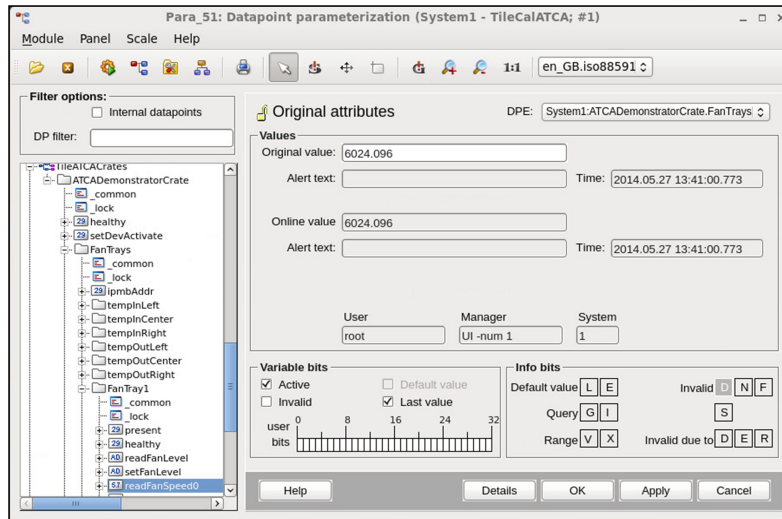


Figure 5. Proposed channel in DCS structure for ATCA Crates.

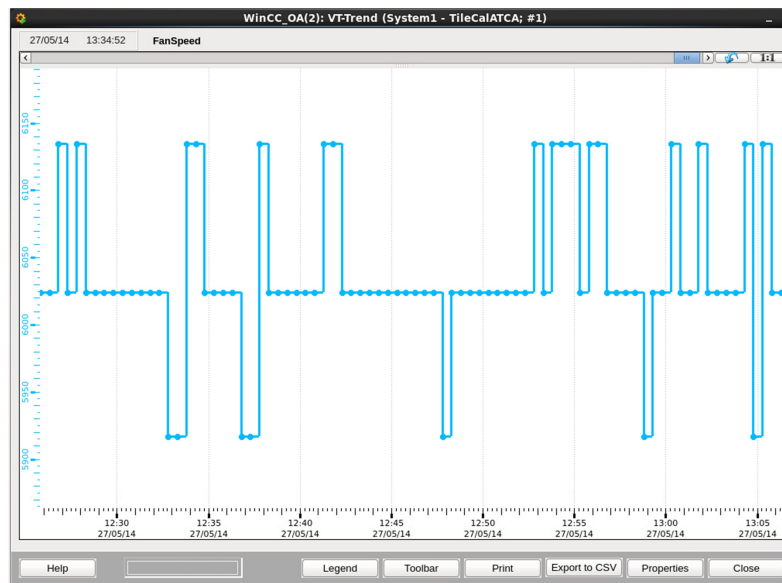
In order to properly integrate the chassis with the DCS system, all variables of interest in the sROD and ATCA chassis must be actively polled to provide the relevant information. The design will use SNMP to retrieve sensor readings via a standard Ethernet connection. This simplifies the infrastructure setup as Ethernet is common, cheap and understood. SNMP is a set of standards for network management which includes a list of data objects to monitor, a database to store and a protocol layer for communication. The device to manage is the ATCA chassis and all modules inserted. The internal control and management of modules is done by the ATCA shelf manager which acts as an SNMP agent. The WinCC software is the network management layer which retrieves information and issues commands to and from the SNMP agent via a standard Ethernet cable. A single variable has been successfully integrated into the WinCC software. The data point structure can be seen in figure 6 (a) which includes all variables of the ATCA. These data points were generated by a script which automates the creation process to make it easier to add more chassis. Figure 6 (b) shows the successful archiving of a single fan speed sensor represented as a trend over a few minutes.

4. Conclusions

The CERN community will be adopting the new ATCA and μ TCA systems to replace the majority of the current VME crates in future upgrades. The work up to now has facilitated the understanding of these new systems. It has been shown that this new SNMP approach does work while using a protocol that is common and already understood. The integration into the DCS system, although not quite plug and play, is possible due to the inbuilt SNMP features in WinCC. The future work will entail script writing to automate a large amount of the work as well as design control scripts to automatically respond to certain events. This work will need to be thoroughly tested before the final integration into the ATLAS DCS.



(a)



(b)

Figure 6. a) Data point structure of ATCA crate and b) Trend graph of single fan speed sensor.

Acknowledgements

The University of the Witwatersrand has facilitated the research by purchasing the required hardware to allow hands on development for this work and for other big science applications. The National Research Foundation (NRF) is thanked for the bursary they provided. The SA-CERN program provided financial assistance for research visits to CERN. I would also like to thank the School of Physics, the Faculty of Science and the Research Office at the University of the Witwatersrand.

References

- [1] The ATLAS Collaboration 2012 Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment URL <http://cds.cern.ch/record/1502664?ln=en>
- [2] Carrió F, Ferrer A, Castillo V, Hernández Y, Higón E, Fiorini L, Mellado B, March L, Moreno P, Reed R, Solans C, Valero A and Valls J A 2013 *ATLAS Note, The sROD Module for the ATLAS Tile Calorimeter Phase-II Upgrade Demonstrator* URL <http://cds.cern.ch/record/1628753?ln=en>
- [3] PICMG PICMG - About Us URL <http://www.picmg.org/v2internal/aboutus.htm>
- [4] Siemens 2014 SCADA System SIMATIC WinCC - HMI Software URL <http://www.automation.siemens.com/mcms/human-machine-interface/en/visualization-software/scada/Pages/Default.aspx>

The ATLAS Tile Calorimeter hybrid demonstrator

Matthew Spoor and Bruce Mellado Garcia

School of Physics, University of the Witwatersrand, Johannesburg, South Africa

E-mail: matthew.spoor@cern.ch, Bruce.Mellado.Garcia@cern.ch

Abstract. The LHC is currently preparing for the Phase II upgrade that is scheduled for 2022. Once upgraded the LHC will begin its high its High Luminosity Phase (HL-LHC) increasing the luminosity by a factor 5-7. This will vastly increase the amount of particles it could detect potentially addressing the unsolved mysteries such as dark matter. A new hybrid demonstrator prototype for the ATLAS TileCal is to be installed during the LS1. The demonstrator system needs to be fully compatible with current detector electronics as well as act as a validation for the new read-out architecture. This report gives overview of the on-detector and off-detector electronics used for the current demonstrator proposed for the Phase II.

1. Introduction

The Phase II upgrade [1] of the ATLAS detector will increase the original design luminosity by a factor of 5 to 7. There are number of upgrades that are planned for the ATLAS detector, with high radiation levels and increased data processing requirements there is a need for a complete redesign and replacement of the read-out electronics. Since the initial design there has been huge advances in technology leading to more advanced chips with greater radiation tolerances.

The current read-out system stores digitised samples in front-end electronics pipeline memories. Only samples that are selected by the L1 trigger are sent to the back-end electronics for further processing. The new electronics system will provide full digital readout of all channels that will be transmitted to the off-detector pipeline memories each bunch crossing (40 MHz) using multi-gigabit optical links. An improved back-end electronics system will now have to supply digital information to the First Level trigger per bunch crossing. This will provide digitally calibrated information with enhanced precision and granularity to the First Level trigger in order to improve both trigger latency and resolution. The new front-end architecture merges some boards to create a more compact 3 board system. Additionally new levels of reliability and redundancy have been introduced with many components being duplicated.

2. ATLAS Tile Calorimeter

The ATLAS (A Toroidal LHC AparatuS) experiment (Figure 1a) at the Large Hadron Collider (LHC) is a general purpose detector located at CERN. ATLAS comprises of several different sub-detector systems. The experiment was designed to record proton-proton collisions that occur every 50ns resulting in data flows in the 100s of GB per second range.

The Tile Calorimeter (TileCal) [2] is the central hadronic calorimeter of the ATLAS experiment. The TileCal is made up of 4 sections: two central barrels in the middle and two extended barrels on either side. Each of the four sections is divided into 64 azimuthal slices (Figure 1b). Within each slice there is a fine lattice of steel plates with plastic scintillating

tiles placed in-between. In total there are 430 000 tiles contained inside the TileCal making up 5000 calorimeter cells. Signals are recorded by 10 000 photomultiplier tubes (PMTs) linked to readout electronics. The Front-end electronics are stored within drawers at the back of every slice.

When particles enter the TileCal system they collide with steel plates which causes them break up into showers of charged and neutral particles. The TileCal has the critical role of measuring the energy and direction of these showers. As particles pass through scintillating materials they deposit tiny amounts of energy. Due to the molecular structure of tiles, this energy is converted into photons. These photons travel through the scintillators into wave length shifting optical fibers. This light is carried along the fibers to PMT modules where the light is converted into an analog electrical signal. The signal is then passed to the front-end electronics for amplification, digitisation and trigger selection after which they are sent to the back-end electronics for data formatting and further L2 trigger selection.

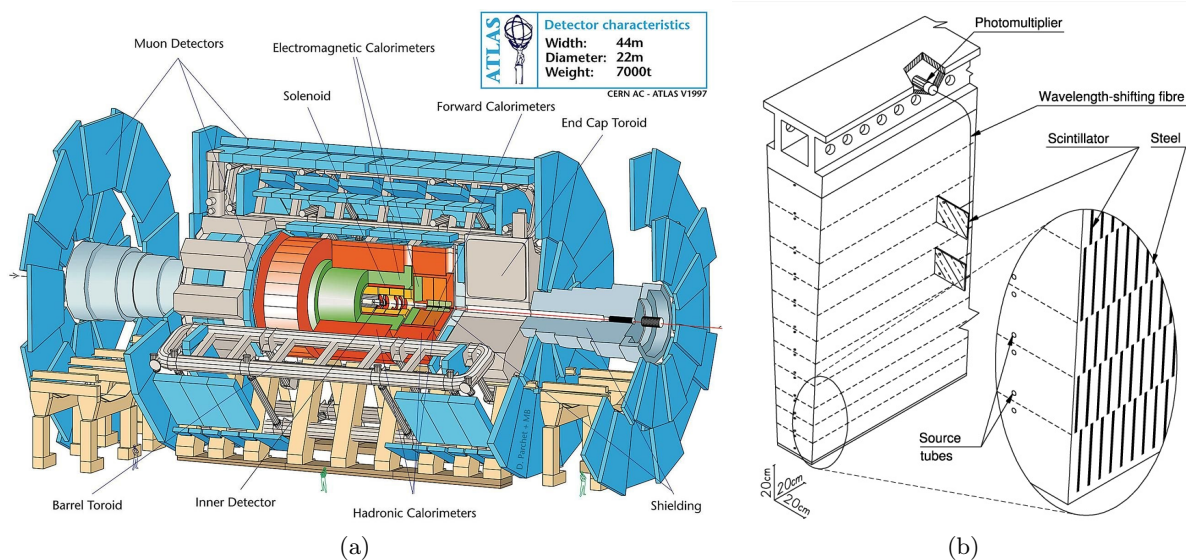


Figure 1: a) Tile Calorimeter Detector b) TileCal module.

3. Present read-out electronics

The present front-end electronics is made up of four board types: 3-in-1 one Front-End board (FEB), 3-in-1 Mainboard (MB), digitiser board and Interface board [3]. The multipurpose 3-in-1 FEBs amplifies and shapes PMT signals into 27 ns pulses as well as provides calibration functions. The main board is responsible for the control and calibration of the 3-in-1 cards. The digitiser samples the analog pulses using both high and low gains and stores them in pipeline memories. Low gain signals are merged in the adder board before being sent to trigger tower to be used for the L1 trigger. Data temporarily waits for trigger validation in the pipeline memory. If data is seen to be interesting it will be accepted by the trigger and sent to the Interface board where it is placed in derandomising buffers and read out. The Interface board will transfer data at 100 KHz to the Read out driver (ROD) to be formatted before being sent to the data acquisition system (DAQ) [2]. Data that is not accepted by the trigger will be written over once new data comes in.

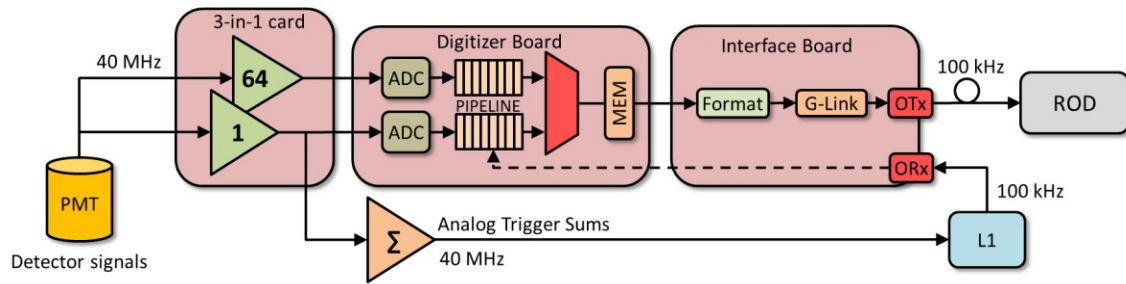


Figure 2: Present Tile Calorimeter readout architecture.

The current housing for the front-end electronics is made up of two 2m in length drawers forming a retractable superdrawer [4]. Each superdrawer can accommodate 4 MBs, 48 FEBs, 48 PMTs, 4 digitiser boards and one Interface board. The MBs of each draw are daisy chained together so all signals and power supplies are shared. This design has a single output to the Interface board for the entire draw.

4. Hybrid demonstrator

During the LS1 a new demonstrator prototype will be installed into an external slice TileCal for testing in Building 175 at CERN. The TileCal Demonstrator is essential for the validation of the performance of the new readout architecture, trigger system interfaces and the implementation of novel trigger algorithms before the complete replacement of electronics in Phase II upgrade [5].

A new front-end electronics architecture was proposed that divides each TileCal module into 4 separate minidrawers that are independent in terms of power, readout and configuration. Each minidrawer is water cooled and houses a Mainboard, daughterboard, 12 PMT modules and a High Voltage card. Four minidrawers can be mechanically connected to form a Superdrawer. The whole module communicates digitised data from each PMT channel through Quad 10 Gbps optical fibers every bunch crossing (40 MHz) to the sROD (Back-end). The downlink (sROD to Front-end) consists of four 4.8 Gbps links mainly used for FEB and HV configuration and DCS information.

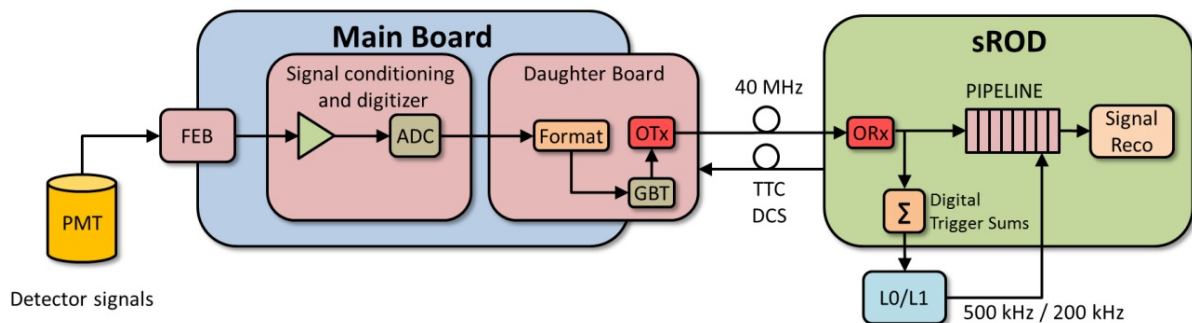


Figure 3: Upgraded Tile Calorimeter readout architecture.

4.1. 3-in-1 Card (FEB)

An improved version of the original 3-in-1 cards was chosen for the FEB in the current demonstrator. Designed by the University of Chicago the new 3-in-1 cards uses modern components allowing for better linearity, radiation hardness and lower noise [5]. As with

the original design the analog trigger output remained as to be compatible with the present system. The FEBs can be calibrated and monitored via commands sent from the daughter board. Commands can be used to charge and discharge capacitors on the 3-in-1 cards in order to produce calibration pulses used for testing.

4.2. Mainboard (MB)

Each minidrawer [6] in the demonstrator has a Mainboard connected to 12 PMT modules. The mainboard is symmetric on either side such that they will be able to function independently of one another. Each side receives 6 analog PMT signals that are digitised using a 12 bit Bi-gain sampling system resulting in two separate high and low digital signals. An additional output of each FEB is used for the analog integrator with is sampled at 50KHz and read out by a I2C bus. The phase of the signal sampling clocks can be adjusted to compensate for path length delays between MB and DB. All electronics on the MB are commercial off the shelf components. Signals are sent to the DB via a 2 bit serial bus at 560 Mbps. The control and board configuration is done using 4 Altera Cyclone IV FPGAs while the digitisation of the FEB analog signals is done by 12-bit linear Technology LTC 2264 ADCs.

4.3. Daughterboard (DB)

The Daughterboard [7] acts as the interface between the front-end electronics and the sROD. While designing the board major aspects focused on were radiation hardness, high speed readout and redundancy [8]. All minidrawer electronics can be divided in two identical sections: sides A and B. Each side receives data from 6 PMTs and is managed by a Kintex 7 FPGA that is directly connected to the sROD through 10 Gbps optical fibre. In the original front-end design a single Interface board was used to communicate with the ROD for an entire superdrawer.

The 3rd generation daughter board currently used on the Hybrid Demonstrator was developed by Stockholm University. The board uses dual Xilinx Kintex 7 FPGAs, two Quad Small Form-Factor Pluggable (QSFP) optical transceivers and two Giga Bit Transceiver (GBTx) chips. The GBTx chips are used to serialise and de-serialise GBT data streams. The DB is able to communicate directly with the MB through a 400 pin FMC mezzanine connector. The DB receives digitised PMT data through dedicated point-to-point differential signals at a rate of 40 MHz from the MB. The Kintex 7 FPGA then collects the PMT data, formats it and transfers it to the sROD every 25ns. The down-link (sROD to DB) is used to send commands received from Detector Control Systems (DCS) to configure, control and monitor the DB, FEBs, MB and high voltage Board.

4.4. sROD

The Super Read-Out Driver (sROD) demonstrator board [8] forms part of the back-end electronics. Each sROD will be connected to a single superdrawer receiving data from four DBs. The boards primary responsibility will be reception and processing of data received from the detector but will also need to send preprocessed data to the trigger system and act as the interface between the DCS and front-end electronics. The sROD will also implement Trigger, Timing and Control (TTC) functionalities for synchronisation of the read out chain.

The board will utilise a Xilinx Virtex 7 and a Kintex 7 FPGA for its processing needs, four QSFP modules, a send a receive Avago MiniPOD and a SFP mezzanine connector. The Board was built to fit into Advanced Telecommunications Computing Architecture (ATCA) using a double mid-size Advanced Mezzanine Card (AMC).

5. Demonstrator progress

On 9th of June 2014 CERN had a ATLAS TileCal week. During this week the experts from all contributing groups of the Demonstrator prototype came to work on the demonstrator. During

this week four minidrawers were fully assembled and connected together (Figure 4). The week allowed the experts to perform the first tests on a fully assembled superdrawer. A total of 45 PMTs, 4 Motherboards, 4 High voltage Cards and 4 daughter boards were needed to be assembled and connected together. The week was very successful with many new problems being identified as well as confirmation that all the components fitted together correctly and that the entire drawer was able to fit in a slice of TileCal.

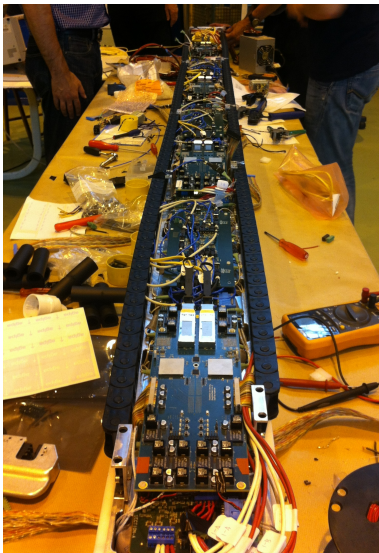


Figure 4: Assembled Superdrawer

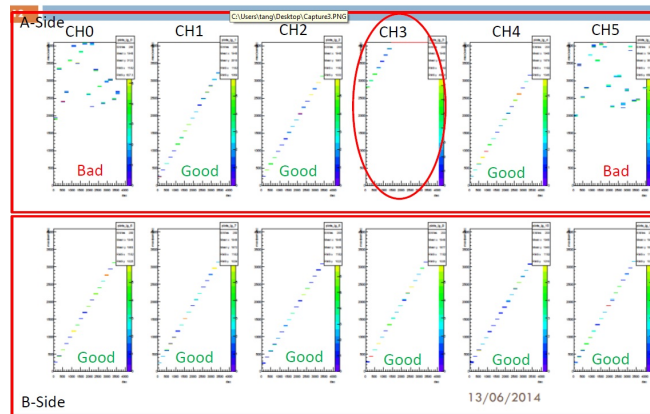


Figure 5: Linearity plots of high-gain channels in mainboard

5.1. Linearity tests

Figure 5 displays linearity plot tests performed on a minidrawer. Here the signal linearity is being checked for the Mainboard high gain channels of one minidrawer. In this test PMTs are turned off and the pedestal setting for each channel are incrementally stepped up. Since there is no signal from the PMTs the plot should be a straight line with minor variations due to noise. Due to firmware problems and calibration issues not all the plots are correct. Currently these issues are in the process of being resolved.

5.2. sROD emulator

The sROD described in section 4.4 is still in development and therefore it was not possible for it to be used for the testing of the minidrawers. The demonstrator team therefore have been using a Xilinx VC707 development board to emulate the sROD as an alternative. The VC707 is similar to the sROD in that it also uses a Virtex 7 FPGA. The high speed QSFP connections to the front-end electronics were possible by attaching additional FMC mezzanine modules to the VC707. The Firmware being developed for the emulator could later be modified to be compatible with the final sROD board.

The current setup used for testing the minidrawers has a VC707 board connected to a single minidraw. One problem faced with this testing setup is that a new IP address and/or MAC address is needed for the sROD when it is connected to a new network. These addresses can be changed by modifying the Firmware but this requires a full firmware rebuild and the reprogramming of the FPGA. To solve this time consuming process a basic interface on the board was created that allows the user to either select a pre-defined address from a list or input a completely new address, all while the board was running. The design uses a LCD module

to display the addresses and four push buttons to act as configuration inputs. The new design saves time and adds a new level of reconfigurability to the demonstrator.

6. Summary

The increase in luminosity predicted after the Phase II upgrade requires a complete redesign of the read-out electronics in the TileCal. The assembly of the Hybrid demonstrator in June greatly assisted with the initial validation of the electronics as well as allowed group members to gain valuable experience. Many more tests still need to be completed and the firmware is still under development but the demonstrator should be ready before the end of LS1.

Acknowledgments

I would like to thank the University of the Witwatersrand and the SA-CERN program for providing the research opportunity to visit CERN. The TileCal upgrade collaboration for significant contributions to this work. I would also like to thank the School of Physics, the Faculty of Science and the Research Office at the University of the Witwatersrand.

References

- [1] Collaboration A 2012 Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment *CERN-LHCC-2012-022*
- [2] Collaboration A 2008 *Journal of Instrumentation* **S08003** 6 URL <http://iopscience.iop.org/1748-0221/3/08/S08003>
- [3] Tang F, Anderson K, Muschter S, Bohm F, Akerstedt H and Carrió F 2013 Upgrade Analog Readout and Digitizing System for ATLAS TileCal Demonstrator *ATL-Tilecal-Proc-2013-019* URL <http://cds.cern.ch/record/1628747>
- [4] Ajaltouni Z and al E 1994 The drawer system concept for the ATLAS Tilecal readout Tech. Rep. 43 URL <http://inspirehep.net/record/1195072/files/tilecal-94-043.pdf>
- [5] Carrió F 2013 *European Physical Journal* **60** 3 URL <http://cds.cern.ch/record/1697775>
- [6] Tang F, Akerstedt H, Anderson K, Bohm F, Hildebrand K, Muschter S and Oreglia M 2014 Design of Main Board for ATLAS TileCal Demonstrator *TL-TILECAL-PROC-2014-004* pp 7–9 URL <http://cds.cern.ch/record/1703525>
- [7] Muschter S, Anderson K, Bohm C, Eriksson D, Oreglia M and Tang F 2013 *Journal of Instrumentation* **C03025** C03025–C03025 ISSN 1748-0221 URL <http://stacks.iop.org/1748-0221/8/i=03/a=C03025?key=crossref.128fb875541fdc384368fe103f3b5c93>
- [8] Carrió F and al E 2013 The sROD Module for the ATLAS Tile Calorimeter Phase-II Upgrade Demonstrator *ATL-TILECAL-PROC-2013-020*

The Higgs as a portal to the hidden sector via an analysis of $H \rightarrow Z_d Z_d \rightarrow 4l$ using the ATLAS detector

D A Unwuchola^{1,2}

¹ University of Johannesburg, Johannesburg, South Africa

² On behalf of the ATLAS Collaboration

E-mail: doomnull@gmail.com

Abstract. The Standard Model (SM) has well known deficiencies, and there is clearly need for new physics beyond the SM. The particles manifesting the new physics would interact at most weakly with the SM particles, and hence they are termed dark. The Higgs boson is potentially a favourable route for the production of the dark particles. There are a large class of theories where couplings or mixings at the Higgs level leads to exotic Higgs decays, which nonetheless do not significantly disturb the known physics below the Higgs level. This is therefore a significant potential discovery opportunity. We present the motivation and progress made in the studies which have been carried out as part of designing the search for the exotic decay of the SM Higgs which proceeds via a dark force back to SM four leptons, $H \rightarrow Z_d Z_d \rightarrow 4l$ from the LHC run 1 data using the ATLAS detector.

1. Introduction

Beyond Standard Model (BSM) phenomenology presents the Higgs boson as a sensitive discovery channel for particles that couple weakly to Standard Model (SM) particles. The potential for a new sector to exist without disturbing the existing good agreement between the SM and all previous data including LHC data arises because the new sector primarily couples to the SM at the Higgs level. This reflects the universality of the Higgs interaction to many theoretical incarnations of "new Physics". The Higgs therefore provides a connection to new physics not charged under SM forces. In addition, the branching ratio of the Higgs in the SM is narrow and the branching ratio of the Higgs to exotic particles can be arranged within the details of the models to be relatively unconstrained by the existing data-theory agreements. This accommodates a significant fraction of its total decay width being available for exotic decays. In a very extensive article, Curtin et al [1] argue that this is therefore an extremely significant window to BSM Physics at the LHC. This motivates the term "Higgs portal" for this type of access to the discovery of possible new dark particles. Essentially it will manifest via an exotic Higgs decay and can be directly searched for. The phenomenology is very rich, and we have chosen a final state signature whereby the discovered Higgs is actually the lightest partner of a Higgs multiplet that arises from the mixing between the BSM dark sector and the SM Higgs. This new Higgs decays via a dark force particle in analogy to the SM, a new dark and light neutral boson, Z_d , which in turn decays back to SM leptons. The coupling of the Z_d to SM leptons is set at the maximum value accommodated by the current agreement between theory and experiment. It is further possible, that even in this case, the phenomenology can be arranged such that the branching ratio of the $Z_d \rightarrow 4l$ is essentially 100%. Our discovery signature is therefore $H \rightarrow Z_d Z_d \rightarrow 4l$. In the rest of the paper, we present an overview of the analysis procedure. The results which include data are will be presented in a future paper.

2. Analysis

The analysis for the exotic Higgs decay parallels that of the four lepton Higgs discovery channel, $H \rightarrow ZZ^* \rightarrow 4l$, since they share the same primary and final states. The major difference arises from the fact that in the intermediate state, we have an unknown Z_d mass, and so we replace the cut for the on-shell Z by a requirement that both the Z_d bosons are on shell so that their masses are equal. There are of course some other differences, some of which are mentioned below.

2.1. Pre-selections, Cuts

The event and object selections, kinematic requirements and trigger matching are the same as in the four lepton Higgs discovery analysis using the ATLAS detector. The isolation requirements for each lepton within the detector, the separation between a lepton pair (ΔR) and the significance for the reconstruction of the interaction point (IP) are also the same. The reader is referred to the Higgs discovery paper [2] and references therein for details. These similarities are shown in green in the table reproduced as figure 1.

A subtlety arises in the consideration of which association to use in the reconstruction of the bosonic parents for the lepton pairs in the final state quadruplet. When all four leptons have the same flavour (but of course there are two opposite sign pairs), then there is an ambiguity in the association of the two di-lepton pairs. This is resolved in the standard analysis by the requirement that one of the dilepton pairs reconstructs to give the correct mass of the SM Z . In this exotic decay analysis, the pairs have the association chosen to minimise the mass difference in the two reconstructed Z_d bosons, which are both required to be on-shell.

2.2. Definition of the Signal Region

The cuts defining the Signal Region (SR), are also different between the exotic analysis and the standard analysis. One can note in the table the following cuts - selection of quadruplets consistent with the discovered Higgs as a parent, $100 < m_{4l} < 150$ GeV, the requirement for the di-lepton pairs to reconstruct to an on-shell Z_d , $m_{4l} < 63$ GeV and the requirement that each reconstructed Z_d has the same mass, $\Delta m < 15$ GeV. The cuts have boundaries which are the results of studies to optimise them, and which recognise systematic effects in the ATLAS detector. The Δm is in fact still being studied, so the value given here is nominal (see section 2.4).

The consideration of various processes which may masquerade as signal, has led to the introduction of vetoes arising from Z s, J/Ψ s and Υ s. Cuts are applied in these cases on the di-lepton events at a late stage in the analysis. As the association procedure of dilepton pairs mentioned above may be ambiguous for these backgrounds, the veto is applied if any of the two di-lepton associations falls foul of it.

These differences in the SR definition are shown in red in the same table. The analysis code can be switched to be consistent with either the standard or the exotic analysis. The cut-flow (measurements of the cut efficiency at each stage) using the analysis code configured for the standard analysis was compared to the cut-flow as established for the initial discovery and current study of the discovered Higgs, and the excellent agreement here is a partial benchmark for validating the analysis code to be deployed in the Z_d analysis.

2.3. Monte Carlo study of the signal and background samples

The process $H \rightarrow ZZ^* \rightarrow 4l$ of the standard analysis, which is the signal for the SM Higgs discovery, is actually a background for the Z_d analysis. In addition electroweak vector boson pair production from $q\bar{q}$ annihilation and loop-induced gluon fusion, $ZZ \rightarrow 4l$ is a further irreducible background. In these cases there is a sizeable correlated production of objects that can manifest in the Z^*Z^* region. These processes are well modelled by Monte Carlo (MC). Other processes considered for the contribution to background include the production and decay of $t\bar{t}$,

Table 1. A summary of the cuts applied in four lepton Higgs discovery channel $H \rightarrow ZZ^* \rightarrow 4l$ and the current analysis $H \rightarrow Z_d Z_d \rightarrow 4l$. The table is explained in the text.

Cuts	$H \rightarrow ZZ \rightarrow 4l$	$H \rightarrow Z_d Z_d \rightarrow 4l$
4 l, kinematics, trigger matching	SAME	SAME
Primary pair mass (m_{12})	$50 < m_{12} < 106$	No cut!
Secondary pair mass (m_{34})	$X < m_{34} < 115 (X = f(m_{4l}))$	No cut!
1 unique quadruplet	m_{12} closest to m_z	$ \Delta m = m_{12} - m_{34} $ minimal
$\Delta R > 0.1[0.2]$ for SS [OS] pairs	SAME	SAME
track and call isolation	SAME	SAME
IP significance	SAME	SAME
Signal Region		$ \Delta M < 10 \text{ GeV}$ $100 < m_{4l} < 150 \text{ GeV}$ $m_{12}, m_{34} < 63 \text{ GeV}$ Z veto

$Zb\bar{b}$ and Z +jets and which produce leptons or jets faking leptons. The J/Ψ and Υ have also been considered as they can be correlated with a Z through the associated quarkonium production mechanism from double parton scattering. These backgrounds have been also modelled in MC, however the statistics available are less than optimal. The Parton Distribution Functions (PDF) sets on which primary scatterings are based, the event generators for the processes mentioned and the description of the full ATLAS simulation based on the GEANT4 framework are mostly described in reference [2].

The signal has been modelled using a particular incarnation of a Higgs Portal model in the $H \rightarrow Z_d Z_d \rightarrow 4l$ exotic decay signature, the Wells Model [3, 4]. This is done at the generator level in order to represent the initial lepton quadruplet kinematic distributions and angular correlations. The actual signal rate has been scaled to the SM rate for $H \rightarrow ZZ^* \rightarrow 4l$ simply for convenient comparison. The use of a model at generator level allows a study of backgrounds and cut optimisation. The final analysis proceeds in some sense rather more model independently, as it is a limit setting procedure based on testing for an excess over the well studied backgrounds in the SR. The left and right parts of figure 1 and the left part of figure 2 indicate the contribution of the signal and the various background distributions viewed as a function of the reconstructed mass for the full lepton quadruplet, the two di-lepton pairs and the mass difference between the two di-lepton pairs. The signal is shown for 2 masses at the limits of our SR for the Z_d . The rest of the description of these plots can be found in the captions. Only the case for the $4e$ final state is shown. There are of course subtleties associated with the cut efficiency and MC distribution validation and therefore the treatment of each different final state, even if in general, there are similarities. These distributions are frozen after the IP level cut, just before the SR cuts are applied. In the di-lepton case of figure 1 at the left, in the case of the ZZ background, a peak is visible at the SM Z boson mass, and a tail at lower mass comes from events containing an off-shell Z boson.

In the case of the lepton quadruplet of figure 1 at the right, the peak at 125 GeV corresponds to $H \rightarrow Z_d Z_d \rightarrow 4l$ events and also to $H \rightarrow ZZ^* \rightarrow 4l$ events. The $ZZ^* \rightarrow 4l$ background is mostly confined to $m_{4l} > 180 \text{ GeV}$ (two on-shell Z bosons). At the region of the SM Z mass we see a peak from the s -channel single on-shell Z production where one of the decay leptons emits an off-shell Z . It is clear that this process still reconstructs to a single Z mass in the m_{4l} spectrum. In the case of the mass difference between the two di-lepton pairs of figure 1 at the left, we see as expected a signal peak in the region of low mass difference, and also a peak in the ZZ background.

In the case of the right section of figure 2, a 2D view of the MC signal and background is

provided, which further illuminates the SR cuts to be applied. Further description of the figure is provided in the caption.

The SR cuts mentioned above clean the background in the SR region very effectively, so that it is dominated by the signal. Evidence for this is not shown here, due to space limitations.

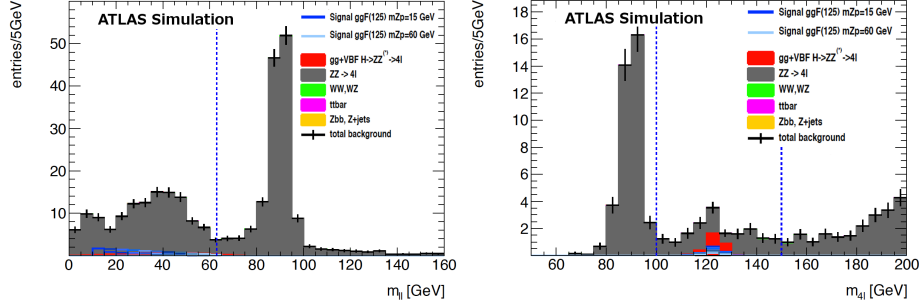


Figure 1. Left : Dilepton invariant mass for all signal and background samples, in the $4e$ final state. All distributions are normalized to the same integrated luminosity (20.7 fb^{-1}) and obtained before applying the Signal Region requirement symbolized by the blue dashed line at 63 GeV. Right : Four-lepton invariant mass for all signal and background samples, in the $4e$ final state. The blue dashed lines show the $100 < m_{4l} < 150$ GeV cut applied on this variable to define the Signal Region. All distributions are normalized to the same integrated luminosity (20.7 fb^{-1}) and obtained before applying the Signal Region requirement.

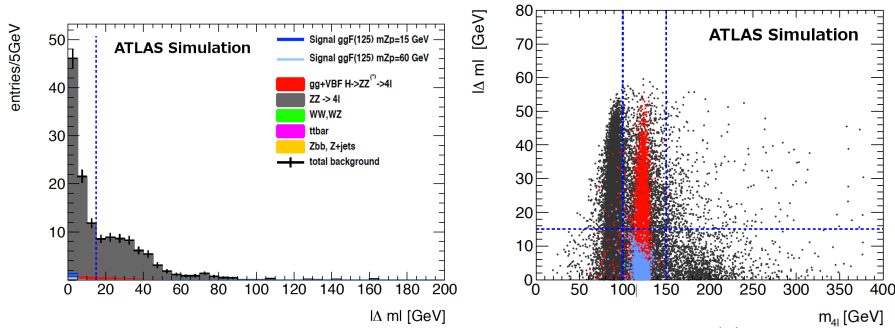


Figure 2. Left : Absolute mass difference for all signal and background samples, in the $4e$ final state. The blue dashed line shows the cut applied on this variable to define the Signal Region. All distributions are normalized to the same integrated luminosity (20.7 fb^{-1}) and obtained before applying the Signal Region requirement. Right : Mass difference Δm as a function of the four-lepton invariant mass m_{4l} for background $ZZ^* \rightarrow 4l$ in dark gray, $H \rightarrow ZZ^* \rightarrow 4l$ in red) and signal $H \rightarrow Z_d Z_d \rightarrow 4l$ with $m_{Z_d} = 50$ GeV, in blue) events, in the $4e$ final states. The blue dashed lines show the limits of the loose Signal Region : $\Delta m < 15$ GeV and $100 < m_{4l} < 150$ GeV.

2.4. Background estimation and Systematics

The $H \rightarrow ZZ^* \rightarrow 4l$ and $ZZ \rightarrow 4l$ processes are estimated by MC simulation and subtracted. Other di-boson process such as WW and WZ which have survived the cuts are also well modelled and can be removed based on MC simulation. The remaining $t\bar{t}$, $Zb\bar{b}$ and Z +jets are estimated with a data driven method, known as the ABCD method. The ABCD method is applied to extrapolate the remaining backgrounds from a more general Control Region (CR) into the SR. In the ABCD method, two uncorrelated variables (observables) are selected. Passing or failing a cut on these variables defines a SR and CR for each of them. In the 2D product space of these two

variables, one will then have a single SR region quadrant and three other CR region quadrants. The backgrounds to be modelled by MC are first removed, and the remaining events (data or MC depending on whether its analysis or a study) will be distributed in these four quadrants. The quadrant representing the SR in both variables can then be estimated by considering that the ratios of events in either horizontally or vertically neighbouring quadrants must be equal (assumption of non-correlation in the two variables). In practice, the non-correlation can be estimated from MC, and this can be applied as a correction. This analysis has now converged on the selection of the two uncorrelated variables as the isolation requirements on the each dilepton pair (the two leptons of the pair pass or fail the cut combining the track-based isolation, calorimeter-based isolation and impact parameter significance requirements). The method is applied to the SR as defined for all cuts. A final procedure to optimise the Δm cut, based on achieving the best possible exclusion limit, is still in process.

The limit setting procedure for discovery or exclusion is based on the best possible knowledge of the statistical and systematic uncertainties in all quantities involved in the quantitative understanding of the data. In the case of the systematic uncertainties, theoretical systematics contribute in the MC evaluation process. These are related to the EW and QCD models themselves as well as to the PDFs which lead ultimately to uncertainties on the cross sections. Then there are the detector systematics relevant to the electron and muon identification efficiency and uncertainty on the signal yields. There is the luminosity uncertainty which affects the scaling of the MC yields and the evaluation of the signal yields. There is also the data driven systematics, as applicable to the ABCD method. This consists in the estimates of the non-correlation and to the uncertainties due MC statistics in the method development. This process is almost completed for this analysis.

Table 2. Acceptance times efficiency for signal and background in the Signal Region (defined with $|\Delta m| < 15$ GeV), for the $4e$ final state. The expected number of signal events is normalized to the SM $H \rightarrow ZZ^* \rightarrow 4l$ rate in the $4e$ channel.

Process	Events processed	Events in the SR	Acc. \times eff.	Expected events ($20.7 fb^{-1}$)
Signal ($m_{Z_d} = 20$ GeV)	29908	1674.5	5.6%	6.16
Signal ($m_{Z_d} = 50$ GeV)	29939	2054.8	6.86%	7.56
(gg+VBF) $H \rightarrow ZZ^* \rightarrow 4l$	119.11	0.42	0.36%	0.42
$ZZ \rightarrow 4l$	$1.81 \cdot 10^5$	0.12	$6.44 \cdot 10^{-5}\%$	0.12
WW, WZ	60170	0	0%	0
$t\bar{t}$	$5.50 \cdot 10^5$	0.037	$6.65 \cdot 10^{-6}\%$	0.037
Zbb, Z +jets	$7.14 \cdot 10^7$	0	0%	0
(Z +) low mass	55577	0	0%	0

2.5. Derivation of a limit for the case of exclusion or discovery.

Table 2 presents the signal and background events expected for the $4e$ channel in the SR where the expected number of signal events is normalized to the SM $H \rightarrow ZZ^* \rightarrow 4l$ rate. The data can also be processed by the analysis code and the number of data events in each SR bin can be obtained. It will then be possible to compare the expected background and the data, considering the uncertainties on each.

The signal strength μ , is defined as the ratio of the $H \rightarrow Z_d Z_d \rightarrow 4l$ rate relative to the SM $H \rightarrow ZZ^* \rightarrow 4l$ rate:

$$\mu = \frac{\sigma \times BR(H \rightarrow Z_d Z_d \rightarrow 4l)}{[\sigma \times BR(H \rightarrow ZZ^* \rightarrow 4l)]_{SM}} \quad (1)$$

A quantitative and robust interpretation of this data as limits for the case of exclusion will be computed from a maximum likelihood fit to the numbers of events in the various signal

regions following the CL_s modified frequentist formalism [5, 6] with the profile likelihood test statistic [7].

As an indication of the type of result that can be expected, the result of Curtin based on the analysis in reference 3 is presented in figure 3.

The case for discovery is closely related and not discussed further here.

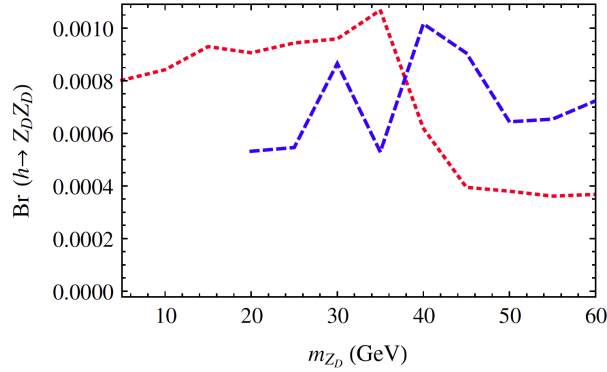


Figure 3. A 95% CL exclusion contour for the exotic Higgs decay $H \rightarrow Z_d Z_d$ based on the assumption of $BR(Z_d \rightarrow 2l) = 30\%$ taken from Curtin [1] figure 30. In the case of this analysis, the limit will be presented on the excluded signal strength μ , defined as the ratio of the $H \rightarrow Z_d Z_d \rightarrow 4l$ rate relative to the SM $H \rightarrow ZZ^* \rightarrow 4l$ rate. This diagram is very important to be referenced as it gives a good picture of what our result would look like.

3. Conclusion

In conclusion, we have presented the motivation for a study of the exotic Higgs decay $H \rightarrow Z_d Z_d \rightarrow 4l$. This is based on using the discovered Higgs as a portal to new physics BSM. The analysis is well progressed, and has been discussed to the stage of limit setting, in the case of either exclusion or discovery.

Acknowledgments

The authors acknowledge the support of the National Research Foundation (NRF) and Department of Science and Technology, both of South Africa, who have supported the research through the SA-CERN programme. The NRF contributed to bursaries and the NRF and the Universities of Johannesburg and Cape Town contributed to post doctoral fellowships.

References

- [1] Curtin D, Essig R, Gori S, Jaiswal P, Katz A, Liu T, Liu Z, McKeen D, Shelton J, Strassler M, Surujon Z, Tweedie B and Zhong Y M 2014 *Phys. Rev. D* **90**(7) 075004 URL <http://link.aps.org/doi/10.1103/PhysRevD.90.075004>
- [2] The ATLAS Collaboration 2012 *Physics Letters B* **716** 1 – 29 ISSN 0370-2693 URL <http://www.sciencedirect.com/science/article/pii/S037026931200857X>
- [3] Kane G and Pierce A (eds) 2008 *Perspective on LHC Physics* (World Scientific) ISBN 978-981-277-975-5
- [4] Gopalakrishna S, Jung S and Wells J D 2008 *Physics Review D* **78** 055002
- [5] Read A L 2002 Presentation of search results: The $cl(s)$ technique Tech. Rep. J.Phys.G G28 (2002) 2693 – 2704
- [6] Cousins R D and Highland V L 1992 *Nucl. Instrum. Meth. A* **320** 331–335
- [7] Cowan G, Cranmer K, Gross E and Vitells O 2011 *Eur. Phys. J. C* **71** 1554 (*Preprint* 1007.1727)

Searching for BSM physics in Higgs to WW coupling with e^+e^- collisions

Stefan von Buddenbrock¹, Alan S. Cornell² and Bruce Mellado³

School of Physics, University of the Witwatersrand, Wits 2050, South Africa

E-mail: ¹stefan.vonbuddenbrock@students.wits.ac.za, ²alan.cornell@wits.ac.za,

³bmellado@mail.cern.ch

Abstract. The study of Higgs production in e^+e^- collisions presents us with an avenue for studying Higgs to WW coupling in the t -channel. Our understanding of the tensor structure of the Higgs boson is furthered by learning the phenomenology of how it couples to the WW pair in these reactions. This can be done by applying effective coupling strength constants to an effective Lagrangian as beyond standard model (BSM) terms and performing Monte Carlo studies with these terms present. The investigation includes a two dimensional analysis of the polar angle and the Higgs boson momentum, such that the correlation between these variables can lead to enhanced sensitivity to new dynamics. We also present an energy scan of cross sections for the processes. A likelihood analysis is performed to show that an electron positron collider operating with an integrated luminosity of 5fb^{-1} would be enough to fully realise an admixture of BSM effects.

1. Introduction

With the July 2012 discovery of what appears to be the Higgs boson [1, 2], we now have a great opportunity to further our knowledge of the particle. We do know that the particle which was discovered resembles the Higgs boson described by the standard model (SM), but a large amount of research is being conducted with the goal of searching for signatures associated with it which can only be described by physics beyond the standard model (BSM).

The Large Hadron Collider (LHC) is the only active provider of data and information pertaining to Higgs production, but we cannot expect it to provide an answer to all of the questions which the scientific community can pose about the Higgs boson. This is why a thorough study of e^+e^- collisions is vital if we wish to find BSM physics intrinsic to Higgs production.

Since e^+e^- collisions are seldom used to study Higgs production in general, it is necessary to first try and evaluate what BSM effects such a collider can produce. In this study, we do this by computationally modelling the channels in e^+e^- which are involved in Higgs production, namely the t -channel and s -channel, incorporated with BSM parameters. Once the hypotheses are created, a likelihood analysis is used to determine the luminosity which an e^+e^- collider should have in order to comfortably detect BSM effects experimentally.

2. Problem Formulation

Electroweak symmetry breaking (EWSB) in the SM is transmitted from the scalar sector to the gauge sector by the use of gauge boson-scalar couplings. The precise couplings of the SM Higgs

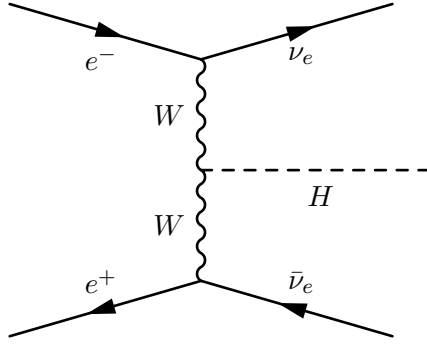


Figure 1. The t -channel process of interest. This is one of three Feynman diagrams (not including diagrams with a Z boson decay in the final state) associated with the t -channel. The vertex in the center of the diagram is where the Higgs to WW coupling is evident. Precise measurements of this vertex are imperative in this study.

to the massive electroweak gauge bosons, W^\pm and Z , come from

$$\mathcal{L}_{int} = -gM_W \left(W_\mu W^\mu + \frac{1}{2\cos^2\theta_W} Z_\mu Z^\mu \right) H. \quad (1)$$

The constants g , M_W and θ_W are all accurately measured, implying that this vertex is fully determined in the SM. In order to fully understand EWSB, independent measurements of these vertices are required.

The virtue of e^+e^- collisions is that they provide a channel of Higgs production which allows for the study of Higgs to WW coupling, namely the t -channel, of which a Feynman diagram associated with the process is shown in figure 1. A good understanding of this coupling will indicate whether or not there is any BSM signature associated with the Higgs boson.

2.1. Background Theory

The $H(k) - W_\mu^+(p) - W_\nu^-(q)$ vertex can be parametrised as [3]:

$$i\Gamma^{\mu\nu}(p, q)\epsilon_\mu(p)\epsilon_\nu^*(q), \quad (2)$$

where deviations from the SM form of $\Gamma_{\mu\nu}^{\text{SM}}(p, q) = -gM_W g_{\mu\nu}$ would indicate the presence of BSM physics. These BSM deviations can be specified in terms of two effective strength constants, λ and λ' , by the equation [3]

$$\Gamma_{\mu\nu}^{\text{BSM}}(p, q) = \frac{g}{M_W} [\lambda(p \cdot q g_{\mu\nu} - p_\nu q_\mu) + \lambda' \epsilon_{\mu\nu\rho\sigma} p^\rho q^\sigma]. \quad (3)$$

The constants λ and λ' are the effective strengths for the anomalous CP-conserving and CP-violating operators respectively. The full vertex is then determined by the sum of these two operators,

$$\Gamma_{\mu\nu} = \Gamma_{\mu\nu}^{\text{SM}} + \Gamma_{\mu\nu}^{\text{BSM}}. \quad (4)$$

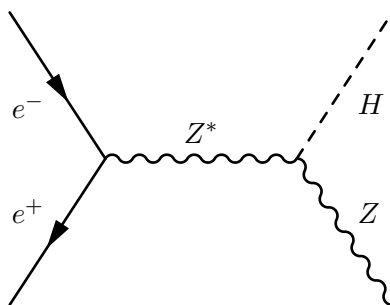


Figure 2. One of the Feynman diagrams associated with the s -channel. This process is also known as *Higgs-strahlung* since it can be likened to the electronic process of bremsstrahlung. Although we are not directly studying this process, its effects must still be accounted for.

A precise identification of the non-vanishing nature of λ and λ' can tell us whether the modification in HWW -couplings are CP-conserving or CP-violating in nature and, if both are present, what their relative proportion is.

2.2. Computational Methodology

MadGraph [4] was used to do the Monte Carlo simulations, where the necessary rules were generated by a Mathematica package called **FeynRules**.

In performing the simulations, the expressions for the BSM parameters λ and λ' are determined from

$$\lambda = -(a + ic) \frac{4M_W^2}{\Lambda_1^2} \quad \text{and} \quad \lambda' = -i(b + id) \frac{8M_W^2}{\Lambda_2^2}, \quad (5)$$

respectively. The parameters a, b, c, d, Λ_1 , and Λ_2 are real values specified in generating the Monte Carlo data. Values of λ and λ' were arbitrarily chosen to be either 1 or -1 in this study, allowing for four different hypotheses to be studied. This required that Λ_1 and Λ_2 be approximately 160 GeV and 226 GeV respectively, while the different hypotheses were built by varying the values of the parameters a, b, c , and d .

The process $e^+e^- \rightarrow H\nu\bar{\nu}$ occurs through two channels. The s -channel features the production of the neutrinos through the decay of a Z boson, given by

$$e^+ + e^- \rightarrow H + Z \rightarrow H + \nu + \bar{\nu}. \quad (6)$$

The t -channel features the production of the Higgs and neutrino pair through vector boson fusion given as

$$e^+ + e^- \rightarrow H + \nu_e + \bar{\nu}_e \setminus Z. \quad (7)$$

Note that the ' $\setminus Z$ ' symbol means that processes contributing to the final state resulting from the decay of the Z boson are excluded.

The t -channel is of primary concern here. It contains information pertaining to the Higgs W^+W^- coupling and has not, in general, been studied as extensively as the s -channel. Even though we focus primarily on the t -channel, the s -channel effects must still be incorporated. To achieve this, the Monte Carlo simulations were run with both the t -channel and s -channel active. The major s -channel contributions were then cut from the distributions.

2.3. Likelihood Formalism

The test statistic used to distinguish hypotheses is incorporated as the logarithm of a profile likelihood ratio.

A profile likelihood ratio is defined as the ratio of two likelihood functions regarding two differing hypotheses. Due to the discrete nature of the probabilities in this analysis, the likelihood functions are created as the product of binned Poisson probabilities over all channels and bins [1]. For Monte Carlo simulation studies, the likelihoods can be parametrised by the probability density function, P_i , of hypothesis i and a set of pseudo-data points, D_j , according to the hypothesis j . The likelihood function is then expressed as $L(P_i|D_j)$.

The profile likelihood ratio, λ_{ij} , is then constructed for two different hypotheses i and j as

$$\lambda_{ij} = \frac{L(P_i|D_i)}{L(P_j|D_i)}, \quad (8)$$

or alternatively as its reciprocal, depending on the analysis required.

As stated earlier, the test statistic used is actually the logarithm of the profile likelihood ratio. This quantity for the test statistic, q_{ij} , is then expressed as

$$q_{ij} = \log \lambda_{ij} = \log \frac{L(P_i|D_i)}{L(P_j|D_i)}. \quad (9)$$

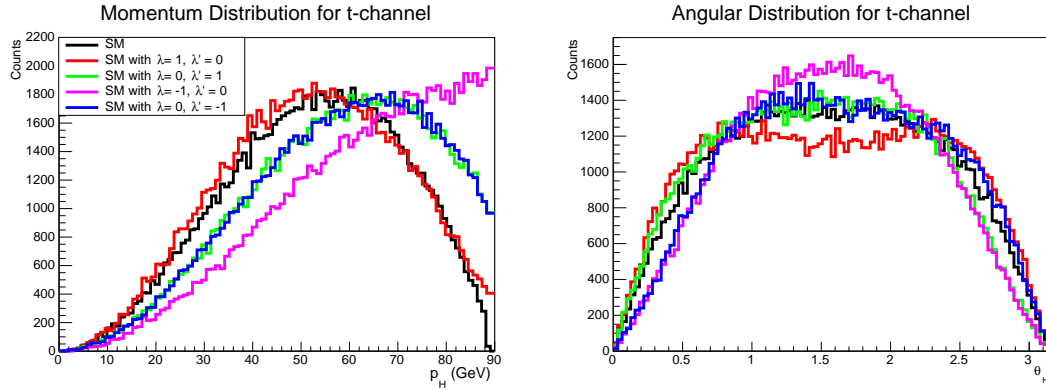


Figure 3. Plots of the Higgs particle properties at $\sqrt{s} = 250\text{GeV}$ for different BSM models and the SM. Left: Higgs momentum (p_H) distribution from the t -channel which features a wide range of momentum values indicating that exclusions near the s -channel peak allow for exploration of the t -channel and variations from the SM. Right: The θ_H distribution from the t -channel. The distribution is indeed affected by all of the BSM parameters.

In Monte Carlo studies, these test statistics emerge as binned peaks. These peaks are built by running pseudo-experiments, each of which compute a value for the test statistic based on a randomly generated set of pseudo-data. The number of pseudo-data points generated is fixed by the cross section of the event being studied.

The test statistics concerned in this analysis are always produced in pairs, in order to discriminate between the SM hypotheses and a BSM hypothesis. This pair of test statistics is represented as

$$q_1 = \log \frac{L(P_{SM}|D_{SM})}{L(P_{BSM}|D_{SM})} \quad \text{and} \quad q_2 = \log \frac{L(P_{SM}|D_{BSM})}{L(P_{BSM}|D_{BSM})}. \quad (10)$$

In this study the test statistic q_1 in equation 10 tends to have a more positive value due to its ordering, and we refer to it as the *upper* test statistic strictly for our purposes; while the opposite is true for q_2 , which we will refer to as the *lower* test statistic.

After test statistics are created, an hypothesis can then be tested for rejection by finding the associated p -value which can be calculated from:

$$p = \int_m^\infty q_2 dq, \quad (11)$$

where the value m is the median of the upper test statistic q_1 .

Knowing the p -value is of utmost importance, since its value is regarded as the percentage of confidence by which one can reject a certain hypothesis. This confidence can further be quantified by knowing the *significance* of the separation between the two test statistics. The median significance, Z_{med} , is defined as the number of standard deviations between the median of the lower test statistic and the left edge of the p -value area, that is, the median of the upper test statistic.

3. Results

The distributions shown in figures 3 and 4 are examples of the results of Monte Carlo simulations performed for determining the emerging Higgs particle momentum, p_H , and scattering angle,

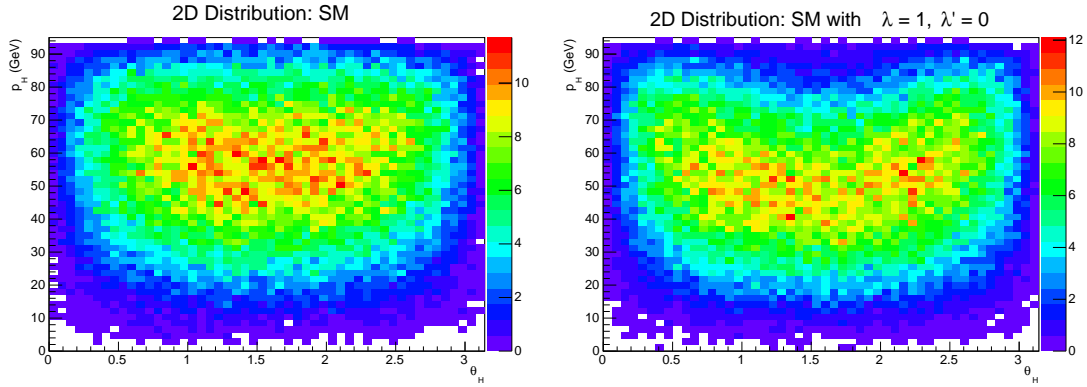


Figure 4. Two dimensional histograms showing the correlation of the t -channel Higgs momentum and θ_H at $\sqrt{s} = 250\text{GeV}$. The z -axis is an indication of the frequency of events and is in arbitrary units.

θ_H , respectively for the t -channel. These were taken from simulations with a specified centre of mass energy (CME) of $\sqrt{s} = 250\text{GeV}$, although various energies were analysed.

As stated earlier, both the s -channel and the t -channel must be considered, but we are only interested in studying the HWW coupling inherent in the t -channel. This was done by excluding a two sigma window on either side of the s -channel momentum peak in the t -channel. Doing this ensures that 95% of the events within the range are cut out, however there will be a 5% migration of some events that were within the window, being detected as having been from without.

The cross sections were scanned as functions of CME, as shown in figure 5, such that we could decouple the kinematics and CME from the total cross section of the processes. Such outputs would have to be modified by the cuts we imposed and other considerations. The cuts on the allowed momentum too will affect the cross section. The fraction of events that survive the cut, multiplied with the cross section signifies the cross section for the process. Finally, the Higgs to $b\bar{b}$ branching ratio of 57.7% must also be considered.

A likelihood analysis for each BSM hypothesis was performed for integrated luminosity values of 1fb^{-1} , 5fb^{-1} , and 10fb^{-1} . The number of pseudo-data points in each analysis was determined from the effective cross section of the SM. The likelihood analysis was performed using a total number of 100,000 pseudo-experiments for each test statistic. The two dimensional distributions, examples of which are shown in figure 4, were also included in the likelihood analysis to demonstrate the effect of the correlation between the two variables, p_H and θ_H . The median significance values were plotted as a function of the CME and are shown in figure 6 for the variation of λ' . A more detailed study and clearer results have been omitted but can be found in reference [5].

4. Conclusions

The results which were obtained from the likelihood analysis provide a good motivation for the role of an electron positron collider in understanding the BSM variations in the HWW coupling which was analysed. The median significance values of the two dimensional analyses were, in general, greater than those done for one dimensional analyses, implying a far greater confidence in rejecting certain hypotheses. An electron positron collider which produces collisions at an integrated luminosity of 5fb^{-1} would be an effective tool in studying e^+e^- collisions with regards to identifying BSM physics.

The correlation of the two dimensional distributions thus carries vital information about the

dynamics of the processes which are studied in e^+e^- collisions. In addition to this, an energy scan of the cross sections is an effective tool in decoupling the kinematics from the total cross sections, as described above. It is also evident from the results tables that an increased integrated luminosity of the particle beams in e^+e^- colliders will, of course, have the effect of providing a better chance of rejecting BSM hypotheses and, therefore, furthering our understanding of the Higgs boson.

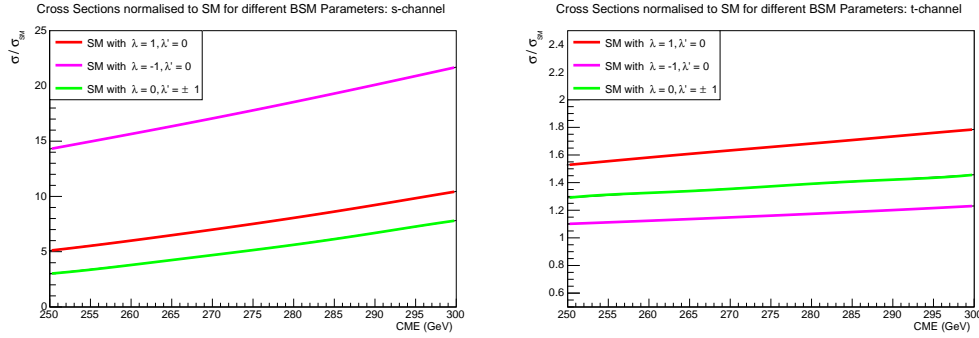


Figure 5. The cross sections for both the s -channel and t -channel produced by MadGraph, normalised to the SM cross section. These cross sections were further altered by the cuts made to the distributions. Note that the λ' parameter produced the result of identical cross sections for both $\lambda' = -1$ and $\lambda' = 1$.

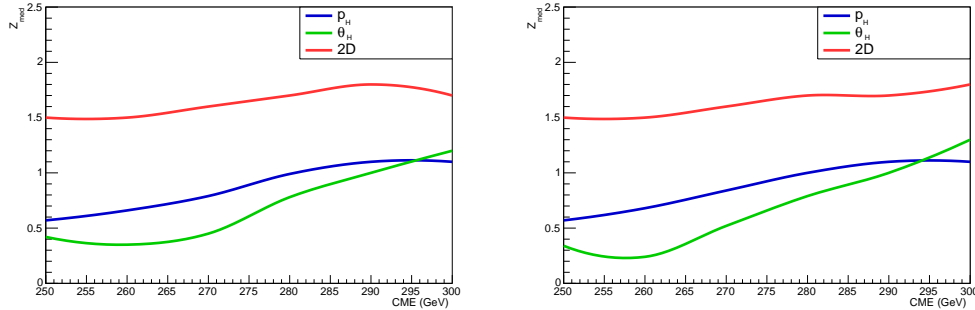


Figure 6. The median significance results from performing likelihood studies at 1fb^{-1} of pseudo data. This not only highlights the power of using two dimensional distributions, but also shows how the correlation of p_H and θ_H is useful in identifying or rejecting BSM hypotheses. These plots pertain to the $\lambda' = 1$ (left) and $\lambda' = -1$ (right) hypotheses.

Acknowledgments

This work is supported by The Claude Leon Foundation as well as the National Research Foundation (NRF).

References

- [1] Aad G *et al* [ATLAS Collaboration] 2012 *Phys. Lett. B* **716** 1 (*Preprint* 1207.7214)
- [2] S Chatrchyan *et al* [CMS Collaboration] 2012 *Phys. Lett. B* **716** 30 (*Preprint* 1207.7235)
- [3] Djouadi A, Godbole R M, Mellado B and Mohan K 2013 *Phys. Lett. B* **723** 307 (*Preprint* 1301.4965)
- [4] Alwall J, Herquet M, Maltoni F, Mattelaer O and Stelzer T 2011 *JHEP* **1106** 128 (*Preprint* 1106.0522)
- [5] Amar G, Banerjee S, von Buddenbrock S, Cornell A S, Mandal T, Mellado B and Mukhopadhyaya B 2014 (*Preprint* 1405.3957)

Memory performance of ARM processors and its relevance to High Energy Physics

T Wrigley, G Harmsen and B Mellado

University of the Witwatersrand, 1 Jan Smuts Avenue, Johannesburg, South Africa 2000

E-mail: thomas.wrigley@cern.ch

Abstract. Projects such as the to-be upgraded ATLAS detector at the Large Hadron Collider at CERN are expected to produce data in volumes which far exceed current system data throughput capacities. In addition, cost considerations for large-scale computing systems remain a source of general concern. A potential solution involves using low-cost, low-power ARM processors in large arrays in a manner which provides massive parallelisation and high rates of data throughput (relative to existing large-scale computing designs). Giving greater priority to both throughput-rate and cost considerations increases the relevance of primary memory performance and design optimisations to overall system performance. Using several primary memory performance benchmarks to evaluate various aspects of RAM and cache performance, we provide characterisations of the performances of three different models of ARM-based SoC, namely the Cortex-A9, Cortex-A7 and Cortex-A15. We then discuss the relevance of these results to high throughput-rate computing and the potential for ARM processors. Finally, applications to the upgrade of the on-line and off-line data processing at the ATLAS detector are also discussed.

1. Introduction and Background

The term Big Data is widespread and its usage appears to be continually increasing [1]. Very large-scale scientific experiments such as the Large Hadron Collider (LHC) at CERN or the under construction Square Kilometre Array (SKA) telescope are increasingly being referred to as belonging to the realm of Big Science. A key challenge of large-scale scientific experiments is the vast amounts of data that they generate requiring far greater data throughput capacity and much larger data storage capacity than existing technologies and infrastructure can feasibly or reasonably accommodate. The ATLAS detector, one of seven within the LHC, makes use of a multi-level triggering system to filter output data, due to *post hoc* processing and long term storage limitations. Even after filtering, the amount of data currently committed to storage is very large and the scheduled upgrade to the LHC will further exacerbate this problem by a significant degree.

A concept known as Data Stream Computing has been proposed as a potential means of addressing these high data throughput challenges [2]. Data Stream Computing (DSC), briefly, has the following three characteristics: very high data throughput rates, severely limited or no offline storage of data, and programming simplicity (i.e. common, easily programmable architectures as opposed to specialised, custom architectures). In addition to the characteristics of DSC, affordability and energy efficiency are two major concerns for large-scale computing in

general. A potential solution involves the use of ARM processors, which are low-power, low-cost and low-energy consumption systems-on-chip (SoC), in large arrays which would provide very high levels of parallelisation. ARM-based SoCs, which are commonly used in mobile devices such as smartphones and tablets, are low-cost, mass-produced and potentially highly energy-efficient [3], all of which bodes well for both affordability and energy efficiency.

Although large scale computing has traditionally placed its primary focus on processor performance, the relevance of memory performance to overall system performance is being increasingly widely acknowledged [4, 5]. Memory performance is a key component of overall system performance and is particularly important for throughput rates, memory bottlenecks could potentially affect energy-efficiency and cost through under-utilisation of existing system hardware. Using ARM-based SoCs in any proposed solution therefore requires that the performance of ARM-based SoCs be properly characterised and understood. Memory performance of existing ARM-based systems and its relevance to DSC for large-scale scientific experiments, particularly ATLAS, is the focus of this paper.

2. Experimental Configuration

The performance of three models of development board containing ARM-based systems-on-chip was evaluated. For practical and financial reasons, commercially available development boards containing ARM-based SoCs were used for the purposes of benchmarking. The technical specifications of these boards are listed in table 1 below.

Table 1. Development board hardware & OS specifications

	Cortex-A7	Cortex-A9	Cortex-A15
Platform	Cubieboard2	Wandboard Quad	Odroid-XU+E
SoC	Allwinner A20	Freescall i.MX6Q	Samsung Exynos 5410
Cores	2	4	4 (+ 4 Cortex-A7)
Max. CPU Clock (MHz)	1008	996	1600
L1 Cache (kB)	32	32	32
L2 Cache (kB)	256	1024	2048
RAM Size (MB)	1024	2048	2048
DDR3 RAM Type	432 MHz 32 bit	528 MHz 64 bit	800 MHz 64 bit
2014 Price (USD)	65	129	169
OS	Ubuntu	Linaro	Ubuntu

A Linux-based distribution was installed on all three types of boards. Three benchmarking software programmes were used to evaluate the memory performance of these three boards, namely the LMBench benchmark suite, the STREAM benchmark and the Parallel Memory Bandwidth Benchmark (*pmbw*).

The LMBench benchmarking suite analyses several aspects of memory performance this study focuses on the measures of memory latency. The STREAM benchmark provides a measure of sustained memory bandwidth. The benchmark works by generating an array of random numbers of a specified size (which is then stored in RAM) and performs four types of operations, namely copy, scale, add and triad. Measures of sustained bandwidth are then produced for each of these four tests. The *pmbw* benchmark is similar to STREAM in that it also provides a measure of sustained memory bandwidth, but does so by means of 14 separate

subtests, each performing a slightly different operation. These variations are: sequential scanning or a random access (permutation walking) test, write or read operation, bit size transferred in each operation, pointer-based iterations vs index-based array access, and number of operations per loop (1 - Simple vs 16 Unroll) [6]. Two of the subtests involve Multiroll Loops and are not analysed here. The benchmark is designed to automatically detect the amount of physical RAM available. It then generates an array and runs one of the subtest routines. The allocated array size is then increased and the subtest routine is then repeated. This is repeated until the highest power of 2 able to fit onto the systems RAM is reached. These steps are repeated for each one of the subtest routines. *pmbw* is useful because it allows both for comparisons to STREAM and will potentially yield deeper insight into memory performance.

3. Results and Discussion

3.1. STREAM and LMBench

For the STREAM benchmark, which measures sustained memory bandwidth, the Cortex-A15 is clearly shown to be the best-performing of the three systems, both in terms of absolute bandwidth and bandwidth efficiency (i.e. percentage of theoretical maximum obtained). The Cortex-A7 displays reasonable bandwidth efficiency, while the Cortex-A9, which is the oldest of the three systems, achieves very low bandwidth efficiency, only reaching 16% of its theoretical maximum. In the case of RAM and cache latencies, the Cortex-A7 and Cortex-A15 both perform well, recording low latencies. The performance of the Cortex-A9 in this regard is also inferior to the A7 and A15 SoCs. For both of these benchmarks, a clear positive correlation can be seen between age of SoC design and performance. Table 2 below summarises the results obtained from both LMBench and STREAM for all three boards.

Table 2. Development board hardware & OS specifications

	Cortex-A7	Cortex-A9	Cortex-A15
Copy (MB/s)	1996	1329	6066
Scale (MB/s)	1444	1110	6114
Add (MB/s)	757	1448	5413
Triad (MB/s)	702	1290	5275
RAM (Theoretical MB/s)	3296	8054	12207
RAM BW Efficiency (%)	37	16	47
L1 Latency (ns)	3.02	4.02	2.51
L2 Latency (ns)	9.2	30.8	13.8
RAM Latency (ns)	58.5	119.8	104.8

3.2. pmbw

The design of the pmbw benchmark means that each subtest routine generates several hundred sets of observations between 200 and 300 observations in the case of the three systems tested here. Because there are several hundred observations per subtest and 12 subtests which are analysed here, the volume of data produced by this benchmark for each system is very large numbering around several thousand observations. For this reason, statistical tools are useful for extracting meaning from these data sets. A statistical test known as analysis of variance (ANOVA) was used for primary analysis of the results of this benchmark. ANOVA is used to compare multiple datasets and determine whether the individual means of these datasets

are equal to one another. More specifically, ANOVA compares the variance within each of these datasets to the variance which is present between these datasets and determines whether statistically significant differences exist between these datasets [7]. If statistically significant differences between these datasets do exist, various *post hoc* tests and analyses can then be used to gain greater insight into the distribution and nature of these differences.

In this case, each subtest (with its 200-300 observations per system) represents a dataset and ANOVA is used to determine whether these individual subtests are statistically similar to one another. A two-way analysis of variance showed that significant differences existed between the subtest groups for all three boards i.e. at least one pair of means was different from one another. Post hoc analysis was then conducted to gain greater insight into the nature and distribution of these results. This analysis revealed the results generated by the 12 subtests appear to be distributed into five general groupings, with each grouping being made up of two, three or four subtests. As each subtest has 5 primary characteristics which vary, the existence of these five groupings gives a greater level of insight into which of these characteristics appear to have the greatest impact on performance insights which allow for memory performance to be better understood. The types of subtests which make up each grouping are briefly detailed in Table 3 below.

Table 3. Constituent subtest types of *pmbw* result groupings

Group No.	Subtest types	Number of subtests in group
1	Random Pointer Permutations (Perm)	2
2	Sequential Reading 32 bit Simple Loop	2
3	Sequential Write 32 bit Simple & Unroll Loop	3
4	Sequential 32 bit Unroll & 64 bit Simple Loop	3
5	Sequential 64 bit Unroll Loop	2

Based on the subtest result groupings determined above, the average of the two/three/four RAM bandwidth results for each of the five groupings was plotted. These bandwidth results are shown in figure 1 below. The first grouping (Random Pointer Permutation) is substantially lower than the other four groupings. This is, however, consistent with expectations, as this benchmark is based on a random pointer permutation and is essentially a measure of raw bandwidth and latency for one memory fetch cycle, while the other four are measures of sustained memory bandwidth for sequential scanning [6]. These results indicate that the Cortex-A7 achieves the lowest performance (approx.. 35 MB/s), the Cortex-A9 produces more than double that rates (approx. 85 MB/s) and the Cortex-A15 is again the best performer (approx.. 127 MB/s). This appears to be inconsistent with the memory latency and sustained memory bandwidth results obtained by LMBench and STREAM, which showed that the newer Cortex-A15 was the best performing of the system, the Cortex-A7 the second best performing and that the Cortex-A9 was the worst performing system by a significant margin. While these two random pointer permutation subtests are not solely dependent on memory latency, this would be expected to have some effect on random memory access performance. It is not immediately clear why the results produced by *pmbw* appear to conflict with the trends implied by the obtained LMBench results, although factors such as the Cortex-A9 SoCs 64 bit RAM bus width compared to the Cortex-A7 SoCs 32 bit RAM bus width may influence this result. This question must be further investigated in future work.

Groupings 2, 3, 4 and 5 are all based on sequential scanning rather than random memory access. This means that these four groupings offer some measure of sustained memory bandwidth. The general profile of all four groups is consistent with the results obtained by the STREAM benchmark, with the Cortex-A15 obtaining the best results by a significant margin, followed by the Cortex-A7 and then finally by the Cortex-A9.

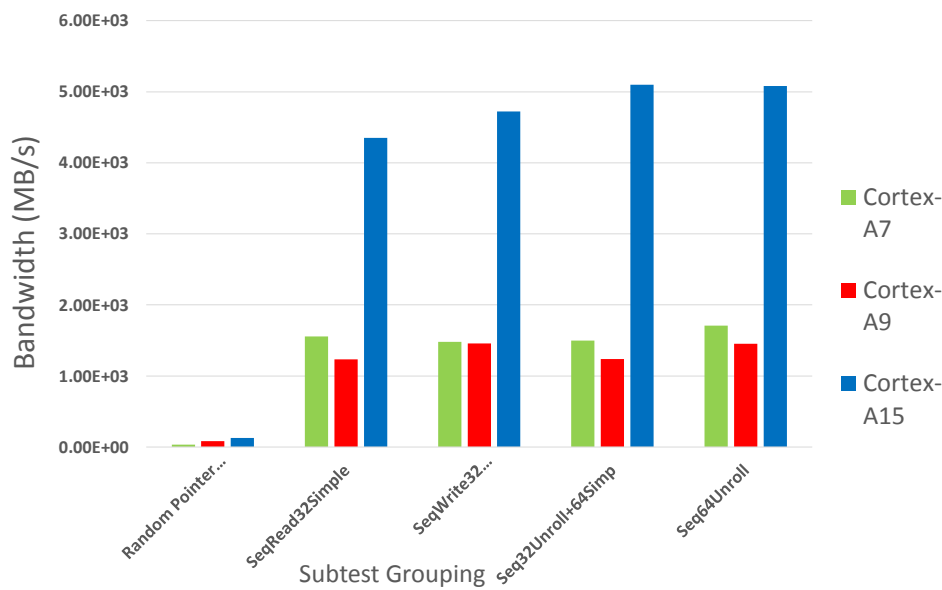


Figure 1. *pmbw* Bandwidth Grouping Results

3.3. Discussion and Analysis

A clear correlation between age of SoC design and bandwidth efficiency (as a percentage) is observed, with the newest SoC, the Cortex-A15 performing the most effectively and the oldest SoC here, the Cortex-A9 performing the least effectively. Preliminary results which are to be presented at SAIP 2014 by Mitchell Cox [8] show that it is possible to obtain I/O connection rates between two Cortex-A9 based SoCs of approximately 300 MB/s. These results suggest that memory performance is not the primary source of throughput rate bottlenecks for relatively simple algorithms (i.e. where CPU performance is not the bottleneck), as this figure is approximately 5 times lower than the sustained memory bandwidth measured for the Cortex-A9. As I/O connection rates continue to improve, this low sustained memory bandwidth may present an obstacle to throughput rates. The Cortex-A9 tested here is, however, the oldest design of the three and the significantly improved memory bandwidth rates and efficiencies measured for the Cortex-A15 in particular means that memory bandwidth remains less likely to be the primary cause of throughput rate bottlenecks than I/O capacity for algorithms which are not highly processor intensive. These improvements are expected to continue as newer ARM-based SoCs are released, particularly with the soon-to-be released ARMv8 architecture 64 bit SoCs. The

potential of ARM-based SoCs for use in Data Stream Computing systems therefore remains strong. In addition to this, it is hoped that new Intel Atom hardware (i.e. development boards) will be procured in due course, allowing for further research and testing to be carried out.

Conclusion

In conclusion, the paradigm or framework underpinning Data Stream Computing and its relevance to high-energy physics has briefly been discussed. The potential role of ARM-based SoCs in providing a solution to excess data production of large scale scientific experiments and the relevance of memory performance to this has also been discussed. The memory performance of three ARM-based SoCs has been evaluated and the implications of these results have been discussed. Finally, potential steps for the continuation of this research and development have been briefly outlined.

References

- [1] Manyika J, Chui M, Brown B, Bughin B, Dobbs R, Roxburgh C and Hung Byers A 2011. *Big data: The next frontier for innovation, competition and productivity* (New York City, NY: McKinsey Global Institute)
- [2] Cox M, Reed R, Wrigley T, Harmsen G, and Mellado B 2014 *Performance Characterisation of ARM Cortex-A7, A9 and A15 System on Chips for Data Stream Computing* (submitted to Journal of Computational Science).
- [3] Aroca RV and Goncalves LMG 2012 Towards green data centers: A comparison of x86 and ARM architectures power efficiency, *Journal of Parallel and Distributed Computing* **72** 1770-80.
- [4] Dongarra, J., and Heroux, M. A. 2013 Toward a new metric for ranking high performance computing systems No. SAND2013-4744 312 (Albuquerque, NM: Sandia National Laboratories).
- [5] Ang, J A, Barrett B W, Wheeler, K B and Murphy R C 2010 *Introducing the graph 500* No. SAND2010-3263C (Albuquerque, NM: Sandia National Laboratories).
- [6] Bingmann, T. 2013 *pmbw Parallel Memory Bandwidth Benchmark/Measurement*. Retrieved from: <http://panthema.net/2013/pmbw/>
- [7] Larson, M G 2008 Analysis of variance *Circulation* **117.1**, 115-121
- [8] Cox, M. 2014 The development of a general purpose Processing Unit for the upgraded electronics of the ATLAS detector Tile Calorimeter, *SAIP 2014* (submitted).

Octupole Correlation and Collective Coupling in the Rare Earth Nucleus ^{154}Dy

G. L. Zimba¹, S. P. Bvumbi¹, L. P. Masiteng¹, P. Jones², S. N. T. Majola^{2,3}, T. S. Dinoko², J. F. Sharpey-Schafer^{2,4}, E. Lawrie², J. J. Lawrie², D. Roux⁵, P. Papka⁶, J. E. Easton⁴, D. Negi², O. Shirinda², S. Khumalo⁴.

¹University of Johannesburg, Physics Department, P.O Box 524, Auckland Park 2006, South Africa.

²iThemba Laboratory for Accelerator Based Sciences, P.O Box 722, Somerset-West, ZA-7129 South Africa.

³University of Cape Town, Physics Department Private Bag X3, Rondebosch 7701, South Africa.

⁴University of Western Cape, Physics Department, P/B X17, Bellville, ZA7535 South Africa.

⁵Rhodes University, Physics Department, P.O Box 94, Grahamstown 6140, South Africa.

⁶Stellenbosch University, Physics Department, P/B X1, Matieland, 7602, South Africa.

E-mail: zimbabwe@yahoo.com

Abstract. The low spin states of ^{154}Dy were studied with the AFRODITE spectrometer array equipped with 9 clover High Purity Germanium (*HPGe*) detectors at iThemba LABS Western Cape. The reaction $^{155}\text{Gd} (^3\text{He}, 4n)^{154}\text{Dy}$ at 37.5 MeV was used to populate these states. A first observation of enhanced E1 transitions in the in the transitional isotones ^{150}Sm and ^{152}Gd from the levels in the first excited 0^+ band to the lowest negative parity band was reported [3]. Here, we report on experimentally observed fingerprint of octupole deformation [19], [20]. And we also give a brief description of our results of the low spin structure of ^{154}Dy .

1. Introduction

The dysprosium (^{154}Dy) nucleus offers an opportunity to study nuclear structures in a transitional region [1], thus the ^{154}Dy nucleus has a rapidly changing shape from spherical to quadrupole deformed. In even-even nuclei, strong deformation is observed at $N = 60$ and lighter isotopes with $N < 59$ are spherical [2]. The $N = 88$ isotones in low-lying spins display a variety of properties, rapidly changing shape, strong E3, and quadrupole and octupole deformations. The strong E1 properties, from the bands built on $0_2^+ \rightarrow 3_1^-$ are mostly caused by octupole vibrations which have been described and explained as due to the proximity of $\Delta j^\pi = 3^-$ shell model orbits to the Fermi surface [19]. This behavior was observed in ^{152}Gd and ^{150}Sm by [3]. The ^{154}Dy nucleus has 6 neutrons and 2 protons outside the closed shell which is nearly spherical and has very common quadrupole phonon states. ^{154}Dy has a quadrupole deformation in the ground state which consists of levels connected by stretched E2 transitions above the 0^+ ground state.

In a microscopic representation, the origin of octupole collectivity can be explained by the interplay of the unique parity orbit in each major shell and a common parity orbit that differ by

angular momentum and total spin $l = j = 3$ [15], [16] and can be described through condensation of single-particle energy level sequence for a harmonic oscillator potential [12]. In certain cases, an orbit is lowered into next lowest major shell by the l^2 and $l.s$ terms. These intruder orbits lie close in energy to orbits with $l = l_{int} = 3$ and $j = j_{int} = 3$. We will refer to the lowest-lying negative parity band as the octupole band. Observations by [9], [10] have shown that the interpretation of the first excited 0_2^+ as β - vibrations are unfounded, He suggests are shown to result from the $[505] \frac{11}{2}$ Nilsson orbit closeness to the Fermi surface hence forming two particle two hole (2p-2h) states. Therefore we will not refer to the bands built on the 0_2^+ as β - vibrations.

2. Experimental Details and Results

The ^{154}Dy nucleus was studied at iThemba LABS Cape Town using the modern state of the art AFRODITE spectrometer array [18] equipped with 9 clover detectors. The low-lying excited states of ^{154}Dy were populated via the $^{155}\text{Gd} (^3\text{He}, 4n)$ reaction at 37.5 MeV, the thickness of the ^{155}Gd target was 3.2 mg/cm^2 . The trigger condition was such that two Compton-suppressed HPGe detectors fired in prompt time coincidence.

The data were sorted into two dimensional matrix using MTsort and analyzed using Radware software package [4]. Calibrations were performed using a standard ^{152}Eu source. The IUrbanWevel scheme deduced from gamma-gamma($\gamma - \gamma$) coincidence relationship is shown in Figure 1.

3. Discussion

3.1. New levels and γ rays observed

A new band-head at 2345.8 keV has been found with spin-parity 6^- . Two new γ rays 340 keV and 419 keV have been observed in the low-lying negative parity band. These two γ rays connect levels $3^- \rightarrow 5^-$ and $7^- \rightarrow 5^-$ respectively. These levels were reported in National Nuclear Data Center [5] database but the γ rays connecting them were not observed. The odd and even γ positive-parity bands built on 3^+ and 2^+ were only reported up to 7^+ and 4^+ respectively [5]. The two bands have since been observed up to 13^+ and 12^+ respectively. The level at 3290.6 keV, 12^+ was the last level on the band built on 0_2^+ is now the new band head for the band with former band head at 3681.1 keV, 14^+ .

There were no observed E3 transitions from the ground state band to the lowest lying negative parity band. It is suggested that no transitions were observed between the two bands because the lowest lying negative parity band in ^{154}Dy has a high excitation energy and the lowest lying negative parity band becomes yrast at a low spin. Figure 2 shows the the E1 transitions in three $N = 88$ isotones and the relative positions of the band head of the octupole band is shown to increase with an increase in proton number. Strong E1 were observed for ^{150}Sm and ^{152}Gd by Bvumbi et al [3], from the first excited 0^+ band to the low lying negative-parity band. In ^{154}Dy , we observed E1 transitions from the lowest lying negative parity band to the first excited 0^+ .

3.2. Alignment properties of bands

The Figure 3(a) and (b) show the excitation energies (minus 7.7 keV) against rotational frequency. In Figure 3(a), it can be seen that the band built on 12^+ cross the ground state band (0_1^+) at $15 \hbar\omega$. This behavior has been previously observed in [11] and suggests that this crossing is caused by an aligned $i_{13} \frac{13}{2} S$ band which crosses the 0_1^+ band. The new extended bands, the bands built on 2^+ and 3^+ are shown to have a split up to spin $10 \hbar\omega$ when the two bands come together, this is shown in the negative parity bands where the bands built on 7^- and 6^- have the same behavior but no splitting is observed, see Figure 3(a). The band built on 6^- is seen to cross both the band built on 7^- and 6^- at $15 \hbar\omega$. The band built on 0_1^+ and 3^- are seen to track the 0_1^+ band.

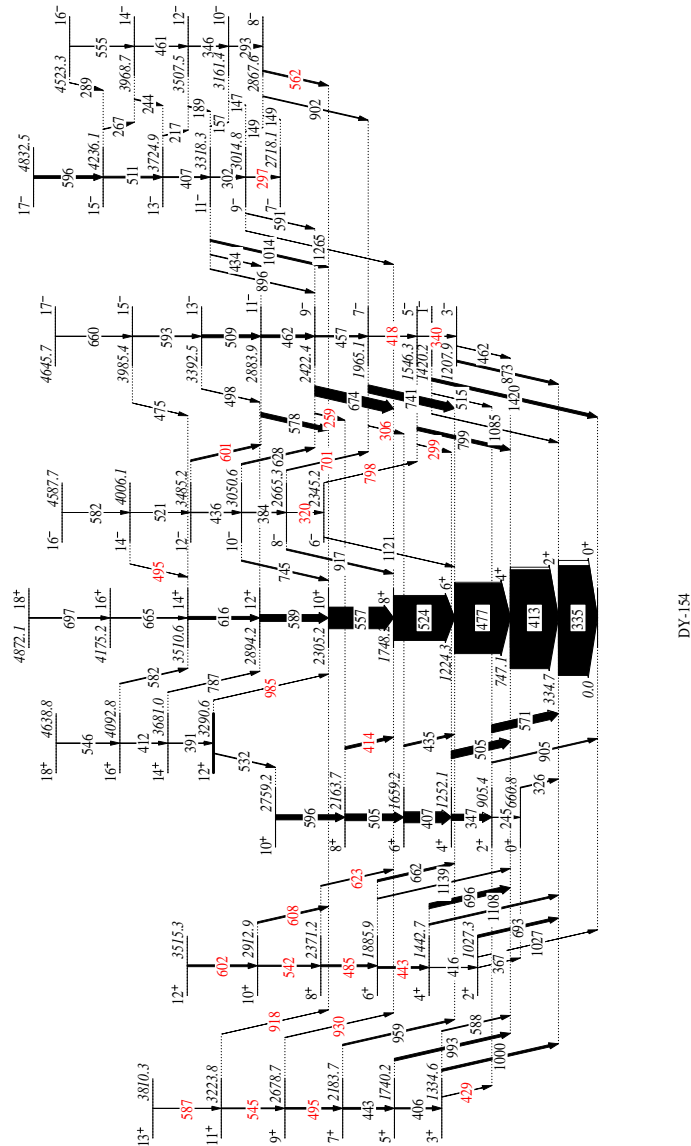


Figure 1. The new level scheme for ^{154}Dy obtained from our $^{155}\text{Gd} (^3\text{He}, 4n)^{154}\text{Dy}$ reaction at 37.5 MeV. The new γ rays are shown in red.

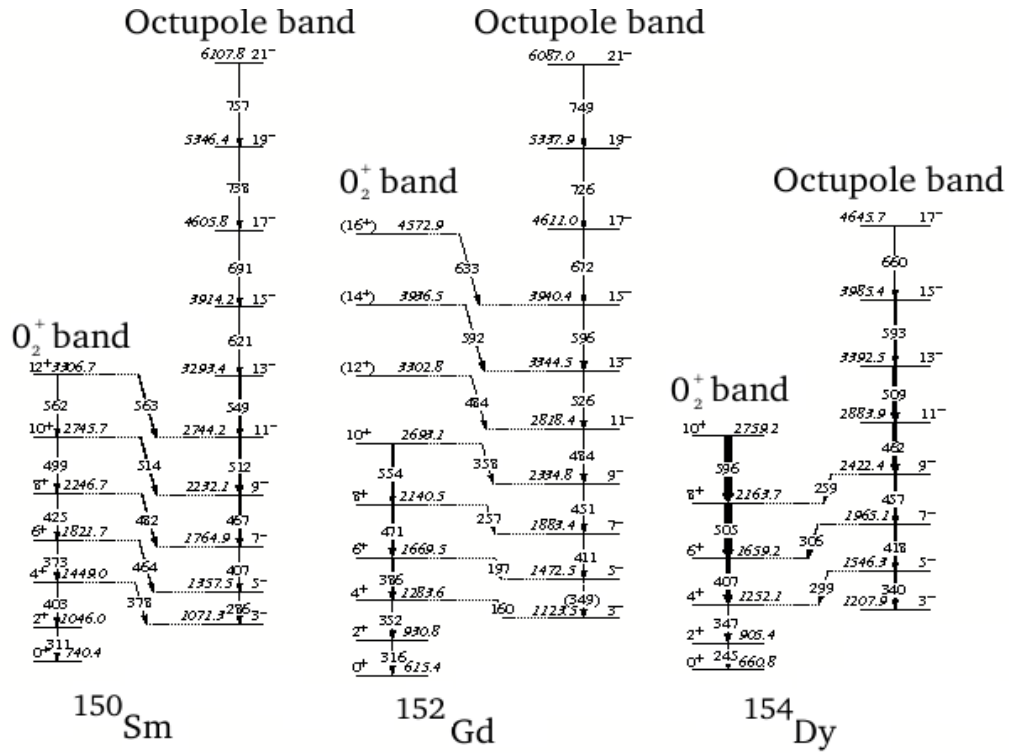


Figure 2. Shows the $0_2^+ \rightarrow 3_1^-$ for ^{152}Gd and ^{150}Sm [3] and the first observation of $3_1^- \rightarrow 0_2^+$ observed in ^{154}Dy .

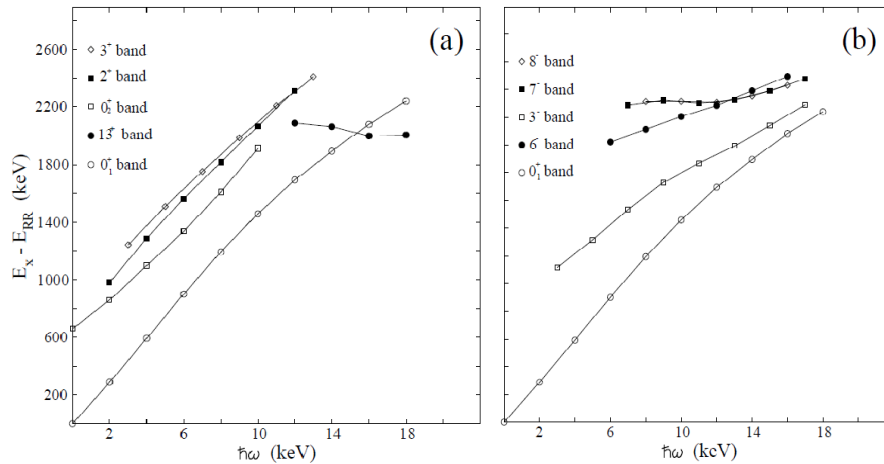


Figure 3. Shows excitation energies minus 7.7 keV against rotational frequency($\hbar\omega$) of the positive-parity bands (a) and the negative-parity bands (b).

The alignment(i_x) of the two new γ bands built on 3^+ and 2^+ positive states and the 0_2^+ positive state bands were plotted as a function of rotational frequency($\hbar\omega$), the ground state

band was also plotted to compare their behavior, see Figure 4(a) and the negative parity bands are shown in Figure 4(b).

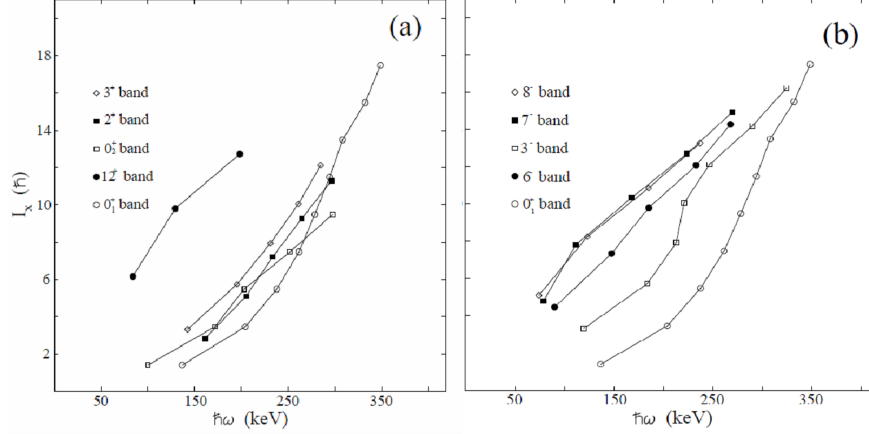


Figure 4. The alignment(i_x) against rotational frequency ($\hbar\omega$) of the positive-parity bands (a) and the negative-parity bands (b).

The change in slopes in Figure 4(a) and 4(b) represents specific positions of different energies at the same rotational frequencies. Rees et al. states that the changes in slope represent a rotational alignment of specific quasi-particle pair [7]. The newly extended odd γ band carries more aligned angular momenta than the even γ band. The 0_2^+ band slope is reduced at rotational frequency of 200 keV which makes it cross the even γ band and ground state bands, 0_1^+ . The ground state band has a gradual aligned angular momenta increase with an increase in rotational frequency up to 310 keV when the slope sharply increases.

4. Conclusion

The experiment showed that $3_1^- \rightarrow 0_2^+$ is present in low-lying states of ^{154}Dy and therefore there is a need to explain why this behavior is opposite to that observed in other rare earth isotones ^{150}Sm and ^{152}Gd as shown in [3]. Spear and Catford [17] showed that there is a maximum in strength in E3 transitions between the ground state band and the lowest lying negative parity band at neutron number $N = 88$ and proton number $Z = 62$ but these E3 transitions were not observed in $N = 88$ Dy. Bands built on the 3^+ , 2^+ and 0_2^+ states have been established and other new transitions. We therefore conclude that calculations along those of [6] and [14] are needed in order to explain the interaction between the 0^+ and 3^- in the $N = 88$ transitional region.

Acknowledgments

We would like to thank the AFRODITE group and students for their participation in the experiment and for their help during the analysis of the data.

References

- [1] R. A. Casten and N. V. Zamfir, Phys. Rev. Lett. **47** 1433 (1981).
- [2] M. Nikaly, A Model Approach to Quadrupole-Octupole Deformation in Atomic Nuclei (2013).
- [3] S. P. Bvumbi et al., Phys. Rev. C **87** 044333 (2013).
- [4] D. C. Radford, Nucl. Meth. A **361** 297 (1995)
- [5] <http://www.nndc.bnl.gov/chart/getdataset.jsp?nucleus=154DY&unc=nds>
- [6] R. R. Chasman, Phys. Rev. Lett. **42** 630 (1979).

- [7] J. M. Rees et al., Phys. Rev. C **83** 044314 (2011).
- [8] A. Khalaf et.al., Spin Assignment and behavior of Superdeformed bands in A 150 mass region.
- [9] J. F. Sharpey-Schafer et al., Eur. Phys. J. A **47** 5 (2011).
- [10] J. F. Sharpey-Schafer et al., Eur. Phys. J. A **47** 6 (2011).
- [11] W. C. Ma et al., Phys. Rev. Lett. **61** 1 (1988).
- [12] P. T. Greenlees *Identification of Excited States and Evidence for Octupole Deformation in ^{226}U* . University of Liverpool PhD thesis (1999).
- [13] S. H. Sharipov, Ukr. J. Phys. **53** (2008).
- [14] S. Frauendorf, Phys. Rev. C **77**, 021304(R) (2008).
- [15] D. R. Inglis, Phys. Rev. **96** 1059 (1954).
- [16] R. Bengtsson and S. Frauendorf, Nucl. Phys. A **327** 139 (1979).
- [17] R. H. Spear and W. N. Catford, Phys. Rev. **C41** R1351 (1990).
- [18] J. F. Sharpey-Schafer, Nucl. Phys. News. Int. **14**, 5 (2004).
- [19] W. Urban, W. Gast, G. Hebbinghaus, A. Krämer-Flecken, K. P.a Blume, and H. Hübel, Phys. Lett. B **185**, 331 (1987).
- [20] W. Urban, J. C. Bacelar, and J. Nyberg, Acta Phys. Pol. B **32**, 2527 (2001).

Division C – Photonics

Propagation characteristics control by variation of PCF structural parameters

James F Jena¹, Peter Baricholo¹, Temba S Dlodlo¹ and Paul K Buah-Bassuah²

¹Department of Applied Physics, National University of Science and Technology, P.O.Box AC939, Ascot, Bulawayo, Zimbabwe

²Department of Physics, University of Cape Coast, Cape Coast, Ghana

james.jena@nust.ac.zw, effjay.jena@gmail.com

Abstract. A photonic crystal slab was designed in COMSOL Multiphysics using gallium arsenide (GaAs) pillars placed equidistant from each other in air. A defect was created by removing some GaAs pillars across the crystal slab geometry to form a 90° bend through the structure. Structural parameters; the pillar diameter and inter-pillar spacing were separately varied and waves were propagated through the created defect with different allowed wavelengths, within and determined by the photonic crystal's bandgap. It was observed for the air filling fraction at constant pitch that, when the factor given by the ratio of the air hole diameter to the pitch is less than 0.63, an increase in air hole diameter requires corresponding increase in the wavelength which can be propagated within the waveguide with minimum loss. When both air hole diameter and pitch were varied the increase in the air filling fraction was also observed to result in a photonic bandgap of lower frequency range as only larger wavelengths would be allowed to propagate. When the air filling fraction was less than 0.63 the photonic crystal waveguides exhibited slightly increased confinement and bend loss. The diameter and pitch affected the core resonance resulting in selected wavelength bands being propagated through the created defect in the waveguide. Only those bands whose value coincides with the photonic bandgap were allowed to propagate.

1. Introduction

Photonic crystals are materials of periodic dielectric media which are able to localise light in specific areas and prevent it from propagating in certain directions. Their structural build up resembles the lattice structures of crystalline materials and they have photonic bandgaps whose optical properties are analogous to the energy bandgaps [1] in electrical semiconductors. Light propagation in a hollow core is possible within photonic crystal fibers (PCFs). These are single material optical waveguides with an array of periodic air holes across the transverse section running down their entire length. In the cladding, the refractive index is modulated in space on a wavelength scale [2]. Very high powers can be propagated without being limited by material threshold [3]. PCFs have application in supercontinuum generation, optical coherence tomography, frequency metrology, microscopy and spectroscopy [4]. Many modelling techniques have been applied to study propagation characteristics [5] of solid core, index guiding PCF's. For hollow core PCF's the control of propagation characteristics which include dispersion, bend loss, confinement loss and effective index are determined by the structural parameters, namely air hole diameter and the pitch. Through simulation, COMSOL Multiphysics [6], a finite element method based software, was used to design and study various photonic crystal structures distinguished by variation in structural parameters and their associated wave propagation characteristics. Confinement loss, bend loss and the allowed photonic bandgap wavelength were monitored for the various air filling fractions.

2. Theory and Simulation Design

In a photonic crystal with periodic dielectric medium of periodicity (Λ), any wave of wavelength comparable to twice the periodicity [7] is not allowed to propagate through the photonic crystal. The wave is within the photonic bandgap, hence it is highly reflected. The partial reflections at each period

are in phase. As they reinforce one another, they form standing waves [8] with the incident wave and as a result such a wave will not be propagated. The other waves with a frequency outside the bandgap experience no scattering as the periodic variation in the refractive index results in the scattering cancelling out coherently [7], so they propagate with minimal attenuation. When a hollow line defect is created in such media the reflecting wave can be allowed to propagate within the defect and the crystal material becomes a waveguide. Varying the periodicity by altering the inter-hole spacing or the air hole diameter affects the wavelength that can be guided as well as the propagation characteristics. Photonic crystal slabs were designed in COMSOL Multiphysics [6] using GaAs pillars placed equidistant from each other in air. A line defect was created by removing some GaAs pillars to form a 90° bend across the slab geometry as shown in figure 1. The structure is a square lattice [9] which has a simpler arrangement than the triangular and hexagonal lattice structures. The bend was chosen to determine with certainty that the defect has guidance to the propagating wave while this motivated study of the effects of structural parameter variation on bend loss.

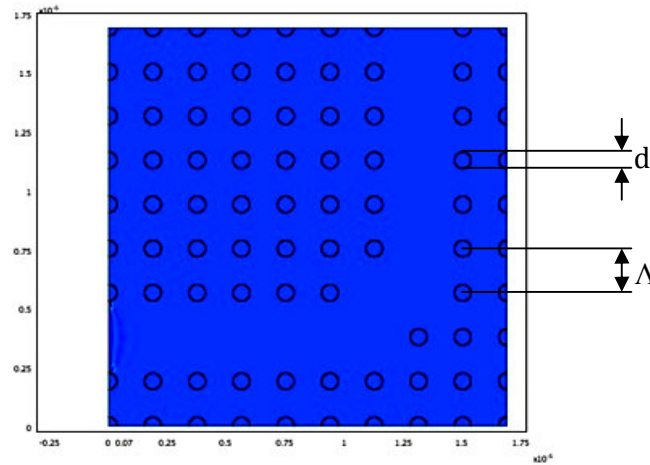


Figure 1. A 2D photonic crystal slab with a 90° bend across its geometry.
(pillar diameter, $d = 0.7 \times 10^{-7}$ m and pitch, $\Lambda = 1.875 \times 10^{-7}$ m).

As determined by the photonic bandgap, different allowed wavelengths within the photonic crystal's bandgap are found and the waves propagated along the 90° bend in different slabs of varying pillar diameter and pitch. The air filling fraction factor (f) given by the ratio in (1) is calculated for each design [10] using this relation;

$$f = (\Lambda - d)/\Lambda \quad (1)$$

where Λ is the pitch given by the inter-pillar spacing and d is the GaAs pillar diameter. $(\Lambda - d)$ represents the air gap separating any adjacent pillars. Low-reflecting boundaries were used on all external boundaries by setting them to the scattering boundary condition and the source at the input set as a plane wave whose electric field in the z component has unit initial amplitude ($E_z = 1 \text{ Vm}^{-1}$).

3. Simulation

The pitch was maintained constant at 3.75×10^{-7} m and the pillar diameter varied within the range 0.7×10^{-7} m to 2.8×10^{-7} m. Results were recorded for the guided wavelength (λ) and the corresponding air filling fraction. The pitch was then varied from 1.875×10^{-7} m to 7.5×10^{-7} m while the diameter remained constant at 1.4×10^{-7} m. The values obtained were used to make plots of normalised wavelength against air filling fraction. Bend loss was determined from measurement of the amplitude of the electric field after the 90° bend through the use of a scaled profile from the cross

sectional plot function. Confinement loss was determined qualitatively from examination of the 3D propagation plots in which modes were seen to leak away from the defect through the boundaries into the photonic crystal. Simulation designs were altered through variation of photonic crystal structural parameters. The same square lattice structure was maintained through all designs as shown in figures 2a and 2b.

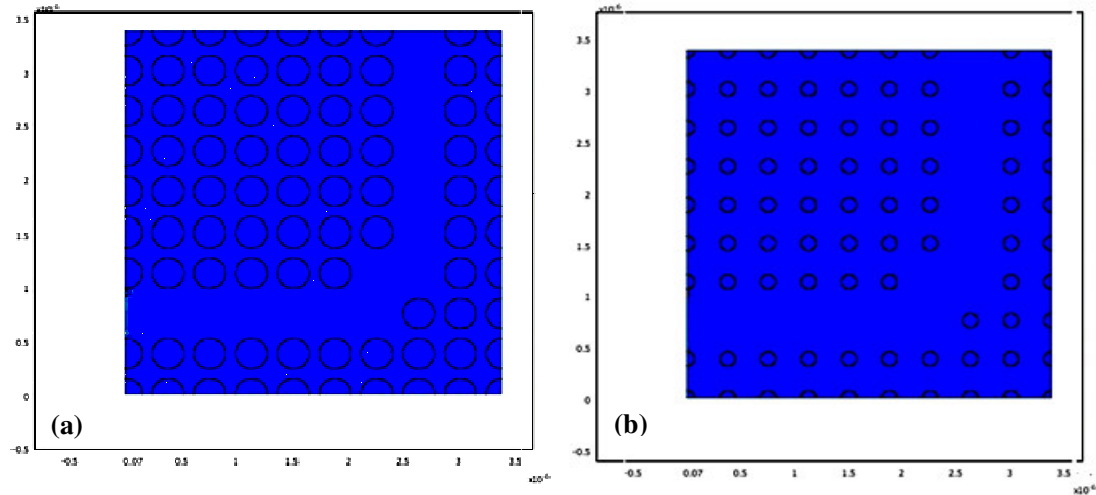


Figure 2. 2D Photonic crystal slabs with same value of pitch $\Lambda = 3.75 \times 10^{-7}$ m. (a) $d = 2.8 \times 10^{-7}$ m, $f = 0.63$ (b) $d = 1.4 \times 10^{-7}$ m, $f = 0.25$

4. Simulation Results and Analysis

Three dimensional views of the wave propagation through the photonic crystal slabs are presented in figures 3, 4 and 5. The guided wavelengths were recorded. A difference in propagation is notably distinguishable in variable designs while identical designs show similar propagation characteristics, only at different wavelengths (figure 5d). The diameter was varied and the air filling fraction factor (f) calculated while pitch was maintained constant at a chosen standard value of $\Lambda_s = 3.75 \times 10^{-7}$ m.

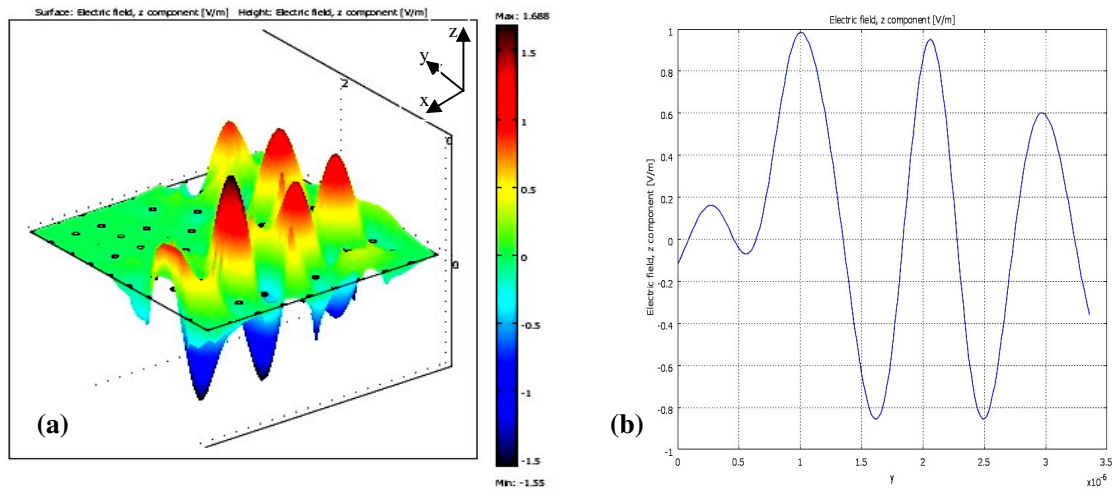


Figure 3. Simulation results for constant pitch $\Lambda_s = 3.75 \times 10^{-7}$ m: (a) 3D propagation for $d = 0.7 \times 10^{-7}$ m, $f = 0.81$, $\lambda = 7.8 \times 10^{-7}$ m. (b) z-y Cross sectional plot of Electric field (E_z) in figure 3a, after the 90° bend, $E_z = 0.98 \text{ Vm}^{-1}$.

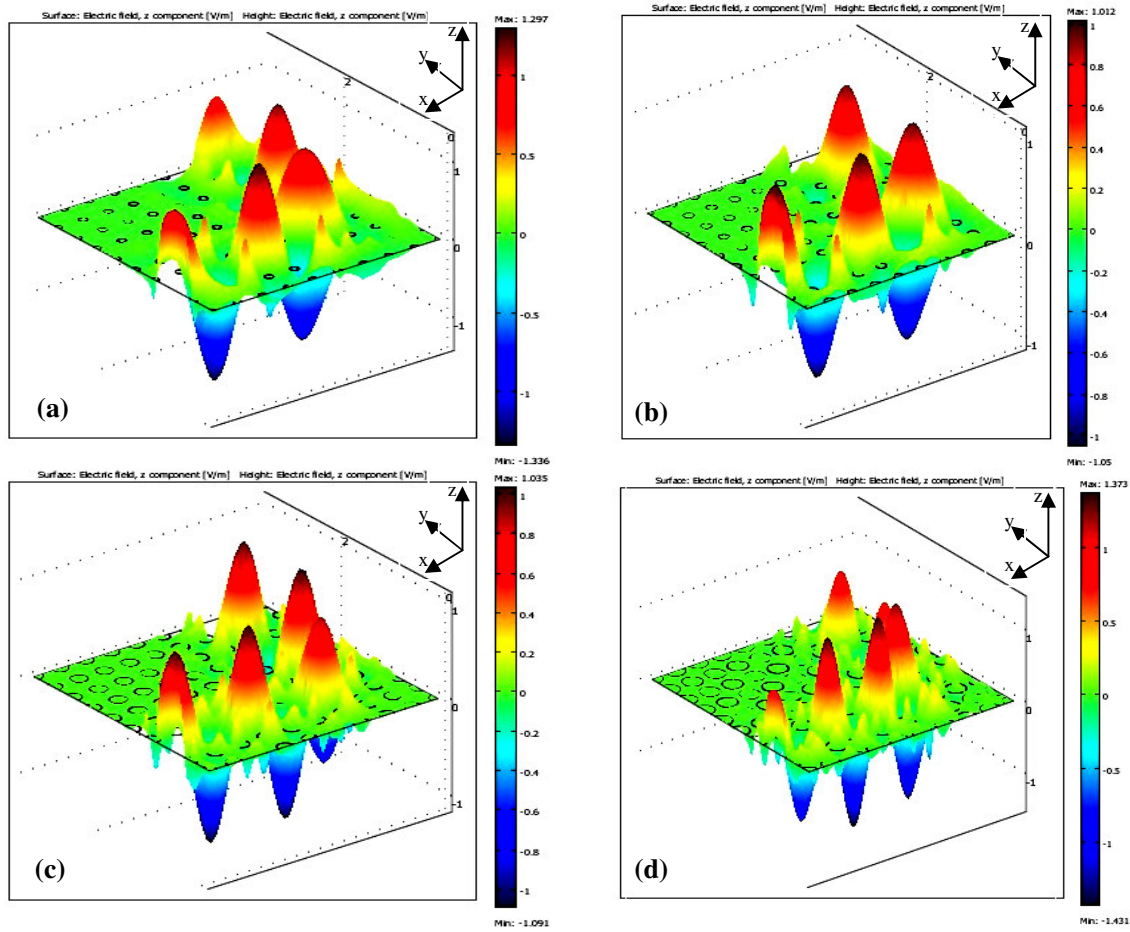


Figure 4. Simulation results for 3D propagation at constant pitch Λ_s : (a) $d = 1.05 \times 10^{-7} \text{m}$, $f = 0.72$, $\lambda = 9.3 \times 10^{-7} \text{m}$, $E_z = 1.00 \text{Vm}^{-1}$. (b) $d = 1.4 \times 10^{-7} \text{m}$, $f = 0.63$, $\lambda = 1.0 \times 10^{-6} \text{m}$, $E_z = 0.60 \text{Vm}^{-1}$. (c) $d = 2.1 \times 10^{-7} \text{m}$, $f = 0.44$, $\lambda = 7.98 \times 10^{-7} \text{m}$, $E_z = 0.90 \text{Vm}^{-1}$. (d) $d = 2.8 \times 10^{-7} \text{m}$, $f = 0.25$, $\lambda = 6.8 \times 10^{-7} \text{m}$, $E_z = 0.55 \text{Vm}^{-1}$.

In figure 3b, the cross sectional plot is used to determine the electric field amplitude ($E_z = 0.98 \text{Vm}^{-1}$) of figure 3a, after the bend. The photonic crystal introduces evanescence to waves whose frequencies are below the bandgap cut-off (figure 5c) while for higher frequencies outside the bandgap it is transparent and incoherently scattering. When compared to figures 4a and 4b, figures 4c, 4d, 5a and 5b show that the smaller wavelengths seem to propagate with increased confinement loss, as they tend to leak away into the cladding. Confinement loss is significant for small air filling fractions [11], where $f < 0.625$ while bend loss is slightly higher for air filling fractions between $0.25 < f < 0.5$ and $0.63 < f < 0.72$ as deduced from the electric field amplitude values of figures 3, 4 and 5. Figure 6b shows that varying the pitch at constant pillar diameter has the effect of decreasing the bandgap frequency with increasing air filling fraction. Larger wavelengths are accommodated as the air holes separating the pillars are increased and so is the air filling fraction (figures 4a and 4b). In figure 5d, when the air filling fraction is the same for similar structural geometry but having proportionately different structural parameter dimensions then the permitted wavelength varies directly as the air hole spacing as evidence in figures 4d and 5a.

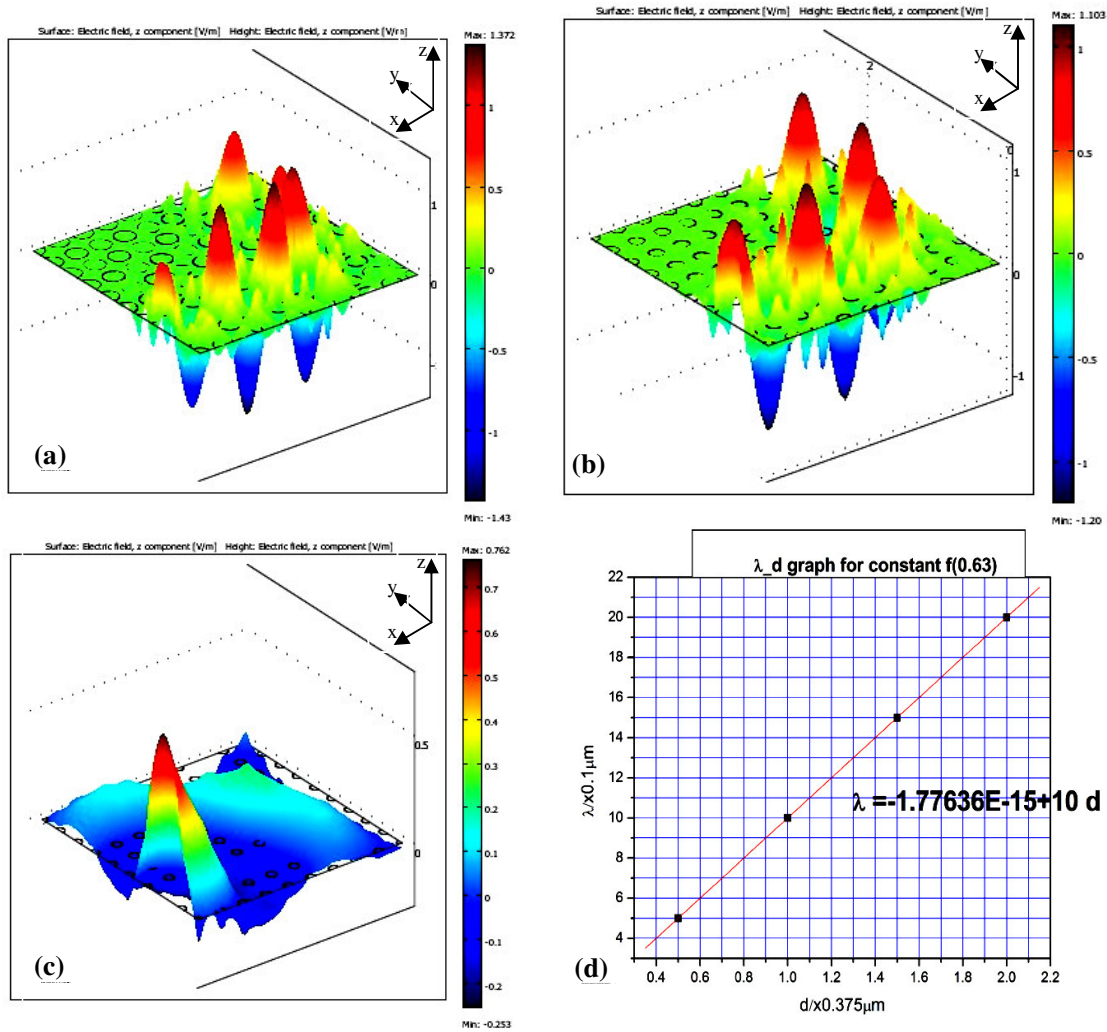


Figure 5. Simulation results for propagation of different wavelengths through the photonic crystal: (a) 3D propagation with pillar diameter constant at $d = 1.4 \times 10^{-7} \text{m}$, $f = 0.25$, $\Lambda = 1.875 \times 10^{-7} \text{m}$, $\lambda = 3.39 \times 10^{-7} \text{m}$, $E_z = 0.60 \text{Vm}^{-1}$. (b) $d = 1.4 \times 10^{-7} \text{m}$, $f = 0.5$, $\Lambda = 2.81 \times 10^{-7} \text{m}$, $\lambda = 5.82 \times 10^{-7} \text{m}$, $E_z = 1.00 \text{Vm}^{-1}$. (c) Evanesence when optical frequency is not within the bandgap $d = 0.7 \times 10^{-7} \text{m}$, $f = 0.63$, $\Lambda = 1.875 \times 10^{-7} \text{m}$, $\lambda = 1.0 \times 10^{-6} \text{m}$. (d) Variation of wavelength with inter-pillar spacing at constant air filling fraction $f = 0.63$.

The simulation results in figures 3 to 5b were used to obtain values of wavelength and air filling fraction to plot the graphs in figures 5d and 6. Increase in air filling fraction at constant pitch requires corresponding increase in the allowed wavelength until the air filling fraction reaches 0.625. After that further increase in air filling fraction results in a corresponding increase in allowed frequency as opposed to the wavelength [9]. The graph in figure 6a resembles a dispersion curve modelled [4] using the fully vectorial effective index method and the scalar effective index method. A change in the air filling fraction consequently alters the effective index of the crystal [11, 12]. At $f = 0.625$ the wavelength change with refractive index is zero, hence $\lambda/\Lambda = 2.68$ is normalised zero dispersion wavelength.

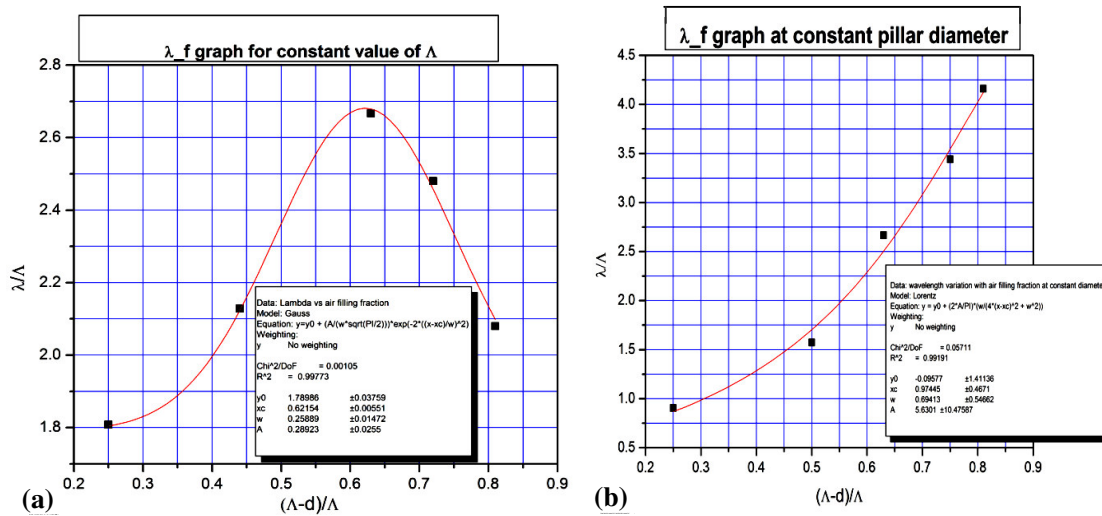


Figure 6. Variation of wavelength with inter-pillar separation: (a) wavelength against air filling fraction determined for constant value of pitch. (b) wavelength against air filling fraction for fixed pillar diameter.

The change in the inverse gradient of figure 6a represents the second derivative of the effective index with respect to wavelength [13, 14]. Similar to the dispersion parameter D , its sign changes from positive to negative as the slope changes from a decreasing to an increasing gradient. This corresponds to the transition from the normal to the anomalous dispersion regime [12].

5. Conclusions

Generally, wavelengths allowed to propagate within the hollow core PCF increase with an increase in the air filling fraction therefore the photonic bandgap is affected by variation of structural parameters. Only frequencies within the bandgap are propagated along the hollow created in the photonic crystal. Confinement loss and bend loss are greater for small air filling fractions. Effective index and dispersion are controlled by variation of air hole diameter and pitch as they are a function of refractive index and wavelength.

Acknowledgements

We thank the National University of Science and Technology for the availing of resources and time with which this research work was carried out and the NUST research board for the funding. We also extend our gratitude to SUPCOM of Tunisia and Prof. Zaghal for availing the software.

References

- [1] Stewart D 2005 Photonic Crystal Bandstructures. Cornell Nanoscale Science and Technology Facility
- [2] Yablonivitch E 1987 Photonic Band-gap Structures. Phys. Rev. Lett. 58,2059
- [3] Khromova I A Melnikov L A 2007 Dispersion Characteristics of hollow – core Photonic Crystal Fibers. T.4 No. 2, 138
- [4] Heidt A Hartung B A Bosman G W Krok Patrizia Rohwer E G Schwoerer H and Hartlet B 2011 Coherent Octave Spanning Near-infrared and Visible Supercontinuum Generation in All-normal Dispersion Photonic Crystal Fibers. Opt . Express, Vol. 19, No. 4, 3775
- [5] Varshney A D and Ravindra K S 2007 *Propagation characteristic of Photonic Crystal Fiber: Scalar Effective index Method and Fully Vectorial Effective Index Method*. Adv. Studies Theo. Phys. Vol 1, no. 2, 75-78

- [6] COMSOL 3.5 2008 COMSOL Multiphysics RF Module Optics and Photonics. CM 010004 COMSOL AB
- [7] Johnson S G 2000 Photonic Crystal: Periodic surprises in Electromagnetism.
- [8] Srivastan B 2002 Photonic Band Gap Crystals. Rensselaer Polytechnic Institute.
- [9] Mohammed M H Jamil M Y and Abdullah A I 2013 Photonic Bandgap Tuning of Photonic Crystals by Air Filling Fraction . Raf. J. Sci. Vol. 24, No. 5, pp 96 – 102.
- [10] Pagnoux D Peyrilloux A Roy P Fevrier S Labonte L and Hilaire S 2010 Microstructured Air-Silica Fibres: Recent Developments in Modelling, Manufacturing and Experiment. Institut de Communications Optiques et Microondes
- [11] Chen M and Yu R 2005 Confinement losses and optimization in Rectangular-Lattice Photonic crystal Fibers.
- [12] Agrawal G P 2001 Nonlinear Optics: *Optics and Photonics*. 3rd Ed. Academic Press, NY.
- [13] Ferrando A Silvestre E and Andres P 2001 Designing the Properties of Dispersion – Flattened Photonic Crystal Fibers. Opt. Exp. Vol.9, No. 13, 687.
- [14] Ferrando A Silvestre E Miret J J and Andres P 2000 Nearly Zero Ultra – Flattened Dispersion in Photonic Crystal Fibers. Opt. Lett. Vol.25, No. 11.

Division D1 – Astrophysics

Radio astronomy in Africa: the case of Ghana

B D Asabere¹, M J Gaylard², C Horellou³, H Winkler¹ and T Jarrett⁴

¹ Department of Physics, University of Johannesburg, P.O. Box 524, 2006, Auckland Park, Johannesburg, South Africa

² HartRAO, P. O. Box 443, Krugersdorp, South Africa (passed away on 14 Aug 2014)

³ Department of Earth and Space Sciences, Chalmers University of Technology, Onsala Space Observatory, SE-439 92 Onsala, Sweden

⁴ Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa

E-mail: bd.asabere@gmail.com

Abstract. South Africa has played a leading role in radio astronomy in Africa with the Hartebeesthoek Radio Astronomy Observatory (HartRAO). It continues to make strides with the current seven-dish MeerKAT precursor array (KAT-7), leading to the 64-dish MeerKAT and the giant Square Kilometer Array (SKA), which will be used for transformational radio astronomy research. Ghana, an African partner to the SKA, has been mentored by South Africa over the past six years and will soon emerge in the field of radio astronomy. The country will soon have a science-quality 32m dish converted from a redundant satellite communication antenna. Initially, it will be fitted with 5 GHz and 6.7 GHz receivers to be followed later by a 1.4 - 1.7 GHz receiver. The telescope is being designed for use as a single dish observatory and for participation in the developing African Very Long Baseline Interferometry (VLBI) Network (AVN) and the European VLBI Network. Ghana is earmarked to host a remote station during a possible SKA Phase 2. The country's location of 5° north of the Equator gives it the distinct advantage of viewing the entire plane of the Milky Way galaxy and nearly the whole sky. In this article, we present Ghana's story in the radio astronomy scene and the science/technology that will soon be carried out by engineers and astronomers.

1. Introduction

In the field of radio astronomy, South Africa (SA) has been the pacesetter on the African continent, with the long established Hartebeesthoek Radio Astronomy Observatory (HartRAO). It has world-class astronomical and space research facilities for cutting-edge radio astronomy research and studies. The current seven-dish MeerKAT precursor array (KAT-7), leading to the 64-dish MeerKAT with first light in late 2016 and the giant Square Kilometer Array (SKA) which construction will start in 2018, will be used for unprecedented radio astronomy researches that will lead to new discoveries. In 1992, Mauritius also appeared on the scene with its meter-wave Fourier Synthesis T-shaped array, the Mauritius Radio Telescope (MRT), designed to survey the southern sky for point sources at 151.6 MHz in the declination of -70° to -10° and sensitivity of 200mJy [1]. The MRT survey produced a southern sky equivalent of the Sixth Cambridge Catalog (6C) of bright radio sources [2]. Although several other African countries had nurtured strategies and plans to host radio astronomy facilities, they are yet to come to fruition. Attempts by Nigeria to build a HartRAO-like facility at Nsukka [3] and Egypt's plans [4, 5] to operate a

radio telescope, in the frequency range of 1.4 to 43 GHz, at Abu Simbel in the southern part of the country are all yet to succeed.

In the mid-2000's, the call on Botswana, Ghana, Kenya, Madagascar, Mauritius, Mozambique, Namibia and Zambia to partner South Africa in its bid to host the SKA, produced new motivation for countries on the continent to speed up their efforts in embracing radio astronomy studies and researches. With the KAT-7 successes [6], the 64-dish MeerKAT project gearing up and the decision that Africa would host a large part of the SKA, the South African SKA Project (SKA-SA) and HartRAO searched in the partner countries to identify redundant satellite communication antennas with potential for conversion of such expensive, but now obsolete, assets for radio astronomy. This type of conversion (e.g. [7, 8, 9]) has been made possible globally owing to the switching over from the data streaming communication satellites to the more efficient and cheaper undersea fibre optic cables for telecommunication signal transport. The goal of the exercise, which led to the discovery of many 30m-class antennas across Africa, is to build the needed capacity in support staff, engineers and scientists in radio astronomy and related disciplines. Ghana has such an asset at its Kuntunse Intelsat Satellite Communication Earth Station (see figure 1 *Left* panel)

In this paper, we give a short briefing on Ghana's radio astronomy prospects and outline the possible science/technology that will soon be done from there.

2. The Ghana Antenna

The Ghana Intelsat Satellite Earth Station at Kuntunse is situated at an elevation of 70m above sea level with position coordinates of $05^{\circ} 45' 01.5''$ N and $00^{\circ} 18' 18.4''$ W [10]. It hosts three antennas of diameters 32m, 16m and 9m, but only the 16m antenna is still operational for satellite communication. The station was commissioned on August 12, 1981 and was operated by the Ghana Telecommunications Corporation until July 03, 2008 when Ghana Vodafone took over, as major shareholder (i.e. purchased 70% shares). Kuntunse is a suburb located off the Nsawam Road, about 25km north-west of the national capital, Accra.

Ghana has warmly embraced the strategy into radio astronomy, according to its radio astronomy development plan incorporated in *The Ghana Science, Technology and Innovation Policy*, and *The Science, Technology and Innovation Development Plan 2011-2015 (GPD15)*. As an SKA Africa partner country, it welcomed and collaborated with the SKA-SA/HartRAO group to access the radio astronomy potential of the redundant satellite communication antennas at Kuntunse. The suitability of the 32m cassegrain antenna and the Kuntunse control station (see figure 1 *Left* panel) for radio astronomy were established by the group after two successive working visits in March and May 2011.

2.1. The Antenna Conversion

To facilitate the agenda envisaged in the GPD15, the Government under the Ministry of Environment, Science, Technology and Innovation (MESTI) on January 01, 2011 established an institution called the Ghana Space Science and Technology Institute (GSSTI), to spearhead all radio astronomy, space science and emerging related technologies programmes, activities and research. GSSTI was established under the Ghana Atomic Energy Commission (GAEC), a national research organisation. It started initially as a Centre under GAEC's Graduate School of Nuclear and Allied Sciences (SNAS), an affiliation of the University of Ghana.

After some protracted negotiations and the official handing over of the station to the state, the antenna conversion exercise by a team of scientists and engineers from SKA-SA/HartRAO and GSSTI started in earnest. Besides the astronomy instrumentation upgrade, major rehabilitative work includes replacing the corroded subreflector quadrupod support legs, replacing the azimuth

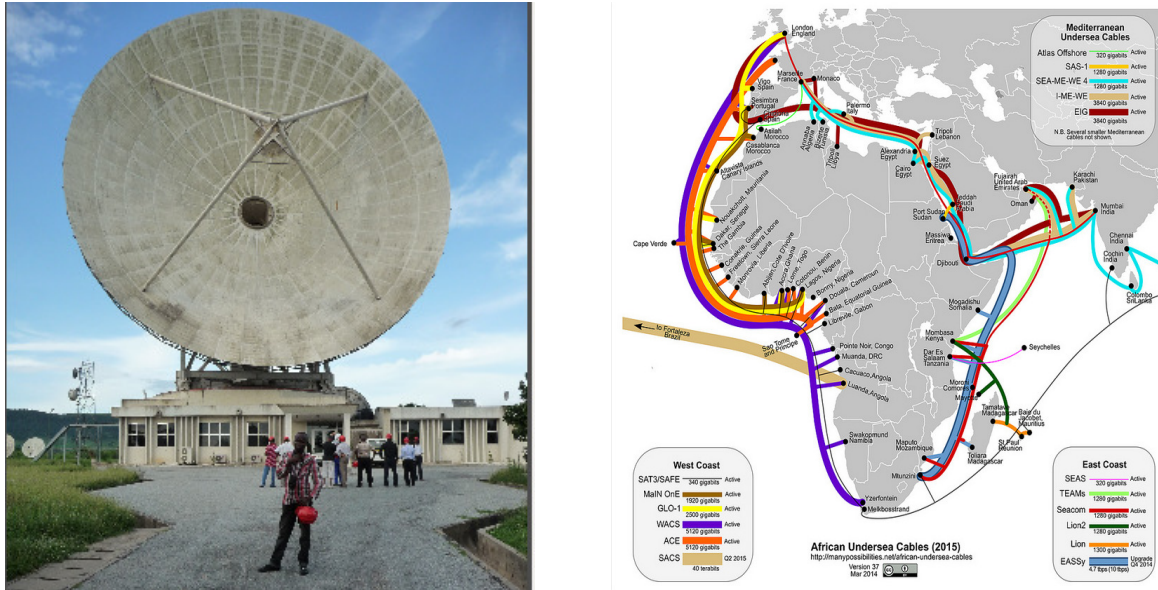


Figure 1. *Left:* The Kuntunse 32m disuse telecommunication antenna currently being converted to radio astronomy facility. Standing in front of the telescope is the first author. *Right:* African Undersea Cables span the coasts of Africa and link the continent to the outside World and represents greater internet bandwidth availability and potential of data transport from the Ghana antenna site [11].

and elevation angle resolvers with more accurate angle encoders, covering the beam waveguide aperture with radome, replacing some rusty hardware on the antenna structure, replacing azimuth and elevation limit switches, flushing and greasing the gearbox systems, changing the elevation and azimuth oil gauges, repainting the whole antenna structure and re-engineering the automatic control/rotation system. The exercise has been running parallel with the requisite local human capital development (HCD) to produce the needed manpower to be custodians and users of the facilities via both formal and informal training interventions. A team of graduates from GSSTI has just completed 6 months training with the South African AVN team at SKA-SA and HartRAO. The latest target for test observations with the 32m antenna is June 2015, with science operations targeted for the end of 2015.

2.2. The Antenna Receivers

The type of observations and science that can best be done with a radio telescope depend on the telescope's location, size, specifications and the receivers and science instruments it is fitted with. During the testing period (Phase-1), the existing telecommunication feed horn in the frequency range 3.8 - 6.4 GHz (C-band) will be used. For the actual science observations (Phase-2), it will initially be fitted with uncooled 5 GHz and 6.7 GHz (C-band) receivers to be followed later by a 1.4 - 1.7 GHz or wider L-band receiver, for which extra funding will be needed. Future receiver developments could include replacing the original C-band feed horn with a wider band design covering more VLBI bands, introducing cryogenic receivers for improved sensitivity and adding more frequency bands.

2.3. Funding for the Conversion

The funds for the antenna conversion and HCD has so far come from the African Renaissance Fund (ARF) of South Africa's Department of International Relations and Cooperation (DIRCO),

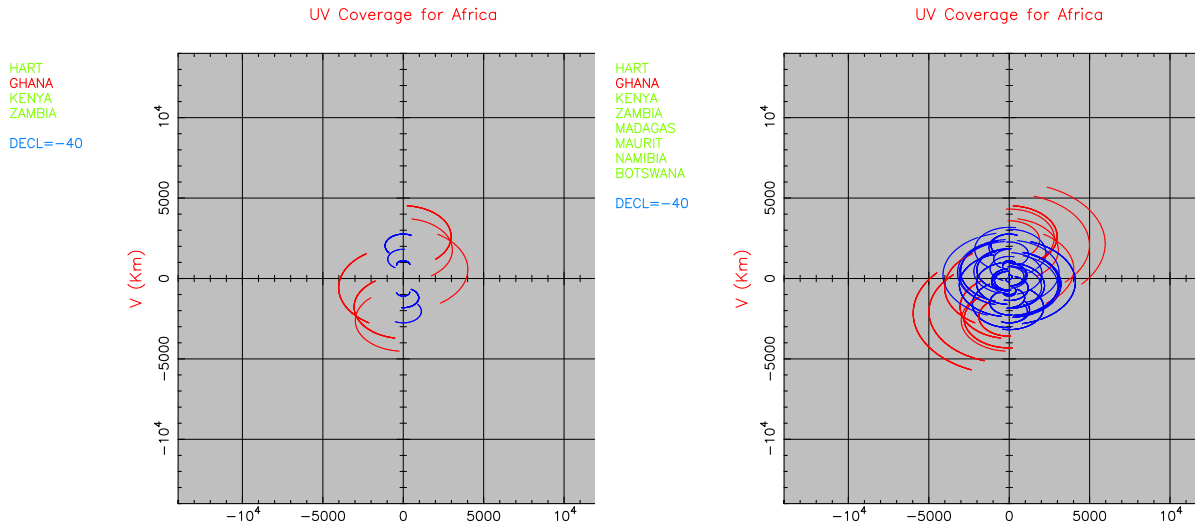


Figure 2. VLBI UV coverages of the developing AVN showing the impacts (in red tracks) of the Ghana 32m antenna in: *Left:* A 4-dish developing AVN, involving the antennas in Ghana (32m), HartRAO (26m), Kenya (32m) and Zambia (30m or 32m), tracking a source at -40° declination. *Right:* A network of antennas in seven SKA-SA partner countries, together with HartRAO, tracking a source at -40° declination.

South Africa's Department of Science and Technology (DST) and the Government of Ghana. The SKA-SA Project, HartRAO and GSSTI are the main facilitating and implementing bodies for the Ghana conversion.

3. Radio Astronomy Projects

On realising a full operational 32m radio astronomy telescope, there are a number of planned astronomical projects:

- Operate as a single dish observatory, to be known as the Ghana Radio Astronomy Facility (GRAF).
- For participation in the developing 4-dish African VLBI Network (AVN) and possible expanded future AVN [12].
- For taking part in the European VLBI Network (EVN).
- With antennas east of South Africa and central Africa, the Ghana antenna would be relevant to both the Australian and American VLBI Arrays. Again, the Ghana antenna, and other AVN telescopes, would be valuable in adding long baseline capability to SKA Phase 1.
- For the possible event of SKA Phase 2, Ghana is earmarked to host some of the single-pixel feed dishes as a remote station.

4. Science with the Ghana Antenna

Ghana's location of 5° north of the Equator (see Section 2) gives it the advantage of viewing the entire plane of the Milky Way galaxy and nearly the whole sky (see figure 2). Another competitive advantage for Ghana, is that it is located close to the African Undersea Cables [11] (see *Right* panel of figure 1) spanning the east and west coasts of the continent and linking Africa to the rest of the World. This close proximity promises greater bandwidth and faster internet connectivity for data transport. With such valuable resources much can be done from Ghana.

In this section we describe some of the science cases of the Kuntunse telescope as a single dish and in VLBI networks, and for the possible long baseline array of SKA Phase 2.

4.1. Single-Dish Science Cases

With the C- and L-bands receivers fitted for Phase-2 operations, one can do the following with the Ghana antenna:

- Radio Continuum Flux measurements (with wideband multi-channel radiometer); use known radio astronomy calibration sources for daily calibrations of receivers and also follow radio emissions from sources such as AGNs emitting gamma-ray flares.
- Pulsar Observations (with wideband multi-channel pulsar timer); monitor the behaviour of pulsars of interest over a long period of time, such as those producing glitches and intermittent pulsars, and hunt for fast radio burst sources.
- Emission Lines Spectroscopy (with narrowband multi-channel spectrometer); maser line monitoring of star forming regions, including hydroxyl masers (1612, 1665, 1667, 1720 MHz) and methanol masers (6668 MHz).

4.2. VLBI Networks Science Cases

With the System Equivalent Flux Density (SEFD) of the dish designed to be better than the typical threshold (i.e. $SEFD < 800\text{Jy}$) for telescopes in current VLBI networks, the Ghana antenna can be a valuable part of the existing VLBI networks. In both the stand-alone AVN, EVN and global VLBI, the inclusion of Ghana will improve imaging and calibration quality and sensitivity in all VLBI astronomy science cases (refer to figures 2 and 3). The VLBI science cases that will be enhanced with the addition of the Ghana telescope equipped for C- and L-band include:

- Mapping Interstellar masers in star-forming regions in the Milky Way.
- Determining the distances to star-forming regions in the Milky Way through methanol maser parallax measurement.
- Using trigonometric parallax measurements to determine accurate pulsar distances as well as pulsar proper motions.
- Imaging active galactic nuclei (AGN).
- Resolving binary systems in extragalactic supermassive black holes.
- Searching for radio transients - long baselines provide discrimination against radio frequency interference.
- Imaging radio emission from X-Ray binary systems and relativistic jets.

4.3. Science at Very High Angular Resolution

In the likely event of Phase 2 of the SKA, Ghana would host a 30-dish single-feed antenna array on the proposed 400m×400m piece of land in the Kwahu mountain range area located in the south-central eastern corridor of the country. This will be in support of the SKA goals of doing science at very high angular resolution (see [13]).

5. Other Relevant Science Instruments

With broad bandwidth and fast internet facilities' availability at the Kuntunse Satellite Earth Station, other instruments such as Meteorological unit (MET-4), Global Navigational Satellite System (GNSS) Receivers, Seismometer, Gravimeters and Magnetometers may be cost-effectively added to enhance the science capability of the station for the benefits of the global community.

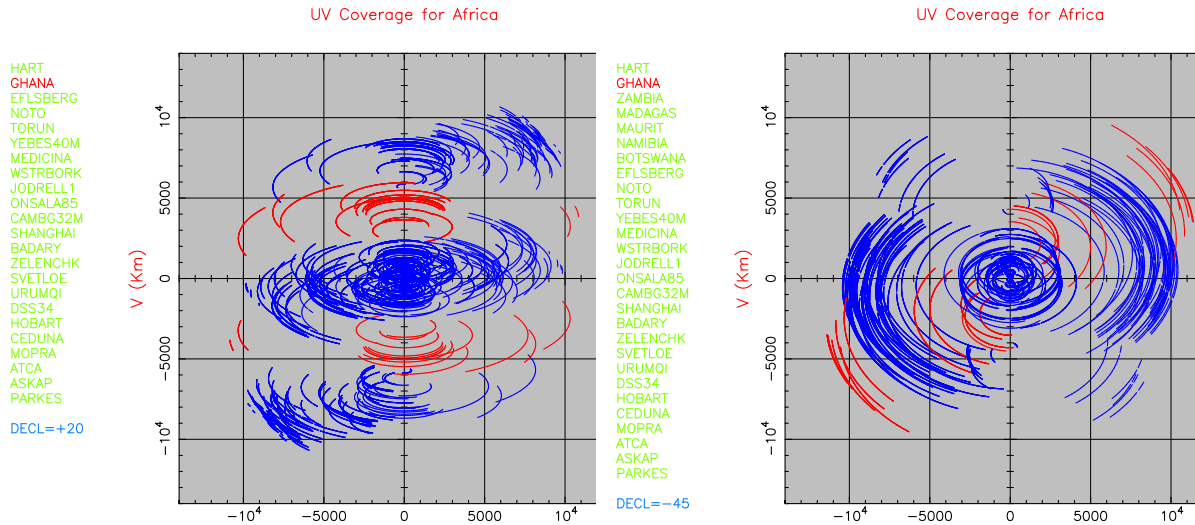


Figure 3. Global VLBI UV coverages of full-track observations showing input from the Ghana 32m antenna in red tracks observing a source: *Left:* At the declination $+20^\circ$ with the existing VLBI antennas. *Right:* At the -45° declination with other five proposed AVN antennas in Africa.

6. Summary

We have unveiled a strong case for the conversion of an obsolete telecommunication facility for radio astronomy use, which presents an excellent prospect of doing radio astronomical science from Ghana. The science cases could be more or less than those outlined in the paper, based on availability of funding and partnerships. The custodians of the dish, GSSTI, warmly welcomes partner support of all forms to realize a befitting facility for radio astronomy and technological science use.

Acknowledgement

The first author acknowledges funding from MIDPREP, an exchange programme between two European Institutes (Chalmers and ASTRON) and three South African partners (University of Stellenbosch, Rhodes University and University of Cape Town). He expresses profound appreciation to SKA-SA/HartRAO and the staff at the Onsala Space Observatory in Sweden for assistance during his regular visits, to mention Jun Jang of the VLBI group for the help.

References

- [1] Somanah R and Udaya Shankar N 2002 *Astrophys. & Space Sci.* **282** 57–67
- [2] Baldwin J E *et al.* 1985 *Mon. Not. R. Astron. Soc.* **217** 717–730
- [3] Okere B and Okeke P N 2011 *African Skies*. **15** 33–4
- [4] Shaltout N 1999 *Radio Sci. Bull. (Belgium)*. **288** 8–12
- [5] Mosalam Shaltout M A 1999 *African Skies*. **14** 10–14
- [6] Woudt P A, Fender R P, Armstrong R P and Carignan C 2013 *South African Journal of Science*. **109** (7/8)
- [7] Fujisawa K, Mashiyama H, Shimoikura T and Kawaguchi N 2002 Proc. IAU 8th Asian-Pacific Regional Meeting, vol II, ed S Ikeuchi, J Hearnshaw and T Hanawa *Astron Soc Japan*. **408** p 3
- [8] McCulloch P M *et al.* 2005 *Astrophysical J.* **129** 2034
- [9] Mansfield I 2010 Telecom New Zealand Converts Satellite Dish into Radio Telescope. <http://www.cellular-news.com/story/46534.php>.
- [10] Ayer J and Fosu C 2008 *Journal of Science and Technology. KNUST, Kumasi, Ghana*. **28** 116–127
- [11] African Undersea Cables (2015) <http://manypossibilities.net/african-undersea-cables/>
- [12] Gaylard M J *et al.* 2011 *SAIP Conference Proceedings*.
- [13] Godfrey L E H *et al.* 2012 *Publications of the Astronomical. Soc. of Australia* **29** 42–53

Imaging & Study of VLBI Reference Frame Sources in the Southern Hemisphere

S. Basu^{1,3}, A. De Witt¹, J. Quick¹, A. Bertarini², L. Leeuw³

¹Hartebeesthoek Radio Astronomy Observatory

²IGG/Max Planck Institute für Radioastronomie

³University of South Africa

E-mail: sayan@hartrao.ac.za, alet@hartrao.ac.za

Abstract. The International Celestial Reference Frame (ICRF) was adopted by the International Astronomical Union (IAU) in 1997. The current standard, the ICRF-2, is based on Very Long Baseline Interferometric (VLBI) radio observations of positions of 3414 extragalactic radio reference sources. The angular resolution achieved by the VLBI technique is on a scale of milliarcsecond to sub-milliarcseconds and defines the ICRF with the highest accuracy available at present. An ideal reference source used for celestial reference frame work should be unresolved or point-like on these scales. However, extragalactic radio sources, such as those that define and maintain the ICRF, can exhibit spatially extended structures on sub-milliarcsecond scales that may vary both in time and frequency. This variability can introduce a significant error in the VLBI measurements thereby degrading the accuracy of the estimated source position. Reference source density in the Southern celestial hemisphere is also poor compared to the Northern hemisphere, mainly due to the limited number of radio telescopes in the south. In order to define the ICRF with the highest accuracy, observational efforts are required to find more compact sources and to monitor their structural evolution. In this paper we show that the astrometric VLBI sessions can be used to obtain source structure information and we present preliminary imaging results for the source J1427-4206 at 2.3 and 8.4 GHz frequencies which shows that the source is compact and suitable as a reference source.

Keywords: Celestial Reference Frame, Quasar, VLBI, IVS, Astrometry, Imaging

1. Introduction

Very Long Baseline Interferometry, or VLBI, radio observations are almost unique in providing us with milliarcsecond or even sub-milliarcsecond resolution, and also allow us to do astrometry with sub-milliarcsecond accuracy (e.g. Walker, 1999; Thompson et al. 2007). VLBI achieves its resolution by using radio telescopes hundreds or thousands of kilometers apart simultaneously to form an interferometer. The received signals are amplified at the participating antennas and are sent to a correlator by storing them on disk packs or by using network links. The received signals from each pair of antennas are then cross-correlated and Fourier transformed in order to determine the sky brightness distribution at the observed radio frequency. Most VLBI applications rely on having reference sources, for example, the imaging of faint radio sources in phase-referencing mode, accurate differential astrometry, spacecraft tracking, space navigation and space geodesy.

Since the late 1970s, VLBI observations have been used to determine the positions of radio sources with milliarcsecond precision. In recent years, the accuracy of the VLBI technique has improved substantially and high precision VLBI measurements of positions of extragalactic radio sources are now used to define and maintain celestial radio reference frames with sub-milliarcsecond precision. Various radio source catalogs have been constructed from a number of high accuracy, dual frequency 8.4 GHz and 2.3 GHz observations which have been collected from different networks for geodetic and astrometric purposes. Ma et al. (1990) produced the first catalog of 182 sources with a positional accuracy of 1 milliarcsecond. All of these sources are located north of -30° declination. The source density was increased significantly in the Northern hemisphere with more observing campaigns and a few sources were also added in the Southern hemisphere (e.g. Reylonds et al. 1994; Russell et al. 1994; Johnston et al. 1995).

The realisation of the radio celestial reference frame was adopted by the 23rd International Astronomical Union (IAU) General Assembly in 1997 to replace the traditional optical fundamental celestial reference frame, the FK5 reference frame. High precision VLBI measurements of positions of extragalactic radio sources now define and maintain the International Celestial Reference Frame (ICRF), which forms the underlying basis for positional astronomy. The current realisation of the ICRF (ICRF-2, Ma et al. 2009) is based on dual frequency (8.4 GHz, 2.3 GHz) VLBI observations of 3414 radio reference sources, including 295 defining sources which determine the orientation of the frame's axes. In 2012, a working group under the IAU was formed with the goal of the realisation of the ICRF-3 by 2018 with specific emphasis on improving the accuracy and coverage in the southern hemisphere.

Extragalactic radio sources that are relatively bright at the frequency of observation, compact or core-dominated on VLBI scales and with no or little detectable motion are well suited as reference sources. The primary sources used as reference sources in VLBI are radio-loud quasars. Quasars being at great distances do not exhibit any measurable proper motion or parallax, making them ideal reference sources. In radio-loud quasars, the radio emission originates with a relativistic jet launched from the vicinity of the black hole. The ones useful as reference sources are those where the scale of this jet is small compared to the resolution such that the radio emission is compact or core-dominated and appears almost point-like. Unfortunately, many of the radio-loud quasars that make up the ICRF exhibit spatially extended intrinsic structures, with VLBI imaging showing jets in addition to compact cores. The extended emission structures in these sources may also evolve significantly over time on scales of months to years, and in addition they also exhibit flux density variations on timescales of years to weeks.

Charlot (1990) showed that the effect of source structure on VLBI astrometric positions can be significant, where any departure from the point source approximation, commonly made in astrometric analysis, introduces errors in the accurate position of the source. Temporal variations of the intrinsic structure of these objects can result in apparent motion when observed at several epochs. Results from Fey, Clegg & Fomalont (1996) and Fey & Charlot (1997), show that structural variations of sources can be extreme, ranging from relatively compact core objects, to compact double sources, to complex core-jet objects. It is therefore important to map the structures of these sources on a regular basis.

Catalogues of compact radio sources, including the ICRF-2 are weak in the south, especially at declinations south of -45° , the limit of the reach of the Very Long Baseline Array (VLBA). Among 295 defining sources, fewer than 30% are in the Southern hemisphere. In the Northern hemisphere, ICRF sources are well distributed and regular imaging of Northern hemisphere ICRF sources are made through the ongoing astrometric and geodetic Research and Development VLBI (RDV) experiments using the VLBA (e.g. Fey & Charlot, 2000; Collioud & Charlot, 2009). The results of the imaging and analysis of these studies prove to show the importance of continual observing and analysis in order to monitor the sources for variability or structural changes so that their astrometric quality can be continuously evaluated. There have been many

efforts in recent years to increase the number of known Southern hemisphere reference source, in particular the astrometric observations from the Long Baseline Array (LBA) Calibrator Survey (LCS), which has already produced a significant improvement at 8.4 GHz (Petrov et al 2011). Dedicated astrometric observations at 8.4 and 2.3 GHz band to density the ICRF in the south are currently underway, as proposed in Lovell et al. (2013). There has also been a few dedicated imaging observations of southern sources (e.g. Hungwe et al. 2011; Ojha et al. 2004; Ojha et al. 2005), and first images from the LCS astrometric experiments have been produced (de Witt & Bietenholz, 2012).

Dedicated imaging observations to map the source structures on a regular basis have proven to be very resource intensive, with availability of antennas being one of the most limiting factors in the south. Based on these considerations we have investigated the possibility of imaging sources from existing geodetic and astrometric observations in the south. Imaging of source structure from the LCS experiments have proven to be successful and we will continue to image more of the LCS experiments. We also have identified the Celestial Reference Frame Deep South (CRDS) astrometric VLBI observations to be potentially suitable for mapping purposes. Present and past surveys (e.g. Pearson & Readhead, 1988; Taylor et al. 1996) show that such surveys are an important tool for astrometry but also for investigation of a wide range of astrophysical phenomena.

2. Observational data:

We have reduced the data for a single epoch, 24-hour session of one LCS (v271m) and one CRDS (CRDS63) experiment.

The LCS¹ is an ongoing VLBI project to observe a list of flat-spectrum radio sources below -30° declination at 8.4-GHz using the LBA array. Observations for LCS experiment v271m were made on 15-16 June 2013 at a central frequency of 8.344 GHz, recording a total bandwidth of 128 MHz, right circular polarization (RCP) only. Eight antennas (see Table 1 with lcs tag) from the LBA participated and a total of 101 sources were observed in this experiment with 2–4 scans of 2–6 minutes duration per source. The *uv*-coverage for LCS experiments is very limited due to the few scans per source and also the array configuration. The data were correlated at the Curtin University of Technology using the DIFX correlator (Deller & Tingay, 2007). Imaging of sources from experiment v271m is still in progress.

The CRDS² astrometric observing sessions are part of the International VLBI Service for Geodesy and Astrometry (IVS) campaign to strengthen the ICRF in the south. Since 2013 observations have been performed using a regular network of six southern stations (see Table 1 with crds tag). A total number of 97 sources have been observed through CRDS sessions with the majority of sources being south of -30° declination. Most sources are observed in at least 2-3 sessions per year with around 2-7 scans of 9 minutes duration per source. As opposed to more typical astrometric sessions where only 2-station scans are required, the majority of scans in the CRDS sessions, observed since 2013, include at least 4 to 6 stations per scan, making these observations more suitable for mapping purposes.

In order to test the suitability of CRDS sessions for imaging of source structure, we have chosen the most recent CRDS session at the time, which is CRDS63. Observations were made on 14-15 January 2013 and a total number of 38 sources were observed. All six telescopes as mentioned in table 3 were scheduled for observations, but unfortunately the Hobart 12m and Wakworth 12m telescopes did not participate. Data were recorded at RCP with 8 IFs at X-band and 6 IFs at S-band and a bandwidth of 4 MHz per IF³. Data were correlated at the Washington Correlator (WACO) in Washington, DC.

¹ Information on the LCS is available on the web at <http://astrogeo.org/lcs/>

² Information on the CRDS sessions are available on the IVS webpage at <http://ivscc.gsfc.nasa.gov>

³ Bandwidth was increased to 8 MHz from CRDS66 in June, 2013.

Table 1. Antennas participating in the LCS & the CRDS experiments.

Station	Location	Diameter(m)	SEFD(2.3 GHz)	SEFD(8.4 GHz)
ATCA ^{lcs} (Array with 6 antennas)	Australia	6×22	106	86
Ceduna ^{lcs}	Australia	30	400	600
HartRAO ^{lcs,crds}	South Africa	26	2929	1978
Hobart ^{lcs,crds}	Australia	26	1019	902
Hobart ^{lcs,crds}	Australia	12	4000	5000
Mopra ^{lcs}	Australia	22	530	430
Parkes ^{lcs}	Australia	64	30	43
Tidbinbilla ^{lcs}	Australia	70	16	25
Yarragedee ^{crds}	Australia	12	3247	5333
Katherine ^{crds}	Australia	12	4337	3286
Warkworth ^{lcs}	New Zealand	12	8000	8000

***System Equivalent Flux Density (SEFD) values are in Jy.*

3. Data reduction & Results:

The data reduction for the imaging of the CRDS and LCS observations was performed using the NRAO Astronomical Processing System (AIPS; Greisen, 1988). The CRDS correlated data were fringe-fitted with *FOURFIT* which is used to produce data suitable for importing into AIPS. The correlated data was then imported into AIPS using *MK4IN* (Alef & Graham, 2002). Amplitude gains were derived from nominal system temperatures. Thereafter, data inspection, initial editing and fringe fitting were done in the standard manner. We did an initial round of fringe-fitting to find approximate residual rates and delays. The main editing of data was carried out using this approximate calibration, and then using the edited data, we proceeded to a second round of fringe-fitting to refine the calibration, with each source being fringe-fitted individually. The visibility data for the source J1427-4206 were Fourier inverted and deconvolved using the CLEAN algorithm, and the amplitude gains were further refined by self-calibrating using CLEAN models. For the final CLEAN image we used the square root of the statistical visibility weights, which increases the robustness of the image.

We present preliminary imaging results (see Figure 1) for the source J1427-4206 at 8.4 GHz and 2.3 GHz. We have chosen J1427-4206 for our initial attempt as this source has the highest flux density among the sample of sources in CRDS63 and was observed during 4 scans that included all 4 antennas. J1427-4206 has also previously been observed with both the VLBA and the LBA which provides an opportunity to compare our results with available images from existing experiments. Figure (a), (b), (c) indicate the source structure, visibility phase and visibility amplitude plots respectively. For any compact source we can expect its amplitude and phase to be steady over all the baselines which is shown by figure (b) and (c) the gap in these two plots is due to the lack of available antenna between HartRAO and Warkworth which has the longest baseline length more than 10,000 km. Figure (b) and (c) conclude that the source has a point like structure which is shown in figure (a). Figure (d), (e) and (f) shows the source structure, visibility amplitude and visibility phase plot respectively of the same source J1427-4206 at 2.3 GHz. All the plots for amplitude and phase strongly indicate that the observed source has a compact structure at both 8.4 and 2.3 GHz frequencies which is shown in figure (a) and (d) respectively.

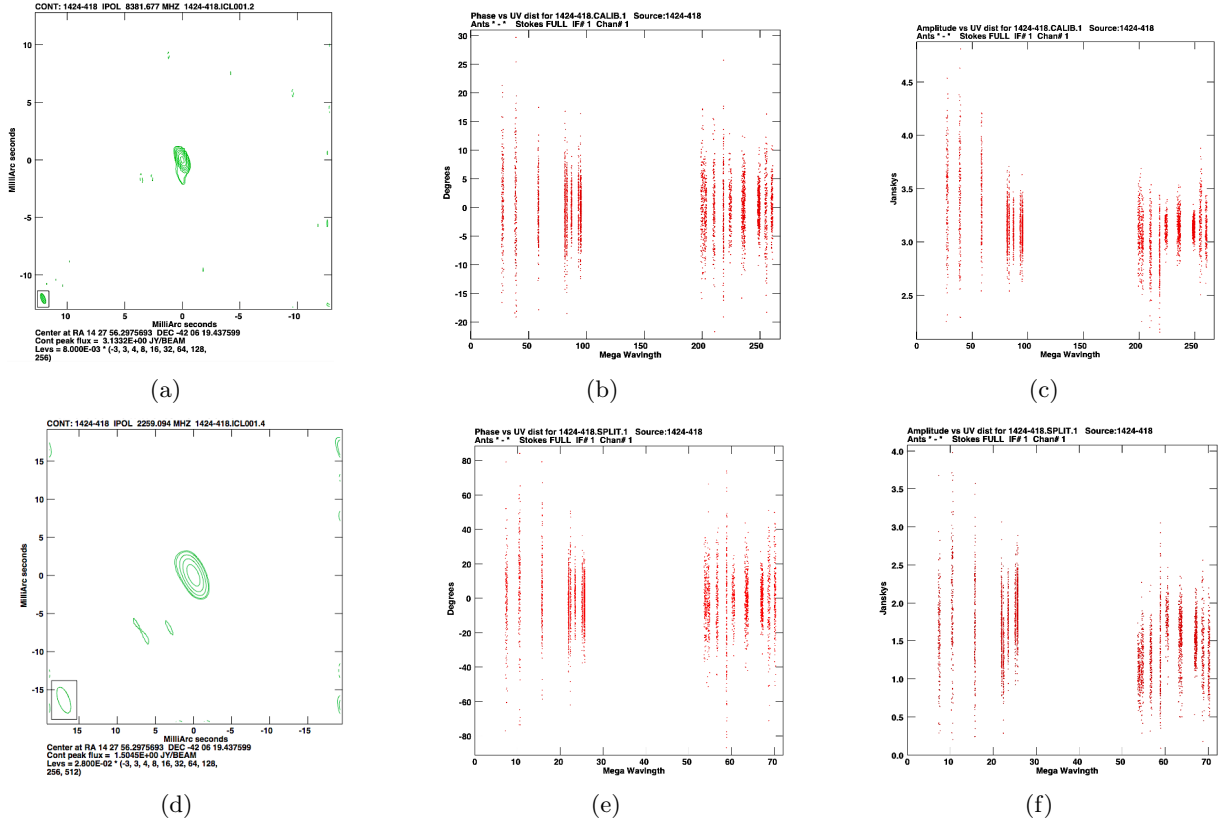


Figure 1. (a) shows a map of the source J1427-4206 from the CRDS63 experiment at 8.4 GHz. The peak flux density is 3.13 Jy/beam, the rms noise is 8 mJy/beam and the contour levels are at -3, 3, 4, 8, 16, 32, 64, 128, 256, 512 times the rms noise. North is up and East is to the left. (b) shows the visibility phase vs uv-distance plot and (c) shows the visibility amplitude vs. uv-distance plot for the source J1427-4206 at 8.4 GHz. (d) shows a map of the source J1427-4206 from the CRDS63 experiment at 2.3 GHz. The peak flux density is 1.50 Jy/beam, the rms noise is 28 mJy/beam and the contour levels are at -3, 3, 4, 8, 16, 32, 64, 128, 256, 512 times the rms noise. North is up and East is to the left. (e) is the visibility phase vs uv-distance plot and (f) is the visibility amplitude vs uv-distance plot of the source J1427-4206 at 2.3 GHz.

4. Summary & Outlook

We have presented here our preliminary imaging results of source J1427-4206 from the CRDS63 astrometric session. We aim is to complete the imaging of all the sources observed in the CRDS63 session in order to evaluate the astrometric quality of these sources for future astrometric and geodetic purposes. Efforts are also underway to image sources from the LCS and to reduce data for more CRDS sessions. In future, we plan to set up a pipeline for automated imaging CRDS sessions. These dual frequency observations will help us to test the frequency-dependence of the sources and to monitor the sources for variability or structural changes on a continuous basis.

Acknowledgements:

We thank Alan Fey, Maria Davis and David Hall from USNO for use of the data from the CRDS observing sessions. We would like to thank HartRAO, NRF-DST and the University of South Africa for financial support for this project. This research has made use of the IVS and the LCS database.

References

- [1] Alef, W. & Graham, D. A. 2002, *Proc. of the 6th European VLBI Network Symposium*. p 31
- [2] Charlot, P. 1990 *AJ*, **99**, 1309
- [3] Collioud, A. & Charlot, P. 2009, in *Proceedings of the 19th European VLBI for Geodesy and Astrometry Working Meeting*, **19**
- [4] Deller, A. T., Tingay, S. J., et al. 2007, *PASP*, **119** 318-36
- [5] De Witt, A. & Bietenholz, M. 2012, in *Proceedings of Science of "11th European VLBI Network Symposium & Users Meeting"*
- [6] Fey, A. L. & Charlot, P. 1997, *ApJS*, **111**, 95
- [7] Fey, A. L., Ojha, R., et al. 2004a, *AJ*, **128**, 2593
- [8] Fey, A. L. & Charlot, P. 2000, *ApJS*, **128**, 17
- [9] Fey, A. L., Clegg, A. W., & Fomalont, E. B. 1996, **ApJS**, *105*, 299
- [10] Greisen, E. W. 1988, AIPS Memo 61, The Astronomical Image Processing System (Socorro: NRAO), <ftp://ftp.aoc.nrao.edu/pub/software/aips/TEXT/PUBL/AIPSMEMO61.PS>
- [11] Johnston, K. J., et al. 1995, *AJ*, **110**, 880
- [12] Lovell, J. E. J. et al. 2013, *Journal of Geodesy*, **87**, 527
- [13] Ma, C., Shaffer, D. B., de Vegt, C. et al. 1990, *AJ*, **99**, 1284
- [14] Ma, C., et al. 2009, *IERS Technical Note*, **35**, 1
- [15] Ojha, R., et al. 2004, *AJ*, **127**, 3609-3621
- [16] Ojha, R., et al. 2005, *AJ*, **130**, 2529-2540
- [17] Petrov, L. et al. 2011, *MNRAS*, **414**, 2518
- [18] Pearson, T. J., & Readhead, A. C. S. 1988, *ApJ*, **328**, 114
- [19] Reynolds, J. E., Jauncey, D. L., et al. 1994, *AJ*, **108**, 725
- [20] Russell, J. L., et al. 1994, *AJ*, **107**, 379
- [21] Taylor, G. B., Vermeulen, R. C., et al. 1996, *ApJS*, **107**, 37
- [22] Thompson, A. R. et al. 2007, in *Interferometry and Synthesis in Radio Astronomy*, John Wiley & Sons, 2007
- [23] Walker, R. C. 1999, in *Astronomical Society of the Pacific Conference Series* 180, ed. G. B. Taylor, C. L. Carilli & R. A. Perley, 433

Weighing Dark Matter in Brightest Cluster Galaxies

H Branken¹, S Loubser¹ and K Sheth²

¹Centre for Space Research, North-West University, Potchefstroom Campus, Private Bag X6001, Potchefstroom, 2520, SA

²National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA

E-mail: 22095306@nwu.ac.za

Abstract. Some authors have noted the peculiarity that a sub-sample of BCGs exhibit a rising velocity dispersion profile, contrary to massive elliptical galaxies not located in the centres of the clusters. This letter highlights novel research that will be conducted on a sample of BCGs in an attempt to shed some light on their peculiar rising velocity dispersion profile. The angle of approach is to determine whether there is a possible correlation between the dark matter (DM) profile and velocity dispersion profile of a galaxy. In particular, a discussion is given on the algorithms generating detailed stellar-mass map estimates and which also constrains the dynamical mass of the galaxy from the stellar velocity dispersion measurements.

1. Introduction

The central galaxies in galaxy clusters have many special properties: they are very massive early-type galaxies (ETGs), have very high luminosities, are diffuse in structure with extended envelopes, and reside at the central locations of clusters. Due to their special location they are exposed to interesting evolutionary phenomena such as dynamical friction, galactic cannibalism, and cooling flows.

An observational aspect difficult to understand, appreciated in the work of [1–3], is the rising velocity dispersion profile of a subsample of BCGs. Murphy [3] summarises the dilemma by asking whether a rising dispersion profile is a true reflection of the gravitational potential of the galaxy (which includes the effects of DM), or if it is simply a snapshot of a dynamical system that has not yet reached equilibrium. Clearly then, there exists an unwanted degeneracy between the contribution of DM, a system that has not yet virialised, or perhaps a combination between the two.

Bahcall & Kulier [4], for example, put forth the argument that the observed cumulative M/L profile rises on small scales, thereby reflecting the increasing M/L of the central brightest galaxy of the cluster, and then flattens to a nearly constant ratio on scales above $\sim 300h^{-1}\text{kpc}$, where light follows mass on all scales and in all environments. They further suggest that most of the dark matter in the universe is located in the large halos of individual galaxies. However, after considering that some BCGs exhibit a positive dispersion slope and the remainder not, this begs the question whether the DM profile truly takes on the same shape in each galaxy, and exactly how universal the distribution is. In other words, the question of ‘*where is the DM*’ still remains.

In this study, a systematic comparison between the dynamical- and stellar-masses of the BCGs will be made for a sample of ~ 20 (from Loubser [1]) of nearby BCGs at relatively low redshift. Of these 20 BCGs, 8 show a clear positive velocity dispersion gradient.

2. Mass profile estimates

A Multi-Gaussian Expansion (MGE) parametrization pioneered by Cappellari [5] is adopted for the stellar density map $[M_\odot \cdot \text{pc}^{-2}]$ due to its accuracy in reproducing the surface brightness profiles of real galaxies. If the x -axis is aligned with the photometric major axis, then the surface brightness Σ $[L_\odot \cdot \text{pc}^2]$ at the location (x', y') on the sky can be written as

$$\Sigma(x', y') = \sum_{k=1}^N \frac{L_k}{2\pi\sigma_k^2 q'_k} \exp \left[-\frac{1}{2\sigma_k^2} \left(x'^2 + \frac{y'^2}{q'_k} \right) \right] \quad (1)$$

where N is the number of the adopted Gaussian components, having total luminosity L_k , observed axial ratio $0 \leq q'_k \leq 1$ and dispersion σ_k along the major axis. Once this parametrization has been obtained for the surface brightness of the real galaxy, then it can be converted into a *stellar surface density map* $[M_\odot \cdot \text{pc}^{-2}]$. Consequently a mass-profile is generated for the baryonic, luminous matter component of the galaxy.

On the other hand, a parametrization for the dark matter mass profile cannot be obtained in this way because DM cannot be observationally detected. Therefore from purely theoretical and simulation arguments, different radial profiles for the dark matter accompanying galaxies have been constructed over the years. Examples of such DM profiles include the Navarro-Frenk-White (NFW) profile [6], the Einasto profile and the Sersic profile [7]. However, the question of which profile is most appropriate for the BCG investigated still remains.

The velocity dispersion measurements of the BCGs play the crucial role in constraining the radial distribution of the DM, as well as constraining the black hole mass and anisotropy profile of the BCG. What the preceding entails for this specific study is as follows: Initially, an arbitrary DM profile is superimposed onto the stellar mass profile (obtained via MGE), thereby representing the *total* mass profile of the BCG. Also, an educated guess is made for the anisotropy profile of the BCG *and* the mass of its black hole. The total mass profile, anisotropy profile, luminosity density and black hole mass are then fed into the Jeans Anisotropic Modelling (JAM) algorithm [8] in order to compute the projected second velocity moment of the BCG coupled to these parametrizations (equation (50) of Cappellari [8]). By deducing the offset between the observed dispersion measurements (extracted from optical spectroscopy) and the predicted JAM dispersion profile, a χ^2 can be computed giving the goodness of the model fit to the data.

This process is then iteratively repeated for numerous other DM profiles, black-hole masses, and anisotropy profiles, thus amounting to 3 free parameters. A final estimate for DM profile, black hole mass, and anisotropy profile for the BCG is obtained depending on which set of parameters yields the minimum χ^2 . This minimum χ^2 corresponds to the most accurate fit possible to the observed velocity dispersion profile. Notice that the luminosity density is not a free parameter, but remains fixed during the procedure. This is motivated by the fact that it is stringently derived from the MGE algorithm.

3. Summary

The work done by Loubser [1] details the spatially-resolved kinematics of the sample of ~ 20 BCGs. In [1] the William Herschel telescope (WHT) and Gemini North and South telescopes were used to perform long-slit spectroscopy on the sample of BCGs. Radial velocity and dispersion profiles were extracted from the long-slit spectroscopy via the ppxf algorithm [9]. Hence, with the available velocity dispersion profiles at hand, it is possible to constrain the dark matter profiles of the 20 BCGs. This of course will be facilitated by the JAM algorithm. Furthermore, luminous stellar mass maps will be obtained by feeding Hubble WFPC2 photometry into the MGE algorithm. The JAM algorithm will also put constraints on the black hole mass, and anisotropy profile of the BCG.

To conclude, the 3 free parameters (DM profile, anisotropy profile, and black hole mass) for each BCG are solved for from the Jeans equations once the minimum χ^2 value has been obtained from the fit. Finally, there will be investigated if there is any connection between the DM profile and the varying dispersion gradients found for different BCGs. It is also instructive to see how the result obtained from this method compares with the values found in literature.

Acknowledgments

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and not necessarily to be attributed to the NRF.

HCB would like to express his sincerest thanks and appreciation to SIL and KS for their continued support, helpful comments and insights.

References

- [1] Loubser S I 2009 *Kinematics and Stellar Populations in Brightest Cluster Galaxies* Ph.D. thesis University of Central Lancashire
- [2] Loubser S I, Sansom A E, Sánchez-Blázquez P, Soechting I K and Bromage G E 2008 **391** 1009–1028
- [3] Murphy J D, Gebhardt K and Cradit M 2014 **785** 143
- [4] Bahcall N A and Kulier A 2014 **439** 2505–2514
- [5] Cappellari M 2002 **333** 400–410
- [6] Navarro J F, Frenk C S and White S D M 1997 **490** 493–508
- [7] Merritt D, Graham A W, Moore B, Diemand J and Terzić B 2006 **132** 2685–2700
- [8] Cappellari M 2008 **390** 71–86
- [9] Cappellari M and Emsellem E 2004 **116** 138–147

The effect of an offset-dipole magnetic field on the Vela pulsar's γ -ray light curves

M Breed¹, C Venter¹, A K Harding² and T J Johnson³

¹ Centre for Space Research, North-West University, Potchefstroom Campus, Private Bag X6001, Potchefstroom, 2520, South Africa

² Astrophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

³ National Research Council Research Associate, National Academy of Sciences, Washington, DC 20001, resident at Naval Research Laboratory, Washington, DC 20375, USA

E-mail: 20574266@nwu.ac.za

Abstract. Over the past six years, the *Fermi* Large Area Telescope has detected more than 150 γ -ray pulsars, discovering a variety of light curve trends and classes. Such diversity hints at distinct underlying magnetospheric and/or emission geometries. We implemented an offset-dipole magnetic field, with an offset characterised by parameters epsilon and magnetic azimuthal angle, in an existing geometric pulsar modelling code which already includes static and retarded vacuum dipole fields. We use these different B -field solutions in conjunction with standard emission geometries, namely the two-pole caustic and outer gap models (the latter only for non-offset dipoles), and construct intensity maps and light curves for several pulsar parameters. We compare our model light curves to the Vela data from the second pulsar catalogue of *Fermi*. We use a refined chi-square grid search method for finding best-fit light curves for each of the different models. Our best fit is for the retarded vacuum dipole field and the outer gap model.

1. Introduction

The discovery of the first pulsar in 1967 by Bell and Hewish [17] gave birth to pulsar astronomy. Pulsars are believed to be rapidly-rotating, compact neutron stars that possess strong magnetic, electric, and gravitational fields [1]. They emit radiation across the entire electromagnetic spectrum, including radio, optical, X-ray, and γ -rays [4]. Since the launch of the *Fermi Gamma-ray Space Telescope* in June 2008, over 150 γ -ray pulsars have been detected, of which the Crab and Vela pulsars are the brightest sources. *Fermi* consists of two parts including the Large Area Telescope (LAT) and the Gamma-ray Burst Monitor. The LAT measures γ -rays in the energy range between 20 MeV and 300 GeV [3]. Over the past six years, *Fermi* has released two pulsar catalogues, both describing the light curve profiles and spectral characteristics of γ -ray pulsars [1, 2]. The light curves show great variety in profile shape, and may be divided into three general classes based on the relative phase differences between the radio and γ -ray pulses [11, 24]. The light curves also show energy-dependent behaviour. Most of the young and millisecond pulsars exhibit two γ -ray peaks whereas some pulsars including Vela display three peaks [2].

Models of pulsar geometry are characterised by the inclination angle (α) between the rotation (Ω) and magnetic (μ) axes, the observer's viewing angle (ζ) between the rotation axis and the observer's line of sight, and the impact angle ($\beta \equiv |\zeta - \alpha|$). Geometric models assume the

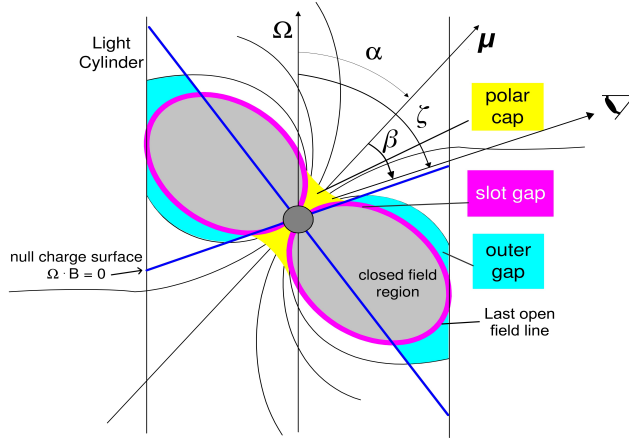


Figure 1. A schematic representation of the different geometric pulsar models. The PC model extends from the neutron star surface up to low-altitudes above the surface (yellow region). The TPC emission region (curved magenta lines) extends from R_{NS} (neutron star radius) up to the light cylinder R_{LC} (where corotation speed equals the speed of light, vertical lines), the OG region (cyan regions) from R_{NCS} (the null charge surface, where the Goldreich-Julian charge density $\rho_{\text{GJ}} = 0$ [12], indicated by the dark blue lines) up to the R_{LC} , and the PSPC from R_{NS} to the R_{LC} , covering the full open volume region. Adapted from [14].

presence of several ‘gap regions’ in the pulsar magnetosphere. These are defined as regions where particle acceleration and emission take place. These geometric models include the two-pole caustic (TPC) [9] (the slot gap (SG) [19] model may be its physical representation), outer gap (OG) [6, 22], and pair-starved polar cap (PSPC) model [15]. All of these models are represented in figure 1. The emissivity ϵ_ν of high-energy photons within this gap region is assumed to be uniform in the corotating frame for geometric models (for physical models the ϵ_ν changes with radial distance when an electric field is assumed). Since the γ -rays are expected to be emitted tangentially to the local magnetic field in the corotating frame [10], the assumed magnetic field geometry is very important with respect to the predicted light curves. Several magnetospheric structures have been studied, including the static dipole field [13], the retarded vacuum dipole field (RVD) [8] and the offset-dipole B -field. The latter is motivated by the fact that retardation of the B -field at the light cylinder causes offset of the polar caps (PCs), always toward the trailing edge of the PC ($\phi_0 = \pi/2$) [16]. The offset is characterised by parameters epsilon (ϵ) and magnetic azimuthal angle (ϕ_0) which represents a shift of the PC away from the magnetic axis, with $\epsilon = 0$ corresponding to the static-dipole case.

In this paper we studied the effect of using different combinations of magnetospheric structures, geometric models, and model parameters on γ -ray light curves. In Section 2 we discuss the implementation of an offset-dipole solution and associated electric field. Section 3 describes how we matched limiting cases of the low-altitude and high-altitude E -fields using a matching parameter η_c . In Section 4 we describe the chi-squared (χ^2) method we used to search for best-fit light curves. Our results are given in Section 5 and the conclusions follow in Section 6.

2. Implementation of an offset-dipole B -field

We implemented [5] an offset-dipole magnetic field [16] into an existing geometric pulsar modelling code [10] which already includes static and RVD fields. The implementation involves a transformation of the B -field from spherical to Cartesian coordinates, rotating both the B -field components and its Cartesian frame through an angle $-\alpha$, thereby transforming the B -field from the magnetic frame ($\hat{\mathbf{z}}' \parallel \boldsymbol{\mu}$) to the rotational frame ($\hat{\mathbf{z}} \parallel \boldsymbol{\Omega}$). We extended the range of ϵ for which we could solve the PC rim (for details, see [10]) by enlarging the colatitude range thought to contain the last open field line (tangent to R_{LC} , see figure 1).

3. Matching parameter

It is important to take the accelerating E -field into account (in a physical model) when such expressions are available, since this will modulate the ϵ_ν in the gap (as opposed to geometric models where we just assume constant ϵ_ν at all altitudes). We use analytic expressions [20] for a low-altitude and high-altitude SG E -field in an offset-dipole magnetic geometry. These are matched at a critical scaled radius $\eta_c = r_c/R_{\text{NS}}$ to obtain a general E -field valid for all altitudes [20]. In previous work we chose $\eta_c = 1.4$. In this paper we solve η_c on each field line. Using the general E -field we could solve the particle transport equation [23, 7] (taking only curvature losses into account) to obtain the particle Lorentz factor γ that is necessary for calculating the curvature emissivity.

4. Finding best-fit light curves

We tested several fitting methods in order to select the most suitable one. We decided to use the standard χ^2 method. For each combination of B -field and geometric model, we calculate χ^2 for each set of free model parameters α , ζ , normalisation (A) and phase shift ($\Delta\phi$). We assume that for the bright Vela pulsar, with a large amount of counts in each bin, the χ^2 will be Gaussian distributed with $N_{\text{dof}} = 96$ the number of degrees of freedom. For the standard Gaussian distribution, assuming very small Gaussian errors, the reduced χ^2_ν values are very large and therefore we needed to scale the χ^2 values with the optimal χ^2_{opt} and multiply by N_{dof} . The scaled χ^2 is denoted by ξ^2 , and the reduced scaled χ^2 will have a value of $\xi^2_\nu = 1$ [21]. We next determined confidence intervals (1σ , 2σ , and 3σ) in (α, ζ) space for these model parameters, using $N_{\text{dof}} = 2$. We only fitted the γ -ray light curve, because we do not want to bias our results with a simplistic radio model.

5. Results

5.1. Phaseplots and light curves

In figure 2, 3 and 4, we show the intensity maps or phaseplots (emission per solid angle versus ζ and pulse phase ϕ) and their corresponding light curves (i.e., cuts at constant $\zeta = \zeta_{\text{cut}}$) for the offset-dipole B -field and the TPC model. The dark circle in figure 2 is the non-emitting PC, and the sharp, bright regions are the emission caustics, where radiation is bunched in phase due to relativistic effects. Figure 2 and 3 represent phaseplots for a fixed $\alpha = 40^\circ$ and $\zeta_{\text{cut}} = 70^\circ$, with ϵ ranging from 0 to 0.18 with increments of 0.03. We contrasted the cases of constant and variable ϵ_ν . We observed a qualitative difference in caustic structure. The caustics seem larger and more pronounced in the constant ϵ_ν case. In figure 4 we chose a fixed value of $\epsilon = 0.18$ for variable ϵ_ν , with α ranging from 0° to 90° and ζ_{cut} from 15° to 90° , both with a resolution of 15° . This shows examples of various light curves that may be obtained in this model.

5.2. Contours and best-fit light curves

In figure 5 we represent our best fits we obtained using the χ^2 method. Our overall best fit is for the RVD field and the OG model, with $\alpha = 78^\circ$, $\zeta = 69^\circ$, $A = 1.3$, and $\Delta\phi = 0$. For the offset-dipole solution and the TPC model, we have a best fit (assuming constant ϵ_ν) for parameters $\epsilon = 0.00$, $\alpha = 73^\circ$, $\zeta = 45^\circ$, $A = 1.3$, and $\Delta\phi = 0.55$. The best fit for the offset-dipole solution and the TPC model, assuming variable ϵ_ν , is for parameters $\epsilon = 0.18$, $\alpha = 73^\circ$, $\zeta = 17^\circ$, $A = 0.5$, and $\Delta\phi = 0.60$. The best-fit parameters for each B -field and geometric model combination are summarised in table 1. The table includes the different model combinations, the optimal χ^2 value (before scaling), the free parameters with errors (found using 3σ connected (α, ζ) contours), a reference fit found using radio polarisation data, and the comparison between models using the difference between the respective optimal values of ξ^2 , represented by $\Delta\xi^2_*$.

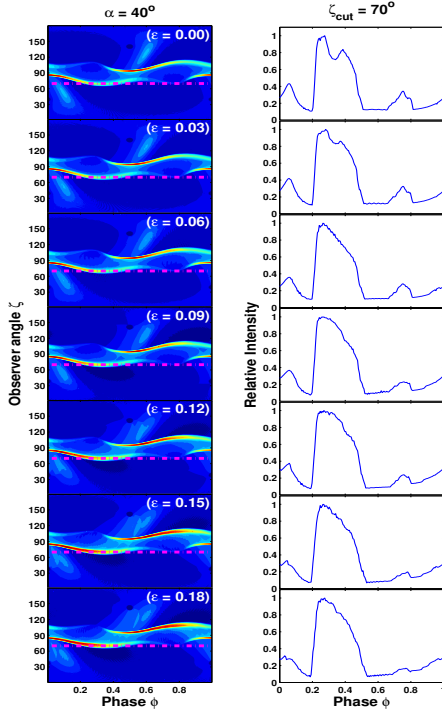


Figure 2. Phaseplots and light curves for different ϵ values, for constant ϵ_ν .

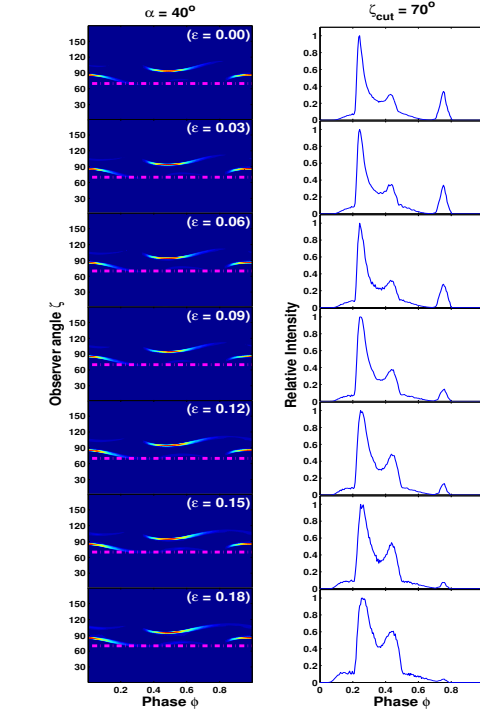


Figure 3. Phaseplots and light curves for different ϵ values, for variable ϵ_ν .

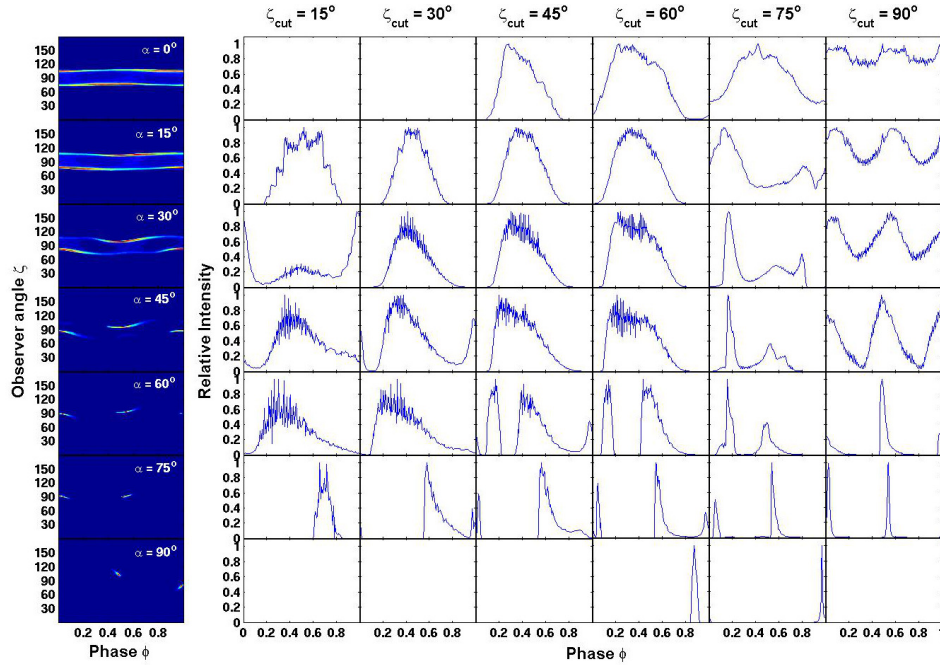


Figure 4. Phaseplots and light curves for different values of α and ζ_{cut} for a fixed value of $\epsilon = 0.18$, for variable ϵ_ν .

6. Conclusions

We have studied the effect of implementing the offset-dipole B -field on γ -ray light curves for the TPC geometry. We find an optimal best-fit for Vela for the RVD B -field and the OG

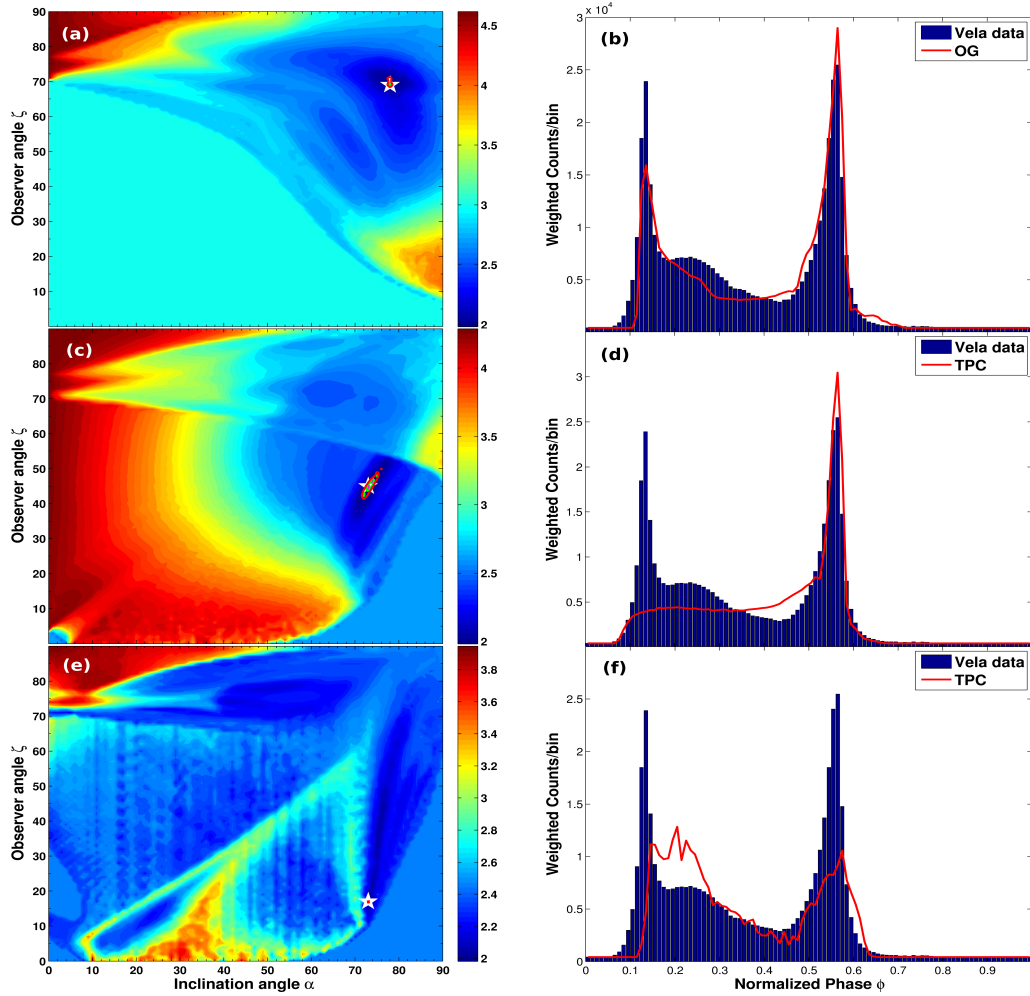


Figure 5. Contour plots (left) and corresponding best-fit light curves (right). Panels (a) and (b) represent the best fit for the RVD and the OG model. Panels (c) and (d) are for the offset dipole and TPC model for constant ϵ_ν and $\epsilon = 0.00$. Panels (e) and (f) are for the offset dipole and TPC model for variable ϵ_ν and $\epsilon = 0.18$. The colour bar of the contour plots represent $\log_{10}\xi^2$. The confidence contour for 1σ (magenta line), 2σ (green line), and 3σ (red line) are also shown. The star indicates the best-fit solution. The blue histogram denotes the observed Vela profile (for energies > 100 MeV) [2] and the red line the geometric model.

model. The OG model displays reduced off-peak emission. We note that the best fits for the offset-dipole B -field for constant ϵ_ν favour smaller values of ϵ and for variable ϵ_ν larger ϵ values. When including an E -field the resulting phaseplots becomes qualitatively different compared to constant ϵ_ν . In future, we want to continue to produce light curves using improved geometric models and B -fields, and also using more data, in order to search for best-fit profiles, thereby constraining the low-altitude magnetic structure and system geometry of several bright pulsars.

Acknowledgments

This work is supported by the South African National Research Foundation. AKH acknowledges the support from the NASA Astrophysics Theory Program. CV, TJJ, and AKH acknowledge support from the *Fermi* Guest Investigator Program.

Table 1. Best-fit parameters for each B -field and geometric model combination.

Model Combinations	ϵ	χ^2 ($\times 10^5$)	Our model parameters				Ref. fit [25]		Radio pol. [18]		$\Delta\chi^2_*$
			α ($^\circ$)	ζ ($^\circ$)	A	$\Delta\phi$	α ($^\circ$)	ζ ($^\circ$)	α ($^\circ$)	ζ ($^\circ$)	
Static dipole:											
TPC	—	0.819	73^{+3}_{-2}	45^{+4}_{-4}	1.3	0.55					0.000
OG	—	0.891	64^{+5}_{-3}	86^{+1}_{-1}	1.3	0.05					8.439
RVD:											
TPC	—	3.278	54^{+5}_{-5}	67^{+3}_{-2}	0.5	0.05	62–68	64			723.5
OG	—	0.384	78^{+1}_{-1}	69^{+2}_{-1}	1.3	0.00	75	64			0.000
Offset dipole - constant ϵ_ν:											
TPC	0.00	0.819	73^{+3}_{-2}	45^{+4}_{-4}	1.3	0.55					0.000
	0.03	0.834	73^{+2}_{-2}	43^{+4}_{-5}	1.3	0.55					1.758
	0.06	0.867	73^{+2}_{-2}	42^{+5}_{-5}	1.3	0.55					5.626
	0.09	0.882	73^{+1}_{-2}	41^{+3}_{-5}	1.3	0.55					7.385
	0.12	1.00	74^{+1}_{-3}	42^{+3}_{-6}	1.4	0.55					21.216
	0.15	0.948	73^{+1}_{-2}	39^{+3}_{-5}	1.4	0.55					15.121
	0.18	0.969	73^{+2}_{-3}	37^{+4}_{-4}	1.3	0.55					17.582
Offset dipole - variable ϵ_ν:											
TPC	0.00	1.46	21^{+2}_{-1}	71^{+1}_{-1}	0.5	0.85					17.032
	0.03	1.63	22^{+3}_{-2}	71^{+1}_{-1}	0.5	0.85					30.194
	0.06	1.68	73^{+1}_{-1}	16^{+1}_{-3}	0.6	0.55					34.065
	0.09	1.65	73^{+1}_{-1}	15^{+1}_{-1}	0.6	0.55					31.742
	0.12	1.59	73^{+1}_{-1}	14^{+2}_{-1}	0.7	0.55					27.097
	0.15	1.52	73^{+1}_{-1}	15^{+3}_{-2}	0.5	0.60					21.677
	0.18	1.24	73^{+1}_{-1}	17^{+1}_{-1}	0.5	0.60					0.000
									53	59.5	

References

- [1] Abdo A A *et al.* 2010 *ApJS* **187** 460–94
- [2] Abdo A A *et al.* 2013 *ApJS* **208** 1–59
- [3] Atwood W B *et al.* 2009 *ApJ* **697** 1071–102
- [4] Becker W, Gil J A, and Rudak B 2007 *Highlights of Astronomy* **14** 109–38
- [5] Breed M, Venter C, Harding A K and Johnson T J 2013 *58th Ann. Conf. of the SA Inst. of Phys.* in press
- [6] Cheng K S, Ho C and Ruderman M 1986 *ApJ* **300** 500–39
- [7] Daugherty J K and Harding A K 1996 *ApJ* **458** 278–92
- [8] Deutsch A J 1955 *Annales d’Astrophysique* **18** 1–10
- [9] Dyks J and Rudak B 2003 *ApJ* **598** 1201–6
- [10] Dyks J, Harding A K and Rudak B 2004 *ApJ* **606** 1125–42
- [11] Espinoza C M *et al.* 2013 *MNRAS* **430** 571–87
- [12] Goldreich P and Julian W H 1969 *ApJ* **157** 869–80
- [13] Griffiths D J 1995 *Introduction to Electrodynamics* (San Francisco: Pearson Benjamin Cummings) p 246
- [14] Harding, A K 2004 *22nd Texas Symp. on Relativistic Astrophysics* ed P Chen, E Bloom *et al.* p 40
- [15] Harding A K, Usov V V and Muslimov A G 2005 *ApJ* **622** 531–43
- [16] Harding A K and Muslimov A G 2011 *ApJ* **743** 181–96
- [17] Hewish A, Bell S J, Pilkington J D H, Scott P F and Collins R A 1968 *Nature* **217** 709–13
- [18] Johnston S 2005 *MNRAS* **364** 1397–412
- [19] Muslimov A G and Harding A K 2003 *ApJ* **588** 430–40
- [20] Muslimov A G and Harding A K 2004 *ApJ* **606** 1143–53
- [21] Pierbattista M *et al.* 2014 *Astronomy & Astrophysics* in press
- [22] Romani R W 1996 *ApJ* **470** 469–78
- [23] Sturmer S J 1995 *ApJ* **446** 292–99
- [24] Venter C, Harding A K and Guillemot L 2009 *ApJ* **707** 800–22
- [25] Watters K P, Romani R W, Weltevrede P and Johnston S 2009 *ApJ* **695** 1289–301

Spectral studies of Flat Spectrum Radio Quasars: constraining models at GeV energies

Richard J G Britto^{1,*}, Soebur Razzaque¹ and Benoit Lott², on behalf of the Fermi-LAT Collaboration

¹ Department of Physics, University of Johannesburg, APK Campus, Auckland Park 2006, South Africa

² Centre d'Études Nucléaires de Bordeaux-Gradignan, Université Bordeaux 1, IN2P3/CNRS, 33175 Gradignan, France

E-mail: * rbritto@uj.ac.za

Abstract. Active Galactic Nuclei (AGNs) at the centres of distant galaxies are understood to be powered by the gravitational energy of supermassive black holes accreting matter from their surroundings and releasing extraordinary amounts of energy through two symmetric jets of ultrarelativistic particles and photons. Blazars are a class of AGNs with one jet directed towards us. A subclass of them are called Flat Spectrum Radio Quasars (FSRQ), which are among the brightest AGNs and have been detected up to redshift $z > 3$. We have analysed 5.5 years of data (2008-2014) from the Large Area Telescope (LAT) of the Fermi gamma-ray space observatory and performed spectral analysis of bright FSRQs in the 100 MeV-200 GeV energy range. Our study focuses on the spectral modification due to $\gamma\gamma$ absorption of $> \text{GeV}$ photons interacting with $\text{Ly}\alpha$, OIV and $\text{Ly}\beta$ photons in the broad line region (BLR). As the sensitivity of this analysis does not constrain the model very much, we can only rule out a very large absorption feature in the gamma-ray spectrum, indicating that the gamma ray emission region would not be located deep within the BLR.

1. Introduction

Blazars, a class of AGN, are the most luminous objects of the Universe, and release a tremendous amount of energy through a pair of jets of plasma ejected at relativistic velocities. Blazar jets produce radiation across the whole electromagnetic spectrum, from radio waves to γ rays. They exhibit a characteristic double-bump spectral energy distribution (SED). The dominant radiation production process is understood to be synchrotron emission for the visible-UV range (first bump), whereas the X-ray/gamma-ray band (second bump) can be modeled by both leptonic and hadronic processes, although the inverse Compton process is most often used in modeling.

Flat spectrum radio quasars (FSRQs), a sub-class of blazars, are generally more luminous and at higher redshift than the BL Lacs which constitute the other subclass of blazars. A specific characteristic of FSRQs is a broad-line emission spectrum, that suggests the presence of clouds and intense radiation fields at a relatively close distance from the supermassive black hole. γ rays can be produced within the jet or/and by the interaction of the particles from the jets with visible and UV photons emitted in the broad-line region (BLR). If γ rays are produced within the BLR, it is expected that they would also undergo $\gamma\gamma \rightarrow e^-e^+$ interactions with BLR photons [10]. The radiation spectrum of the BLR would suggest that this absorption effect

could be measured in the 10-100 GeV band. The Large Area Telescope (LAT) onboard the Fermi Gamma Ray Space Observatory is sensitive to γ rays between 20 MeV and 300 GeV [3], and is the appropriate instrument for studying this effect and constraining the γ -ray emission region within the BLR by fitting Fermi-LAT spectral data. Traditionally, the BLR was modeled as a spherical thick shell-type structure while more modern models assume a flat BLR [11, 8].

We present the 5.5 yr Fermi-LAT data sample and the analysis procedure in Section 2. In Section 3 we present our modeling of the opacity $\tau_{\gamma\gamma}(E, z)$. We present our results and discussion in Section 4 and give perspectives on further studies.

2. Data sample and analysis

We selected a sample of 7 bright FSRQs from the upcoming “Third LAT AGN catalog” (3LAC): PKS 1510-089 (redshift $z=0.360$), 4C +21.35 (also known as PKS 1222+216, $z=0.434$), 3C 279 ($z=0.536$), 3C 454.3 ($z=0.859$), 4C +55.17 ($z=0.899$), PKS 0454-234 ($z=1.003$), PKS 1502+106 ($z=1.839$). We derived the BLR luminosity for each source, using [9, 13]. We displayed these values in units of $10^{44} \text{ erg s}^{-1}$, in the order of the source list: 5.62, 15.80, 3.10, 33.00, 3.80, 3.70 and 17.50. For each source, we analysed about 5.5 years of data (from August 2008 up to at least January 2014). PKS 1510-089 was also studied during its Feb-Apr 2012 flare, and showed along with the data taken by MAGIC during the same period [2]. The brightest flare ever detected in gamma was seen from 3C 454.3 in Sep-Dec 2010. As flaring activity affects the shape of the SEDs, we studied this source separately during its Sep-Dec 2010 giant flare [1], during the first two years of data before the flare, and during the last 3 years of data after the flare when it was mostly quiescent.

Our data analysis was performed using the public release of the Fermi Science Tools *v9r32p5-fssc-20130916*¹ and the instrument response function *PASS7_V15_SOURCE*. For each data sample, we considered all the photons from a region of 10 degree radius around the source of interest, and we estimated the signal contribution from other sources in a 20 degrees field of view, using an unbinned likelihood algorithm (*gtlike* science tool).

3. Modeling of the opacity of the BLR

We assume a model for the BLR luminosity and radius as $L_{BLR} = 0.1L_{Disc}$ and $R_{BLR} = 10^{17}(L_{disc}/10^{45} \text{ erg/s})^{0.5} \text{ cm}$, respectively, where L_{Disc} is the accretion disc luminosity [4, 6]. We also assume that BLR luminosity is dominated by Ly α and OIV+Ly β lines with relative flux ratio 1 : 0.18 [12]. We model the Ly α and OIV+Ly β lines as Breit-Wigner functions with peak energies $\epsilon_1 = 10.2 \text{ eV}$ and $\epsilon_2 = 12.04 \text{ eV}$, with widths $\omega_1 = 0.8 \text{ eV}$ and $\omega_2 = 0.2 \text{ eV}$, respectively. The $\gamma\gamma \rightarrow e^-e^+$ process has a threshold $E \approx 25.6/(1+z) \text{ GeV}$ for Ly α photons, where E is the γ -ray energy.

The differential opacity per unit distance for γ rays passing through the BLR will then be written as:

$$\frac{d\tau_{\gamma\gamma}(E, z)}{dx} = \frac{r_0^2}{2} \left[\frac{m^2 c^4}{E(1+z)} \right]^2 \sum_{i=1}^2 \left(n_i \omega_i \int_{\frac{m^2 c^4}{E(1+z)}}^{\infty} \frac{\bar{\varphi} \left[\frac{\epsilon E(1+z)}{m^2 c^4} \right] d\epsilon}{[(\epsilon - \epsilon_i)^2 + (\omega_i/2)^2] \epsilon^2} \right)$$

where $\bar{\varphi}$ is defined in [7, 5]. Here $n_i \simeq 1.66 \times 10^{11} \left(\frac{L_{i,45}}{\epsilon_{i,eV} R_{BLR,17}^2} \text{ cm}^{-3} \right)$, with $L_1 = L_{BLR}/(1+f)$ for Ly α photons and $L_2 = f L_{BLR}/(1+f)$ for OIV+Ly β photons, where $f = 0.18$.

The integrated opacity we obtain is $\tau_{\gamma\gamma}(E, z) = R_{BLR} \times \frac{d\tau_{\gamma\gamma}}{dx}(E, z)$.

¹ <http://fermi.gsfc.nasa.gov/ssc/data/analysis/>

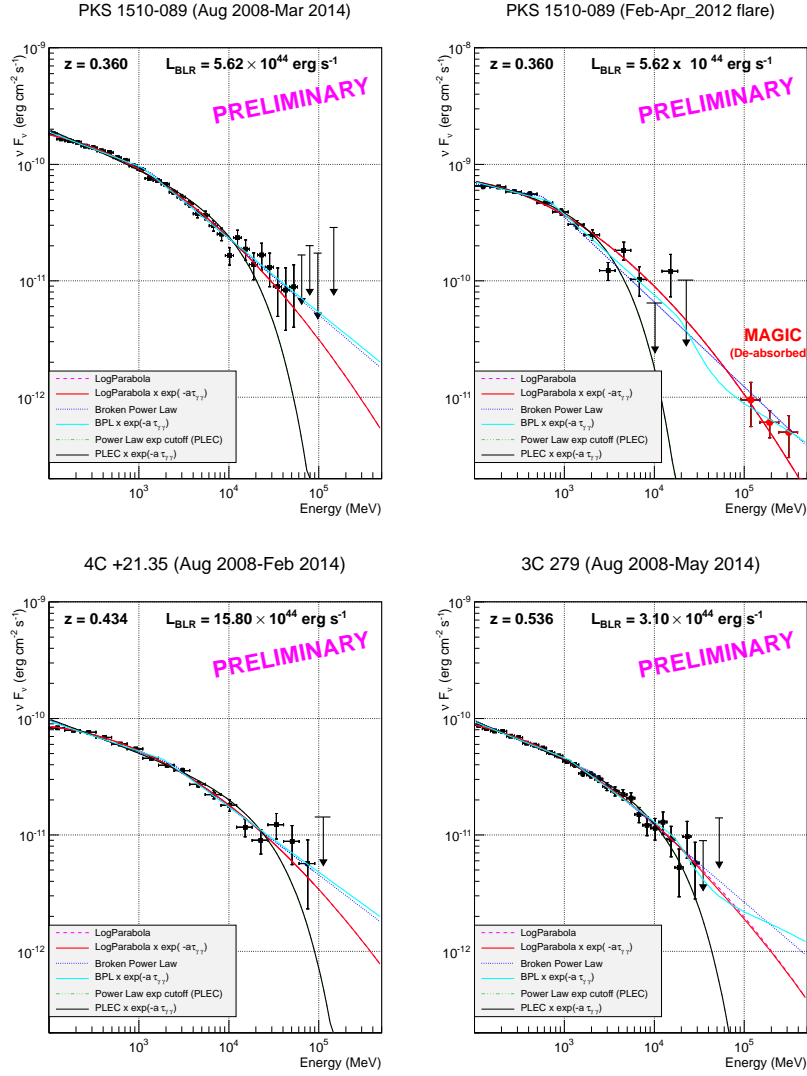


Figure 1. νF_ν SEDs of PKS 1510-089 (all 5.5 yr of data and flare), PKS 1222+216 (4C +21.35) and 3C 279. *PLEC* and *LP* fits are hidden beneath *PLEC* τ and *BPL* τ .

4. Preliminary results and discussion

We produced the SEDs of the 7 bright sources we presented earlier between 100 MeV up to a few tens of GeV. In figure 1 are presented the three TeV sources PKS 1510-089, 4C +21.35 (also called PKS 1222+216), and 3C 279. PKS 1510-089 is also presented during its Feb-Apr 2012 flare, when data were taken by the MAGIC gamma-ray telescope. In figure 2 (top) is presented 3C 454.3, during three different periods that represent three different states: the first two years of data (Aug 2008-Aug 2010), the giant flare of Autumn 2010, the last three years of data (till January 2014). Before the giant flare, 3C 454.3 was often in a high activity state, whereas after December 2010, flaring episodes were rare and did not last for a long period. In figure 2 (bottom) are presented three additional sources, each one at a redshift $z \geq 0.9$: 4C +55.17, PKS 054-234 and PKS 1502+106. The LAT data points are not corrected for the absorption of extragalactic background radiation (*i.e.* not deabsorbed).

Most of the FSRQ SEDs are usually well fitted using at least one of the three following

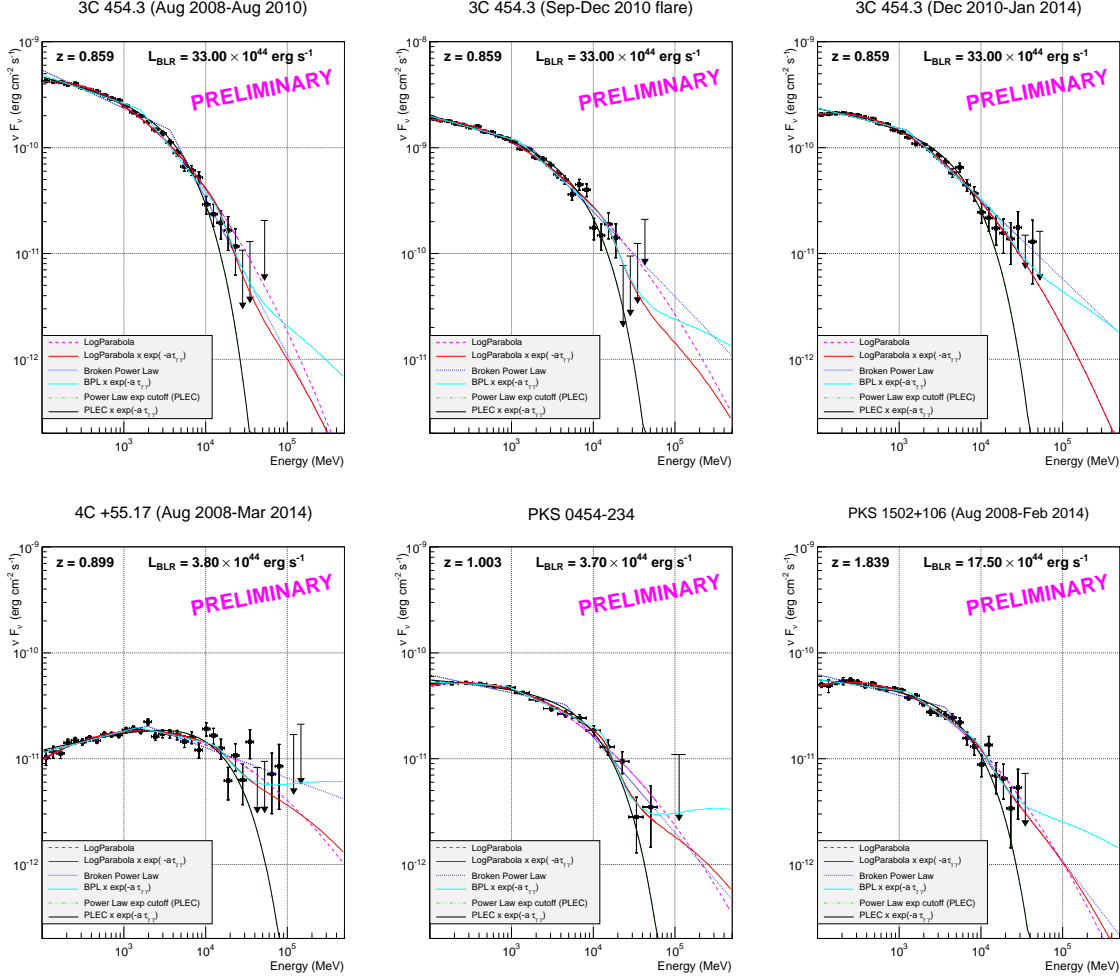


Figure 2. νF_ν SEDs of 3C 454.3 during three observation periods and for the three higher redshift sources. *PLEC* and *LP* fits might be hidden beneath *PLEC* τ and *BPL* τ .

functions in the MeV-GeV range: a *LogParabola* (*LP*), defined in this paper as $N(E) = N_0 (E/1000 \text{ MeV})^{-\alpha-\beta \log(E/1000 \text{ MeV})}$; a *broken power law* (*BPL*), defined as $N(E) = N_0 (E/E_b)^{-\Gamma_i}$, with $i = 1$ if $E < E_b$ and $i = 2$ if $E > E_b$; a *power law with an exponential cutoff* (*PLEC*), defined as $N(E) = N_0 (E/1000 \text{ MeV})^{-\Gamma} \exp(-E/E_c)$. We use these functions to fit the SEDs presented in this paper from 100 MeV up to the last data point preceding an upper limit².

We also performed another set of three fits, by a convolution of each of the previous functions with $\exp(-a \tau_{\gamma\gamma}(E, z))$, where a is an additional parameter that represents the fraction of the BLR in which the gamma-ray photons are traveling, assuming the hypothesis of the previous section. In this case, the observed flux would be $F_{obs}(E) = e^{-\tau_{\gamma\gamma}(E, z)} F_{int}(E)$, and these new fit functions will be referred hereafter as *LP* τ , *BPL* τ and *PLEC* τ .

We show the fit parameters of all six fits of each SED in tables 1, 2, and 3, corresponding to the plots of figure 1 and 2. When we consider the fits performed with the *LP*, *BPL* and *PLEC*

² except for the PKS 1510-089 flare SED because we added the MAGIC points, and also considered the LAT point lying between two upper limits (see Fig. 1)

Table 1. Fitting parameters and derived significances, for the three TeV sources. We indicate with (*) when a reached the lower edge of the interval for fitting.

Param	Model	PKS 1510-089	PKS 1510-089 flare	4C +21.35	3C 279
functions (no absorption)	LP α	2.442 ± 0.010	2.450 ± 0.018	2.337 ± 0.014	2.424 ± 0.013
	LP β	0.063 ± 0.006	0.069 ± 0.008	0.056 ± 0.010	0.055 ± 0.008
	BPL Γ_1	2.321 ± 0.015	2.146 ± 0.050	2.278 ± 0.026	2.341 ± 0.016
	BPL Γ_2	2.655 ± 0.041	2.728 ± 0.026	2.587 ± 0.080	2.662 ± 0.076
	BPL E_{break}	1392 ± 226.9	570 ± 60	2465 ± 896.0	2088 ± 506.9
	PLEC Γ	2.325 ± 0.019	2.155 ± 0.049	2.283 ± 0.032	2.311 ± 0.020
	PLEC E_{cutoff}	16023 ± 2966.8	3324 ± 760	33678 ± 11941	16369 ± 3490.4
	PLEC α	2.442 ± 0.010	2.450 ± 0.018	2.337 ± 0.014	2.423 ± 0.014
functions (absorption)	LP β	0.063 ± 0.006	0.069 ± 0.008	0.056 ± 0.010	0.054 ± 0.009
	BPL Γ_1	2.295 ± 0.016	2.199 ± 0.036	2.257 ± 0.028	2.309 ± 0.020
	BPL Γ_2	2.630 ± 0.042	2.691 ± 0.076	2.560 ± 0.065	2.561 ± 0.053
	BPL E_{break}	1092 ± 180.0	684 ± 111	1862 ± 614.8	1128 ± 274.6
	PLEC Γ	2.325 ± 0.019	2.155 ± 0.049	2.284 ± 0.032	2.311 ± 0.020
	PLEC E_{cutoff}	16022 ± 2967.3	3324 ± 760	33695 ± 11961	16370 ± 3491.6
	LP α	0.00001^*	0.00010^*	0.00001^*	0.0014 ± 0.066
	BPL β	0.0003 ± 0.037	0.00010^*	0.00001^*	0.0143 ± 0.020
a	PLEC	0.00001^*	0.01157 ± 0.01163	0.00001^*	0.00001^*
	LP	14.700 (28)	23.67 (12)	4.894 (14)	7.306 (25)
	BPL	23.265 (27)	16.75 (11)	8.759 (13)	15.286 (24)
χ^2 (ndf) (no absorption)	PLEC	42.140 (28)	51.90 (12)	15.093 (14)	13.670 (25)
	LP	14.702 (27)	23.69 (11)	4.898 (13)	7.301 (24)
	BPL	16.377 (26)	13.80 (10)	8.025 (12)	11.127 (23)
χ^2 (ndf) (absorption)	PLEC	42.144 (27)	51.90 (11)	15.098 (13)	13.671 (24)
	LP	$2.669\text{e-}03$	0.0276	$4.146\text{e-}03$	$5.168\text{e-}03$
	BPL	$6.888\text{e+}00$	2.945	$7.338\text{e-}01$	$4.159\text{e+}00$
$\Delta\chi^2$	PLEC	$3.737\text{e-}03$	0.000041	$4.342\text{e-}03$	$9.954\text{e-}04$
	LP	$9.588\text{e-}01$	$8.680\text{e-}01$	$9.487\text{e-}01$	$9.427\text{e-}01$
	BPL	$8.680\text{e-}03$	$8.614\text{e-}02$	$3.917\text{e-}01$	$4.141\text{e-}02$
p-value	PLEC	$9.513\text{e-}01$	$9.949\text{e-}01$	$9.475\text{e-}01$	$9.748\text{e-}01$

Table 2. Fitting parameters and derived significances for the three periods of 3C 454.3.

Param	Model	3C 454.3 (2008-2010)	3C 454.3 (Sep-Dec 2010 flare)	3C 454.3 (2010-2014)
functions (no absorption)	LP α	2.509 ± 0.012	2.415 ± 0.012	2.412 ± 0.012
	LP β	0.121 ± 0.006	0.086 ± 0.006	0.112 ± 0.007
	BPL Γ_1	2.360 ± 0.013	2.249 ± 0.013	2.189 ± 0.014
	BPL Γ_2	3.478 ± 0.093	2.801 ± 0.073	2.739 ± 0.050
	BPL E_{break}	3650 ± 0.4	1749 ± 241	1252 ± 139
	PLEC Γ	2.183 ± 0.019	2.184 ± 0.017	2.149 ± 0.021
	PLEC E_{cutoff}	5144 ± 472	7221 ± 811.9	6412 ± 762
	PLEC α	2.501 ± 0.013	2.410 ± 0.014	2.412 ± 0.012
functions (absorption)	LP β	0.118 ± 0.007	0.084 ± 0.007	0.112 ± 0.007
	BPL Γ_1	2.270 ± 0.013	2.218 ± 0.014	2.189 ± 0.014
	BPL Γ_2	2.968 ± 0.077	2.693 ± 0.062	2.707 ± 0.058
	BPL E_{break}	1567 ± 174	1245 ± 164	1200 ± 148
	PLEC Γ	2.183 ± 0.019	2.184 ± 0.017	2.149 ± 0.021
	PLEC E_{cutoff}	5143 ± 472	7222 ± 812	6412 ± 761
	LP	0.007 ± 0.007	0.007 ± 0.012	0.00001^*
	BPL	0.007 ± 0.007	0.008 ± 0.013	0.004 ± 0.005
a	PLEC	0.00001^*	0.00001^*	0.00001^*
	LP	16.598 (24)	22.034 (23)	9.693 (25)
	BPL	85.579 (23)	34.690 (22)	51.369 (24)
χ^2 (ndf) (no absorption)	PLEC	32.384 (24)	25.094 (23)	52.935 (25)
	LP	15.080 (23)	21.541 (22)	9.695 (24)
	BPL	41.220 (23)	26.682 (21)	50.514 (23)
χ^2 (ndf) (absorption)	PLEC	32.388 (23)	25.096 (22)	52.940 (24)
	LP	$1.518\text{e+}00$	$4.928\text{e-}01$	$2.048\text{e-}03$
	BPL	$4.436\text{e+}01$	$8.008\text{e+}00$	$8.547\text{e-}01$
$\Delta\chi^2$	PLEC	$3.230\text{e-}03$	$2.377\text{e-}03$	$4.397\text{e-}03$
	LP	$2.180\text{e-}01$	$4.827\text{e-}01$	$9.639\text{e-}01$
	BPL	$2.733\text{e-}11$	$4.658\text{e-}03$	$3.552\text{e-}01$
p-value	PLEC	$9.547\text{e-}01$	$9.611\text{e-}01$	$9.471\text{e-}01$

functions, the χ^2/ndf of LP is good for all the sources (~ 1). It gives better fits than BPL and PLEC, except in the case of 4C +55.17 where χ^2/ndf of BPL is similar to the LP one. Adding the extra parameter a by using the $LP\tau$, $BPL\tau$ and $PLEC\tau$ containing the opacity $\tau_{\gamma\gamma}$ allows better fits in some cases. Best fit is obtained by optimising a within a search interval between 0.00001 and 1 in our fitting programs, as a representation of the fraction of the BLR in which the γ -ray absorption may occur. In many cases, a takes the value 0.00001 ('at limit'), and in those case the corresponding fit is matching the fit without the $\exp(-a \tau_{\gamma\gamma})$ factor (fitting line

Table 3. Fitting parameters and derived significances for the three high redshift sources.

Param	Model		4C +55.17	PKS 0454-234	PKS 1502+106
func (no absorption)	LP	α	1.925 ± 0.016	2.233 ± 0.015	2.300 ± 0.014
	LP	β	0.087 ± 0.011	0.089 ± 0.009	0.110 ± 0.009
	BPL	Γ_1	1.810 ± 0.027	2.168 ± 0.018	2.197 ± 0.014
	BPL	Γ_2	2.286 ± 0.049	2.903 ± 0.099	3.011 ± 0.152
	BPL	E_{break}	1931 ± 250	4569 ± 5	3664 ± 572
	PLEC	Γ	1.815 ± 0.024	2.063 ± 0.025	2.066 ± 0.023
	PLEC	E_{cutoff}	14831 ± 2354.0	11658 ± 2141.7	7349 ± 913.0
func (absorption)	LP	α	1.921 ± 0.016	2.212 ± 0.016	2.288 ± 0.017
	LP	β	0.078 ± 0.013	0.072 ± 0.011	0.101 ± 0.012
	BPL	Γ_1	1.794 ± 0.031	2.042 ± 0.023	2.104 ± 0.018
	BPL	Γ_2	2.169 ± 0.055	2.363 ± 0.046	2.547 ± 0.074
	BPL	E_{break}	1476 ± 245	880 ± 173	1317 ± 206
	PLEC	Γ	1.815 ± 0.024	2.064 ± 0.025	2.066 ± 0.023
	PLEC	E_{cutoff}	14833 ± 2354	11660 ± 2140	7350 ± 914
a	LP		0.009 ± 0.009	0.020 ± 0.009	0.004 ± 0.003
	BPL		0.015 ± 0.008	0.031 ± 0.009	0.007 ± 0.003
	PLEC		0.00001*	0.00001*	0.00001*
χ^2 (ndf) (no absorption)	LP		33.211 (26)	7.993 (13)	16.056 (25)
	BPL		31.252 (25)	35.170 (12)	69.639 (24)
	PLEC		42.552 (26)	16.315 (13)	32.431 (25)
χ^2 (ndf) (absorption)	LP		32.134 (25)	3.451 (12)	14.622 (24)
	BPL		29.304 (24)	3.777 (11)	26.473 (23)
	PLEC		42.554 (25)	16.317 (12)	32.440 (24)
$\Delta\chi^2$	LP		1.078e+00	4.542e+00	1.434e+00
	BPL		1.948e+00	3.139e+01	4.317e+01
	PLEC		2.342e-03	2.003e-03	9.256e-03
p-value	LP		2.993e-01	3.307e-02	2.311e-01
	BPL		1.628e-01	2.107e-08	5.030e-11
	PLEC		9.614e-01	9.643e-01	9.234e-01

superimposed on the figure and hiding the previous line). We estimate that a has large errors due to large error bars in the $> 10 \text{ GeV}$ regime where the value of a become relevant in our studies. Note that fitting with the $\exp(-a \tau_{\gamma\gamma})$ factor leads to improvements in the χ^2/ndf values in many cases. However, while comparing each $LP\tau$, $BPL\tau$ and $PLEC\tau$ model with the corresponding model (LP, BPL and PLEC respectively), and if both fits have a $\chi^2/ndf \leq 1$, we obtain a p -value which indicates no possible preference to be given to the model with BLR opacity (always below 5σ). Only the higher redshift sources PKS 0454-234 and PKS 1502+106 have a parameter a with a error bar less than half the parameter value for $BPL\tau$ and a relatively small χ^2/ndf . Though we cannot compare these $BPL\tau$ fits with their corresponding BPL fits which have a bad χ^2/ndf , we can still consider this result as showing a potential absorption that should be confirmed by more statistics. PKS 0454-234 has also good χ^2/ndf 's for both LP and $LP\tau$ fit, and $a = 0.02 \pm 0.01$. We would prefer this model to the $LP\tau$ that implies a very hard spectrum beyond 50 GeV.

We are planning to study more sources and to produce more sub-data sets showing flaring activities to continue this study. 3C 454.3 is a good candidate to pursue quiescent *versus* flaring state comparison, as we were starting to do. Still we see no evidence of a strong γ -ray absorption, which indicates that the production of γ rays should not happen deep within the BLR, where the optical/UV photon flux density is higher.

Acknowledgments

The *Fermi*-LAT Collaboration acknowledges support for LAT development, operation and data analysis from NASA and DOE (United States), CEA/Irfu and IN2P3/CNRS (France), ASI and INFN (Italy), MEXT, KEK, and JAXA (Japan), and the K.A. Wallenberg Foundation, the Swedish Research Council and the National Space Board (Sweden). Science analysis support in the operations phase from INAF (Italy) and CNES (France) is also gratefully acknowledged.

References

- [1] Abdo A A et al 2011 ApJ L **733**:L26
- [2] Aleksic J et al 2014 A&A 569 A46

- [3] Atwood W B et al 2009 ApJ **697** 1071
- [4] Baldwin J A and Netzer H 1978 ApJ **226** 1-20
- [5] Brown R W Mikaelian K O and Gould R J 1973 ApJ L **14** L203-L205
- [6] Ghisellini G and Tavecchio F 2009 MNRAS **397** 985
- [7] Gould R J and Schröder G P 1967 Physical Review **155** 5
- [8] Lei M and Wang J 2014 Publications of the Astronomical Society of Japan **66** 1 id.7 (arXiv:1308.4210v1)
- [9] Pacciani L et al 2014 arXiv:1312.3998v2
- [10] Poutanen J and Stern B 2010 ApJ L **717**:L118
- [11] Tavecchio F and Ghisellini G 2012 submitted to MNRAS (arXiv:1209.2291v1)
- [12] Telfer R C Zheng W Kriss G A and Davidsen A F 2002 ApJ **565**:773-785
- [13] Xiong D R and Zhang X 2014 MNRAS **441** 4 p.3375-3395

DFT, LS or GLS? A ‘road test’ of three period-finding techniques applied to stellar photometry

C A Engelbrecht^{1,3} & F A M Frescura²

¹Department of Physics, University of Johannesburg, Kingsway, Auckland Park, Johannesburg 2092, South Africa

²School of Physics, University of the Witwatersrand, Private Bag 3, WITS 2050, Johannesburg, South Africa

E-mail: chrise@uj.ac.za

Abstract. The mathematical properties of harmonic functions have made Fourier-based algorithms very popular in searches for periodic behaviour in stellar light curves and other astronomical data. The Discrete Fourier Transform and the Lomb-Scargle periodogram have been in common use for many decades, as methods particularly suitable for the non-equally-spaced time data that are typical of astronomical measurements. The “Generalised Lomb-Scargle” or GLS method is a recent refinement of the Lomb-Scargle method. Recent tests of these methods have emphasised the strengths and deficiencies of each. These test results demonstrate that certain methods are distinctly unsuitable for certain types of data and could lead to erroneous conclusions about the types of periodic behaviour present in measured data. Given the rapidly growing body of time-domain data in astronomy and the considerable importance of some of the conclusions that have been made on the basis of these data, the recent developments in the study of period-finding algorithms are significant. This paper reports on a comparative study of the above-mentioned methods applied to simulated stellar photometry compiled for three observing scenarios: space-based observation, ground-based observer-driven campaigns, and ground-based survey programs.

1. Introduction

In recent years, asteroseismology has made enormous contributions to the depth and scope of our understanding of stellar structure and evolution. For example, a very powerful correlation between the spacing of g-modes in the frequency domain and the dominant nuclear energy source (i.e. helium-core or hydrogen-shell fusion) in red giants was reported in [1]. Accurate determination of pulsation frequencies, amplitudes and phases is an essential component of such breakthroughs in asteroseismology. Two widely applied ‘workhorses’ of period determination are the Discrete Fourier Transform (DFT) adapted for non-equally-spaced time series, and the Lomb-Scargle (LS) periodogram. A number of diverse new approaches for determining the features of periodic behaviour in stellar magnitudes have been reported in recent years. One such approach is the “Generalised Lomb-Scargle” (or GLS) method [2]. This reference will be referred to as ZK09 in what follows. Following on earlier work by [3], who introduced a floating mean into the LS periodogram calculation, ZK09 added a weighting procedure to the floating mean procedure to accommodate observations that do not all have exactly the same precision. They concluded that, compared with the

³ To whom any correspondence should be addressed.

classical LS periodogram, the GLS periodogram provides a more accurate frequency determination, is less susceptible to aliasing, and gives a much better determination of the spectral intensity. A recent, comprehensive comparison of various period-finding methods by [4] provides strong endorsement for the accuracy of the GLS algorithm.

2. Simulated light curves

In order to explore the relative merits of the DFT, LS and GLS methods, three different observing scenarios were simulated. The following parameters are common to the three simulated time series: the total time-span of the observations is 110 days and a total of 1000 observations are taken during this time-span. These choices were informed by the authors' research interest in pulsating B stars, for which the chosen values are typical of actual observing scenarios that produce useful asteroseismological results for such stars.

The following three realistic observing scenarios were chosen:

- i) Scenario 1: A strictly equally-spaced time sampling over the 110 days, as would be possible to schedule on a space-based telescope capable of 24-hour observation of the same object. The observing protocol of the original mission design of the *Kepler* space telescope is similar to this scenario (although *Kepler* – in its 'Long-Cadence' observing mode - would obtain approximately 5200 equally-spaced observations over a 110-day period, instead of the chosen value of 1000 in this simulation). With the chosen numbers, the equal time-spacing between successive observations is 0.11 days, or 2.64 hours. This implies a Nyquist frequency of just over 4.5 cycles per day (c/d). All of the calculated periodograms in this study were therefore terminated at a frequency value of 4.5 c/d. This equally-spaced observing scenario is perfectly suited to the DFT method and we expected the DFT periodograms to perform just as well as the LS and GLS periodograms in this scenario.
- ii) Scenario 2: A typical ground-based observing campaign, consisting of five full weeks (typical of the small-telescope observing allocations of the South African Astronomical Observatory at its Sutherland site in the Northern Cape) spread over the 110-day time-span and consisting of varying numbers of photometric nights per week. In the simulation, a total of 17 photometric nights were chosen among the 35 days of allocated observing time. Again, this is typical of actual conditions at Sutherland. The actual sampling times were adapted to account for the advance of sidereal time over the 110 days of calendar time, which accounts for the earlier rising of target stars at the end of the observing campaign than at the beginning. Within a night, the observations were taken at equal time intervals of 7.49 minutes. This number was determined by requiring 1000 observations to be taken in 17 photometric nights. Again, this time-spacing is typical of ground-based, observer-driven observations of pulsating B stars.
- iii) Scenario 3: A ground-based, pre-programmed survey project with no observer input, observing the target star a few times each day (and moving to other survey locations in between). A random component was introduced in selecting the actual times of observation on any particular day. On most nights, the total number of observations taken per night was 8, 9 or 10. The spacing between successive observations was *not equal* and varied by more than an order of magnitude.

The simulated light curves shown in figure 1 illustrate the nature of the sampling regimes for the three scenarios.

Scenarios 1, 2 and 3 determined the *time* vectors in the simulated light curves. The *signal* vectors were calculated as follows: As a first step, three sinusoids with amplitudes, frequencies and phase

differences typical of pulsating B stars were added together. The value of this cumulative ‘clean’ signal was evaluated at the calculated time values obtained in each of the respective scenarios. Secondly, an underlying mean signal level as well as a random error (different for each individual signal value (i.e. for each individual ‘observation’)) were added to these ‘clean’ signal values. The magnitudes of the random errors were chosen to scatter around typical mean values for small-telescope photometry at Sutherland.

The calculated light curves were then fed into the mathematical algorithms for calculating DFT, LS and GLS periodograms and the results were compared. These results are discussed in the following section.

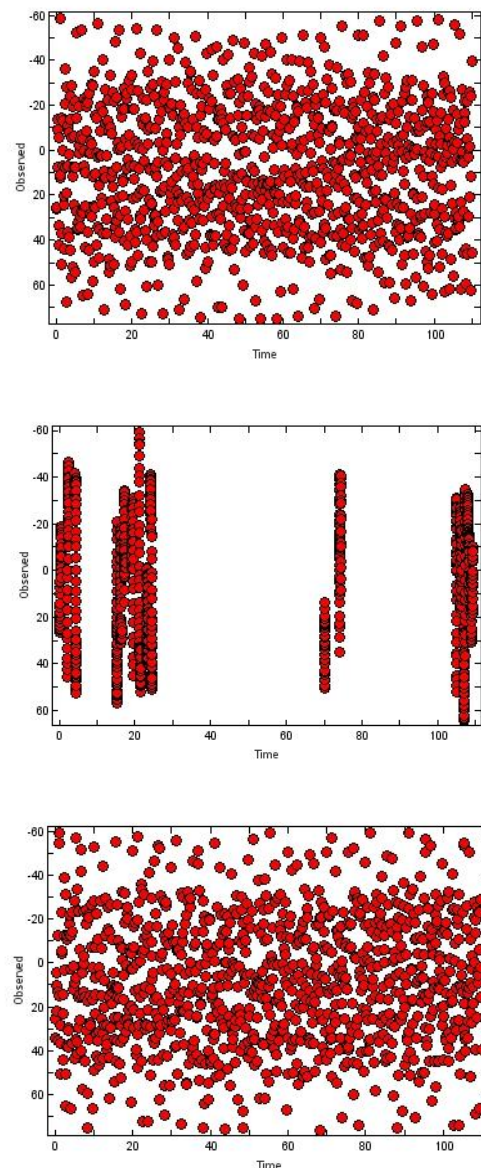


Figure 1. A sample of light curves generated for Scenarios 1, 2 and 3 (from top to bottom) respectively. Since the signal has a random noise component, the light curves will appear slightly differently every time they are calculated.

3. Retrieving periods

The ability of the DFT, LS and GLS periodograms to retrieve the input parameters of the periodic variations present in the simulated light curves was explored for various signal-to-noise (SNR) ratios and differing degrees of homogeneity of the signal errors. A brief summary of results is presented in this section.

Result A: Firstly, we considered typical SNR values for ground-based asteroseismology of B stars and a homogeneous error distribution across the entire observing epoch. As expected, the three methods all succeeded in retrieving the input frequencies and amplitudes of harmonic variation. Initial periodograms (i.e. before prewhitening) for Result A are shown in figure 2.

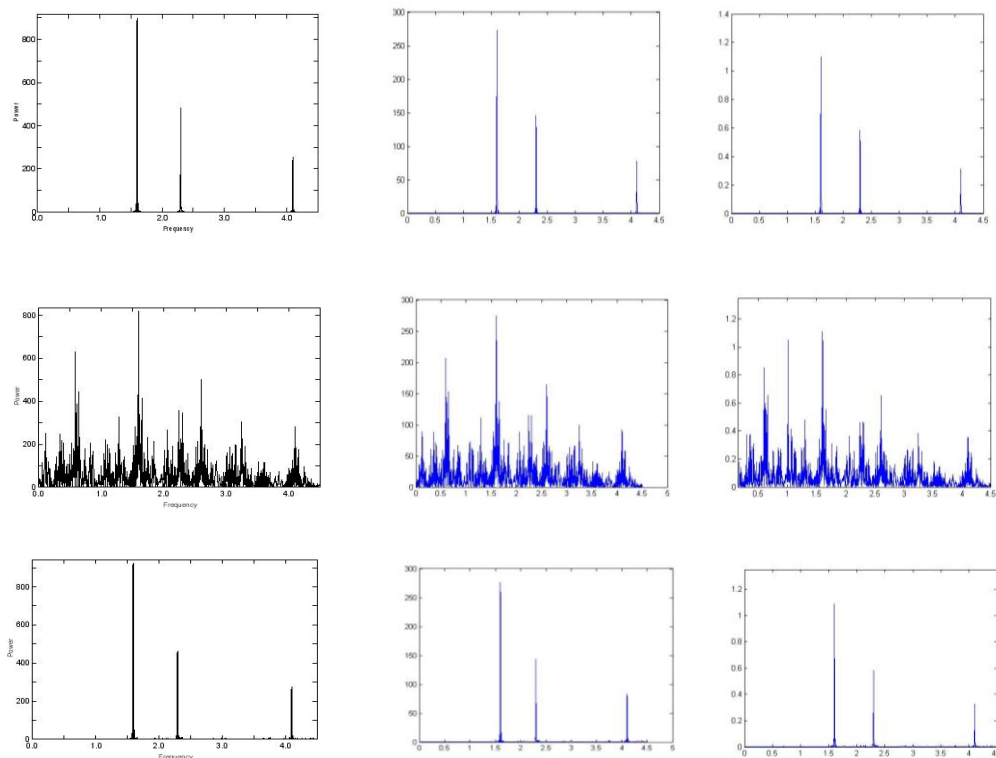


Figure 2. Periodograms obtained with DFT, LS and GLS algorithms respectively, for result A. First row: scenario 1 (equal time spacing); second row: scenario 2 (ground-based observer); third row: scenario 3 (ground-based survey). In each row, the DFT, LS and GLS periodograms appear from left to right respectively. In the second row, the daily gap aliasing is clearly seen. There are (mostly) subtle differences between the periodograms in this row, but all three input frequencies are successfully retrieved. One strange feature is the 1 c/d peak (with no aliases) in the GLS periodogram in the second row.

Result B: As one of the stated advantages of the GLS periodogram is its diminished vulnerability to aliasing, we deliberately constructed a time series that produces severe aliasing in the DFT periodogram by simply removing every second day's data from the equally-spaced time series (scenario 1). The time series for the other two scenarios were also reduced by 50% in a quasi-regular way. Initial periodograms (i.e. before prewhitening) for Result B are shown in figure 3.

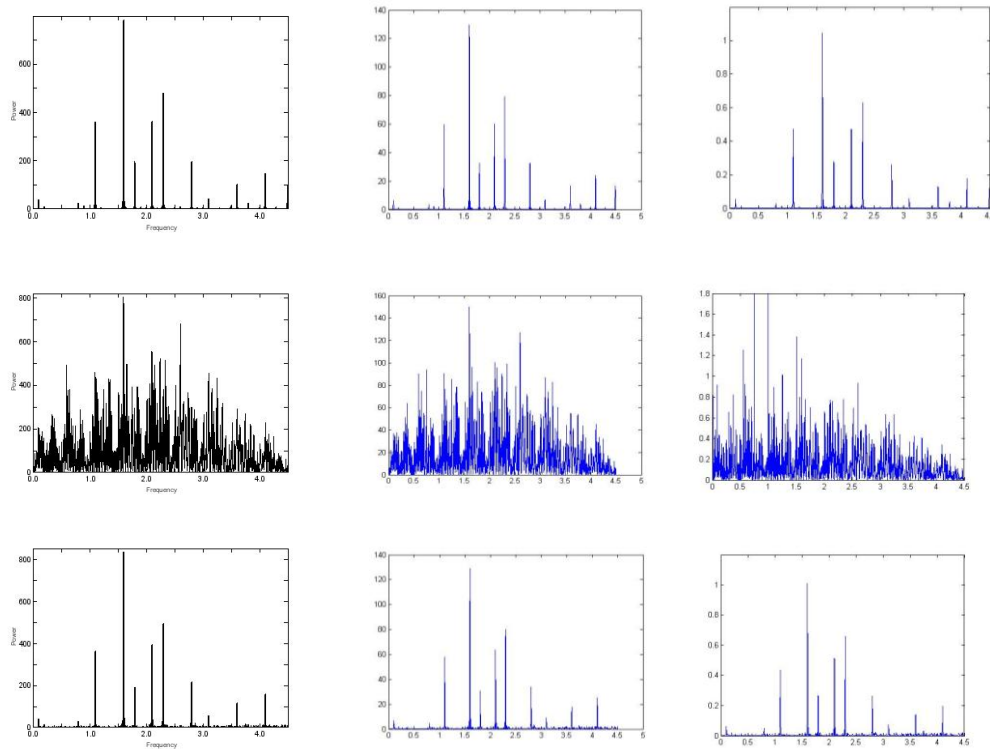


Figure 3. Periodograms obtained with DFT, LS and GLS algorithms respectively, for Result B. First row: scenario 1 (equal time spacing); second row: scenario 2 (ground-based observer); third row: scenario 3 (ground-based survey). In each row, the DFT, LS and GLS periodograms appear from left to right respectively. In the first (DFT) plot in the first row, we do indeed see a lot of aliasing, as expected (eight very prominent peaks appear in the periodogram, whereas only three periodic functions are used to construct the time series). However, the LS and GLS periodograms alongside are equally badly affected. In this instance, the GLS algorithm does not appear to avoid aliasing any better than the DFT and LS periodograms do. The same argument applies to the second and third rows. The GLS periodogram in row 2 actually appears more affected by aliasing than the DFT and LS periodograms are.

Result C: As the GLS periodogram is specifically designed to deal with varying amounts of signal noise, a third experiment was run with the time series divided into three blocks - with the mean noise levels chosen to lie in the ratio 4:2:1 among the blocks. The highest noise level was chosen to be larger than two of the three input amplitudes of the harmonic signal in the simulated time series. Initial periodograms (i.e. before prewhitening) for Result C are shown in figure 4.

4. Discussion and conclusions

The three experiments described as Result A, Result B and Result C in the previous section have not supported the strong claims made in favour of the GLS algorithm by ZK09. For the various time series subjected to the tests reported here, there is no advantage in choosing the increased processing time involved in calculating GLS periodograms, as far as period detection is concerned. In fact, in the experiment which was expected to favour the GLS periodogram (Result C), it was of no practical use. It is accepted that there might be time series formats for which GLS does offer a distinct advantage. The test protocol described in this paper may be adapted to other time series constructs and tested in similar fashion.

It will be of value to the practitioners of asteroseismology to be informed of the relative successes of the methods considered here in other scenarios.

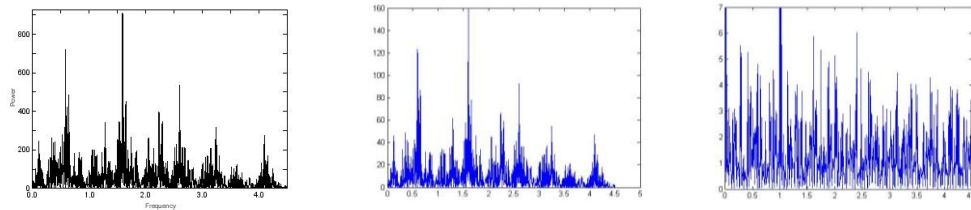


Figure 4. Periodograms obtained with DFT, LS and GLS algorithms respectively, for Result C. Only scenario 2 is shown, as the other scenarios add nothing new. Whereas the DFT and LS periodograms retrieve the input signal amongst some aliasing, the GLS periodogram is effectively of no use in this scenario.

References

- [1] Bedding T S et al. 2011 *Nature* **471** 608
- [2] Zechmeister M. and Kuerster M 2009 *A&A* **496** 577
- [3] Cumming A, Marcy G W and Butler R P 1999 *ApJ* **526** 890
- [4] Graham M J, Drake A J, Djorgovski S G, Mahabal A A, Donalek C, Duan V and Maker A 2013 *MNRAS* **434** 3423

A Different Approach to the Perturbation of Astrophysical Fluids

F A M Frescura¹ and C A Engelbrecht²

¹ School of Physics, U Witwatersrand, 1 Jan Smuts Avenue, Braamfontein, Johannesburg

² Department of Physics, U Johannesburg, Kingsway Ave, Auckland Park, Johannesburg

E-mail: fabio.frescura@wits.ac.za

Abstract. We show that Eulerian and Lagrangian perturbations can be interpreted as finite differences that arise from suitably defined differential operators. These operators lead to exact rather than to approximate perturbation relations and equations. The equations obtained are more general than those usually encountered, and are de facto linear. No approximation is thus needed to linearise them. We also explore the possibility of extending this formalism to the description of stellar pulsations of arbitrary amplitude without using power series expansions.

1. Introduction

Perturbation theory considers two flows, called respectively the ‘unperturbed’ and ‘perturbed’ flows, and uses these to define two variations, called respectively Eulerian and Lagrangian variations. The Eulerian variation describes changes seen by an observer who is at rest in a given frame of reference. The Lagrangian variation describes changes as seen by an observer who moves with the fluid. Both are finite differences. They therefore make perturbation theory approximate at the outset and intrinsically propagate error into all results.

In this paper, I replace these variations by derivatives and show how these may be used to develop a perturbation theory that is free from error.

2. Eulerian and Lagrangian Variations

Let Q be any property of the flow (density, pressure, etc.). Values in the unperturbed flow are subscripted with zero. Values in the perturbed flow are not subscripted.

The Eulerian variation δQ is the difference $\delta Q = Q - Q_0$. Evaluated at \vec{x} and t , it gives the difference in the two flows at the same space point \vec{x} and time t of the property Q .

Laws that govern material behaviour must be applied to a given element of fluid material. Accordingly, the Lagrangian variation ΔQ in Q is defined as the difference at time t in the value of Q in the two flows of one and the same element of material. Suppose that at time t a given fluid element is at position \vec{x} in the unperturbed flow, and at $\vec{x} + \delta\vec{x}$ in the perturbed flow. The vector $\delta\vec{x}$ is called the Lagrangian displacement of the element. The Lagrangian variation of Q is then defined to be the difference

$$\Delta Q = Q(\vec{x} + \delta\vec{x}, t) - Q_0(\vec{x}, t) \quad (1)$$

Unlike the Eulerian variation, which depends only on \vec{x} and t and so is a field on the spacetime occupied by the fluid, the Lagrangian variation is a function also of $\delta\vec{x}$. It is thus a field on the tangent bundle of the spacetime.

Eulerian and Lagrangian variations are not independent. They are related by the equation

$$\Delta Q = \delta Q(\vec{x}, t) + [Q(\vec{x} + \delta\vec{x}, t) - Q(\vec{x}, t)] \quad (2)$$

Correct to first order in $\delta\vec{x}$, this relation becomes

$$\Delta Q \approx \delta Q + \delta\vec{x} \cdot \nabla Q + \dots \quad (3)$$

Extensive use is made of this approximate relation in standard perturbation theory.

Both Eulerian and Lagrangian variations are finite differences. This leads to a number of difficulties, all of which are avoided in the theory presented below.

3. One parameter families of perturbed flows

Problems introduced by the finite-difference nature of Eulerian and Lagrangian variations are avoided by taking their as $|\delta\vec{x}| \rightarrow 0$. This limit cannot be performed if only two flows, perturbed and unperturbed, are considered. We must replace these by a new concept in which the two flows are replaced by a one-parameter family of flows, $\vec{F}(\vec{a}, t, \lambda)$, in which one flow corresponds to each value of the family parameter λ . Each of these flows is characterised by its own set of fluid variables (pressure, density, etc.). Thus, each variable will be a function, not only of \vec{x} and t , but also of the parameter λ . A quantity Q will thus be represented by a function $Q(\vec{x}, t, \lambda)$. The unperturbed flow is represented by parameter value $\lambda = 0$. Non-zero values of λ correspond to perturbed flows. Small values of λ correspond to small perturbations; larger values, to larger perturbations.

The flow $\vec{F}(\vec{a}, t, \lambda)$ is interpreted as follows. Consider the unperturbed flow at time $t = 0$. Each fluid point is at a well defined position \vec{a} in space. We use these coordinates to identify the fluid point at all subsequent times. The values $\vec{a} = (a^1, a^2, a^3)$ are called the Lagrangian coordinates of the material point. As time advances, the fluid point changes its position, which is given at time t by $\vec{F}(\vec{a}, t, 0)$. As t changes, the pathline traced out by this fluid point is given by the function $\vec{F}(\vec{a}, t, 0)$ where \vec{a} is held fixed.

In the perturbed flow with parameter value λ , the fluid point with Lagrangian coordinates \vec{a} occupies position $\vec{F}(\vec{a}, t, \lambda)$ at time t . As t changes, the function $\vec{F}(\vec{a}, t, \lambda)$, where \vec{a} and λ are held fixed, traces out the pathline of this fluid point in the perturbed flow defined by λ .

The set of all possible positions of a given fluid point \vec{a} defines a two dimensional surface whose Gaussian coordinates are t and λ , given by $\vec{F}(\vec{a}, t, \lambda)$. The parameters t and λ define a surface grid in which the $\lambda = \text{constant}$ curves are the pathlines of the fluid point \vec{a} for different perturbed flows, while the $t = \text{constant}$ curves represent the locus of positions assumed by the fluid point \vec{a} at time t as the flow is perturbed or deformed. Interestingly, the function $\vec{F}(\vec{a}, t, \lambda)$ with fixed t is also a flow. It is the flow at time t that deforms the unperturbed fluid into its various perturbed configurations at time t . I will call it the *deformation flow*.

I will now relate this formalism to that of standard perturbation theory. The unperturbed flow is defined by the function $\vec{F}(\vec{a}, t, 0) = \vec{F}_0(\vec{a}, t)$ and has fluid properties represented by functions $Q(\vec{x}, t, 0) = Q_0(\vec{x}, t)$. The perturbed flow of standard theory is then a flow corresponding to parameter $\delta\lambda$. It is represented by flow function $\vec{F}(\vec{a}, t) = \vec{F}(\vec{x}, t, \delta\lambda)$ with properties $Q(\vec{x}, t) = Q(\vec{x}, t, \delta\lambda)$. More correctly, \vec{F} and Q should be subscripted by $\delta\lambda$, but standard theory does not use subscripts for perturbed quantities.

The Lagrangian displacement of the fluid point \vec{a} is given by

$$\delta\vec{x} = \vec{F}(\vec{a}, t, \delta\lambda) - \vec{F}(\vec{a}, t, 0) \approx \frac{\partial \vec{F}}{\partial \lambda}(\vec{a}, t, 0) \delta\lambda \quad (4)$$

This vector is the finite difference version of the tangent vector to the constant t curves on the deformation surface for the fluid point \vec{a} at parameter value $\lambda = 0$. It is clear that a similar procedure can be implemented for any two adjacent perturbations λ and $\lambda + \delta\lambda$. We thus obtain a field of Lagrangian displacements on each deformation surface, which I will call the *generalised Lagrangian displacements*.

Rather than work with a field of finite differences, it is better to use the vector field from which they are defined. This vector field, which I denote by $\vec{\xi}$ and call the *deformation field*, is defined by the identity

$$\vec{\xi}(\vec{F}(\vec{a}, t, \lambda), t, \lambda) = \frac{\partial \vec{F}}{\partial \lambda}(\vec{a}, t, \lambda) \quad (5)$$

This identity shows that the constant- t curves on the deformation surface are the integral curves (field lines) of the deformation field. The field $\vec{\xi}$ thus fully defines the 1-parameter family of deformations at each time t .

These observations reduce the problem of perturbation to that of finding the vector field $\vec{\xi}$, and then constructing its integral curves to yield the one parameter family of flows \vec{F} . This theory involves no approximations. It is exact from beginning to end. The only approximations that might arise are those needed for numerical solution of the exact equations.

The velocity field of the flow for parameter value λ is defined by

$$\vec{v}(\vec{F}(\vec{a}, t, \lambda), t, \lambda) = \frac{\partial \vec{F}}{\partial t}(\vec{a}, t, \lambda) \quad (6)$$

These are the tangent vectors to the constant λ curves on the deformation surface.

The t - λ curves form a Gaussian coordinate grid on the deformation surface. The Lie bracket of the vector fields \vec{v} and $\vec{\xi}$ is thus necessarily zero,

$$[\vec{x}, \vec{\xi}] = \mathcal{L}_{\vec{v}} \vec{\xi} = -\mathcal{L}_{\vec{\xi}} \vec{v} = 0 \quad (7)$$

This is a key property of this theory.

4. Eulerian and Lagrangian derivatives

The previous section contains the theoretical infrastructure needed for replacing the finite differences ΔQ and δQ by derivatives. In terms of this formalism, the Eulerian variation is defined by

$$\delta Q = Q(\vec{x}, t, \lambda) - Q(\vec{x}, t, 0) \quad (8)$$

It is clear that this variation can be extended to the entire deformation surface by defining

$$\delta Q_\lambda = Q(\vec{x}, t, \lambda + \delta\lambda) - Q(\vec{x}, t, \lambda) \quad (9)$$

I will call this the *generalised Eulerian variation*. We can now define an associated derivative by forming the differential quotient

$$\lim_{\lambda \rightarrow 0} \frac{Q(\vec{x}, t, \lambda + \delta\lambda) - Q(\vec{x}, t, \lambda)}{\delta\lambda} = \frac{\partial Q}{\partial \lambda}(\vec{x}, t, \lambda) \quad (10)$$

I will call this the *Eulerian derivative* of the 1-parameter family of flows. It is the derivative from which Eulerian variations are defined. Interestingly, it is just the ordinary partial derivative of the 1-parameter family of flows with respect to the family parameter λ .

Similarly, the Lagrangian variation is defined in this formalism by

$$\Delta Q = Q(\vec{F}(\vec{a}, t, \lambda), t, \lambda) - Q(\vec{F}(\vec{a}, t, 0), t, 0) \quad (11)$$

This too can be extended to the entire deformation surface. I therefore define the *generalised Lagrangian variation* to be

$$\Delta Q_\lambda = Q(\vec{F}(\vec{a}, t, \lambda + \delta\lambda), t, \lambda + \delta\lambda) - Q(\vec{F}(\vec{a}, t, \lambda), t, \lambda) \quad (12)$$

The Lagrangian derivative of Q is then defined by forming the limit

$$\frac{DQ}{D\lambda} = \lim_{\lambda \rightarrow 0} \frac{\Delta Q_\lambda}{\lambda} = \lim_{\lambda \rightarrow 0} \frac{Q(\vec{F}(\vec{a}, t, \lambda + \delta\lambda), t, \lambda + \delta\lambda) - Q(\vec{F}(\vec{a}, t, \lambda), t, \lambda)}{\lambda} \quad (13)$$

We can evaluate $DQ/D\lambda$ explicitly from the definition using (5) get

$$\frac{DQ}{D\lambda} = \frac{\partial Q}{\partial \lambda} + \vec{\xi} \cdot \nabla Q \quad (14)$$

If we interpret the function $\vec{F}(\vec{a}, t, \lambda$ for fixed t as a flow (the deformation flow), then the Lagrangian derivative is seen to be the material derivative for this flow. It measures the rate of change of Q at fixed time observed by an observer who moves with the fluid as the fluid deforms.

Interestingly, equation (14) relates the Eulerian to the Lagrangian derivative. This result resembles that between Lagrangian and Eulerian variations given in standard theory by $\Delta Q = \delta Q + \delta \vec{x} \cdot \nabla Q$. Note, however, that the standard result is approximate and correct only to first order in $\delta \vec{x}$, whereas result (14) is exact.

5. Some results

Using (7), the material derivatives D/Dt and $D/D\lambda$ can be shown to commute. De facto, the derivatives $\partial/\partial t$ and $\partial/\partial \lambda$ commute. These results are exact and do not need the elaborate proofs found, for example, in Lynden-Bell and Ostriker (1967).

Using these commutation relations and identities (5) and (6), it follows almost trivially that

$$\frac{D\vec{\xi}}{Dt} = \frac{D\vec{v}}{D\lambda} \quad (15)$$

Result (15) is the differential version of the result usually stated in finite difference form as¹

$$\delta \vec{v} = \vec{v}(\vec{r}, t) - \vec{v}_0(\vec{r}_0, t) = \frac{d(\delta \vec{r})}{dt} \quad (16)$$

Conservation of mass in the perturbed flows is described by the integrated mass equation (Cox, 1980, p 54-55). This result requires considerable effort to prove for a general perturbation using standard theory. The method developed above delivers this result as a trivial consequence of the fact that deformation of the fluid is a flow. Since a fluid element does not change its mass in any flow, it cannot do so in the deformation flow. Hence, we must have

$$\frac{\partial \rho}{\partial \lambda} + \nabla \cdot (\rho \vec{\xi}) = 0 \quad (17)$$

¹ See, for example, Cox (1980), p 53, eq. (5.20).

or, equivalently,

$$\frac{D}{D\lambda} \ln \rho + \nabla \cdot \vec{\xi} = 0 \quad (18)$$

The momentum and energy equations are also easily obtained, but with a little more algebra. Essentially, we need to differentiate the flow equations with respect to λ using either the Lagrangian or the Eulerian derivative. This procedure delivers equations for the λ -derivatives of the fluid variables. These derivatives take the place of their Lagrangian and Eulerian variations in ordinary perturbation theory. The equations obtained are de facto linear because of the nature of the operation of differentiation. They therefore do not need to be linearised in a separate step by neglect of higher order terms. In fact, no approximations need be performed to obtain the required perturbations equations. Furthermore, the equations obtained contain all the terms required to describe the effects of large perturbation. As the fluid deformation becomes larger, we expect the forces that act on a fluid element to differ from those that act on it when the distortion is small. The method developed above delivers without additional work the equations of motion of the fluid for all deformations, large or small. If only small perturbations are of interest, the equations of standard perturbation theory can be recovered by setting $\lambda = 0$ and by using the properties of the equilibrium flow to simplify the equations.

To illustrate the procedure, differentiate Euler's equation with respect to λ to obtain

$$\frac{\partial}{\partial t} \left(\frac{\partial \vec{v}}{\partial \lambda} \right) + \frac{\partial \vec{v}}{\partial \lambda} \cdot \nabla \vec{v} + \vec{v} \cdot \nabla \left(\frac{\partial \vec{v}}{\partial \lambda} \right) = \frac{1}{\rho^2} \frac{\partial \rho}{\partial \lambda} (\nabla p) - \frac{1}{\rho} \nabla \left(\frac{\partial p}{\partial \lambda} \right) + \frac{\partial \vec{f}}{\partial \lambda} \quad (19)$$

This is the perturbed momentum equation. the derivative $\partial \vec{v} / \partial \lambda$ plays the role of the Eulerian variation \vec{v}' of ordinary perturbation theory, $\partial \rho / \partial \lambda$ that of ρ' , $\partial p / \partial \lambda$ that of p' , and $\partial \vec{f} / \partial \lambda$ that of \vec{f}' . To express this equation in terms of the deformation field $\vec{\xi}$, it is better to apply the Lagrangian derivative to Euler's equation, followed by (15). This gives

$$\frac{D^2 \vec{\xi}}{Dt^2} = \frac{1}{\rho^2} \left(\frac{D\rho}{D\lambda} \right) (\nabla p) - \frac{1}{\rho} \frac{D}{D\lambda} (\nabla p) + \frac{D\vec{f}}{D\lambda} \quad (20)$$

The energy equation for the perturbed fluid is treated similarly. For example, if the energy equation for the unperturbed flow is written in the form

$$\frac{D}{Dt} \ln p = \Gamma_1 \frac{D}{Dt} \ln \rho + (\Gamma_3 - 1) \frac{\rho}{p} \left(\varepsilon - \frac{1}{\rho} \nabla \cdot \vec{\Phi} \right) \quad (21)$$

where Φ is the heat current density in the fluid and the Γ 's are Chandrasekhar's adiabatic indices given by $\Gamma_1 = (d \ln p / d \ln \rho)_s$ and $\Gamma_3 - 1 = (d \ln T / d \ln \rho)_s$, then the corresponding equation for the perturbed flow becomes

$$\begin{aligned} \frac{D}{Dt} \left(\frac{D}{D\lambda} \ln p \right) &= \Gamma_1 \frac{D}{Dt} \left(\frac{D}{D\lambda} \ln \rho \right) + \frac{D\Gamma_1}{D\lambda} \frac{D}{Dt} \ln \rho \\ &+ (\Gamma_3 - 1) \frac{\rho}{p} \frac{D}{Dt} \left[\frac{D}{D\lambda} \left(\varepsilon - \frac{1}{\rho} \nabla \cdot \vec{\Phi} \right) \right] + \frac{D\Gamma_3}{D\lambda} \frac{\rho}{p} \frac{D}{Dt} \left(\varepsilon - \frac{1}{\rho} \nabla \cdot \vec{\Phi} \right) \\ &+ \frac{\Gamma_3 - 1}{p} \frac{D\rho}{D\lambda} \frac{D}{Dt} \left(\varepsilon - \frac{1}{\rho} \nabla \cdot \vec{\Phi} \right) - \frac{(\Gamma_3 - 1)\rho}{p^2} \frac{Dp}{D\lambda} \frac{D}{Dt} \left(\varepsilon - \frac{1}{\rho} \nabla \cdot \vec{\Phi} \right) \end{aligned} \quad (22)$$

The equations given above are insufficient to determine the Eulerian or Lagrangian derivatives of all of the fluid parameters. Additional constraints must be supplied. These are usually given in the form of a process equation. For example, if the deformation of the fluid is a relatively rapid

process, then the entropy density s of the fluid is conserved during the deformation procedure. This gives rise to adiabatic perturbations. The condition for the perturbation to be adiabatic is easily implemented in the new formalism and takes the form

$$\frac{Ds}{D\lambda} = 0 \quad (23)$$

Non-adiabatic perturbations, in which the perturbation follows some other process equation, are described with the same ease: differentiate the process equation with respect to λ using either the Euler or the Lagrangian derivative. This yields an additional equation to be satisfied by the λ -derivatives of the fluid parameters. Equations that determine the rates of change of properties of fluid material are best formulated in terms of the Lagrangian derivative. They can then be converted to Eulerian derivatives using (14).

After the Eulerian derivatives of the fluid variables have been obtained as functions of \vec{x}, t and λ , these expressions can be integrated with respect to λ to obtain them as functions of \vec{x}, t and λ . This provides a full solution of the perturbation problem without approximation. Of course, it is unlikely that the equations obtained will be integrable by analytical methods. Approximation procedures may need to be used for the integration.

6. Conclusion

The principal advantage of the method described above over the standard method is that the equations obtained are themselves free of approximation, unlike their counterparts in the standard theory. Approximations need only be introduced, if necessary, when solving the final equations.

An added advantage of the above method is clarity. The conceptual framework outlined in Sections 3 and 4 is sufficiently well defined to bypass completely and discussion about which flow, perturbed or unperturbed, should be used for defining the material derivatives of interest. It also provides clear proofs of the central results on which the theory is based.

References

Cox, J. P., 1980, *Theory of Stellar Pulsation*, (Princeton University Press, Princeton, New Jersey).

Lynden-Bell, D., and Ostriker, J.P., 1967, MNRAS, **136**, 293.

Comparing the results of two CCD image reduction packages.

R Kgoadi and C A Engelbrecht¹

Department of Physics, University of Johannesburg, Kingsway, Auckland Park, Johannesburg, South Africa

E-mail: rgkgoadi@gmail.com, chrise@uj.ac.za

Abstract. One of the most critical procedures in observational studies of stellar physics is photometric data reduction. In this study, two software packages for this purpose are compared using a large volume of raw CCD data from the 1.0-m telescope based at the Sutherland site of the South African Astronomical Observatory (SAAO). The *Image Reduction and Analysis Facility* (IRAF) is developed by the National Optical Astronomy Observatory (NOAO) in the USA and runs on a Linux operating system. IRAF (version 2.16) is open source software that is commonly used for data reduction. Astronomical Image Processing for Windows Software (AIP4WIN, version 2.4.1), on the other hand, is designed for Windows and is obtained with a purchase of a textbook titled *The Handbook of Astronomical Image Processing* by Richard Berry and James Burnell. These packages are compared on the basis of the results that they produce from the same dataset. Reductions in IRAF resulted in fewer CCD frames images being rejected as compared to AIP4WIN and average magnitudes are approximately 14 with AIP4WIN extractions exhibiting a smaller variance.

1. Introduction

Obtaining accurate stellar magnitudes may be considered as one of the key procedures in stellar physics. It is an essential component of the extraction of information on periodic behaviour in a star and consequently, of the deduction of internal structure of a star. Hence, it is important that one tries to be as critical as possible during the reduction of astronomical photometry. A photometric reduction analysis of two chosen stars (referred to as Star 1 and Star 2 in this paper) was conducted using CCD frames obtained during three unevenly spaced observation runs at South African Astronomical Observatory (SAAO) in Sutherland to extract a time series of stellar magnitudes. The magnitudes of the target stars analysed are unknown due to data availability. This paper investigates whether the two mentioned software packages can be used irrespective of the nature of the FOV. By this we mean that, it is unavoidable that AIP4WIN and IRAF might centre the aperture differently. A crowded field of view (FOV) centred at $\alpha = 77.7500^\circ$ and $\delta = -68.8417^\circ$ was observed using the 1.0 m telescope with a V Johnson filter. Figure 1 shows the FOV and target stars analysed labelled 1 and 2, respectively, from one of the calibrated CCD frames used. Results show that there is some similarity from the results obtained from these packages. Further comparison of the magnitudes was done by extracting frequencies from light curves with the aid of the Lomb Scargle periodogram as part of the analysis.

2. Methodology

Raw CCD frames contain not only scientific data but also systematic noise from components surrounding the detector and pixel variations during observation. Hence, CCD frames must be calibrated before instrumental magnitude extraction. A typical calibration procedure involves Bias and Dark subtraction and flat fielding [1]. CCD calibrations result in frames consisting of signal mainly from celestial objects and the sky background. Even though Dark subtraction was performed during the analysis, it was not of practical benefits since the CCD camera used in the study has a liquid nitrogen cooling system attached to the detector, reducing the level of Dark current during observation to practically nil [2]. Master Bias, Dark and Flat frames were created using the median algorithm with both packages. The following subsections describe how each package was confirmed to reduce the data.

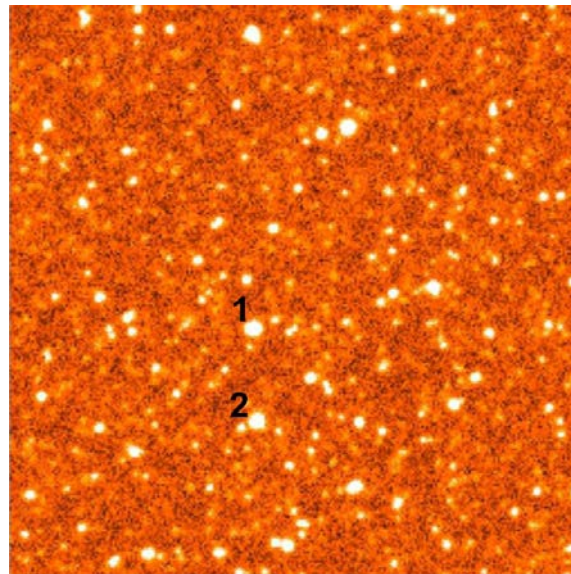


Figure 1. A calibrated CCD frame showing the field of view that enclosed stars 1 and 2. This CCD frame was calibrated by subtracting bias and dark levels as well as flat fielding a raw frame.

Data reduction was performed per observation session, taking climatic conditions into consideration. For consistency a number of parameters were kept constant for both IRAF and AIP4WIN during reductions, such as *zmag* and the aperture size. The value of the aperture size was based on the full width half maximum (FWHM) values from IRAF radial plots.

2.1. Image processing using IRAF (<http://iraf.noao.edu>)

Calibration steps followed were similar to those described in [3] and the SAOimage ds9 viewer was used to display frames. The IRAF system sub-package IMRED was used for general image processing and two NOAO optical astronomy sub - sub-packages - CCDRED and DAOPHOT - were used for calibrations and photometry respectively. The task *ccdproc* was used to calibrate science CCD frames and remove the over scan region as described in [4]. Frames were aligned using *imalign* task. Photometry was only performed to calibrated frames with an average value of the FWHM not greater than 4 for Star 1. The value of the FWHM used was determined by taking a sample of the batch that was distributed evenly throughout the night. As a result, certain nights/observations were rejected for the purposes of this study and in order to remain consistent the same observations were omitted while using AIP4WIN. Aperture photometry was performed using the *digiphot.daophot.phot* task.

2.2. Image processing using AIP4WIN (<http://www.willbell.com/aip4win/aip.htm>)

A tutorial guide provided with the AIP4WIN textbook was used to perform calibrations and photometry. Since we were working with a large volume of data frames, the Multi-Image Photometry Tool (MPT) was used. Compared to IRAF, AIP4WIN is more automated: data frames can be calibrated during photometry and master calibration frames were created using chosen CCD frames from the dataset. Using a built-in display tool, an image's grey level can be adjusted to improve the visibility of the stars [5].

3. Results

One of the obvious arguments that may be stipulated is that IRAF requires a fair knowledge of the Linux operating system and installation maybe be to some extent a challenge as well. However AIP4WIN, is an easy installation either with the CD provided or downloading software from publishers. Figure 2 shows results of aperture photometry for one week of the sample data for the two packages respectively for star 2. There exists a similar pattern in both packages indicating that star 2 was fainter in the first two nights of observation and brighter on the third night. The change in the brightness of the star may be due to the atmospheric conditions for the specific night. AIP4WIN extracted magnitudes tend to vary less compared to IRAF magnitudes. Simple descriptive statistics was performed on the entire data; as shown in table 1. Data from AIP4WIN have a sample variance of 0.002 compared to that of 0.039 for IRAF.

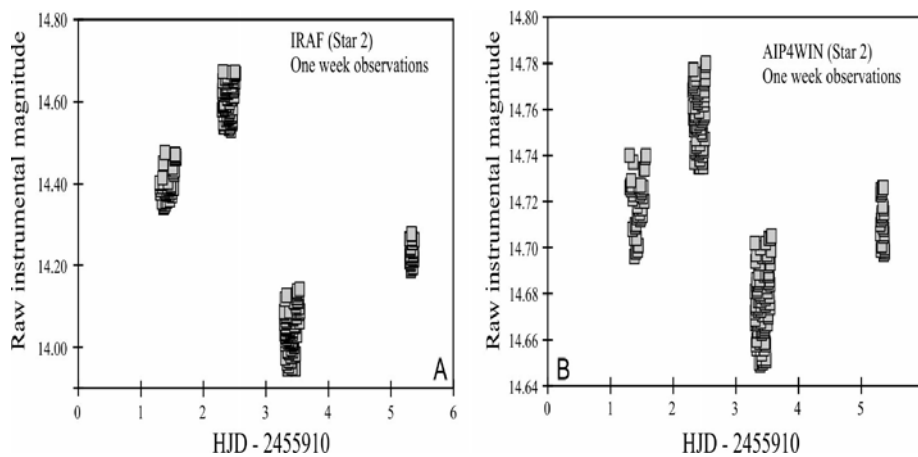


Figure 2. Raw instrumental magnitudes extracted from IRAF and AIP4WIN respectively for star 2

Table 1 shows basic statistics done on the overall data for respective packages. It is evident that light curves in figure 2 fully represents how the magnitudes deviate from their respective averages. As much as AIP4WIN possesses a low variance, these results may not be considered conclusive since parameters such as the airmass and psf fitting were not considered during extractions. Note that the value of N in table 1 differs for IRAF and AIP4WIN reductions, this is because the respective packages have distinct algorithms for rejecting frames.

Table 1. Statistics of extracted magnitudes for star 2, these data are from IRAF and AIP4WIN respectively.

IRAF				AIP4WIN			
N	\bar{x}	range	s	N	\bar{x}	range	s
531	14.321	0.865	0.039	451	14.694	0.158	0.002

N represents the size of the data, \bar{x} the average and s is the sample variance.

The variation in the results can also be seen when light curves are plotted using normalized daily averages, as shown in figure 3. Normalized magnitudes were obtained by subtraction of the daily mean from respective magnitudes. The light curve plotted using AIP4WIN reduced data shows that the scatter is more clustered around the mean as compared to IRAF light curve. This is most probably caused by the AIP4WIN package rejecting far more frames that IRAF did.

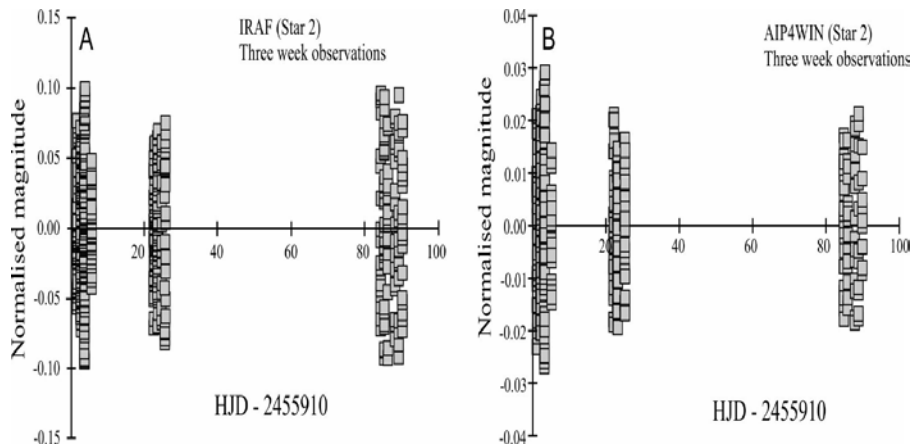


Figure 3. Normalized magnitudes for star 2 for both software packages.

The raw instrumental magnitudes obtained from the two packages were used to detect periods in the respective light curves. Periodic variations were extracted using the Lomb-Scargle periodogram and gave the following results: strongest frequencies of 3.18 c/d and 4.01 c/d were obtained for stars 1 and 2 respectively, for both AIP4WIN and IRAF light curves. Figure 4 only shows an output for star 2 LSP analysis.

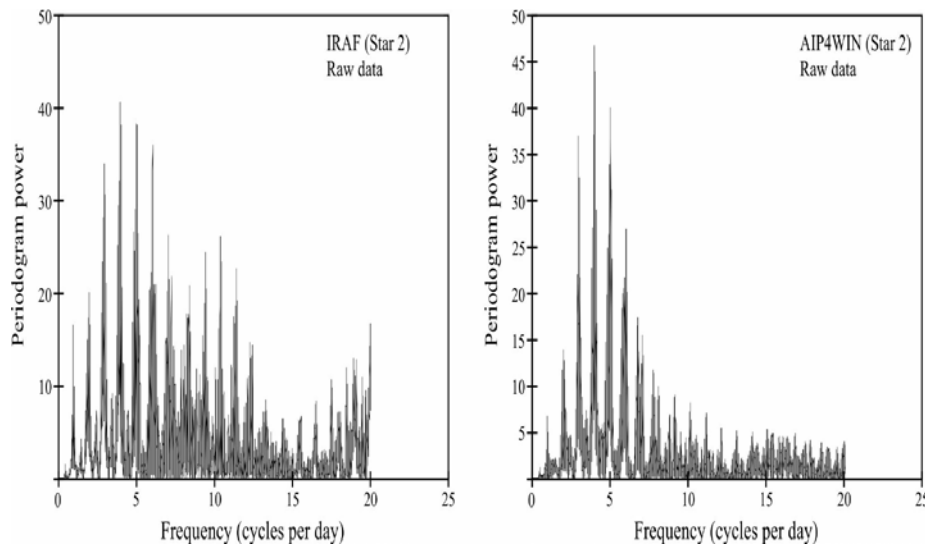


Figure 4. Lomb-Scargle analysis of star 2, using light curves from raw instrumental magnitudes.

4. Conclusion

Although the packages produce broadly similar results (especially in the important matter of identifying the strongest peak in the periodogram), there are rather striking differences: the IRAF-generated periodogram clearly contains more structure than the 'clean-looking' AIP4WIN-generated one. This is most probably the result of AIP4WIN rejecting more frames than IRAF did, causing to a tighter scatter shown as in figures 3 and 2. The extra structure in the left-hand periodogram of figure 4 could be the result of either a more sensitive periodicity detection (which would favour IRAF) or a greater amount of noise (which would favour AIP4WIN). However, given the remarkable discrepancies between the magnitudes calculated by AIP4WIN and IRAF respectively, as displayed in figures 2 and 3, it is more likely that AIP4WIN is not suitable for the crowded field in this experiment. However, further work is necessary to make a final statements on the relative merits of the tested packages for reductions of the type of photometry presented here. In the case of crowded FOV, IRAF appears to be more suitable since further analysis like psf photometry can be done to account for fainter stars and nearby stars. AIP4WIN may be more suitable for uncrowded fields and as it is a developing package, parameters such as the airmass corrections may be considered in later developments.

Acknowledgments

Financial support from the NRF and the University of Johannesburg is gratefully acknowledged. The South African Astronomical Observatory is thanked for generous amounts of observing time on the 1.0-m telescope and for helpful assistance in various operational procedures.

References

- [1] Tody D 1986 The IRAF Data Reduction and Analysis System *Instrumentation in astronomy VI (Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series vol 627)* ed Crawford D L p 733
- [2] Howell S B 2006 *Handbook of CCD Astronomy* 2nd ed (New York: Cambridge University)
- [3] Massey P 1997 A user's guide to ccd reductions with iraf Tuscon AZ USA National Optical Astronomy
- [4] Valdes F 1988 The iraf ccd reduction package - ccdred *Instrumentation for Ground-Based Optical Astronomy* ed Robinson L B p 417
- [5] Berry R and Burnell J 2000 *Briefly Noted: The Handbook of Astronomical Image Processing* 2nd ed (Richmond VA USA Willmann-Bell)

Multi-wavelength classification of unidentified AGN in the *Fermi*-2LAC catalogue

L Klindt, P J Meintjes, and B van Soelen

Department of Physics, University of the Free State, Bloemfontein, 9301, South Africa

E-mail: lizelkeklingdt@gmail.com

Abstract. Certain selection criteria have been applied to sources listed in the *Fermi*-2LAC catalogue in order to construct a target sample of twenty unidentified sources at high galactic latitudes with possible blazar characteristics. Blazars constitute the most violent and active astronomical objects which emit radiation at all wavelengths. Therefore multi-wavelength analysis provides the opportunity to construct a Spectral Energy Distribution (SED), which will allow the identification and modelling of the candidate sources. Preliminary spectroscopic results yield spectra of four observed targets which resemble that of BL Lac or flat spectrum radio quasars (FSRQs) with optical counterparts. Estimated redshifts are obtained within the range $0.11 < z < 0.37$. Potential Ca II H&K, MgIb and/or NaD spectral lines have been detected in the spectra of four targets, which motivates further observations in order to identify spectral lines that are possibly present.

1. Introduction

The *Fermi* Gamma-ray Space Telescope spacecraft with the Large Area Telescope (LAT) and the Gamma-ray Burst Monitor (GBM) have been in operation since August 2008. The LAT is the main instrument on the spacecraft and covers 20% of the sky at any time in the energy range 20 MeV to 300 GeV [1]. The *Fermi*-LAT 2-year Source catalogue (2FGL) consists of a clean sample of 866 sources of which 81% are blazars and 18% are candidate blazars of unknown type [2]. The aim is to classify possible extragalactic Active Galactic Nuclei (AGN) through multi-wavelength analysis within the 95% error circle of the *Fermi*-LAT, and to undertake spectral modelling of these systems.

AGN are extremely energetic objects powered by accretion onto a supermassive blackhole [3]. When the relativistic jet propagates close to our line of sight the AGN is classified as a blazar, and displays rapid variability, high polarization and non-thermal emission from the jet at multi-wavelengths. Blazars are classified into BL Lacartae (BL Lac) and Flat-spectrum Radio Quasars (FSRQs) according to the strength of their emission lines (see e.g., [4]). The spectra of BL Lac consist of no or weak emission lines while quasars have strong narrow and broad emission lines. Spectral energy distributions (SEDs) of blazars are characterised by two-peaks, namely a low-energy peak extending from radio to UV/X-rays, and a high-energy peak that extends from X-rays to GeV/TeV gamma-rays. The low-energy component is due to synchrotron emission, whereas the processes producing the high-energy component are still under debate, with models suggesting either leptonic or hadronic processes (see e.g., [5]). Depending on the peak frequency of the low energy component, BL Lac objects can be classified as either low-, intermediate-, or high-synchrotron peaked (LSP, ISP, or HSP) sources [6], with peak frequencies of $\nu_{peak} < 10^{14}$ Hz, $10^{14} \leq \nu_{peak} < 10^{15}$ Hz, or $\nu_{peak} > 10^{15}$ Hz, respectively.

Multi-wavelength observations will be undertaken in order to construct a SED which will allow modelling and classification of twenty unclassified 2 Year *Fermi*-LAT AGN Catalogue (2LAC) sources. These observations will be undertaken with the SAAO 1.9-m telescope (SpCCD and SHOC),

the Southern Africa Large Telescope (SALT), the Boyden/UFS 1.5-m telescope, the Watcher Robotic Telescope and possibly the HartRAO 26-m radio telescope, which will be complemented with archival data obtainable at the NASA/IPAC Extragalactic Database (NED). The selection is based on criteria that consider properties such as high galactic latitude, photon spectral indices, redshifts, radio brightness and gamma-ray variability [7]. Spectroscopic observations will be utilized to determine the redshifts and spectral lines present of the targets and photometric observations will be used to determine the variability of the sources, particularly the intra-night variability. For BL Lac and FSRQs the intra-night variability may be an indicator of the non-thermal emission in shocked regions of the jets (see e.g., [7]). The optical observations will contribute to multi-wavelength observations with the aim to model the target sample.

2. Source Selection

Counterparts in other wavelengths of the blazar candidate sources in the *Fermi* 2LAC were selected by considering a certain selection criteria as discussed in Section 2.1.

2.1. Selection Criteria

2.1.1. *High galactic latitude sources.* Near the galactic plane the source density is high and therefore to eliminate source confusion and to exclude galactic background diffuse emission, only sources at high galactic latitudes were selected e.g. $|b| > 10^\circ$.

2.1.2. *Gamma-ray Photon Spectral index.* Assuming a power law $dN/dE = N_0(E/E_0)^{-\Gamma}$ for blazars, the 2FGL blazar-type sources show spectral indices in the range of $1.2 < \Gamma < 3$ (see e.g., figure 17 in [2]).

2.1.3. *Error circle.* The counterpart has to be within the 95% (2 sigma) error circle that is associated with the *Fermi*-LAT object. This strategy was applied on the *Fermi*-LAT data which is available within the ASI Science Data Center (ASDC).

2.1.4. *Radio brightness.* Radio brightness was used to select sources within the error circles which are bright enough in the radio band for single dish observations e.g. the HartRAO 26-m telescope. However some sources that are faint in the radio band were still selected to contribute to this study. The VizieR database was used to obtain the radio flux densities at 4.85 GHz. The data were mainly catalogued in the GB6 [8]. Sources with radio flux densities above 100 mJy were considered to be radio bright and would be appropriate candidates for radio observations.

2.1.5. *Observability.* The declination of the sources had to be chosen such that multi-wavelength observations were possible from South Africa. The SAAO 1.9-m telescope can observe sources between $-90^\circ < \text{dec} < 20^\circ$, while HartRAO can reach northern declinations up to 45° . The sources are also faint and therefore an upper optical magnitude limit of 21 mag was applied, based on the limiting magnitudes of the telescopes we propose to use.

2.1.6. *Gamma-ray variability.* Blazars exhibit strong variability over various times scales over all frequencies. Sources in the *Fermi*-2LAC are classified as being variable in gamma-rays in 24 time bins (one per month) if they have a variability index $VI > 41.6$ (at a 99% confidence level) [1]. The detection of such variability is limited by the time bins and the brightness of the sources. The candidate sources we have selected have a range of variability indices, but 5 are above a 95% confidence level for gamma-ray variability. Further multi-wavelength observations are required to establish variability for the other sources

2.1.7. *Redshift.* For this study sources were selected which have no determined redshifts (with the exception of one source, see table 1). This, therefore, provides further motivation to observe the

targets and compare the measured redshifts with that of previously obtained results for blazars (see e.g. figure 12 in [2]).

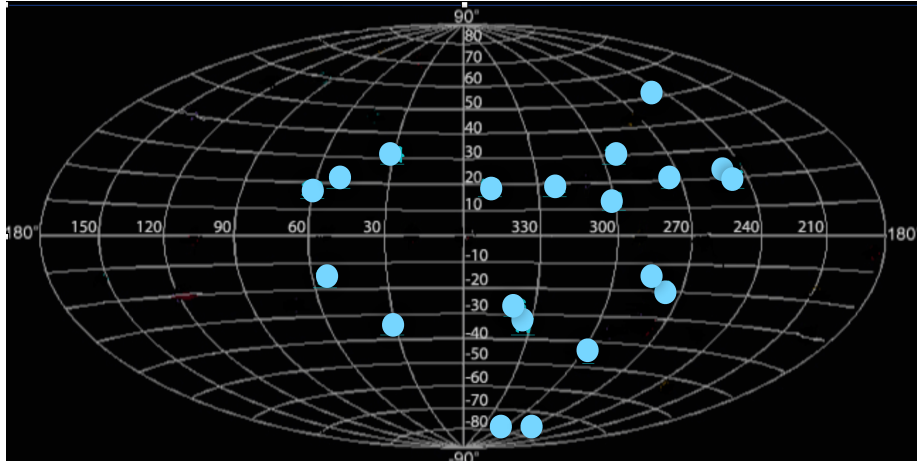


Figure 1. The galactic positions of the candidate sources that are tabulated in Table 1. All the sources are at high galactic latitude ($|b| > 10^\circ$) which ensures that the selected sources are outside the Galactic plane.

2.2. Target list

Twenty radio and optical counterparts have been selected within the 95% error boxes of the unidentified blazar-like *Fermi* 2LAC sources. The candidate sources are defined as Active Galactic Nuclei of unknown type (AGU) and are all located at high galactic latitudes ($|b| > 10^\circ$). The properties by which the targets have been selected are displayed in table 1, while the galactic distribution is shown in figure 1. The gamma-ray photon spectral indices correlate well with the blazar range given in *Fermi*-LAT observations [2].

3. Multi-wavelength follow-up studies

Radio, optical spectroscopic and photometric follow-up observations are proposed for the target list using the HartRAO 26-m telescope at Hartebeesthoek, the SAAO 1.9-m telescope and/or Southern African Large Telescope (SALT) at the South African Astronomical Observatory (SAAO) near Sutherland. Spectroscopic observations of a subset of targets have been performed with the SAAO 1.9-m telescope during June 2014, however since the targets are faint ($V \sim 20$ mag.) further spectra are required which will be proposed for SALT. Intended photometric observations of the sources will possibly be proposed for the Sutherland High-speed Optical camera (SHOC) attached to the SAAO 1.9-m telescope. The optical photometric measurements will be used to determine the variability of the targets which can be classified into three groups namely intra-day/night, short-term and long-term variability. Fast photometric observations with SHOC are proposed, with the aim to detect possible intra-night variability which is related to shocks that propagate down the relativistic jet and interacts with the surrounding medium. The observations proposed to detect the intra-night variability will be combined with a long-term monitoring campaign using the Watcher Robotic telescope at the UFS-Boyden Observatory. In previous studies it has been found that the magnitude variability of blazars ranges from 0.3 mag over a few hours to 1.2 mag within a single night [9]. Similar studies of selected unidentified counterparts of the EGRET sources have been undertaken, where one of the targets PKS 0820-5705 yielded an intra-night variability of 1.2 magnitude in the B-filter [7].

The *Fermi* data ranges from 20 MeV - 300 GeV and, therefore, to determine whether the targets peak in the TeV energies (such as high synchrotron peak BL Lac, HSP) multi-wavelength analysis is required to construct a Spectral Energy Distribution (SED). Blazars have characteristic SEDs which contain two peaks caused by processes within the systems namely synchrotron radiation (lower-frequency peak) and inverse Compton scattering (higher-frequency peak). The presence of the two peaks suggest blazar sources and the frequency at which the peak is located will allow one to determine whether it is a FSRQ or BL Lac.

Table 1. Twenty blazar candidates selected among the unidentified *Fermi* 2LAC objects.

No.	2LAC name	Counterpart	RA (hh mm ss)	Declination (° ' ")	Vmag	Spl	ePos (deg)	SED	Radio Flux (mJy)	z
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	2FGL J0044.7-3702	PKS J0045-3705	00 45 12	-37 05 48	19,6	2,57	0,153	-	330	-
2	2FGL J0113.2-3557	PKS 0110-361	01 12 39	-35 51 28	20,58	2,16	0,19	-	78	-
3	2FGL J0201.5-6626	PMN J0201-6638	02 00 53	-66 36 43	20,56	2,25	0,18	LSP	168	-
4	2FGL J0644.2-6713	PKS 0644-671	06 44 28	-67 12 57	20,69	2,16	0,05	-	218	-
5	2FGL J0730.6-6607	CRATES J073047-660226	07 30 50	-66 02 19	15,13	1,34	0,092	HSP	82	-
6	2FGL J0855.1-0712	3C 209	08 55 10	-07 15 07	19,78	2,62	0,213	-	1157	-
7	2FGL J0919.3-2203	NVSS J091922-220757	09 19 26	-22 00 45	19,95	2,00	0,163	LSP	26	-
8	2FGL J1059.0+0222	PMN J1058+0225	10 59 06	+02 25 04	-	2,29	0,151	-	97	-
9	2FGL J1106.3-3643	PMN J1106-3647	11 06 35	-36 46 59	19,4	2,2	0,14	-	53	-
10	2FGL J1154.1-3242	PKS 1151-324	11 54 32	-32 37 51	18,88	2,03	0,10	-	212	-
11	2FGL J1218.8-4827	CRATES J121901-482624	12 19 02	-48 26 27	17,53	2,4	0,144	-	65	-
12	2FGL J1407.5-4257	PKS 1404-427	14 07 40	-43 02 32	17,47	1,91	0,088	LSP	149	-
13	2FGL J1617.6-2526	PMN J1617-2537	16 17 21	25 37 23	-	2,52	0,168	-	120	-
14	2FGL J1624.4+1123	MG1 J162441+1111	12 24 55	11 12 28	17,64	2,65	0,306	-	113	-
15	2FGL J1803.6+2523	NVSS J180312-252118	18 03 12	+25 21 19	14,19	2,83	0,29	-	166	-
16	2FGL J1955.0-5639	1RXS J195503.1-564031	19 55 03	-56 40 30	17,25	1,88	0,076	HSP	9	-
17	2FGL J2040.2-7109	PKS 2035-714	20 40 08	-71 14 52	17,47	2,03	0,123	HSP	481	0,162
18	2FGL J2049.8+1001	PKS 2047+098	20 49 46	+10 03 14	-	2,38	0,139	-	295	-
19	2FGL J2108.6-1603	NVSS J210833-160724	21 08 33	-16 07 24	-	2,59	0,214	-	7	-
20	2FGL J1848.6+3241	IVS B1846+326	18 48 34	32 44 00	17,77	2,43	0,116	-	1015	-

- (1) *Fermi*-LAT name (from 2FGL catalogue; Ackermann et al. 2011).
- (2) Possible radio counterpart within the 95% error circle of the unidentified *Fermi*-LAT sources.
- (3) Right ascension for counterpart.
- (4) Declination for counterpart.
- (5) V band magnitude for the 2FGL object.
- (6) The spectral index alpha; $dN/dE = N_0(E/E_0)^{-\Gamma}$.
- (7) The spectral index alpha; *Fermi*-LAT 95% error radius.
- (8) Spectral energy distribution; Low-synchrotron peak (LSP) associated with LBLs; High-synchrotron peak (HSP) associated with HBLs.
- (9) Radio flux densities (in mJy) at 4.85 GHz [8].
- (10) Redshift of the unidentified source.

4. Spectroscopic Results

Spectroscopic observations of four targets undertaken with the SAAO 1.9-m telescope during May/June 2014 are reported here. The preliminary spectra of the targets are mainly featureless showing potential absorption lines of the host galaxy as shown in figure 2. Ca II H&K lines ($\lambda_{\text{rest,H}} = 3935 \text{ \AA}$; $\lambda_{\text{rest,K}} = 3970 \text{ \AA}$) are present in the spectra of 2FGL J1218.8-4827, 2FGL J1407.5-4257 and 2FGL J2049.8+1001, and MgIb ($\lambda_{\text{rest}} = 5174 \text{ \AA}$) and NaD ($\lambda_{\text{rest}} = 5894 \text{ \AA}$) absorption lines are detected in the spectrum of 2FGL J0730.6-6607. Using the preliminary results a rough estimate of the redshifts have been determined (see table 2) with

$$z = (\lambda_{\text{obs}}/\lambda_{\text{rest}}) - 1, \quad (1)$$

where λ_{obs} is the observed wavelength and λ_{rest} is the rest wavelength of the corresponding absorption line. No other lines have been determined in the spectra since the sources are faint ($V \sim 20$ mag.) therefore one needs to consider further observations of the targets in order to confirm the presence of other possible spectral lines. When considering the estimated redshifts and the photon spectral indices of the sources, a comparison can be made with the redshifts determined in [2]. These preliminary redshifts resemble that of BL LAC or FSRQ objects which comprise a significant fraction of the *Fermi*-LAT gamma-ray sources [10; 11].

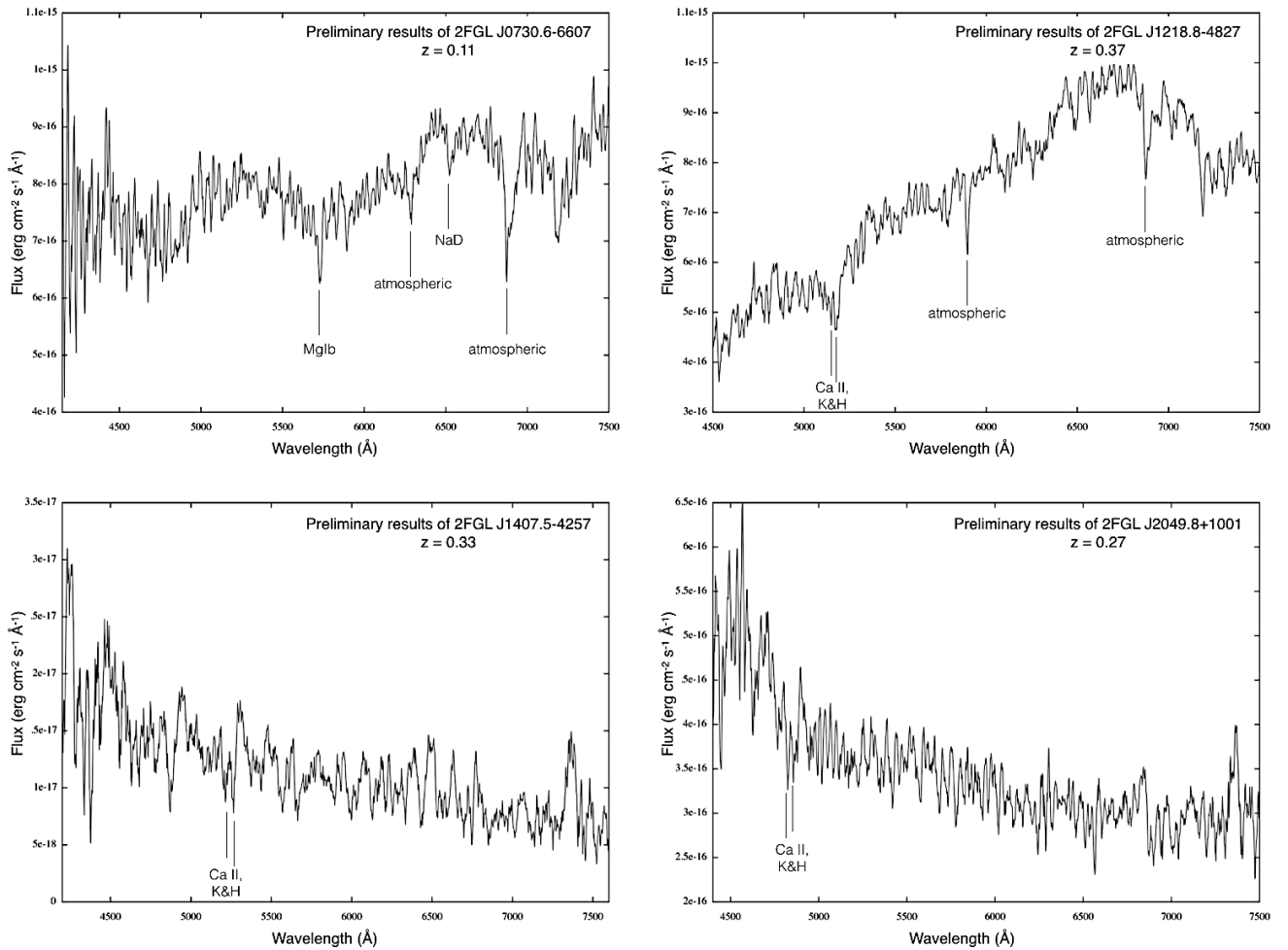


Figure 2. Spectra of four observed optical counterparts of the *Fermi*-2LAC targets with the SAO 1.9-m telescope during May/June 2014. Potential Ca II H&K, MgIb and NaD lines are present. Estimated redshifts have been determined from the absorption lines giving a redshift range of $0.11 < z < 0.37$.

Table 2. Measured absorption lines in the spectra of four unclassified *Fermi*-2LAC sources. Estimations of the redshifts are obtained from Ca II H&K, MgIb and NaD absorption lines that featured in the spectra.

2LAC name	Ca II, H (Å)	Ca II, K (Å)	MgIb (Å)	NaD (Å)	Redshift, z
2FGL J0730.6-6607	-	-	5727	6521	0.11
2FGL J1218.8-4827	5397	5443	-	-	0.37
2FGL J1407.5-4257	5210	5264	-	-	0.33
2FGL J2049.8+1001	4825	4863	-	-	0.23

5. Conclusion

The selection criteria which were used to identify possible blazar-candidates in the 2LAC have been applied to construct a target list comprising of twenty 2LAC sources. Only sources which are at high galactic latitude were included to ensure that the galactic diffuse emission was excluded and to limit confusion since the galactic plane is densely packed with sources. Spectroscopic observations for four of the unidentified *Fermi*-2LAC target sources undertaken compare favourably to that of blazars with non-thermal emission. We have identified Ca II H&K, MgIb and/or NaD absorption lines in the spectra of these sources which allowed for an estimation of the redshifts. The preliminary redshifts range $0.11 < z < 0.37$ allows one to conclude that the candidates lie in a similar redshift range as reported for the classified blazars [2]. Further multi-wavelength observations are underway to establish variability and the SEDs for these sources.

Acknowledgements

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. This paper uses observations made at the South African Astronomical Observatory (SAAO).

References

- [1] Atwood W B et al. 2009 *ApJ* **697** 071
- [2] Ackermann M et al. 2011 *ApJ* **171** 37
- [3] Urry C M and Padovani P 1995 *PASP* **107** 803
- [4] Landt H Padovani P Perlman E S and Giommi P 2004 *MNRAS* **351** 83
- [5] Böttcher M et al. 2013 *ApJ* **768** 54
- [6] Padovani P and Giommi P 1995 *ApJ* **444** 567
- [7] Nkundabakura P and Meintjes P J 2012 *MNRAS* **427** 859
- [8] Gregory P C Scott W K Douglas K and Condon J J 1996 *ApJ* **103** 427
- [9] Fan J H et al. 2004 *ChJA&A* **4** 133
- [10] Galbiati E et al. 2005 *A&A* **430** 927
- [11] Abdo A A et al. 2010 *ApJ* **710** 1271

Neutrino Events at IceCube and the Fermi Bubbles

C Lunardini¹, S Razzaque², and L Yang^{1,2,3}

¹ Department of Physics, Arizona State University, Tempe, AZ 85287-1504, US

² Department of Physics, University of Johannesburg, PO Box 524, Auckland Park 2006, South Africa

³ Laboratory for Astroparticle Physics, University of Nova Gorica, Vipavska, 5000 Nova Gorica, Slovenia

E-mail: lyang54@asu.edu

Abstract. The IceCube Neutrino Observatory announced thirty-seven candidate events observed with deposited energies above ~ 30 TeV with three-year dataset, more than expected from atmospheric backgrounds. We discuss the detectability of the Fermi Bubbles (FB) by IceCube and show that up to 6 – 7 of the 37 events could originate from the FB. If the observed gamma rays from the FB are created due to baryonic mechanism, high-energy (> 100 GeV) neutrinos should be emitted as a counterpart. These neutrinos should be detectable as shower- or track-like events at a Km^3 neutrino detector. For a hard primary cosmic-ray proton spectrum, $E^{-2.1}$, and cutoff energy at or above 10 PeV, the FB flux substantially exceeds the atmospheric backgrounds. For a steeper spectrum, $E^{-2.3}$, and/or lower cutoff energy, detection with high significance will require a longer running time.

1. Introduction

The IceCube Neutrino Observatory located at the South Pole, was designed to search for high energy neutrinos from extraterrestrial origin. IceCube recently found strong evidence for high energy astrophysical neutrino flux at the level of 10^{-8} $\text{GeV}/\text{cm}^2/\text{s}/\text{sr}$ in three years of data [1]. Except the first 28 neutrinos announced before [2, 3], an additional 9 neutrino candidates from the third year of data has been reported [1]. These total 37 events with observed energy of 30 TeV – 2 PeV more than expected background of 15 events from atmospheric muons and neutrinos, indicate a 5.7σ rejection of purely atmospheric hypothesis. Of these events, 9 are identified with visible muon tracks. The remaining 28 events are cascades or showers, caused by neutrino interactions other than ν_μ charged current.

The origin of these high energy neutrinos at IceCube is still mysterious [4]. Although the prompt atmospheric neutrino may contribute significantly with energy between 10 and 100 TeV [5, 6], the isotropic distribution of these events indicates astrophysical sources are the most natural explanation. Many scenarios have been discussed in the literature, including cores of active galactic nuclei [7, 8], γ -ray bursts [9] and their lower-powered counterparts [10, 11], galaxies with intense star-formation [12–14], active galaxies [15], flat-spectrum radio quasars [16] and intergalactic shocks [17]. Due to the weak cluster of events near the Galactic Centre, some analyses implied that Galactic sources may be responsible for a fraction of the signals [18, 19]. The IceCube events may originate from the Galactic Centre (GC) [20], the FB [21],

Galactic Halo [22], or Galactic plane [23, 24]. Some studies suggested that the origination could also be from PeV dark matter decay as well [25–27].

In this paper, we investigate the possibility that IceCube might be observing signals from the FB. The two gamma-ray bubbles, observed by Fermi-LAT [28], are symmetrically located above and below the Milky Way's galactic center, extending up to 50° in Galactic latitude. These bubbles are mostly made of energetic gamma rays emanating from the galactic core. The origin of the FB is still unknown. There are two main hypothesis, the supermassive black hole activity [28] or high rate of star formation in the GC [29, 30]. The emitted gamma rays from the FB are produced either via leptonic model, which is the Inverse Compton scattering of low energy photons by highly relativistic electrons, or via hadronic model, i.e., decays of π^0 created from interactions of energetic baryons with gas in the FB. In the hadronic model, a neutrino counterpart with similar magnitude as gamma ray flux from the FB is produced [29, 31], which could be detected as muon tracks or showers at KM3NeT, ANTARES, or IceCube [31–33]. The ANTARES collaboration has placed an upper limit for this neutrino flux, as seen in figure 3 [32]. Later, we will quantitatively present the down-going muon track as well as shower events from the FB at IceCube, not only as a possible interpretation of IceCube events, but also as a prediction for future searches.

2. Neutrino Fluxes

The Fermi Bubbles are seen as extended sources in the southern sky, see figure 1, subtending a 0.808 sr solid angle [28]. Their gamma ray flux is fairly uniform over the extent of the bubbles [28], and it is expected the neutrino flux has the same feature [31]. We compare the coordinates of the 37 IceCube all-sky search events and their median angular errors [2] with the bubbles coordinates, see figure 1. Note that event 32 produced by a pair of background muons is not labeled in figure 1, because it cannot be reconstructed with a single direction and energy [1]. It appears at least 5 events, number 2, 12, 14, 15, 36 are strongly correlated with the FB whose central position values are inside the bubbles. In addition, events 17, 22, 24, 25, are weakly correlated with the FB whose values, within errors, are compatible with the bubbles. Among these 9 correlated events, event 14, has ~ 1 PeV deposited energy [2]. figure 2 displays the time correlation of these 8 events with errors on the observed energy.

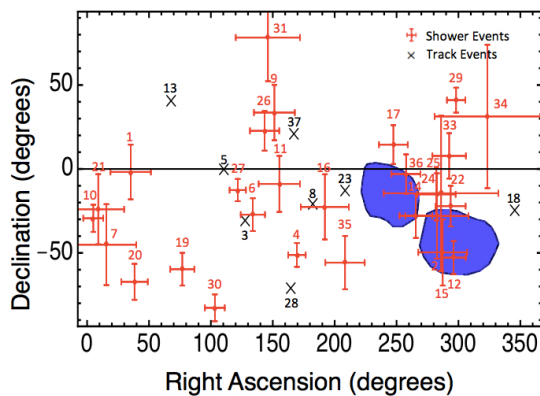


Figure 1. IceCube events distribution in equatorial coordinates, with their median angular errors, from [2]. The corresponding FB regions are shown with shaded contours.

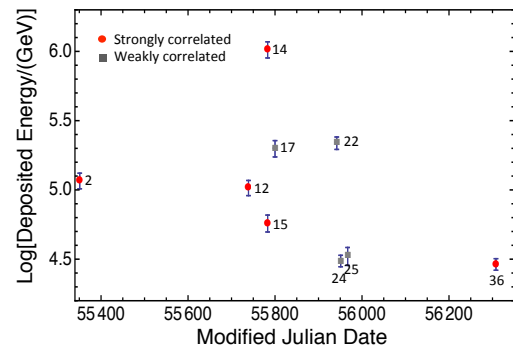


Figure 2. Events 2, 12, 14, 15, 17, 22, 24, 25, 36 (spatially correlated with bubble) are shown with errors on energy in modified Julian time and deposited energy distribution.

We adopt the FB neutrino fluxes from Ref. [31], to calculate the expected event rate at IceCube due to FB. In the model, the cosmic ray protons from supernova remnants are driven by the high velocity bipolar wind from galactic center to fill the bubbles. The neutrino fluxes are produced from the interactions of these cosmic rays with the ambient gas. This proton spectrum has the form of $dN/dE \propto E^{-k} \exp(-E/E_0)$, where E_0 is the cutoff energy, motivated by the maximum energy that accelerated protons may reach, varying from 1 - 100 PeV [34]. By fitting the gamma ray data of the FB, the spectral index $k = 2.1$ is taken as our default model; meanwhile, due to the limited number of gamma ray data points, a steeper spectrum with $k = 2.3$ is also compatible with observation, as seen in figure 3. These two sets of fluxes depending on E_0 differ significantly above ~ 200 GeV. The flux with cutoff energy of $E_0 = 30$ PeV is the most optimistic one, which is approximately 20% higher than the normalization of the whole flux, allowed by the uncertainty in the gamma ray data. The IceCube best-fit astrophysical flux of $E^2 dN/dE = 0.95 \times 10^{-8}$ GeV/cm²/s/sr is shown as well [1], see figure 3.

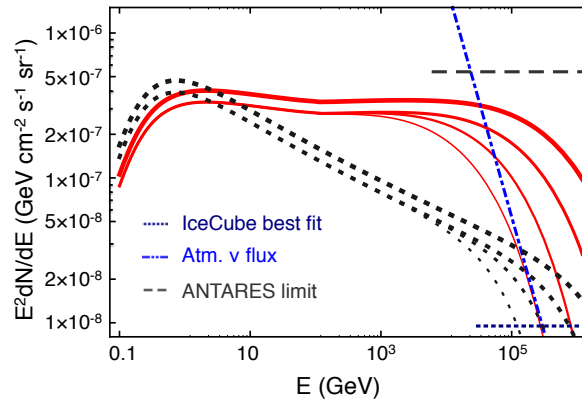


Figure 3. The pre-oscillated ν_μ and $\bar{\nu}_\mu$ flux from the FB, as a function of the energy, normalized to the gamma ray flux, for spectral indices $k = 2.1$ (solid, red) and $k = 2.3$ (dotted, black). For each index, the curves from thick to thin correspond to $E_0 = 30, 10, 3, 1$ PeV. The dash-dot curve presents the atmospheric neutrino flux [35] averaged over 25° - 95° zenith angle. The dotted line is the ANTARES upper limit [32]. The dashed line is the estimated IceCube best-fit diffuse flux expected for the FB region [2]. This figure has been taken from Ref. [36]

The neutrino fluxes initially produced at the FB of all flavors are in a composition of ν_μ : ν_e : $\nu_\tau = 2 : 1 : 0$. After oscillations, the flavor ratios are close to ν_e : ν_μ : $\nu_\tau = 1 : 1 : 1$, with deviations up to $\sim 30\%$ at $E \sim 1$ PeV.

The atmospheric muons and neutrinos are the two main backgrounds at IceCube, with an expectation of 8.4 and 6.6 events respectively for a 988-day running time [1]. The atmospheric muon level is depending on the detector veto, more details can be found in [2]. For the atmospheric neutrinos, we adopt the ν_μ flux model from Ref. [35], which is a good fit of IceCube's atmospheric data [37], extrapolated at high energies. In our calculation, we consider the flux to be symmetric in $\cos\theta_z$ [38], take the ν_μ / ν_e ratio of 14 [39], and neglect flavor oscillation due to the short propagation distance [40], as shown in figure 3.

3. Event Rate

The expected number of atmospheric neutrino backgrounds and signals for $k = 2.1, 2.3$ in 10 years, above the energy threshold E_{th} is estimated, as shown in figure 4. Taking into account of the angular resolution 1° and 10° for tracks and showers respectively at IceCube [2], we calculate

the number of atmospheric shower events over a larger solid angle than that of FB (0.808 sr). To do this, we encase each bubble in a rectangle in the θ and ϕ coordinates (see figure 1), and enlarge the area by 15° on each side. On the other hand, we take FB solid angle for track events, due to the angular resolution less than 1° . As seen in figure 4, the atmospheric shower (thick, dashed) and track (thin, dashed) events are comparable. This can be explained that the predominant atmospheric ν_μ flux compensates for the smaller effective area for tracks. For a harder spectrum ($k = 2.1$), the signal is above the background with cutoff energy $E_0 \gtrsim 10$ PeV. With $E_0 = 30$ PeV, the signal is approximately 2 orders of magnitude higher. For $E_{th} = 40$ TeV, we find 23 signals and 3 background events, indicating a $\sim 4\sigma$ excess due to the FB. It would be quite promising to discover the FB with the use of detailed statistical analyses of the spatial correlation. For $E_0 \sim 1$ PeV, the signal is even lower than the background for some thresholds, and with lower event rate as well. For a steeper spectrum ($k = 2.3$), even with $E_0 = 30$ PeV, the signal is still not significant, comparable to the background.

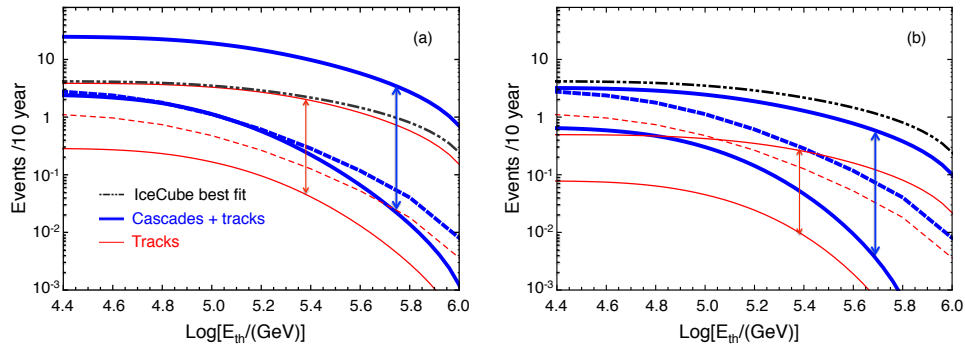


Figure 4. Expected number of events at IceCube for 10 years, as a function of energy threshold E_{th} , for the spectral indices of $k = 2.1$ (a) and $k = 2.3$ (b), from Ref. [36]. The total FB (atmospheric neutrino) signals of showerlike and tracklike events are presented as thick solid (dashed) curves. The tracklike FB (atmospheric neutrinos) signals are thin solid (dashed) curves. The arrows indicate the variation of the primary spectrum cutoff in the interval $E_0 = 1 - 30$ PeV. The estimated total showerlike and tracklike events from IceCube best-fit diffuse flux.

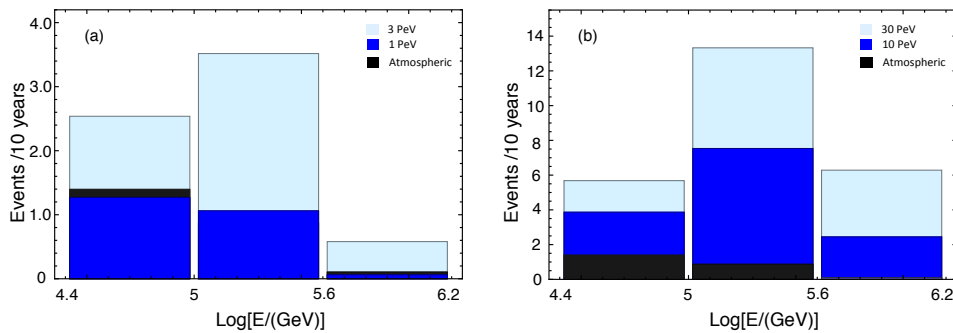


Figure 5. The expected number distribution of shower and track events per decade in each energy bin for atmospheric background and signal with $k = 2.1$ and cutoff energy $E_0 = 1, 3$ PeV (a) and $E_0 = 10, 30$ PeV (b), from Ref. [36].

The distribution of number of signal and background events per decade in neutrino energy bins is given in figure 5. As confirmed from figure 4(a), for $E_0 = 1$ PeV, the signal is comparable

Table 1. The expected number of shower and track events from the atmospheric background and from the FB in three neutrino energy bins, with different cutoff energy E_0 , for a 988-day IceCube livetime. The number of tracklike events are in brackets.

E (GeV)	$10^{4.4} - 10^5$	$10^5 - 10^{5.6}$	$10^{5.6} - 10^{6.2}$	Total
$E_0 = 1\text{PeV}$	0.34 [0.03]	0.29 [0.04]	0.02 [0]	0.65 [0.08]
$E_0 = 3\text{PeV}$	0.69 [0.07]	0.95 [0.14]	0.16 [0.03]	1.8 [0.24]
$E_0 = 10\text{PeV}$	1.05 [0.1]	2.04 [0.32]	0.66 [0.14]	3.75 [0.56]
$E_0 = 30\text{PeV}$	1.54 [0.15]	3.61 [0.57]	1.7 [0.35]	6.85 [1.08]
Background	0.38 [0.11]	0.24 [0.09]	0.03 [0.01]	0.64 [0.21]

to the background, and the most events are distributed in the energy of $10^5 - 10^{5.6}$ GeV due to the rise in the effective area below ~ 1 PeV [2].

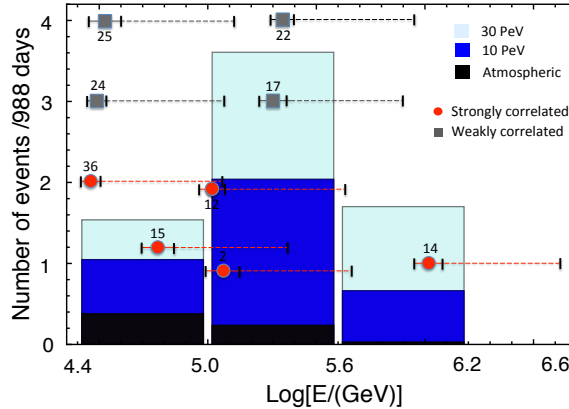


Figure 6. The same as figure 5(b) for the IceCube 988-day running time. Four strongly (red dot) and four weakly (grey square) correlated with the FB events at IceCube are shown. The solid and dashed error bars present the errors on the observed energy and a factor of $\sim 3-4$ difference between neutrino energy and deposited energy for neutral current events [41], respectively. Their coordinates on the vertical axis present the number of events for which the central value of the observed energy falls in the corresponding bins.

For comparison with IceCube data, we rescale the event rate with a 988-day running time [1], as seen in table 1 and figure 6. The total shower and track background events is less than one in the energy of $10^{4.4} - 10^{6.2}$ GeV. For $E_0 > 3$ PeV, it is expected more than 2 signals from FB with a livetime of 988 days. Especially, for $E_0 = 30$ PeV, $N \sim 5$ and $N \sim 2$ events are estimated below and above $E = 10^{5.6}$ GeV $\simeq 400$ TeV of neutrino energy respectively. Which is intriguingly close to the number of 5 events that has strongly spatial correlation to the FB.

4. Discussion

In this work, we present the possibility that IceCube may have detected the FB. With a hard spectrum and the most optimistic neutrino flux model $E_0 = 30$ PeV, we predict that up to 5 of IceCube observed events may due to the FB. Especially, 9 IceCube events are spatially correlated with the FB (figure 1), which might be the conservative upper limit for the number of events. To have a significant detection of the FB, the statistics not only depends on the backgrounds, but also on the level of the other neutrinos sources. However, as seen in figure 4, if with the most promising flux model, in seven to ten running years, the signals will be well identified.

The observation of a neutrino flux from the FB will provide clues to the mechanism of the FB, which will strongly support the hadronic origin of gamma rays from these structure. It will give a hint to the maximum limit of particle acceleration in supernova remnants, and to the time scale of the activity of the Galactic center as well.

Acknowledgments

We thank Albrecht Karle, Mariola Lesiak-Bzdak, Jakob van Santen and Nathan Whitehorn. C.L. and L.Y. acknowledge the National Science Foundation grant number PHY-1205745. K.T. acknowledges the ASU/NASA Space Grant 2013 for partial support. S.R. acknowledges support from the National Research Foundation (South Africa) grants CPRR 2014 number 87823 and BS 2014 number 91802.

References

- [1] Aartsen M *et al.* (IceCube Collaboration) 2014 *Phys.Rev.Lett.* **113** 101101 (*Preprint* 1405.5303)
- [2] Aartsen M *et al.* (IceCube) 2013 *Science* **342** 1242856 (*Preprint* 1311.5238)
- [3] Aartsen M *et al.* (IceCube Collaboration) 2013 *Phys.Rev.Lett.* **111** 021103 (*Preprint* 1304.5356)
- [4] Aartsen M *et al.* (IceCube Collaboration) 2014 (*Preprint* 1406.6757)
- [5] Lipari P 2013 (*Preprint* 1308.2086)
- [6] Aartsen M *et al.* (IceCube Collaboration) 2015 *Phys.Rev.* **D91** 022001 (*Preprint* 1410.1749)
- [7] Stecker F W 2013 *Phys.Rev.* **D88** 047301 (*Preprint* 1305.7404)
- [8] Winter W 2013 *Phys.Rev.* **D88** 083007 (*Preprint* 1307.2793)
- [9] Razzaque S 2013 *Phys.Rev.* **D88** 103003 (*Preprint* 1307.7596)
- [10] Razzaque S, Meszaros P and Waxman E 2003 *Phys.Rev.* **D68** 083001 (*Preprint* astro-ph/0303505)
- [11] Murase K and Ioka K 2013 *Phys.Rev.Lett.* **111** 121102 (*Preprint* 1306.2274)
- [12] Loeb A and Waxman E 2006 *JCAP* **0605** 003 (*Preprint* astro-ph/0601695)
- [13] Stecker F W 2007 *Astropart.Phys.* **26** 398–401 (*Preprint* astro-ph/0607197)
- [14] Liu R Y, Wang X Y, Inoue S, Crocker R and Aharonian F 2014 *Phys.Rev.* **D89** 083004 (*Preprint* 1310.1263)
- [15] Kalashev O E, Kusenko A and Essey W 2013 *Phys.Rev.Lett.* **111** 041103 (*Preprint* 1303.0300)
- [16] Dermer C D, Murase K and Inoue Y 2014 *JHEAp* **3-4** 29–40 (*Preprint* 1406.2633)
- [17] Murase K, Ahlers M and Lacki B C 2013 *Phys.Rev.* **D88** 121301 (*Preprint* 1306.3417)
- [18] Fox D, Kashiyama K and Mszars P 2013 *Astrophys.J.* **774** 74 (*Preprint* 1305.6606)
- [19] Gonzalez-Garcia M, Halzen F and Niro V 2014 *Astropart.Phys.* **57-58** 39–48 (*Preprint* 1310.7194)
- [20] Razzaque S 2013 *Phys.Rev.* **D88** 081302 (*Preprint* 1309.2756)
- [21] Ahlers M and Murase K 2014 *Phys.Rev.* **D90** 023010 (*Preprint* 1309.4077)
- [22] Taylor A M, Gabici S and Aharonian F 2014 *Phys.Rev.* **D89** 103003 (*Preprint* 1403.3206)
- [23] Anchordoqui L A, Goldberg H, Lynch M H, Olinto A V, Paul T C *et al.* 2014 *Phys.Rev.* **D89** 083003 (*Preprint* 1306.5021)
- [24] Joshi J C, Winter W and Gupta N 2013 **3414** 3419 (*Preprint* 1310.5123)
- [25] Feldstein B, Kusenko A, Matsumoto S and Yanagida T T 2013 *Phys.Rev.* **D88** 015004 (*Preprint* 1303.7320)
- [26] Esmaili A and Serpico P D 2013 *JCAP* **1311** 054 (*Preprint* 1308.1105)
- [27] Barger V and Keung W Y 2013 *Phys.Lett.* **B727** 190–193 (*Preprint* 1305.6907)
- [28] Su M, Slatyer T R and Finkbeiner D P 2010 *Astrophys.J.* **724** 1044–1082 (*Preprint* 1005.5480)
- [29] Crocker R M and Aharonian F 2011 *Phys.Rev.Lett.* **106** 101102 (*Preprint* 1008.2658)
- [30] Fujita Y, Ohira Y and Yamazaki R 2013 *Astrophys.J.* **775** L20 (*Preprint* 1308.5228)
- [31] Lunardini C and Razzaque S 2012 *Phys.Rev.Lett.* **108** 221102 (*Preprint* 1112.4799)
- [32] Adrian-Martinez S *et al.* (ANTARES Collaboration) 2014 *Eur.Phys.J.* **C74** 2701 (*Preprint* 1308.5260)
- [33] Adrian-Martinez S *et al.* (KM3NeT Collaboration) 2013 *Astropart.Phys.* **42** 7–14 (*Preprint* 1208.1226)

- [34] Ptuskin V, Zirakashvili V and Seo E 2010 *Astrophys.J.* **718** 31–36 (*Preprint* 1006.0034)
- [35] Honda M, Kajita T, Kasahara K, Midorikawa S and Sanuki T 2007 *Phys.Rev.* **D75** 043006 (*Preprint* astro-ph/0611418)
- [36] Lunardini C, Razzaque S, Theodoseau K T and Yang L 2014 *Phys.Rev.* **D90** 023016 (*Preprint* 1311.7188)
- [37] Abbasi R *et al.* (IceCube Collaboration) 2011 *Phys.Rev.* **D83** 012001 (*Preprint* 1010.3980)
- [38] Sajjad Athar M, Honda M, Kajita T, Kasahara K and Midorikawa S 2013 *Phys.Lett.* **B718** 1375–1380 (*Preprint* 1210.5154)
- [39] Sinigovskaya T, Ogorodnikova E and Sinigovsky S 2013 (*Preprint* 1306.5907)
- [40] Gaisser T and Stanev T 1998 *Phys.Rev.* **D57** 1977–1982 (*Preprint* astro-ph/9708146)
- [41] Aartsen M *et al.* (IceCube Collaboration) 2014 *JINST* **9** P03009 (*Preprint* 1311.4767)

Tidal effects on pulsation modes in close binaries

M Predieri¹, F A M Frescura¹ and C A Engelbrecht²

¹ School of Physics, University of the Witwatersrand, private bag 3, Wits 2050, Johannesburg, South Africa

² Department of Physics, University of Johannesburg, Kingsway, Auckland Park, Johannesburg, South Africa

E-mail: lotmass@gmail.com

Abstract. Light curve data from the *Kepler* satellite on pulsating eclipsing Algol-type binary systems display a peculiar feature: the primary shows preferential excitation of pulsation modes with frequencies resonant with the orbital frequency of the binary system. A proposed explanation of this phenomenon is tidal driving of pulsations by the secondary. This paper presents a preliminary calculation of the effects of linear representations of tides on the pulsation frequencies of a polytropic primary.

1. Introduction

The *Kepler* space telescope has produced a large number of time-domain observations of unprecedented regularity and precision. A few hundred close eclipsing binary systems have emerged from the *Kepler* database and many of these have been studied intensively since the first *Kepler* data became available in 2009. Due to the precision (noise levels around a few micromagnitudes in white light) of the data, weak pulsation signatures have been detected in many close (semi-detached) systems with circularised orbits. These pulsation signatures frequently point to a resonant coupling of pulsation frequencies with orbital motion. Initial expectations are that resonance should not occur in a circularised orbit, since there is no substantial difference between the periastron and apastron separations of the binary components that could drive such a resonance.

However, given that these are close systems, tidal effects might be able to produce the apparent observed resonances. Since mass transfer between binary components is not considered, attention is given specifically to detached systems where negligible mass transfer occurs [1]. We present a simple model of the effects of tides on a polytropic star and indicate how a numerical model may be used to test this conjecture. We use the formula for the tidal potential as presented in Reynier's papers [2, 3]. We then Fourier transform the tidal potential with respect to the time variable. This yields an equation containing three delta functions that select a triplet of frequencies indicating a splitting of the eigenstates into three levels. The eigenstates are evenly spaced in the frequency domain. The spacing between them is determined by the angular velocity of rotation of the primary and the mean motion of the secondary. This splitting does not occur in a phase-locked circularised orbit, since no tidal dragging is present.

2. Equations of motion

We begin by considering a close binary system. Suppose the primary star has mass M , radius R and constant angular velocity Ω with respect to an inertial frame. The axis of rotation of the binary system is assumed to be perpendicular to its orbital plane. The companion star, with mass m , is modelled as a point mass. We will use a reference frame, with spherical coordinates (r, θ, φ) , that co-rotates with the primary whose center is at its origin with the frames z-axis being parallel to the primary's axis of rotation. In this reference frame the equations of motion for an ideal fluid incorporating the effects of a tidal interaction are presented below. The conservation of momentum equation is given by:

$$\frac{\partial \vec{v}}{\partial t}(\vec{r}, t) + (\vec{v}(\vec{r}, t) \cdot \vec{\nabla}) \vec{v}(\vec{r}, t) = -\vec{\nabla} \phi(\vec{r}, t) - \frac{1}{\rho(\vec{r}, t)} \vec{\nabla} P(\vec{r}, t) - \epsilon_T \vec{\nabla} W(\vec{r}, t) \quad (2.1)$$

The quantity ϵ_T is a small, dimensionless parameter describing the magnitude of the tidal interaction and is defined as:

$$\epsilon_T = \left(\frac{R}{a}\right)^3 \frac{M}{m}$$

where a is the semi-major axis of the orbit of the companion. The tidal potential $W(\vec{r}, t)$ is given by [2] as:

$$W(\vec{r}, t) = \frac{GM}{R} \left(\frac{r}{R}\right)^2 \left\{ \frac{1}{2} P_2(\cos \theta) - \frac{1}{4} P_2^2(\cos \theta) \cos[2\varphi + 2t(\Omega - n)] \right\} \quad (2.2)$$

where G , n , P_2 and P_2^2 represent respectively Newton's gravitational constant, the mean angular velocity of the orbit of the companion, the second order Legendre polynomial and the second order associative Legendre polynomial with azimuthal number 2. The first term in equation (2.2) represents the time-independent static tide, characteristic of a synchronous orbit, in which $\Omega = n$. The second term represents the time-dependent dynamic tide, characteristic of an asynchronous orbit.

In addition to the above equations we have three more equations completing the set, they are:

$$\frac{\partial \rho}{\partial t}(\vec{r}, t) + \vec{\nabla} \cdot [\rho(\vec{r}, t) \vec{v}(\vec{r}, t)] = 0 \quad (2.3)$$

which is the continuity equation,

$$\nabla^2 \phi(\vec{r}, t) = 4\pi G \rho(\vec{r}, t) \quad (2.4)$$

which is Poisson's equation, and lastly

$$P(\vec{r}, t) = \kappa(\rho(\vec{r}, t))^\gamma \quad (2.5)$$

which is the polytropic equation of state. Collectively, equations (2.1) - (2.5) constitute a complete set of equations for the fluid variables.

3. Perturbation equations

To determine the perturbation equations we first consider an "unperturbed" equilibrium configuration, which we then perturb and linearise. The initial configuration is assumed to be an hydrostatic, spherically symmetric configuration. Linearisation is achieved by assuming that all perturbations are small so that all products of perturbations can be neglected. In pulsation

theory, two types of variation are considered, called Eulerian and Lagrangian respectively. The first defined as the perturbation of physical fields, denoted by a prime, and the second the perturbation suffered by a given material fluid element, denoted by a δ [4, 5]. Given an arbitrary physical quantity, the Eulerian and Lagrangian perturbations of Q are defined respectively by:

$$Q' = Q(\vec{r}_0, t) - Q_0(\vec{r}_0) \quad (3.1)$$

$$\delta Q = Q(\vec{r}, t) - Q_0(\vec{r}_0) \quad (3.2)$$

where \vec{r}_0 and \vec{r} represent the current position of the material fluid element in the equilibrium configuration and the perturbed configuration respectively. The vector quantity relating these two positions is called the Lagrangian displacement and is defined as:

$$\vec{\xi} = \vec{r} - \vec{r}_0 \quad (3.3)$$

Using equations (3.1) - (3.3) it can be shown that the relationship between the Eulerian and Lagrangian perturbations is given, to first order in $\vec{\xi}$, by:

$$\delta Q(\vec{r}, t) = Q'(\vec{r}, t) + \vec{\xi} \cdot \vec{\nabla} Q_0(\vec{r}) \quad (3.4)$$

If the unperturbed system is assumed to be in hydrostatic equilibrium, the velocity field \vec{v} is related to $\vec{\xi}$ by:

$$\vec{v}(\vec{r}, t) = \frac{\partial \vec{\xi}}{\partial t} \quad (3.5)$$

and

$$\vec{v}(\vec{r}, t) = \delta \vec{v}(\vec{r}, t) = \vec{v}'(\vec{r}, t) = \frac{\partial \vec{\xi}}{\partial t} \quad (3.6)$$

Using equations (3.1) - (3.6) with equations (2.1) and (2.3) - (2.5) the perturbation equations can be shown to be:

$$\rho'(\vec{r}, t) + \vec{\nabla} \cdot (\rho_0(r) \vec{\xi}) = 0 \quad (3.7)$$

$$\frac{\partial^2 \vec{\xi}}{\partial t^2}(\vec{r}, t) + \frac{1}{\rho_0(r)} \vec{\nabla} P'(\vec{r}, t) + \vec{\nabla} \phi'(\vec{r}, t) + \frac{\rho'(\vec{r}, t)}{\rho_0(r)} \frac{\partial \phi_0}{\partial r}(r) + \frac{\epsilon_T}{\rho_0(r)} \vec{\nabla} W(\vec{r}, t) = 0 \quad (3.8)$$

$$\nabla^2 \phi'(\vec{r}, t) - 4\pi G \rho'(\vec{r}, t) = 0 \quad (3.9)$$

$$P'(\vec{r}, t) - \gamma \left(\frac{P_0(r)}{\rho_0(r)} \right) \rho'(\vec{r}, t) = 0 \quad (3.10)$$

where the primed quantities represent the Eulerian perturbation of that physical quantity. The tidal term is not perturbed as it is considered to be an applied external force which does not undergo perturbation.

4. Eigenvalue problem

To determine the pulsation frequencies of the system, we search for normal mode solutions to equation (3.8). Because the perturbation equations are linear, the general solution can be constructed from a linear superposition of the normal mode solutions. Taking the Fourier

transform of equations (3.7) - (3.10), we obtain the following equations:

$$\rho'(\vec{r}, \omega) + \vec{\nabla} \cdot \left(\rho_0(r) \vec{\xi}(\vec{r}, \omega) \right) = 0 \quad (4.1)$$

$$-\omega^2 \vec{\xi}(\vec{r}, \omega) + \frac{1}{\rho_0(r)} \vec{\nabla} P'(\vec{r}, \omega) + \vec{\nabla} \phi'(\vec{r}, \omega) + \frac{\rho'(\vec{r}, \omega)}{\rho_0(r)} \frac{\partial \phi_0}{\partial r}(r) + \frac{\epsilon_T}{\rho_0(r)} \vec{\nabla} W(\vec{r}, \omega) = 0 \quad (4.2)$$

$$\nabla^2 \phi'(\vec{r}, \omega) - 4\pi G \rho'(\vec{r}, \omega) = 0 \quad (4.3)$$

$$P'(\vec{r}, \omega) - \gamma \left(\frac{P_0(r)}{\rho_0(r)} \right) \rho'(\vec{r}, \omega) = 0 \quad (4.4)$$

where $W(\vec{r}, \omega)$ can be calculated using:

$$\begin{aligned} W(\vec{r}, \omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} W(\vec{r}, t) e^{-i\omega t} dt \\ &= r^2 \left(\frac{GM}{R^3} \right) \left(\sqrt{\frac{\pi}{8}} P_2(\cos \theta) \delta(\omega) - \sqrt{\frac{\pi}{32}} P_2^2(\cos \theta) \right. \\ &\quad \left. \{ e^{2i\varphi} \delta(2(\Omega - n) - \omega) - e^{-2i\varphi} \delta(2(\Omega - n) + \omega) \} \right) \end{aligned} \quad (4.5)$$

Here δ represents the Dirac delta function. Given a set of boundary conditions, the values of ω satisfying these boundary conditions can be found numerically.

If we remove the tidal term, assume the Cowling approximation ($\Phi(\vec{r}, t) \approx \Phi_0(r)$) and consider the perturbation of a spherically symmetric equilibrium configuration that admits only radial oscillations we obtain the eigenvalue equation for the radial oscillations of a polytropic star:

$$\alpha \frac{\partial^2 \rho'}{\partial r^2}(r, \omega) + \beta \frac{\partial \rho'}{\partial r}(r, \omega) + \lambda \rho'(r, \omega) = 0 \quad (4.6)$$

with

$$\begin{aligned} \alpha &= \gamma \sigma \\ \beta &= 2\gamma \frac{\partial \sigma}{\partial r} + \frac{\partial \phi_0}{\partial r} + \frac{2\gamma \sigma}{r} \\ \lambda &= \omega^2 + \gamma \frac{\partial^2 \sigma}{\partial r^2} + \frac{\partial^2 \phi_0}{\partial r^2} + \frac{2\gamma}{r} \frac{\partial \sigma}{\partial r} \end{aligned}$$

where $\sigma = \frac{P_0(r)}{\rho_0(r)}$. All quantities contained in the coefficients α , β and λ are known, except for ω . By implementing finite difference expressions for the derivatives and introducing the boundary conditions for ρ' we can search for values of ω numerically.

5. Concluding remarks

The naive model described in this paper can be extended to a more realistic model, which takes into account in a better way the several effects that need to be incorporated. We have considered a polytrope, which does not cover all stellar types and configurations. Using a more realistic equation of state will allow for the final results to be coded in such a way that they can be included in well developed models of stellar structure such as MESA and ASTEC. Non-radial oscillations have not been considered. These can be included in the calculation by not requiring spherical symmetry in the perturbed configuration. In standard pulsation theory the pulsation equations can be separated into radial and angular parts. This separation may not be possible when we include the tidal term, leading to a coupling of radial and angular modes. Assuming that the oscillations are radial is very unrealistic, especially since the external forces, resulting from the inclusion of tides, have non-radial as well as radial components. We are currently coding our preliminary model and will be extended to include non-radial pulsations.

References

- [1] Kopal, Z. *Dynamics of close binary systems*. D. Reidel Publishing Company, Dordrecht, Netherlands, 1914.
- [2] Reyniers, K. & Smeyers, P. Tidal perturbations of linear, isentropic oscillations in components of circular-orbit close binaries: Synchronously rotating components. *Astronomy & Astrophysics*, 404:1051–1065, 2003.
- [3] Reyniers, K. & Smeyers, P. Tidal perturbations of linear, isentropic oscillations in components of circular-orbit close binaries: Validity of the perturbation method applied to equilibrium tides. *Astronomy & Astrophysics*, 409:677–688, 2003.
- [4] Unno, W., Osaki, Y., Ando, H., Saio, H., & Shibahashi, H. *Nonradial oscillations of stars*. University of Tokyo Press, Tokyo, 1989.
- [5] Cox, J.P. *Theory of stellar pulsation*. Princeton University Press, Princeton, New Jersey, 1980.
- [6] Smeyers, P. Dynamic theory of tides in close binaries, and effects on the apsidal motion and stellar oscillations. *Third Granada Workshop on Stellar Structure, ASP Conference Series*, 333, 2005.
- [7] Polfiet, R. & Smeyers, P. Dynamic tides in stars as forced isentropic oscillations and their effects on free oscillations. *Astronomy & Astrophysics*, 237:110–124, 1990.
- [8] Kallrath, J. & Milone, E.F. *Eclipsing binary stars: modeling and analysis*. Springer-Verlag, New York, New York, 2009.

Comparison of photometric and spectroscopic parameters of eclipsing contact binary stars

Patricia Skelton

Department of Mathematical Sciences, PO Box 392, UNISA, 0003, South Africa

E-mail: skeltp1@unisa.ac.za

Abstract. To model an eclipsing contact (EC) binary star requires the temperature of at least one of the components, usually T_1 , and the mass ratio q . Other parameters are determined by minimizing residuals between the model and phase-magnitude data. Rucinski *et al.* (2005) have pointed out that model solutions of EC stars obtained from photometric data are unreliable because the photometrically determined mass ratios are different to those determined from spectroscopic data. The temperatures determined from colour indices are also found to differ from those determined spectroscopically. Clearly, in order to produce reliable models of these stars requires a combination of photometric and spectroscopic data. Using the SpCCD spectrograph on the 1.9m telescope at the South African Astronomical Observatory in Sutherland, spectroscopic data were obtained for selected EC stars. The results of the observations and a comparison of the photometrically and spectroscopically determined temperatures and mass ratios are presented.

1. Introduction

Eclipsing contact binary stars, or EC stars, have orbital periods of less than 1 day [1]. The component stars of an EC system range in spectral type from mid-A to late-K and each is assumed to be a main sequence star. Binnendijk [2] divided EC stars into two subclasses which he called A-type and W-type. In the A-type systems, the larger component has a higher temperature than its companion, whereas in the W-type systems the smaller component has a higher temperature than its companion. Observationally it has been found that the A-type systems tend to have low mass ratios $q < 0.3$ and spectral type from A to F [3]. The W-type systems tend to have mass ratios $q > 0.3$ and spectral types of G or K.

Modelling to determine the parameters of as many EC stars as possible may help to answer the many questions we have about the formation and evolution of EC stars. To determine the absolute parameters of an EC star requires a combination of photometric and spectroscopic data. Due to the large number of automated sky surveys, there is no shortage of photometric data of EC stars [4]. One example is the All Sky Automated Survey (ASAS). Since the project began, over 50,000 variable stars have been detected and of those, over 5,000 have been classified as EC stars. A large portion of the variable stars discovered by the ASAS have not been classified previously as variable stars. For the newly discovered EC stars, absolute parameters have yet to be determined for these stars due to the lack of spectroscopic data.

It is in general agreement that physical parameters of EC stars obtained from the use of photometric data only are often unreliable. Terrell & Wilson [5] have shown that this is especially true of photometric mass ratios determined for partial eclipsing systems. For total eclipsing

systems, the situation is different because the durations of the eclipses constrain the ratio of the radii of the component stars and hence the mass ratio q [6]. This implies that for a total eclipsing system, the photometric mass ratio should be very close, if not equal to, the spectroscopically determined mass ratio.

In a series of papers, Rucinski and collaborators have published the results of radial velocity studies of close binary stars (see *Radial Velocity Studies of Close Binary Stars*, Papers I – XV). The authors have pointed out that for most of the stars in their programme, the photometrically determined mass ratios are not equal to the spectroscopically determined mass ratios [7]. XY Leo is an example of just how difficult it is to determine a reliable mass ratio for a partial eclipsing system using photometric data only. Rucinski and Lu [8] determined a spectroscopic mass ratio $q_{\text{sp}} = 0.348(29)$ which is substantially smaller compared to the photometric mass ratio $q_{\text{ph}} = 0.726$ as determined by Niarchos *et al.* [9]. For most total eclipsing systems, Rucinski and Lu have found that the spectroscopic mass ratio is in good agreement with the photometric mass ratio. UZ Leo is an example of a total eclipsing system whose photometric mass ratio differs from the spectroscopic mass ratio. Vinko, Hegedues & Hendry [10] determined two possible photometric mass ratios, $q_{\text{ph}} = 0.2329$ and $q_{\text{ph}} = 0.2266$, for UZ Leo. These two values differ substantially from the spectroscopic mass ratio, $q_{\text{sp}} = 0.303(24)$, determined by Rucinski and Lu. These results demonstrate the importance of determining the spectroscopic mass ratios for EC stars.

For modelling and for EC subclass classification, it is important to know the spectral type. For systems that do not have spectroscopic data, estimates of the temperatures are obtained from photometric colour indices. Due to uncertainties regarding reddening corrections and the effect of interstellar extinction, infrared colour indices like $J - K$ and $H - K$ provide more reliable estimates of temperatures compared to $B - V$ colour indices [4]. In the series of papers by Rucinski and collaborators, cases are presented where the spectral type determined from different photometric colour indices do not match and these also differ from the spectral type obtained from spectroscopic data.

Spectroscopic data for two total eclipsing EC stars ASAS 120036–391536 and ASAS 093818–675524 were obtained. The results of the spectroscopic data analysis for the two stars are presented.

2. Observations

Observations were carried out using the SpCCD spectrograph on the 1.9m telescope at the South African Astronomical Observatory (SAAO) Sutherland. Both stars were observed during two runs in 2013 (February and May). Grating 4 with 1200 lines mm^{-1} and resolution of 1 Å was used. The grating has a useful range of 800 Å and a blaze wavelength of 4600 Å. The grating was centred close to the blaze angle in order to have a spectral range coverage of around 4200 – 5000 Å. This wavelength coverage includes the two Balmer lines, $H\gamma$ and $H\delta$, as well as a plethora of iron and other ‘metal’ lines in the range 4300 – 4500 Å. The intrinsic shape of the iron and ‘metal’ lines is symmetric and narrow [11], and due to the large number of these lines, they have proven to be useful in measuring radial velocities via cross-correlation.

Observations of both ASAS 093818–675524 and ASAS 120036–391536 were targeted at and around the quadratures with a few spectra obtained at phases 0 and 0.5. The latter observations assist with determining the relative system velocity which is required in order to determine the spectroscopic mass ratio. For both stars, the spectra obtained close to phase 0.5 correspond to the total eclipse portions of their light curves (see Figure 1). Since the light from one component is visible during this phase, these spectra were used to determine the spectral type of the eclipsing component. A total of 46 spectra were obtained for ASAS 093818–675524 and 34 spectra for ASAS 120036–391536. Spectra for two radial velocity standard stars, HD 126053 and HD 154417, were obtained in order to be used for the radial velocity measurements of ASAS

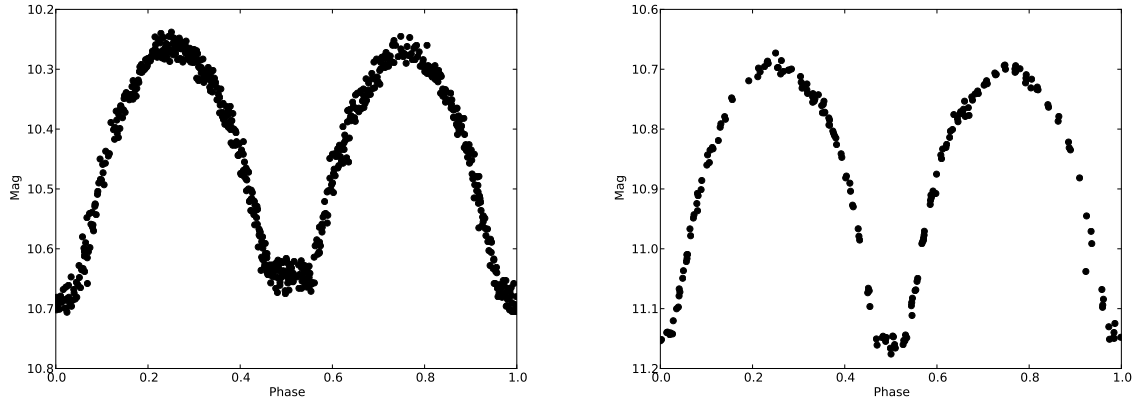


Figure 1. Phase-Magnitude diagrams for ASAS 093818–675524, left, and ASAS 120036–391536, right. Both stars display total eclipses at phase 0.5.

093818–675524 and ASAS 120036–391536.

Standard CCD reductions were performed using the relevant routines in IRAF. Spectra were extracted using the APALL task and wavelength calibrated using CuAr arc lamp exposures. Radial velocities were measured by cross-correlation with the spectra of the radial velocity standard stars, using the FXCOR task in IRAF. The radial velocity standard star HD 154417 (F9V, $V_r = -16.70 \text{ km s}^{-1}$) was used for ASAS 093818–675524 and HD 126053 (G1.5V, $V_r = -19.40 \text{ km s}^{-1}$) for ASAS 120036–391536.

PHOEBE, an eclipsing binary modelling software package [12] based on the Wilson-Devinney code [13], was used to determine the relative system velocity and spectroscopic mass ratio for both EC stars. To determine the spectral types of ASAS 093818–675524 and ASAS 120036–391536, the phase 0.5 spectra of both stars were cross-correlated with the spectra of main-sequence stars of spectral types from mid-A to late-K using FXCOR. The heights of the cross-correlation peaks were used to identify which spectral types best matched the phase 0.5 spectra.

3. Results

3.1. Radial Velocities

Photometric mass ratios for both stars were determined by using the iterative procedure described by Wadhwa and Zealey [14]. The software package PHOEBE was used to model the stars and the Differential Corrections (DC) routine of the programme was used to refine the parameters. Photometric mass ratios $q_{\text{ph}} = 0.170$ and $q_{\text{ph}} = 0.256$ were determined for ASAS 093818–675524 and ASAS 120036–391536 respectively.

The radial velocity curves for both stars are shown in Figure 2. Filled circles correspond to the radial velocities of the primary component and the filled squares correspond to the radial velocities of the secondary component. The DC routine in PHOEBE returned a spectroscopic mass ratio $q_{\text{sp}} = 0.210$ for ASAS 093818–675524 and $q_{\text{sp}} = 0.315$ for ASAS 120036–391536. For both stars, the spectroscopic mass ratio does not equal the photometric mass ratio. Although both systems are total eclipsing systems, the fact that $q_{\text{ph}} \neq q_{\text{sp}}$ highlights the importance of using spectroscopic data to determine the mass ratios of EC stars.

During the calculation of the phase values for ASAS 093818–675524, it became clear that the given ASAS period was incorrect as the maximum radial velocities did not correspond to phases 0.25 and 0.75. The period was adjusted to correct for this but during the analysis in

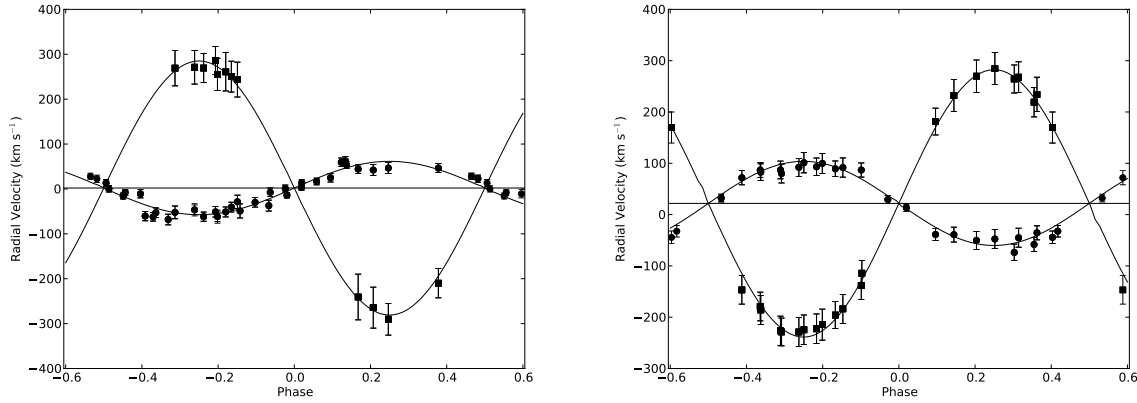


Figure 2. Radial velocity curves for ASAS 093818–675524, left, and ASAS 120036–391536, right.

PHOEBE, a phase shift of 0.5 had to be introduced so that the synthetic radial velocity curve matched the observed radial velocity curve. The ASAS data for ASAS 093818–675524 suggests that the period of the system is changing and so a period analysis will need to be performed in order to refine the period and determine a period change rate.

3.2. Spectral Types

Table 1 lists the V magnitude, $B - V$ colour indices, 2MASS colour indices and corresponding spectral types for both ASAS 093818–675524 and ASAS 120036–391536. The $B - V$ colour indices were taken from the SIMBAD astronomical database and the 2MASS colour indices were taken from the ASAS site. The spectral types corresponding to the $B - V$ colour indices were taken from Fitzgerald [15] and those corresponding to the 2MASS colour indices were taken from Straižys and Lazauskaitė [16].

Table 1. V magnitude, $B - V$ and 2MASS colour indices and corresponding spectral types for ASAS 093818–675524 and ASAS 120036–391536.

ASAS ID	V	$B - V$	$J - H$	$H - K$	Sp. Type ($B - V$)	Sp. Type ($J - H$)	Sp. Type ($H - K$)
093818–675524	10.26	0.73	0.30	0.08	G7V	G5V	G5V
120036–391536	10.45	0.80	0.41	0.07	K0V	K0/1V	G0V

Running the phase 0.5 spectra for both stars against the template spectra resulted in a spectral type classification of F7V for ASAS 093818–675524 and G6V for ASAS 120036–391536. Figure 3 shows the spectrum of ASAS 093818–675524 along with the spectrum of an F7V star. Figure 4 shows the spectrum of ASAS 120036–391536 along with the spectrum of a G6V star. The spectroscopically determined spectral types of both stars differs from the spectral types determined from the photometric colour indices.

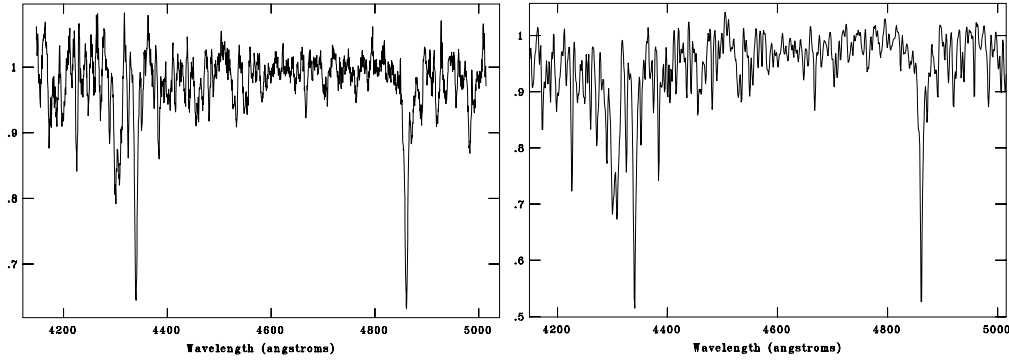


Figure 3. Spectrum of ASAS 093818–675524, left, and the best matching spectrum, F7V, on the right.

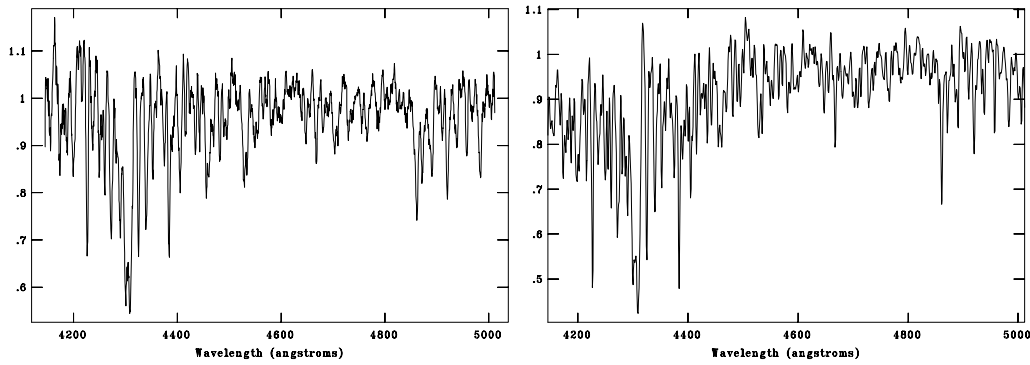


Figure 4. Spectrum of ASAS 120036–391536, left, and the best matching spectrum, G6V, on the right.

4. Discussion

Spectroscopic mass ratios and spectral types for two total eclipsing EC stars, ASAS 093818–675524 and ASAS 120036–391536, have been determined from analysis of spectroscopic data. Table 2 lists the results of the analysis. For both stars, the spectroscopically determined mass ratios are different to the photometrically determined mass ratios. The spectral types for both are also found to differ to the spectral types associated with the $B - V$ and 2MASS colour indices of the stars. The results illustrate the importance of using spectroscopic data to determine mass ratios and spectral types.

The spectroscopic mass ratios and spectral types have been used to determine the subclass type for both stars. For ASAS 093818 – 675524, the mass ratio and spectral type are consistent with the A-type subclass. For ASAS 120036 – 391536, the mass ratio and spectral type are consistent with the W-type subclass.

Acknowledgments

This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France and the SAO/NASA Astrophysics Data Stream. This research was funded by the NRF Thuthuka Bursary programme.

Table 2. Mass ratios and spectral types determined for ASAS 093818–675524 and ASAS 120036–391536.

ASAS ID	q	Sp. Type
093818–675524	0.210	F7V
120036–391536	0.315	G6V

References

- [1] Shore S N 1994 *Saas-Fee Advanced Course 22: Interacting Binaries* ed Shore S N, Livio M and van den Heuvel E P J
- [2] Binnendijk L 1970 *Vistas. Astron.* **12** 217
- [3] Csizmadia S and Klagyivik P 2004 *Astron. Astrophys.* **426** 1001
- [4] Deb S and Singh H P 2011 *Mon. Not. R. Astron. Soc.* **412** 1787–1803
- [5] Terrell D and Wilson R E 2005 *Astrophys. Space Sci.* **296** 221–230
- [6] Wilson R E 1978 *Astrophys. J.* **224** 885–891
- [7] Rucinski S M, Pych W, Ogłóza W, DeBond H, Thomson J R, Mochnacki S W, Capobianco C C, Conidis G and Rogoziecki P 2005 *Astron. J.* **130** 767–775
- [8] Rucinski S M and Lu W 1999 *Astron. J.* **118** 2451–2459
- [9] Niarchos P G, Hoffmann M and Duerbeck H W 1994 *Astron. Astrophys.* **292** 494–500
- [10] Vinko J, Hegedus T and Hendry P D 1996 *Mon. Not. R. Astron. Soc.* **280** 489–497
- [11] Hilditch R W 2001 *An Introduction to Close Binary Stars* (Cambridge University Press)
- [12] Prša A and Zwitter T 2005 *Astrophys. J.* **628** 426
- [13] Wilson R E and Devinney E J 1971 *Astrophys. J.* **166** 605
- [14] Wadhwa S S and Zealey W J 2005 *Astrophys. Space Sci.* **295** 463
- [15] Fitzgerald M P 1970 *Astron. Astrophys.* **4** 234
- [16] Straižys V and Lazauskaitė R 2009 *Balt. Astron.* **18** 19–31

Characterization of variable stars using the ASAS and SuperWASP databases

Derck P. Smits

Dept of Mathematical Sciences, PO Box 392, UNISA, 0003, South Africa

E-mail: smitsdp@unisa.ac.za

Abstract. A photographic survey of an area of the sky centred on RA (1900) $12^h 40^m$ and Dec (1900) $-32^\circ 00'$ made by Harvard Observatory in the 1930s identified 108 variable stars. Many of these variables have not been studied since their identification, and therefore there are no published light curves or accurate periods for these stars, and the data in the General Catalogue of Variable Stars (GCVS) is based on the original survey. Sources brighter than $V = 14$ mag can be found in the ASAS and SuperWASP databases, and can be used to check the classifications and parameters of the variable stars. In particular, in the SuperWASP archive there are over 25 000 observations for many of the sources, making it possible to determine accurate periods, maximum and minimum magnitudes and correctly classify these objects. The methods used to determine the parameters of these variable stars will be described, and results presented of the stars that have been studied.

1. Introduction

High Latitude Field (HLF) 356 for epoch B1900 is centred on right ascension $\alpha_{1900} = 12^h 40^m$ and declination $\delta_{1900} = -32^\circ 00'$. The field covers the range of RA between $12^h 20^m$ and $13^h 00^m$ and Dec between $-27^\circ 00'$ and $-37^\circ 00'$, an area of about 80 square degrees. J2000.0 coordinates for this position are $\alpha_{2000} = 12^h 45^m 23^s$, $\delta_{2000} = -32^\circ 32' 49.7''$ with Galactic coordinates $l = 301.4540^\circ$ and $b = 30.307^\circ$. This region lies way above the Galactic plane in the Hydra/Centaurus region and contains no obvious clusters (open or globular) or any nebulae (emission, reflection, dark or planetary).

Huruhata [1] used 116 plates of HLF 356 taken between 1935 and 1938 on two telescopes to identify variable stars. He reported 108 new variables, and a further 18 stars with ranges of less than half a magnitude as suspected variables. The numbers of each type of star in five classifications are listed in Table 1. More than 50% of the 108 variables are classified as RR Lyrae (hereafter RR) stars.

Many of the stars in Huruhata's list of 108 stars have been given GCVS designations, all others, including the 18 suspected variables, have entries in the catalogue of New Suspected Variables (NSV) [2]. Many of these stars have never been studied in any detail, and some of those that have been subjected to further analysis have been found to be classified incorrectly. Accurate periods for most of this sample of variables have not been determined previously.

2. Data

Data for these stars have been looked for in the All Sky Automated Survey (ASAS) [3] and the Wide Angle Search for Planets (SuperWASP, hereafter SW) [4]. Photometric observations for

Table 1. Numbers of new variable stars reported by Huruhata [1].

RR Lyr	Eclipsing	Short	Mira	Irregular	Total
59	23	6	11	9	108
0	7	10	0	1	18

the brighter stars can be found in the ASAS Catalogue of Variable Stars (ACVS), together with an estimated period and epoch for the light curve, as well as a suggested classification. For some of the fainter stars measurements are available in the ASAS All Stars Catalogue (AASC), but no periods or classifications are provided for these data.

For the SW data, periods are determined but no classifications for the type of variable are provided. The SW telescopes are fitted with filters giving a passband from 400 – 700 nm. The spatial resolution of the data is about 30 arcsec for the smallest aperture used; stars closer than this are blended into one light curve which can lead to the change in magnitude being diluted by a brighter star in the aperture. Because each SW telescope consists of several separate cameras, each with its own response, there can be small differences between data from different cameras. The SysRem algorithm [5] has been used to identify and remove patterns of correlated noise in the data. This produces detrended data which has been used to investigate many of the objects discussed here.

Towards HLF 356 there are about 25 000 SW observations per object. For variables with regular periods of hours to days, the data can be folded on the period to produce phase-magnitude diagrams (PMD), from which it is possible to classify the star. By binning the data at phase intervals of 0.01 there are $\sim 200 - 300$ points per bin, which allows an average curve with small uncertainties to be generated. There are at least four seasons of SW data and mean curves can be fitted to each season. By looking at the relative phase of the maxima and/or minima in the PMD from each season, the period of regular variables can be determined to an accuracy of 10^{-6} d.

At this stage no attempt has been made to clean the detrended data although it is quite clear that there are days when the data is so scattered that it should be discarded. The detrended SW data were run through the programme Period04 [6] to determine if there are any periodicities present in the data. The periods from this analysis were then used to fold the data. In most cases the periods were reliable, but in a few instances, particularly when the source is faint, adjustments were made by looking at the phase over different seasons.

For long period systems, such as Miras and semi-regular pulsators, the detrended data is distorted by the SysRem algorithm and it is better to use the archived SW (which has not been SysRem corrected) or ASAS data. Because neither the periods nor the amplitudes in these systems are constant from one cycle to the next, the data cannot be folded to produce PMDs. For bright sources light curves can be generated but for faint sources it is difficult to determine the reliable data from noisy signals so they have not been studied beyond noting that they are irregular or semi-regular pulsators.

3. Results

The detection limit of the Huruhata [1] survey is around magnitude 16. This is similar to the limit of the SW data, but fainter than the ASAS limit which is ~ 14 mag. Figure 1 shows an example of an SW PMD in which the curve goes below a magnitude of 16. This is the faintest source detected in the SW data. The photometry is not that accurate, but it is possible to

determine a period by folding the data, and identify that this is an eclipsing contact binary (EW in the GCVS classification scheme).

There is only one star for which there are ASAS data but no SW data. A PMD of the ASAS data is shown in figure 2. The plot was made using 601 cleaned data points that cover a baseline of 8.7 years. The period has been adjusted from that determined by ASAS to produce the very sharply defined upward phase of the light curve. From the plot it is clear that this star can be classified as R Rab.

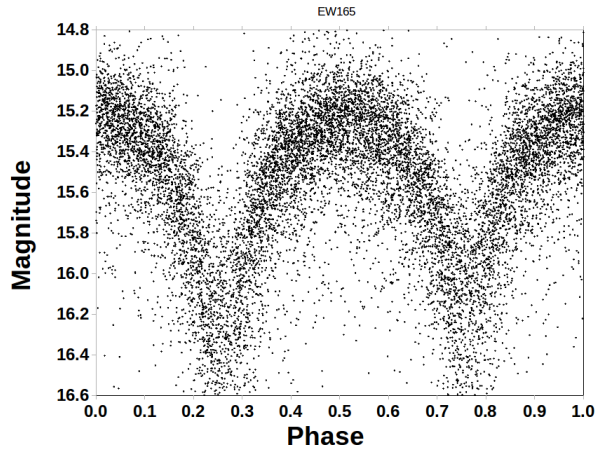


Figure 1. A W UMa-type variable (EW) that was classified as RR in [1].

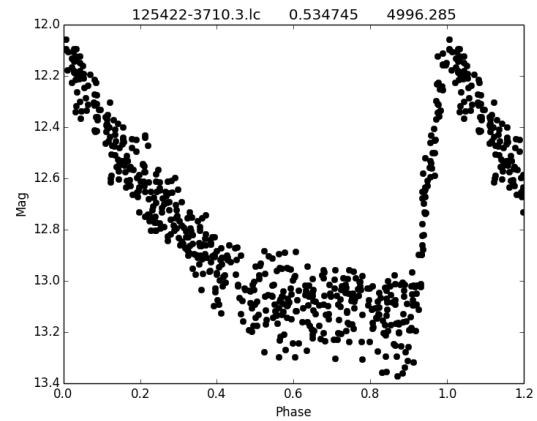


Figure 2. The only star in the sample for which there are ASAS data but no SW data.

For 31 of the 108 variables and 12 of the 18 suspected variable stars there are no data in either ASAS or SW. Their distribution in terms of classification are listed in Table 2. Independent searches have identified one RR and two Irregulars from this list.

Table 2. Numbers of variables for which there is no data in ASAS or SW.

RR Lyr	Eclipsing	Short	Mira	Irregular	Total
23/59	2/23	2/6	1/11	3/9	31/108
0	6/7	5/10	0	1/1	12/18

A few targets had stars that show no obvious variability. In these instances, stars close to the target were looked for in SW to see if any of them showed variability. Figure 3 shows a $10' \times 10'$ field of a POSS2/UKSTU Red image retrieved from the Digitized Sky Survey around one of the target stars. According to Huruhata [1] the target RR star varies between 14.4 – 15.2 mag. As indicated on the chart, only a non-varying pair of stars is at the specified position. Huruhata [1] makes no mention of this star being a binary (as noted for other stars). Nearby stars for which there are SW data are circled in blue and their SW magnitudes indicated. Stars that are not recognised by SW are indicated by red circles.

The 70 systems for which positive identifications have been made and periods determined include 35 RR, 21 eclipsing and 5 Miras. Light curves of stars showing variability but for which Period04 could not identify any reasonable periods have been classified as Irregulars. Two of

these are known dwarf novae and one of them appears to be a long period W Vir-type star. The GCVS classification of this star is CWA. For long period variables and semi-regular pulsators (Irregulars) the ASAS data often provide a better light curve than the SW data. An example of the ASAS light curve of one of the Miras in this sample is shown in figure 4.

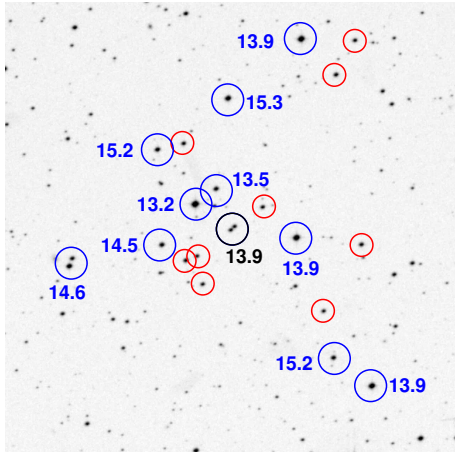


Figure 3. A $10' \times 10'$ field around a target star. The numbers indicate the SW magnitude of stars showing no variability.

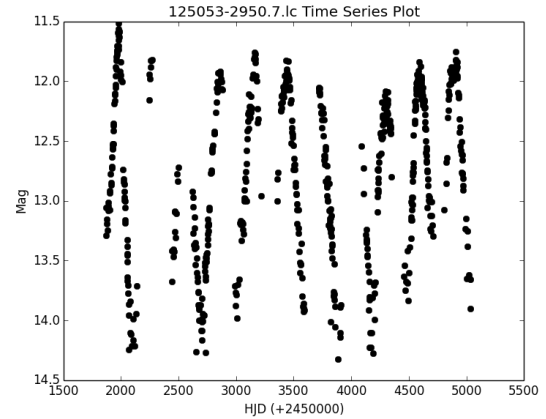


Figure 4. The light curve of a Mira made using ASAS data. The amplitude of the peaks vary from one cycle to the next, and the period is not constant.

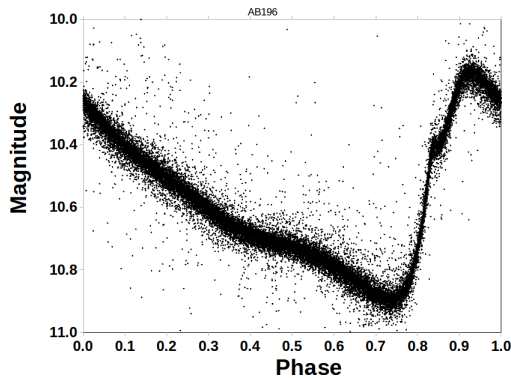


Figure 5. The PMD of an RRab type showing a rapid rise and slow decline with a hump at phase 0.7 and a bump at phase 0.9.

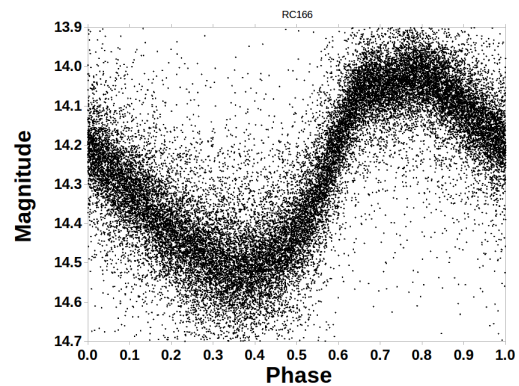


Figure 6. The light curves of RRc types are nearly sinusoidal. There is a bump in the light curve just before maximum.

Of the 35 RRs identified, 29 are type RRab and 6 are RRc types. The light curves of RRab types have an asymmetrical shape with a rapid rise and a slow decline. Features such as humps (at $\phi \sim 0.7$) and bumps (at $\phi \sim 0.95$) can be seen in the PMD for many of the RRab systems (see figure 5), and several RRc types also show bumps just before the maximum, as shown in figure 6. The nearly sinusoidal shape of the light curve together with its period identifies this as RRc.

The eclipsing binaries are made up of 7 detached (EA), 1 semi-detached (EB) and 13 contact (EW) systems. One of the EW stars was not part of Huruhata's original list, but it lies within

the field studied and, with a magnitude of 9.6, should have been detected.

In figure 7 the light curve of a detached eclipsing binary system is shown. The SysRem data were used to generate the PMD shown on the left, and ASAS data for that on the right. From the plot on the left, the magnitude of the system appears to be varying from one season to the next. The period of the system is close to 1 day and therefore less than half the cycle can be observed during the course of a night. Because the phase changes slowly during the course of a season, only a part of a cycle can be observed. For example, in one season only the secondary eclipse was visible but in other seasons part of the primary eclipse was observed. The SysRem algorithm tries to correct the light curve for changes from night to night and season to season using a mean value of the curve. Because the mean changes with the season, the corrections have shifted the curves so that they do not match up.

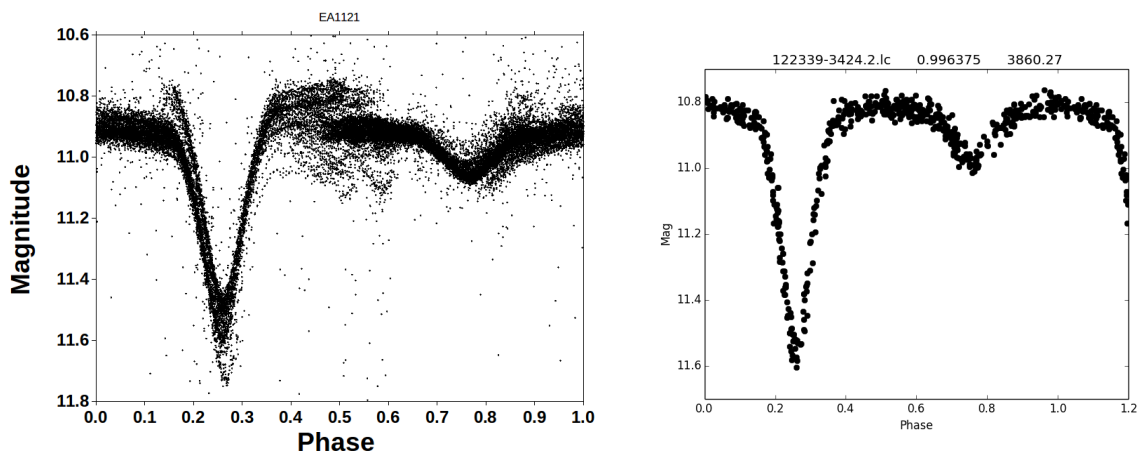


Figure 7. PMDs made from SW (left) and ASAS (right) data for an EA system with a period $P \sim 1$ day. The SysRem algorithm has affected the SW data.

The light curve amplitude of contact binaries is a maximum when the orbital inclination is edge-on ($i = 90^\circ$), with a typical value of $\Delta V = 0.8$. The shape of the light curve changes with inclination, becoming more sinusoidal as the inclination decreases. Most of the EW systems in this sample have ΔV in the range $0.8 - 0.3$ mag. A light curve with a small ΔV but a characteristic EW shape suggests a system with a third body whose light is not part of the variable system. An example of a system whose variations are anomalously small is shown in the left-hand panel of figure 8. The light curve has been folded on a period of 0.362716 d to produce the PMD shown in the diagram. This is clearly an EW system but for a star with a magnitude of 12, the variations go from a maximum of 12.22 to a minimum of 12.32, i.e. $\Delta V \sim 0.1$. According to the NSV catalogue the star labelled NSV on the finder chart is the variable. A Fourier analysis of the SW data for the NSV star does indicate a periodicity of 0.362716 but the amplitude is less than 0.1 mag. The folded SW data for the star labelled 2MASS has a larger amplitude than NSV, but the variations are small and not sinusoidal, suggesting contamination from a third star rather than an inclined system.

To investigate which star is varying, observations were made through a V-band filter with an SBIG ST9E CCD camera attached to the 35 cm Schmidt-Cassegrain telescope in the Unisa Observatory. The system has a field of view of about $9' \times 9'$ which is slightly smaller than the finder chart shown in figure 8. Differential photometry was done on several of the stars, in particular on those marked NSV, 2MASS and the one labelled with the red arrow. From the Unisa Observatory observations it can be shown that the variable is the star labelled with the red arrow, which is fainter than the other two stars. It is not possible with the SW system to

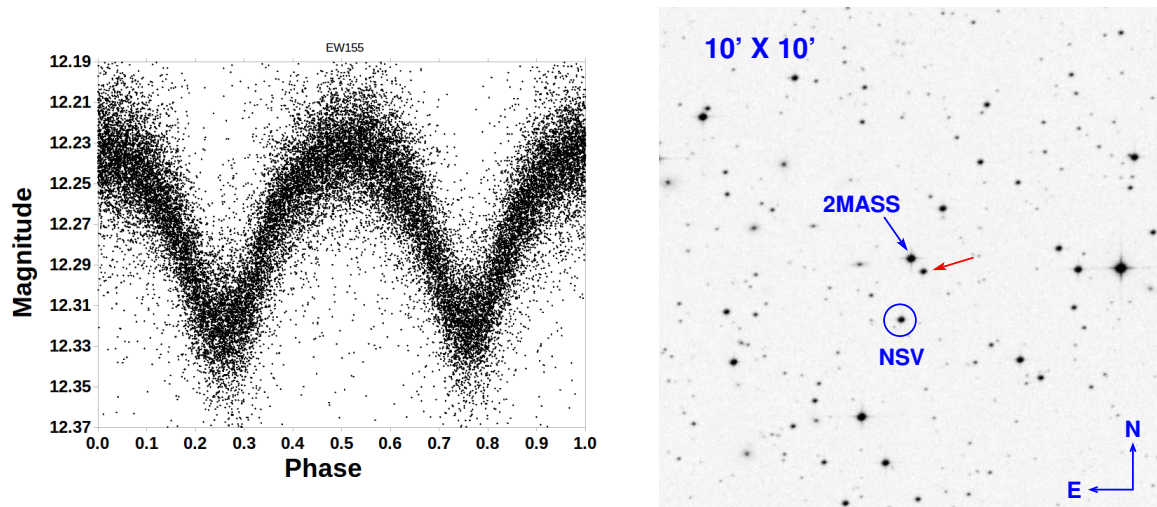


Figure 8. The amplitude of the EW variations are too small for a 12th magnitude star. The coordinates given in the NSV catalogue for this variable are incorrect. The variable is indicated by the red arrow on the finder chart.

resolve the 2MASS and variable star spatially, but it appears that some of the flux from the variable is even creeping into the NSV star observations, producing the small variation that was detected.

Several other stars in the sample display similar properties to this, i.e. the amplitudes of the variations are too small relative to the magnitude of the star, suggesting that the variable is a faint star in the aperture being swamped by light from a bright star. Further observations will need to be done on these systems to identify the variable stars and determine their apparent magnitudes.

4. Conclusion

Using data from the ASAS and SW projects the light curves of 108 variable stars and 18 suspected variables in HLF 356 have been investigated. For 70 systems data are available in ASAS and/or SW, making it possible to determine the period and classification of RR, eclipsing and Mira variables. Data for several irregular variables have also been retrieved, but this data have not been processed yet.

References

- [1] Huruhata M 1940 *Harvard College Observatory Bulletin* **913** 14–17
- [2] Kukarkin B V and Kholopov P N 1982 *Moscow: Publication Office “Nauka”*
- [3] Pojmanski G 2002 *Acta Astronomica* **52** 397–427
- [4] Pollacco D L *et al* 2006 *PASP* **118** 1407–18
- [5] Tamuz O, Mazeh T and Zucker S 2005 *MNRAS* **356** 1466–70
- [6] Lenz P and Breger M 2005 *CoAST* **146** 53–136

The excitation of pulsation modes in rapidly rotating main sequence B-stars

P van Heerden^{1,2}, C Engelbrecht¹ and P Martinez²

¹ Department of Physics, University of Johannesburg, Kingsway, Auckland Park, Johannesburg 2092, South Africa

² South African Astronomical Observatory, Observatory Rd, Observatory, Cape Town 7925 South Africa

E-mail: pierre@sao.ac.za

Abstract. This article summarises the current progress of a project to characterise the internal rotation characteristics of hot, young stars using pulsating members of this group. So far, multi-colour photometry has been collected for 11 stars using the 0.5 – m telescope at the Sutherland station of the SAAO on 81 nights during 2013 and 2014. For 6 of these stars, sufficient photometry has been collected to potentially allow the identification of the star's pulsation frequencies and to find the quantum numbers (n , l and m) of the associated pulsation modes. A preliminary frequency analysis was performed on the photometry of one of the stars, HD 81589, which yielded 4 frequencies that are probably due to pulsations of the star. In order to simplify and expedite the reduction and analysis of photometry, a database system was developed along with a set of data reduction and analysis tools that communicate directly with the database. This article reports the results of an initial period analysis of data reduced with these newly developed tools.

1. Introduction

All stars rotate - some quite rapidly, others quite slowly - which causes a breakdown in the spherical symmetry of stars, which (in turn) complicates and reduces the accuracy of stellar models, seeing as most stellar models in use today rely on the assumption of spherical symmetry to simplify the analysis to a manageable level. Stellar rotation also gives rise to various fluid dynamical phenomena, which result in large uncertainties in the rates of stellar evolution and physical stellar parameters in general. Treating rotation consistently in stellar models is unfortunately very difficult and has (until recently) been mostly neglected in studies of the structure and evolution of stars. Good accounts of the scientific problem appear in [1] and in [2].

An intensive observational study of the internal rotation dynamics of stars can therefore provide a wealth of information to support further theoretical work. Asteroseismology is a unique tool for probing stellar interiors - in this context to determine the internal rotation dynamics of stars using the pulsations that are excited deep within the stellar interior. The Beta Cephei class of pulsating stars are ideal for such a study, since they typically exhibit multiple non-spherical pulsation modes and fairly rapid rotation (typically 100 – 300 km/s). The rapid rotation makes the frequency spacing between rotationally split pulsation modes sufficiently large to be distinguished in photometry collected over approximately 1 year [3].

Table 1. Stars selected for this study.

Name	RA (J2000.0)	Dec (J2000.0)	Spect. Type	V_{mag}	No. freq[5]
HD 37115	05:35:54	-05:37:42	B7Ve	7.07	2
HD 47416	06:38:51	+07:58:18	B5V	7.80	2
HD 81589	09:24:43	-53:00:43	A3/4	8.90	2
HD 298682	09:49:37	-51:52:21	B0.5 II	9.06	2
HD 303142	10:40:00	-57:32:59	B0 (V)	9.73	2
HD 133823	15:09:55	-65:30:23	B2 IV	9.62	4
CPD -53 6701	15:48:40	-54:14:38	B D	10.37	2
HD 326305	16:52:56	-41:35:08	B0 V	9.97	3
HD 324369	17:48:08	-42:17:12	B1 II	10.00	3
HD 314893	18:01:18	-23:56:43	B3	9.88	3
CPD -31 6271	20:28:20	-31:27:06	B2	10.30	2

Even one strong detection of differential rotation rates would add great value to current theoretical efforts to model internal rotation, as the sample of observationally-detected rotation rates in B stars is very sparse. As there are eleven stars in our sample, we anticipate a strong likelihood of providing very significant results for this purpose. It will be of particular interest to compare pulsationally-determined rotational splitting with the predictions of different theoretical studies, using the spectroscopically-determined photospheric rotation velocities (via spectral line broadening).

Section 2 lists the candidates selected for this study as well as the criteria by which they were selected. Section 3 gives a summary of the observations made to date. Section 4 describes the software systems employed in managing, reducing and analysing the data. Section 5 gives the preliminary results for one of the candidate stars. Section 6 describes the work that needs to be completed during the remainder of this project.

2. Candidates

Over 100 new Beta Cephei pulsating stars have been identified from the All-Sky Automated Survey data - they are listed in [4] and in [5]. Several dozen of these show multiple pulsation modes, which makes them excellent candidates for this study, since the probability of finding two (or more) rotationally split modes is higher for these stars. Since pulsation modes of different order (quantum number n) originate from different depths of the radiative envelope of these stars, finding two rotationally split pulsation modes of different order in one of the candidates enables the determination of the rotation rate at two depths within the star (and similarly for stars with more than two rotationally split modes of different order).

The accessibility through the year of the stars at Sutherland and the apparent magnitude of the stars were also considered. Most of the stars listed by Pigulski & Pojmanski have right ascensions between 7 and 20 hours, making the winter months in Sutherland ideal for observing them.

Eleven stars were selected from the lists of Pigulski & Pojmanski - their details are given in table 1. A magnitude cut-off of $V = 10.5\text{mag}$ was used, since such bright stars are still easily observed with the 0.5 – m telescope (in terms of centering the star in the photometric aperture and the required integration times in order to obtain a sufficient signal-to-noise ratio) and there are many stars in the lists of Pigulski & Pojmanski that satisfy this criterion.

Table 2. Summary of observations done to date.

Name	No. Obsns	Hours observed	First JD	Last JD	Last-First JD
HD 37115	1243	17.32	2456527	2456621	93.771
HD 47416	202	2.80	2456579	2456621	41.906
HD 81589	2299	47.96	2456407	2456753	346.195
HD 298682	1192	24.14	2456407	2456783	376.101
HD 303142	285	10.40	2456407	2456832	425.020
HD 133823	819	31.23	2456407	2456780	373.004
CPD -53 6701	565	29.49	2456407	2456831	424.053
HD 326305	377	11.38	2456407	2456778	371.186
HD 324369	335	11.09	2456407	2456579	171.662
HD 314893	337	14.81	2456454	2456832	378.003
CPD -31 6271	625	28.43	2456499	2456831	332.158

3. Summary of Observations

The stars listed in section 2 have been observed on 81 nights, resulting in 283 hours of photometry, some of which may be unusable due to environmental conditions (specifically atmospheric transparency variations). A summary of the observations is given in table 2.

Approximately 1000 data points are required in each filter in order to get a sufficient signal-to-noise ratio in the frequency domain and so determine the pulsation frequencies of the star. Furthermore, the time between the first and last observation must be on the order of 1 year in order to get sufficient resolution in the frequency domain. So far there are 3 stars that match these criteria: HD 81589, HD 298682 and HD 133823.

In addition to the photometry collected to date, one spectrum has been recorded for each of HD 81589 and HD 298682 using the Cassegrain spectrograph on the 1.9 – m telescope in Sutherland. Spectra for the remaining stars will be collected during an observing run scheduled on the 1.9 – m telescope during quarter 3 of 2014. All the candidates that are accessible by the Southern African Large Telescope during the current semester have also been added to the queue to be observed with SALT using the Robert Stobie Spectrograph.

4. Database System and Associated Software Tools

Software tools were developed to upload the data to the database from the text-based file format produced by the photometer control software on the 0.5 – m telescope, to calculate the geocentric and heliocentric Julian days of the observations, to calculate and apply the dark current and dead time corrections, to perform the sky background subtraction and finally to calculate the corrections to the colour equations (most notably the coefficients of air-mass dependent terms) from the standard star measurements on a particular night and apply them to the data in order to obtain a standardised magnitude for each datum.

Naturally, these new software tools had to be vetted by comparing the results to the results given by established data reduction software or tabulated data. The entire vetting procedure was quite lengthy and meticulous and so falls beyond the scope of this document, figure 1 shows the differences between the results produced by the new software tools developed as part of this project with those of the XReduce software package that is currently used at the SAAO to reduce photometry.

Figure 1 shows a comparison of the new software developed as part of this project with the standard reduction software in use at the SAAO. The fact that all the corresponding points in

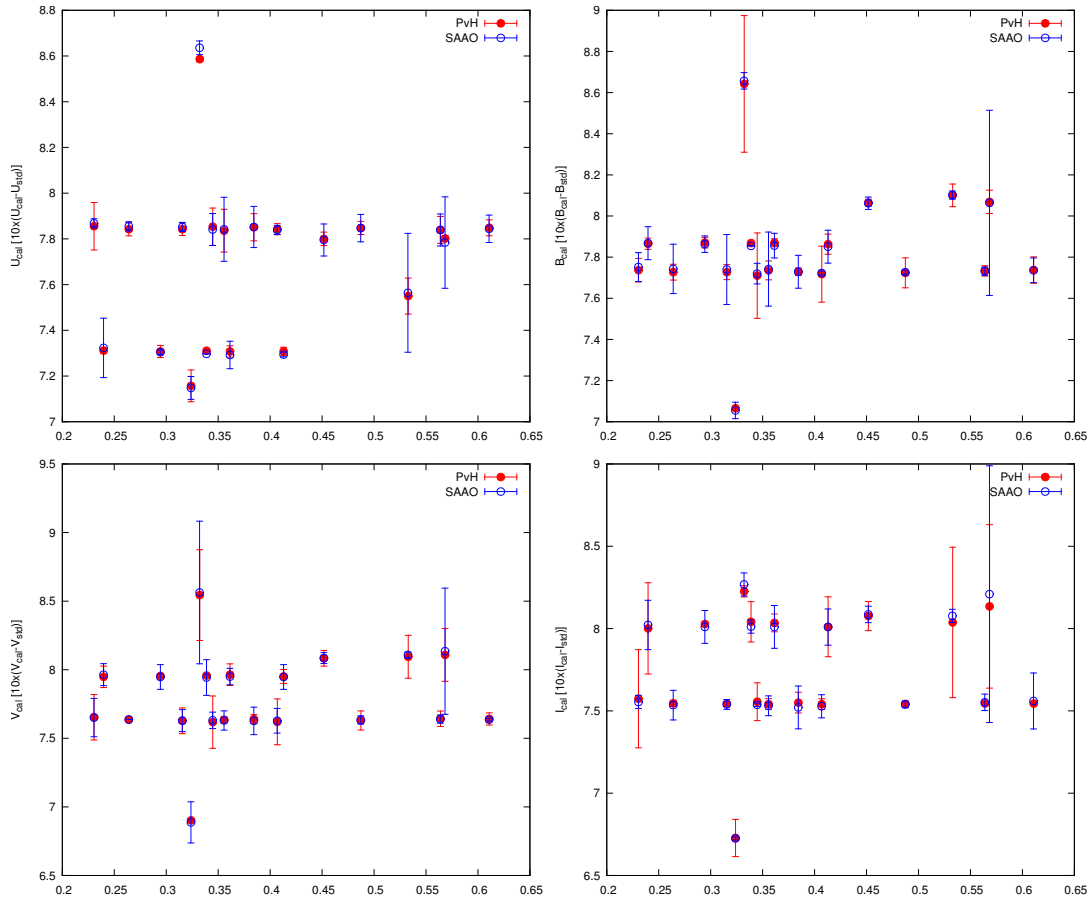


Figure 1. Comparison between the standardised magnitudes produced by the new software developed as part of this project (“PvH”) with those produced by the current standard photometry reduction software in use at the SAAO (“SAAO”). All the standard star measurements on two nights (JD 2456408 and JD 245669, two nights of acceptable photometric quality) are combined in these plots. The x-axis shows the fractional HJD (with the integral JDs 2456408 and 2456699, respectively, removed) and the y-axis shows the calibrated magnitudes in the Johnson-Cousins filters used during this study. The error bars show the difference between the standardised magnitudes produced by the two software programmes and the true magnitudes of the standard stars - the difference is typically under 10mmag and so is multiplied by a factor of 10 in these plots for display purposes.

the two sets are virtually coincident indicates that the calibrated (or standardised) magnitudes produced using the new software agree well with those produced using the existing software. The error bars, which indicate an exaggeration of the magnitude difference between the calibrated magnitudes of the observed standard stars and their true magnitudes, further indicate that although there are a small number of cases in both system where this difference is large, neither system is clearly more accurate than the other. Figure 1 therefore indicates that the new reduction system is at least as accurate as the existing one.

5. Preliminary Results

The Johnson U filter photometry of HD 81589 was put through the data reduction system, in the process removing any erroneous data that could be attributed to poor photometric conditions.

Table 3. Frequencies identified in the HD 81589 U filter photometry. Frequency number 1 (marked with an asterisk) is probably a result of atmospheric variations. The specified phases are relative to the epoch JD 2450000.

No.	Frequency (cycles/day)	Amplitude (mag)	Phase (rad)
1*	0.387	0.0377 ± 0.001	2.573 ± 0.026
2	3.060	0.0208 ± 0.001	6.216 ± 0.055
3	8.34	0.0140 ± 0.001	5.461 ± 0.078
4	6.7	0.0147 ± 0.001	4.577 ± 0.076
5	5.15	0.0130 ± 0.001	4.623 ± 0.094
6	14.47	0.0083 ± 0.001	1.025 ± 0.127

In total, 1347 data collected over 19 nights were found to be viable, which were used in a frequency analysis process using the Lomb-Scargle algorithm.

Figure 2 shows the power spectrum produced from the photometry and table 3 lists the frequencies identified from it. Please note that this is simply a preliminary analysis. Note also that the lower frequency variations present in the power spectrum (most notably the peak at 0.387 cycles per day) are probably due to atmospheric transparency variations and do not represent pulsations within the star.

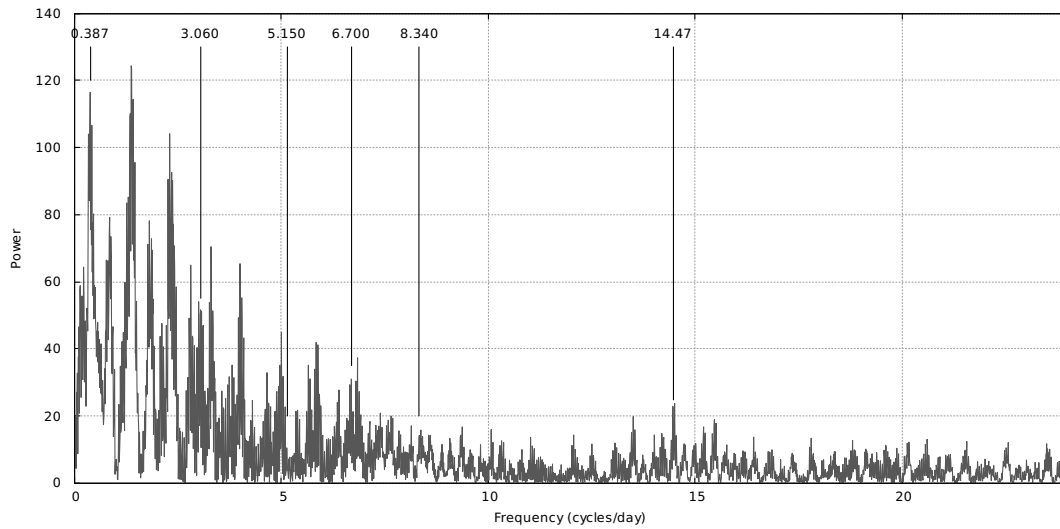


Figure 2. Power spectrum of the Johnson U filter photometry of HD 81589 with the identified frequencies indicated. Once the low-frequency signals are removed from the data, these peaks become prominent in the periodogram.

It is clear from figure 2 that sufficient data have been collected for this star - the focus of the observing campaign may now be shifted to the other candidates while a comprehensive frequency analysis can be done on the HD 81589 photometry in all available photometric filters. This is likely to be the case for HD 298682 and HD 133823 as well.

6. Future Work

The photometric component of the observing campaign for the candidates listed in table 2 is ongoing and is likely to continue until the 0.5 – m facility in Sutherland is closed down. At this time the observational burden will be shifted to one of the other manned (“small”) telescopes in Sutherland (e.g. the 1.0 – m) or to the Alan Cousins Telescope, which is expected to be ready for regular automated operation within the next few months. Spectroscopic observations of the candidates will (as mentioned in section 3) be continuing on the 1.9 – m telescope using the Cassegrain spectrograph and on SALT using RSS for the remainder of 2014.

The data reduction and analysis efforts are continuing and, as indicated in section 5, are likely to produce several positively identified pulsation frequencies for a number of stars in the near future.

Once the pulsation frequencies of a star are identified, the quantum numbers (n , l , m) corresponding to the pulsation frequencies are found by comparing the frequencies and relative amplitudes in a number of photometric filters to those predicted by stellar pulsation models for a model star similar in spectral type to the target star. This can be achieved by, for example, the FAMIAS software package (see [6]), provided that the effective temperature and metallicity parameters are fairly well-known. Pulsation modes with the same n and l , but different m are degenerate in the absence of rotation, but in the case of a rotating star can be distinguished from one another provided that there is sufficient resolution in the frequency domain. The frequency spacing between modes with the same n and l can then be used to calculate the rotation rate at the level within the star where the pulsations originate.

References

- [1] Maeder A and Meynet G 2000 ARA&A **38** 143–190 (*Preprint astro-ph/0004204*)
- [2] Zahn J P 2008 *Communications in Asteroseismology* **157** 196–202
- [3] Pamyatnykh A A 2003 Ap&SS **284** 97–107
- [4] Pigulski A and Pojmański G 2008 A&A **477** 917–929 (*Preprint 0711.2530*)
- [5] Pigulski A and Pojmański G 2010 *Private communication*
- [6] Zima W 2008 *Communications in Asteroseismology* **157** 387

Spectral modelling of a H.E.S.S.-detected pulsar wind nebula

Carlo van Rensburg¹, Paulus Krüger¹ and Christo Venter¹

¹Centre for Space Research, North-West University, Potchefstroom Campus, Private Bag X6001, Potchefstroom, 2520

E-mail: 21106266@nwu.ac.za

Abstract. In the last decade ground-based Imaging Atmospheric Cherenkov Telescopes have discovered roughly 30 pulsar wind nebulae at energies above 100 GeV. We present first results from a leptonic emission code that models the spectral energy density of a pulsar wind nebula by solving the Fokker-Planck transport equation and calculating inverse Compton and synchrotron emissivities. Although models such as these have been developed before, most of them model the geometry of a pulsar wind nebula as that of a single sphere. We have created a time-dependent, multi-zone model to investigate changes in the particle spectrum as the particles diffuse through the pulsar wind nebula, as well as predict the radiation spectrum at different positions in the nebula. We calibrate our new model against a more basic previous model and fit the observed spectrum of G0.9+0.1, incorporating data from the High Energy Stereoscopic System as well as radio and X-ray experiments.

1. Introduction

In the last decade ground-based Imaging Atmospheric Cherenkov Telescopes (IACT_s) have discovered almost 150 very-high-energy (VHE; $E > 100$ GeV) γ -ray sources. Roughly 30 of these are confirmed pulsar wind nebulae (PWNe), while other source classes include supernova remnants, active galactic nuclei, or unidentified sources¹. A subset of the unidentified sources may eventually turn out to be PWNe. H.E.S.S. is currently the world's largest IACT and is located near the Gamsberg mountain in Namibia. H.E.S.S. consists of five telescopes which were built in two phases. In December 2003 Phase I consisting of four 12-m telescopes was completed. The fifth 28-m H.E.S.S. II telescope has been operational since July 2012, making H.E.S.S. the largest and most sensitive ground-based γ -ray telescope in the world.

A PWN is a bubble of shocked relativistic particles, produced when a pulsar's relativistic wind interacts with its environment [1]. The conversion of a fast-moving particle wind into electromagnetic radiation happens by two main processes. The leptons in the particle wind interact with the magnetic field of the nebula and this causes synchrotron radiation (SR) up to several keV. The second process is when the high-energy leptons interact with low-energy photons through inverse Compton scattering (IC), boosting the photon energies up to $E > 100$ GeV. Due to these two effects the radio, X-ray, and VHE γ -ray emissions are tightly linked, as all three emerge from the same lepton population. The following are typical characteristics of a PWN [7]:

¹ tevcat.uchicago.edu

- (i) A filled-centre morphology which is brightest at the centre and dimming in all directions towards the edges;
- (ii) A flat radio photon spectrum with an index between 0 and -0.3;
- (iii) A well organised internal magnetic field structure;
- (iv) A high level of linear polarisation at high radio frequencies;
- (v) Evidence of particle re-acceleration;
- (vi) Evidence of synchrotron cooling which means that the size of the PWN decreases with increasing energy.

There are many unanswered questions in PWN physics. Surprisingly, it was noted [3] that the measured γ -ray luminosity (1-10 TeV) of the PWNe does not correlate with the spin-down energy of their embedded pulsars. On the other hand, it was also found [4] that the X -ray luminosity is correlated with the pulsar spin-down energy. Furthermore, it is currently unknown whether there is any correlation between the TeV surface brightness of the PWNe and the spin-down energy of their pulsars. Due to this reason it is necessary to create a spatially dependent model to calculate the spectral energy density (SED) from the PWN. The spatial dependence will yield the flux as a function of the radius. We calibrate our model against a more basic previous model and fit the spectrum of a known PWN. We describe our model in Section 2, and its calibration and spectral fitting to G0.9+0.1 in Section 3. We discuss our conclusions in Section 4.

2. PWN Model

2.1. Transport equation and injection spectrum

We model the transport of charged particles in a PWN by solving a Fokker-Planck-type equation similar to the Parker equation [5]. This equation includes diffusion, convection, energy losses, as well as a particle source. If we neglect spatial convection, we are left with [6]

$$\frac{\partial N_e}{\partial t} = Q - \frac{\partial}{\partial E_e}(\dot{E}_e N_e) + \nabla \cdot (\vec{\kappa} \cdot \nabla N_e), \quad (1)$$

with Q the particle injection spectrum, N_e the particle spectrum, \dot{E}_e the particle energy loss rate, and $\vec{\kappa}$ the diffusion tensor. We used a broken power law for the particle injection spectrum following [2],

$$Q(E_e, t) = \begin{cases} Q_0(t) \left(\frac{E_e}{E_b} \right)^{\alpha_1} & E_e < E_b \\ Q_0(t) \left(\frac{E_e}{E_b} \right)^{\alpha_2} & E_e \geq E_b. \end{cases} \quad (2)$$

Here Q_0 is the normalisation constant, E_b the break energy, E_e the lepton energy and α_1 and α_2 the spectral indices. To obtain Q_0 we use a spin-down luminosity $L(t) = L_0 / (1 + t/\tau_0)^2$ of the pulsar [8], with τ_0 the initial spin-down time scale of the pulsar, and L_0 the initial spin-down luminosity. We normalise Q as follows:

$$\epsilon L = \int_{E_{\min}}^{E_{\max}} Q E_e dE_e, \quad (3)$$

with ϵ the conversion efficiency of the spin-down luminosity to particle power.

2.2. Radiative energy losses in the PWN

Particle energy is dissipated from the system due to radiation. We incorporated SR and IC scattering, similar to calculations done by [6] in their globular cluster model. SR losses are given by [9]

$$\left(\frac{dE_e}{dt}\right)_{\text{SR}} = -\frac{\sigma_T c}{6\pi} \frac{E_e^2 B^2}{(m_e c^2)^2}, \quad (4)$$

with σ_T the Thompson cross section and B the PWN magnetic field. The IC scattering energy loss rate is given by

$$\left(\frac{dE_e}{dt}\right)_{\text{IC}} = -\frac{g_{\text{IC}}}{E_e^2} \sum_{p=1}^3 \int \int n_{\varepsilon,p}(r, \varepsilon, T_p) \frac{E_\gamma}{\varepsilon} \hat{\zeta}(E_e, E_\gamma, \varepsilon) d\varepsilon dE_\gamma, \quad (5)$$

with $g_{\text{IC}} = 2\pi e^4 c$, ε the soft photon energy, $n_{\varepsilon,p}$ the blackbody photon number density, T_p the photon temperature of component p , E_γ the TeV up-scattered photon energy, and $\zeta(E_e, E_\gamma, \varepsilon) = \zeta_0 \hat{\zeta}(E_e, E_\gamma, \varepsilon)$ the collision rate with $\zeta_0 = 2\pi e^4 E_0 c / \varepsilon E_e^2$, and $\hat{\zeta}$ given by

$$\hat{\zeta}(E_e, E_\gamma, \varepsilon) = \begin{cases} 0 & \text{if } E_\gamma \leq \frac{\varepsilon E_0^2}{4E_e^2}, \\ \frac{E_\gamma}{\varepsilon} - \frac{E_0^2}{4E_e^2} & \text{if } \frac{\varepsilon E_0^2}{4E_e^2} \leq E_\gamma \leq \varepsilon, \\ f(q, g_0) & \text{if } \varepsilon \leq E_\gamma \leq \frac{4\varepsilon E_e^2}{E_0^2 + 4\varepsilon E_e}, \\ 0 & \text{if } E_\gamma \geq \frac{4\varepsilon E_e^2}{E_0^2 + 4\varepsilon E_e}. \end{cases} \quad (6)$$

Here, $E_0 = m_e c^2$, $f(q, g_0) = 2q \ln q + (1 - q)(1 + (2 + g_0)q)$, $q = E_0^2 E_\gamma / (4\varepsilon E_e (E_e - E_\gamma))$, and $g_0(\varepsilon, E_\gamma) = 2\varepsilon E_\gamma / E_0^2$. The two radiation loss rates can be added to find \dot{E}_e used in (1).

2.3. Calculation of the particle spectrum

If we use spherical coordinates and assume spherical symmetry, i.e. $\frac{\partial}{\partial \theta} = 0$ and $\frac{\partial}{\partial \phi} = 0$, the only change in particle (lepton) spectrum will be in the radial direction, so $\nabla^2 = \frac{1}{r^2} \left(\frac{\partial}{\partial r} \left(r^2 \frac{\partial N}{\partial r} \right) \right)$. For the diffusion scalar coefficient κ we consider Bohm diffusion

$$\kappa = \kappa_B \frac{E_e}{B}, \quad (7)$$

with $\kappa_B = c/3e$, and c and e denote the speed of light and the elementary charge. We can now discretise (1):

$$(1 - z + \beta)N_{i,j+1,k} = (1 + z - \beta)N_{i,j-1,k} + (\beta + \gamma)N_{i,j,k+1} + (\beta - \gamma)N_{i,j,k-1} + \frac{2}{\delta E_2 + \delta E_1} \left(r_a \dot{E}_{i+1,j,k} N_{i+1,j,k} - \frac{1}{r_a} \dot{E}_{i-1,j,k} N_{i-1,j,k} \right) + Q_{i,j,k} dt, \quad (8)$$

with i the time index, j the energy index, k the radial index, $\beta = 2\kappa dt / \Delta r^2$, $\gamma = 2\kappa dt / (r \Delta r)$ with Δr the size in the spatial bins. Also $r_a = \Delta E_{i+1} / \Delta E_i$ with

$$z = \left(\frac{1}{\Delta E_{i+1} - \Delta E_i} \right) \left(\frac{1}{r_a} - r_a \right) \dot{E}_{i,j,k}. \quad (9)$$

We initially approached the discretisation process by using a simple Euler method. It soon became clear that this method was not stable. We then decided to use a DuFort-Frankel scheme to solve (1). In solving this equation, we calculate the electron spectrum of the PWN due to

the injected particles from the embedded pulsar, taking into account their diffusion through the PWN and the IC and SR energy losses. We use the following parametrised form for $B(t)$ [2]

$$B(t) = \frac{B_0}{1 + \left(\frac{t}{\tau_0}\right)^{\alpha_B}}, \quad (10)$$

with B_0 the birth magnetic field, t the age of the PWN, and α_B the magnetic field parameter. We can calculate L_0 and B_0 by using the present-day luminosity and the present-day magnetic field together with the age of the PWN. We limit the particle energy using $E_{\max} = \frac{c}{2} \sqrt{\frac{L(t)\sigma}{c(1+\sigma)}}$ [2], with σ the ratio of electromagnetic to particle luminosity. Particles with $E_e > E_{\max}$ are assumed to escape.

To solve (8) numerically, our multi-zone model divides the PWN into shells. The particles are injected into zone one close to the centre and are allowed to diffuse through the different zones. We assumed as initial condition, that all zones were devoid of any particles. We used a reflective boundary condition at the inner boundary of the PWN, and on the outer rim we set N_e equal to zero to allow the electrons to escape from the PWN.

2.4. Calculation of radiation spectrum

The time-dependent photon spectrum of each zone can now be calculated, given the electron spectrum solved for each zone. For IC we have [6]

$$\left(\frac{dN_\gamma}{dE_\gamma}\right)_{IC} = \frac{g_{IC}}{A} \sum_{p=1}^3 \int \int n_{\varepsilon,p}(r, \varepsilon, T_p) \frac{\mathcal{N}_e}{\varepsilon E_e^2} \hat{\zeta}(E_e, E_\gamma, \varepsilon) d\varepsilon dE_e, \quad (11)$$

where $A = 4\pi d^2$, d the distance to the source, and \mathcal{N}_e is the number of electrons per energy in a spherical shell around r . We consider three components of target photons, cosmic background radiation (CMB) with $T_1 = 2.76$ K and $u_1 = 0.23$ eV/cm³, Galactic background infra-red photons with $T_2 = 35$ K and $u_2 = 0.5$ eV/cm³, and starlight with $T_3 = 4\,500$ K and $u_3 = 50$ eV/cm³. For SR we have

$$\left(\frac{dN_\gamma}{dE_\gamma}\right)_{SR} = \frac{1}{A} \frac{1}{hE_\gamma} \frac{\sqrt{3}e^3 B(r)}{E_0} \int \int_0^{\pi/2} \mathcal{N}_e(E_e, r) \tilde{\kappa} \left(\frac{\nu}{\nu_{cr}(E_e, \theta, r)} \right) \sin^2 \theta d\theta dE_e, \quad (12)$$

Table 1. Values of model parameters.

Model Parameter	Symbol	Value
Braking index	n	3
B -field parameter	α_B	0.5
Present day B -field	$B(T)$	40.0 μG
Conversion efficiency	ϵ	0.6
Age	T	1900 yr
Initial period of the pulsar	P_0	43 ms
Current period of the pulsar	P	52 ms
Pulsar spin-down rate	\dot{P}	1.5557×10^{-13}
Distance	d	8.5 kpc
Q index 1	α_1	-1.0
Q index 2	α_2	-2.6
Maximum energy parameter	σ	0.2
Soft photon component 1	T_1 and u_1	2.76 K, 0.23 eV/cm ³
Soft photon component 2	T_2 and u_2	35 K, 0.5 eV/cm ³
Soft photon component 3	T_3 and u_3	4500 K, 50 eV/cm ³

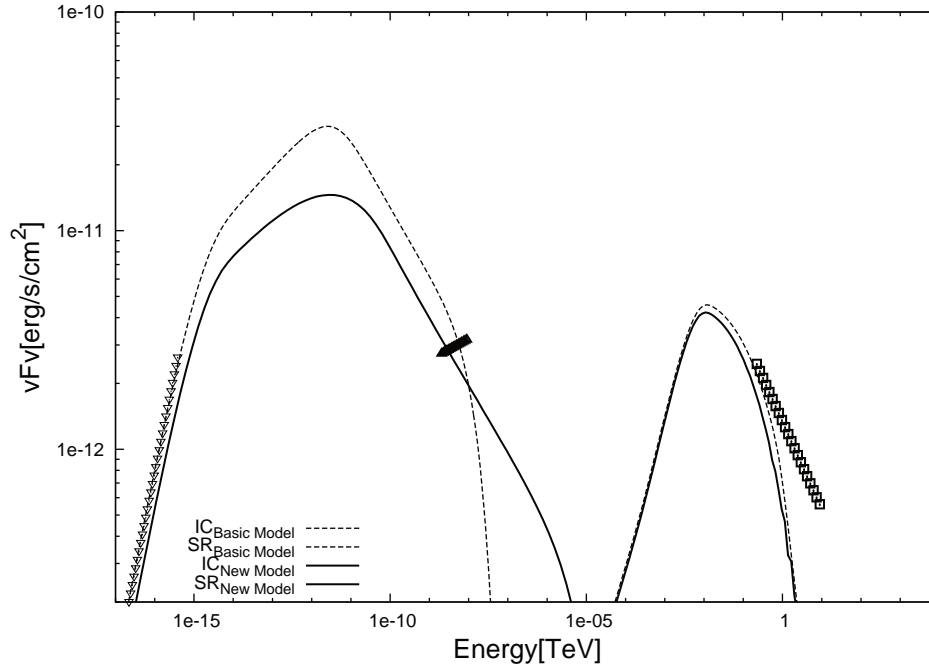


Figure 1. Calibration of our model against the model of [2]. The radio data are from [11], the X-ray data are from [12], and the γ -ray data are from [13].

with ν_{cr} the critical frequency (with pitch angle θ)

$$\nu_{\text{cr}}(E_e, \theta, r) = \frac{3ec}{4\pi E_0^3} E_e^2 B(r) \sin \theta \quad (13)$$

and $\tilde{\kappa}$ given by

$$\tilde{\kappa}(x) = x \int_x^\infty K_{5/3}(y) dy, \quad (14)$$

where $K_{5/3}$ a modified Bessel function of the second kind of order $5/3$.

2.5. Calculation of the line of sight flux

In order to do the line of sight (LOS) integration of the flux from the PWN we need to use the flux per unit volume calculated in the previous section and multiply it with the volume in a particular LOS as viewed from Earth. Since our model is spherically symmetric and since the source is very far from Earth, we used cylinders intersecting the spherical zones. We chose the cylinders and the spheres to have the same radii. We then used the intersection volume to calculate the flux in a particular LOS. We will present the radial dependence of the flux elsewhere.

3. Model calibration and spectral modelling of G0.9+0.1

To calibrate our model we compared our results to those of [2] as shown in figure 1. Pulsar J1747–2809 has now been discovered in PWN G0.9+0.1 [10] with $P = 52$ ms and $\dot{P} = 1.5557 \times 10^{-13}$. We used these values to calculate $\tau_0 = P_0/2\dot{P}_0$ assuming $P_0 = 43$ ms and no decay of the pulsar B -field. In the figure, the radio data are from [11], the X-ray data are from [12], and the γ -ray data are from [13]. We list our model parameters in table 1.

4. Conclusions

We created a time-dependent, multi-zone SED model to model the radiation spectrum observed from PWNe. We calibrated this model using a previous model. We then fitted the data of G0.9+0.1. The new model does not exactly reproduce [2], but the results are quite close. The new SR component is lower due to the fact that the old model did not consider IC losses while performing the particle-transport. We will next perform a population study to probe a potential relationship between the TeV surface brightness and the spin-down luminosity of the embedded pulsar. We also intend to investigate the spatial dependence of the different fluxes.

References

- [1] Gaensler B M and Slane P O 2006 *ARA&A* **44** 17–47
- [2] Venter C and de Jager O C 2007 *WE-Heraeus Seminar on Neutron Stars and Pulsars 40 years after the Discovery* ed Becker W and Huang H H p 40
- [3] Kargaltsev O and Pavlov G G 2010 *X-ray Astronomy 2009; Present Status, Multi-Wavelength Approach and Future Perspectives* **1248** 25–28
- [4] Mattana F, Falanga M, Götz D, Terrier R, Esposito P, Pellizzoni A, De Luca A, Marandon V, Goldwurm A and Caraveo P A 2009 *ApJ* **694** 12–17
- [5] Parker E N 1965 *ApJ* **142** 1086
- [6] Kopp A, Venter C, Büsching I and de Jager O C 2013 *ApJ* **779** 126
- [7] de Jager O and Venter C 2005 *Towards a Network of Atmospheric Cherenkov Detectors VII (astro-ph/0511098)* ed Degrange B and Fontaine G (*Preprint arXiv:astro-ph/0511098*)
- [8] Reynolds S P and Chevalier R A 1984 *apj* **278** 630–648
- [9] Blumenthal G R and Gould R J 1970 *Reviews of Modern Physics* **42** 237–271
- [10] Camilo F, Ransom S M, Gaensler B M and Lorimer D R 2009 *ApJ* **700** L34–L38
- [11] Becker R H and Helfand D J 1987 *ApJ* **316** 660–662
- [12] Sidoli L, Bocchino F, Mereghetti S and Bandiera R 2004 *MmSAI* **75** 507
- [13] Porquet D, Decourchelle A and Warwick R S 2003 *A&AS* **401** 197–203

Search for Very High Energy candidate sources using South African observatories

B van Soelen¹, L Klindt¹, IP van der Westhuizen¹, PJ Meintjes¹, M Böttcher², L Hanlon³, M Topinka³ and P Väisänen⁴

¹Department of Physics, University of the Free State, Bloemfontein, 9301, South Africa,

²Centre for Space Research, North-West University, Potchefstroom, 2520, South Africa,

³School of Physics, University College Dublin, Belfield, Dublin 4, Ireland, ⁴South African Astronomical Observatory, PO Box 9, Observatory 7935, Cape Town, South Africa

E-mail: vansoelenb@ufs.ac.za

Abstract. The multi-wavelength South African observatories are ideally located to complement the very high energy (VHE) observations undertaken with the H.E.S.S. telescope located in Namibia. We are undertaking a long-term multi-wavelength campaign and a literature search to identify potential Very High Energy extra-galactic sources which may be observable with Imaging Air Cherenkov Telescopes such as the H.E.S.S. telescope. The early stages of this project have focussed on identifying candidate sources and undertaking optical photometric observations with the UFS/Boyden 1.5-m and Watcher telescopes located at the Boyden observatory, and optical spectroscopy observations with the South African Large Telescope (SALT) and the SAAO 1.9-m telescope located at the South African Astronomical Observatory (SAAO). We present an overview of the proposed observational programme, the different possibilities available for multi-wavelength observations and initial results from this project.

1. Introduction

Astronomy is currently in a special position as we are able to observe the universe from radio up to γ -ray energies. More and more, multi-wavelength astronomy is required to understand the physical processes occurring in a multitude of sources. This extends from, for example, γ -ray pulsars which show both radio and γ -ray pulsations [1], γ -ray binary systems which show varying multi-wavelength behaviour [2], Active Galactic Nuclei (AGN) which show variations over the entire spectrum of energies, including rapid < 1 day flaring in γ -rays (for example 3C 454.3 [3]) to γ -ray bursts (GRBs) and their afterglows spanning from GeV energies to radio [4]. All these examples point towards the importance of multi-wavelength support to complement the γ -ray observations at both GeV and TeV energies.

South Africa is in a good position to contribute towards multi-wavelength support of these observations. Given our similar geographical location to the H.E.S.S. γ -ray telescope in Namibia and the possibility that the southern station of the Cherenkov Telescope Array (CTA) may be placed in Namibia (currently the two candidate sites are Aar, Namibia and the European Southern Observatory, Chile), we would argue that South Africa should become more involved with the High Energy (HE) and Very High Energy (VHE; $E > 100$ GeV) sources and continue to contribute to multi-wavelength follow-ups and observations of HE/VHE sources.

Here we present a short overview of the proposed observational programmes we are undertaking. These have, thus far, focussed on optical spectroscopic observations of unclassified sources within the Second *Fermi*-LAT AGN Catalogue (2LAC) [5], as well as optical photometric monitoring of candidate blazar sources as part of multi-wavelength campaigns. It is intended that this will expand to include radio observations of these sources. Since this proceedings will focus mainly on AGN, below we briefly present an overview of AGN and our search strategy, before reviewing the available optical and radio observatories within South Africa. We conclude with a discussion of preliminary observations undertaken thus far, as well as some related observations of known VHE sources.

2. Active Galactic Nuclei

AGN represent some of the most energetic sources in the sky. They are power by accretion of material onto a supermassive black hole in the centre of the galaxy and often produce relativistic jets. The observed properties of AGN depend on the angle between our line of sight and the direction of flow of these jets. In sources where the direction of flow lies close to our line of sight the Doppler boosted non-thermal emission being produced in the jet can dominate over the thermal emission from the galaxy. Such sources are collectively known as blazars [6].

The non-thermal emission of blazars extends from radio to γ -rays and the Spectral Energy Distributions (SEDs) of these sources exhibit two different components: a low energy component extending from radio to UV/X-ray; and a high energy component extending from UV/X-ray to γ -rays. While the low energy component is produced through (leptonic) synchrotron emission, the high energy component could be produce through either leptonic or hadronic processes. In the leptonic scenario the high energy component is produced through inverse Compton scattering, either of external photons originating from, for example, the accretion disc or the broad line region, while in the hadronic scenario the emission could be produced through, for example, proton synchrotron or π^0 decay (see e.g. [7] and references therein).

Blazars are sub-divided based based on the frequency, ν_{pk} , at which the peak in the lower energy component of the SED occurs. These sources are classified as Low Synchrotron Peaked (LSP), Intermediate Synchrotron Peaked (ISP) and High Synchrotron Peaked (HSP) if $\nu_{\text{pk}} \leq 10^{14}$ Hz, 10^{14} Hz $< \nu_{\text{pk}} \leq 10^{15}$ Hz and $\nu_{\text{pk}} > 10^{15}$ Hz, respectively. It can, to a first approximation, be assumed that HSP blazars are better candidates for VHE emission sources since, if both component are produced through leptonic processes, the high energy component should also peak towards higher energies. See for example [8], and references therein, for a discussion. Therefore, in our search for VHE sources we are establishing the multi-wavelength SED of these sources to determine whether there is an indication that the sources should have emission extending to VHE γ -rays.

The degree to which the Extragalactic Background Light (EBL) will attenuate any VHE emission must also be considered when searching for extra-galactic VHE candidates. VHE γ -rays emitted from AGN can interact with the EBL resulting in electron-positron pair production [9, 10]. This should introduce a redshift dependence on the opacity of AGN at VHE, and places a limited on the observability of blazars at VHE.

3. Search strategy

In order to identify new VHE candidate AGN, it is necessary to establish the SED of these sources, to classify them as LSP, ISP or HSP and to establish their redshift, as VHE γ -rays will be attenuated for high redshift sources. For the initial phase of this project we are focussing on multi-wavelength observations of a selection of unclassified extra-galactic HE sources identified in the *Fermi*-LAT catalogue in order to establish whether they present good candidates for VHE observations. Various factors were included in the selection of these candidate sources (as discussed in section 5.1).

In order to establish the SEDs a broad wavelength coverage from radio to γ -ray observations is required. While X-ray and γ -ray observations are available from archival data (e.g. *Fermi*), South African observatories can help to confirm/establish the radio and optical flux. Such observations, combined with initial modelling, can indicate sources which present good potential for VHE emission at a detectable flux. In addition we are undertaking optical spectroscopy to establish the redshifts of these targets as most detectable VHE sources lie at low redshifts, which therefore places a constraint on the observability of these sources at VHE.

In addition, photometric observations are being undertaken of the best candidate sources to search for short (~ 1 day) and long (> 1 day) term variability. The detection of such variability will help to confirm the counter parts association with the γ -ray sources, particularly if optical/ γ -ray correlation can be detected between the long-term photometric observations being undertaken with Watcher and data available with *Fermi*-LAT. Three of the candidate sources appear to exhibit γ -ray variability in *Fermi*-LAT observations since they have a variability test statistic $TS_{\text{var}} \gtrsim 41.6$. Here, TS_{var} is a test statistic established in [5] and $TS_{\text{var}} > 41.6$ indicates a 99% probability of variability over the 2 year period of the catalogue.

4. Optical and radio telescopes within South Africa

There are a number of telescopes available in South Africa and more detailed presentations are presented elsewhere at this meeting. Some of these telescopes have already been involved in multi-wavelength/TeV studies [11]. Here we only very briefly review the telescopes that we have used or are planning to use.

4.1. Optical telescopes

4.1.1. Boyden observatory. The Boyden observatory is located approximately 25 km North-East of Bloemfontein, South Africa. The main science telescope at the Boyden observatory is the 1.5-m Boyden reflector. The telescope is in a Cassegrain configuration with photometry capabilities. The system is equipped with standard U, B, V, R, I filters and an Apogee U55 Back-illuminated CCD camera. The second science telescope on-site is the 40-cm Watcher Robotic Telescope [12]. This telescope is operated remotely and has the main science aim of observing GRB afterglows. However, non-alert time can be used for on-going monitoring of other systems and ,for example, [13] reported on on-going blazar monitoring.

Both the Boyden 1.5-m and Watcher telescopes focus on on-going long term photometric monitoring of candidate sources.

4.1.2. South African Astronomical Observatory. The South African Astronomical Observatory (SAAO) is host to a number of optical/near-infrared telescopes run both by the South African National Research Foundation (NRF) and external organizations. Here we mention only the Southern African Large Telescope (SALT) and the SAAO 1.9-m telescopes. Both systems are equipped for optical photometric and spectroscopic observations. The grating spectrograph on the SAAO 1.9-m telescope allows for low to medium spectral resolution. SALT, on the other hand, is a 10-m class telescope, which can undertake low to medium resolution spectroscopy ($R = \lambda/\Delta\lambda \sim 500 - 9000$) with the Robert Stobie Spectrograph (RSS) and high resolution spectroscopy with the High Resolution Spectrograph (HRS) which is currently being commissioned. In addition, both the SAAO 1.9-m telescope and SALT are capable of very rapid photometry with the Sutherland High Speed Optical Cameras (SHOC) and the Berkeley Visible Image Tube (BVIT), respectively.

Optical observations with SALT and SAAO 1.9-m telescope focus on obtaining spectroscopic observations to establish the distance to these sources (redshift).

4.2. Radio telescopes

4.2.1. *HartRAO.* The Hartebeesthoek Radio Astronomy Observatory (HartRAO) is located near Hartbeespoort, South Africa and operates (among others) a 26-m single-dish radio telescope which operates between 1.67 to 23 GHz. The system is regularly involved with Very Long Baseline Interferometry (VLBI) observations and contributes to AGN monitoring observations. Radio observations of our candidate sources will be proposed to establish and/or confirm the radio brightness of these sources as well as searching for variability. These observations will be, in general, limited to bright (> 100 mJy) sources.

4.2.2. *KAT and SKA.* The expansion towards the development of Square Kilometre Array (SKA) in South Africa (see e.g. [14]) has lead to the development of Karoo Array Telescope (KAT) which currently consists of 7 dishes (KAT-7) which also regularly contributes towards the AGN monitoring and scientific observations. The next phase of development is the construction of the MeerKAT 64 dish array (see e.g. [15]). The first telescope was installed in March 2014 with a planned completion by 2017. The development of KAT-7/MeerKAT opens new possibilities for observing fainter radio sources, which may increase the number of candidate sources which can be investigated.

5. Initial results

5.1. Classification of AGU within the 2LAC catalogue

The Fermi/2LAC catalogue contains 886 identified sources in the “clean sample”, classified as extragalactic due to their location 10° above the galactic plane, of which 157 had no clear classification (“AGU” sources) [5]. Follow-up observations reported by, for example [16], have classified a number of these sources, but a large number remain unclassified. We have begun a processes to identify sources we wish to follow-up with South African telescopes from this sample. This is an extension of previous work undertaken to identify the unknown sources included within the 3rd EGRET catalogue by [17, 18], where 13 candidate sources were investigated. The main aim is to search for sources which could be TeV emitters. Furthermore, only (in general) bright radio sources (> 100 mJy) were selected to allow us to undertake radio observations with the HartRAO 26-m telescope. Follow-up optical spectroscopic observations were undertaken with the SAAO 1.9-m telescopes for two weeks during during the end of May 2014. Preliminary analysis shows, mainly featureless spectra as expected from blazars. This is an ongoing analysis and higher signal to noise observations are proposed for SALT. Please see [19] for a detailed discussion.

Long term observations of the blazar candidates identified as part of this project are being undertaken with the Watcher Telescope, to search for variability. In addition, observations to search for short term variability with SHOC on the SAAO 1.9-m telescope have been submitted.

6. Observations of known VHE sources

In addition to the campaign to investigate new VHE candidates we have also undertaken observations of known VHE sources. Two such campaigns are briefly mentioned below.

6.1. Optical monitoring of Blazars candidates with Watcher

The Watcher Robotic Telescope has been used as part of monitoring of campaigns of blazars and blazar candidates. Optical monitoring of the BL Lac 1ES 0229+200 was undertaken with the Watcher Robotic Telescope during October 2013 as part of a multi-wavelength campaign which coincided with NuSTAR and H.E.S.S. observations. Further observations were undertaken with the Boyden 1.5-m telescope, but the low altitude of the source made observations difficult as the telescope was very near its visibility limits.

6.2. PSR B1259-63

We have undertaken optical spectroscopic observations of the γ -ray binary system PSR B1259-63 during April-June with SALT and the SAAO 1.9-m telescope around the time the system went through periastron. This binary star system consists of a Be star and 48 ms radio pulsar and has been detected at TeV energies during previous periastron passages by H.E.S.S. (see e.g. [20]). The optical results may have important consequences as this can be used as an indication of the size and strength of the Be stars circumstellar disc. The disc will influence the shock which forms between the pulsar and stellar winds which is believed to be the production site of the non-thermal emission. The disc is known to vary around periastron and a disruption of the disc will influence the bow shock which will in turn influence the non-thermal emission. Variations in the optical spectrum have previously been reported by [21] who also suggested that the change in the disc mass may be responsible for the GeV flare report during 2011 [22].

7. Conclusion

South Africa is well placed to offer multi-wavelength support to complement TeV observations, in particular the H.E.S.S. γ -ray telescope located in Namibia and potentially CTA. Such multi-wavelength observations have already been undertaken. Here we briefly presented our on-going project to identify potential TeV emitting sources among the Fermi/2LAC catalogue as well as ongoing observations focussed on optical observations of known TeV sources. We have identified a subset of *Fermi* 2LAC AGU for which multi-wavelength SED are being constructed and optical spectroscopic and photometric observations are being undertaken, using the SAAO 1.9-m, SALT and Watcher Robotic Telescopes. Preliminary redshift distance have been established for 4 of these sources [19]. Further follow-up observations are planned.

Acknowledgments

This work is based on the research supported in part by the National Research Foundation of South Africa. Any opinion, finding and conclusion or recommendation expressed in this material is that of the authors and the NRF does not accept any liability in this regard.

References

- [1] Abdo AA et al. 2013 *ApJS* **208** 17
- [2] Dubus G 2013 *A&ARv* **21** 64
- [3] Ackermann M et al. 2010 *ApJ* **721** 1383
- [4] Mszros P 2013 *APh* **43** 134
- [5] Ackermann M et al. 2011 *ApJ* **743** 171
- [6] Urry CM, Padovani P 1995 *PASP* **107** 803
- [7] Böttcher M, Reimer A, Sweeney K and Prakash A 2013 *ApJ* **768** 54
- [8] Böttcher M 2012 Modeling the Spectral Energy Distributions and Variability of Blazars *Preprint arXiv/1205.0539*
- [9] Gould RJ and Schröder GP 1967 *Phys. Rev.* **155** 1408
- [10] Jelley JV 1966 *Phys. Rev. Lett.* **16** 479
- [11] Abramowski A et al., 2012 *A&A* **539** A149
- [12] Topinka M et al. 2013 *EAS Publications Series* **61** 487
- [13] Tisdall P et al., 2013, *Third Workshop on Robotic Autonomous Observatories*, Revista Mexicana de Astronomía y Astrofísica, eds JC Tello, A Riva, D Hiriart, AJ Castro-Tirado, in-press
- [14] Jonas JL 2014 *South African Institute of Physics 2014 Conference* these proceedings
- [15] Booth RS and Jonas JL 2012 *African Skies* **16** 101
- [16] Shaw et al. 2013 *ApJ* **764** 135
- [17] Meintjes PJ and Nkundabakura P 2013 *Acta Polytechnica* **53** 63
- [18] Nkundabakura P and Meintjes PJ 2012 *MNRAS* **427** 859
- [19] Klindt L, Meintjes PJ and van Soelen B 2014 *South African Institute of Physics 2014 Conference* these proceedings
- [20] Abramowski A et al. 2013 *A&A* **551** A94
- [21] Chernyakova M et al. 2014 *MNRAS* **439** 432
- [22] Abdo AA et al. 2011 *ApJ* **736** L11

Examining the behaviour of PKS 1424-418 during flaring

Pfesesani van Zyl^{1,2}, Michael Gaylard¹ and Sergio Colafrancesco²

¹Hartebeesthoek Radio Astronomy Observatory, Farm 502JQ, Broederstroom road, Krugersdorp, 1740

²WITS School of Physics, 1 Jan Smuts Avenue, Braamfontein, Johannesburg, 2000

E-mail: pfesesani@hartrao.ac.za, mike@hartrao.ac.za, sergio.colafrancesco@wits.ac.za

Abstract. The Fermi Large Area Telescope (Fermi-LAT) is a gamma-ray (γ -ray) telescope designed to observe high energy (HE) γ -ray events with energies between 20 MeV to just over 300 GeV. In October 2012, Fermi-LAT observed an energetic flare from the blazar source PKS 1424-418 reaching an average daily flux of $1.4 \pm 0.2 \times 10^{-6}$ ph cm⁻²s⁻¹. A flare is considered as This event triggered radio follow up observations with the Hartebeesthoek Radio Astronomy Observatory (HartRAO) 26 m radio telescope. The source was observed at frequencies between 2.3 GHz and 12.2 GHz using the drift scan technique. Soon after the follow up campaign began the source showed sharp increases in the radio flux density at all frequencies, with the 8.4 GHz data rising from about 6.1 Jy to 7.0 Jy within a month in January 2013. In this study we investigate the relationship between the γ -rays and the radio emission and search for time-lags using the discrete correlation function (DCF). We also study how the spectral index varied over time during the flaring event at weekly, 2 weekly and 4 weekly epochs. A Lomb-Scargle periodicity search was performed to investigate whether some periodic modulation was present in the γ -ray data. We find that there is a strong γ -ray/radio correlation with the γ -rays leading the radio with exception of one correlation showing the radio leading the γ -rays. An 86 day period was also discovered in the γ -ray data flares.

1. Introduction

Blazars are amongst the brightest known AGN and emit large amounts of energy across the entire electromagnetic spectrum [1]. They are known to exhibit multi-wavelength variability from radio frequencies to γ -rays [2], polarizations (in the radio and optical wavelengths), and show superluminal motion in their jets which are pointed towards the observer. They are traditionally divided into two main groups, BL Lacertae (BL-Lac) and Flat Spectrum Radio Quasars (FSRQ), blazars that show no emission lines or have broad emission lines in their spectra respectively.

The spectral energy distribution (SED) of blazars is characterized by two broad peaks. The first peak, located between the radio and optical frequencies, is thought to be caused by synchrotron emitting particles in the jet. The second peak, located in the high energy γ -ray region, is due to Inverse Compton scattering of relativistically boosted photons. The nature of the particles responsible for the different emissions is still under debate with some suggesting leptons and others hadrons as the main emission particles. This is one of the many reasons multi-wavelength correlation studies are important. They provide a platform for studying the

underlying physics of violent emission regions from the behavior of different components within the blazar jets ([3],[4]).

In this study we observed the source PKS 1424-418. This is an FSRQ that recently underwent a series of flaring events from October 2012 to September 2013. These events were detected by Fermi-LAT with the lowest flare having energies of about $1.4 \pm 0.2 \times 10^{-6}$ ph cm⁻²s⁻¹ and the highest reaching $2.3 \pm 0.1 \times 10^{-6}$ ph cm⁻²s⁻¹ for $E > 100$ MeV. The radio emission detected by the HartRAO 26-m telescope shortly after the first flare showed an increased flux density trend at all wavelengths revealing that there might be a connection between the two energy regimes. Using the DCF, we investigated the correlation between the γ -ray light-curve and the radio light curves of the source. We also investigated the relationship between the radio light curves themselves. We further conducted a periodicity search in the γ -ray data to investigate the presence of a possible periodic modulation in the individual flares.

This paper is arranged as follows: Section 2 describes the observations and the data reduction processes, Section 3 describes the data analysis methods applied to the data, and Section 4 discusses the results.

2. Observations & data reduction

This work is based on observations collected by Fermi-LAT and the HartRAO 26-m telescope during the observing epochs October 2012 to September 2013. The Fermi-LAT data was obtained pre-processed in standard flux density units by Roopesh Ojha¹ from the NASA Goddard Space Flight Center, and therefore this paper will only discuss observations done using the HartRAO 26-m telescope.

2.1. HartRAO 26-m observations

Following the Astronomers telegram (ATel) #4494², HartRAO 26-m began follow-up observations on PKS 1424-418 from 17 October 2012. We observed PKS 1424-418 with 2.3-GHz (cryogenically cooled) and 12.2-GHz (uncooled) single-beam receivers, and with 4.8-GHz and 8.4-GHz (cryogenically cooled) dual-beam Dicke-switched receivers. The telescope collected radio emission from the source in both left and right circularly polarized orientations using the drift scan technique. During observations, the telescope is parked a little west of the current source position, then as the source “drifts” over the dish due to the earth’s rotation it collects radio emission incident on the dish surface. On average, observations consisted of conducting two scans per day with the exception of days when the telescope was used for Very Long Baseline Interferometry (VLBI) observing.

2.2. Data reduction

To get useful flux density estimates of the source, the received emission from the source had to be calibrated against a known calibrator source of unvarying flux density. PKS 1424-418 was calibrated against Hydra A (3C218) using the flux density equation provided by Ott et al. [8]. The Hydra A measured antenna temperature was corrected for pointing errors that can occur during observing due to ground or atmospheric effects by taking scans of the source at the half-power points to the North and South of the source. Each scan was visually inspected for any irregularities that might make it an outlier, through a series of data quality checks. Radio telescopes are highly temperature sensitive instruments and the observations collected can be affected by contributions from their surroundings.

¹ roopesh.ojha@gmail.com

² <http://www.astronomerstelegam.org/?read=4494>

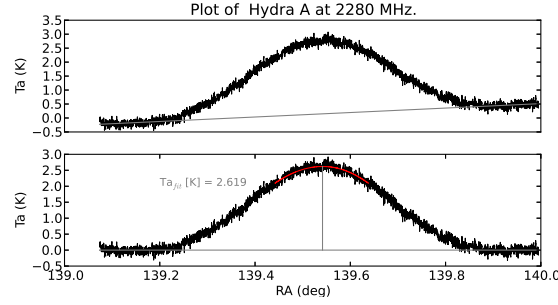


Figure 1. The figure shows two drift scan plots of the source Hydra A. The top figure is the raw drift scan data before reductions. The bottom figure is the raw drift scan after reduction. The red line shows an estimate parabolic peak fit to the top 20% of the data. Ta_{fit} is the maximum antenna temperature for the fit in Kelvins.

2.2.1. Quality checks for outlier detections. As a first step all data containing any one of the following: an RFI (Radio Frequency Interference) signal, step changes in the data (possibly due to receiver instabilities), monotonically increasing data and bumps in the data (due to poor weather conditions that distort the expected bell shaped beam pattern) were considered outliers and removed from observations. These quality checks ensured that all data used in further analysis conformed to the same standard. Under good observing conditions, the system temperature (T_{sys}) of the observed calibrator source is expected to remain constant throughout the observing campaign. All data scans were examined for deviations in the average T_{sys} to ensure that none of the elements in the telescope had over-heated during observing. The recorded T_{sys} per scan had to fall within 2.7σ of the average T_{sys} .

3. Data analysis

3.1. Data fitting.

In our second step of the data reduction first order baselines were subtracted from the data to remove any instabilities in the drift scan that may have been caused by the changes in the telescope's surroundings. To find the peak antenna temperature, second order polynomials ($y = ax^2 + bx + c$) were fitted to the top 20% of the scans. Figure 1 shows a typical beam pattern for a drift scan using Hydra A at 2.3 GHz as an example.

3.2. Flux density calculations.

In the last step of our data reduction, we calculated the flux density of the target source using the point source sensitivity (PSS) of the telescope calculated from the calibrator source Hydra A [8]. The calibrator flux density for frequencies < 24 GHz was calculated using,

$$S_{hyd_\nu} = 10 \exp(0.013 \times \log(\nu)^2 - 1.025 \times \log(\nu) + 4.729) \quad [Jy] \quad (1)$$

where S_{hyd_ν} was the Hydra A flux density at a frequency ν . The PSS was calculated separately for each polarization using the equation,

$$PSS = \frac{S_{hyd_\nu}/2}{K_s T_{a_{hyd_pol_\nu}}} \quad [KJy^{-1}] \quad (2)$$

where $K_s = 1$ was the size correction factor and $T_{a_{hyd_pol_\nu}}$ was the antenna temperature of the calibrator. The factor of 2 results from the fact that each feed collects half the flux density due to circular polarization. From the PSS, we then calculated the total flux density of the source

(see eq. 3) as the sum of the flux density contributions from the two polarizations. The flux density estimates for each observing day were used to create the light curves in figure 2.

$$S_{src_pol_\nu}[Jy] = 2 \times PSS_{hyd_pol_\nu} Ta_{src_pol_\nu} \quad (3)$$

where $S_{src_pol_\nu}$ and $Ta_{src_pol_\nu}$ are the flux density and antenna temperature of the target source PKS 1424-418.

3.3. Correlation analysis

We examined the correlation between γ -rays and radio frequencies using the Discrete Correlation Function (DCF) [9]. The method estimates the time-lag between two discrete light curves. Designed to work with unevenly sampled light curves, the DCF makes no prior assumptions about the data and only calculates the correlation and time-lags for the data sample as is without the interpolation of missing data. Given two light curves x and y , the DCF calculates the cross-correlation function of the light curves using

$$DCF_{ij} = \frac{1}{N} \left[\sum \frac{(x_i(t) - \bar{x})(y_j(t) - \bar{y})}{\sqrt{(\sigma_x^2 - e_x^2)(\sigma_y^2 - e_y^2)}} \right] \quad (4)$$

where \bar{x} and \bar{y} are the means, σ_x and σ_y the standard deviations, and e_x and e_y are the errors of the respective light curves. The DCF estimates an unbinned cross-correlation function for all measured pair points (x_i, y_j) , then averages it over a pairwise lag $\Delta t_{ij} = t_j - t_i$ within the range $\tau - \Delta\tau/2 < \Delta\tau < \tau + \Delta\tau/2$ where τ is the time-lag, $\Delta\tau$ is the chosen bin width and N is the number of the unbinned entries. The errors were then estimated as the scatter in the unbinned entries of the DCF terms. We applied Gaussian fits to the top 20% of the peaks to get the lag values in days. The results are shown in Section 4.1. The rapidly varying γ -ray flux density at face value appeared as if it was flaring periodically over time (figure 2). This prompted an investigation into the possible periodicity that might exist within the γ -ray data of the source PKS 1424-418. Periods have been found to exist within γ -ray data [5], but the topic of γ -ray periodicity has been poorly studied and should be approached with caution. The periodicity was analyzed using the period detection method [6], designed for period searches in unevenly sampled data. The significance of the detection was analyzed using the false alarm probability method [7].

4. Results & concluding remarks

Figure 2 displays the 2.3-GHz to 12.2-GHz radio light curves and the γ -ray light curve of the source observed from October 2012 to September 2013. The source shows rapid variability at the higher energies, and a slower more gradual change as you go down in frequency, typical of blazar variability. The 12.2-GHz and 8.4-GHz frequencies showed a steep increase and decrease during the observing period, whereas the 4.8 GHz and 2.3 GHz still showed gradually increasing trends. This is expected as longer wavelengths generally take longer to show variation in their light curves than shorter wavelengths. At 8.4 GHz, there appears to be a sudden upturn towards the end of the light curve, suggesting that something interesting was happening but we require more data to investigate whether this effect was real.

4.1. Correlation

Results from the DCF correlations found that the γ -ray light-curve correlated well with the radio light curves. Table 1 shows a summary of the results obtained for all the correlations. Some DCF correlations were highly scattered especially at longer wavelengths, showing broad peaks and here we give only two examples of the best correlated fits. Figure 3 shows the result

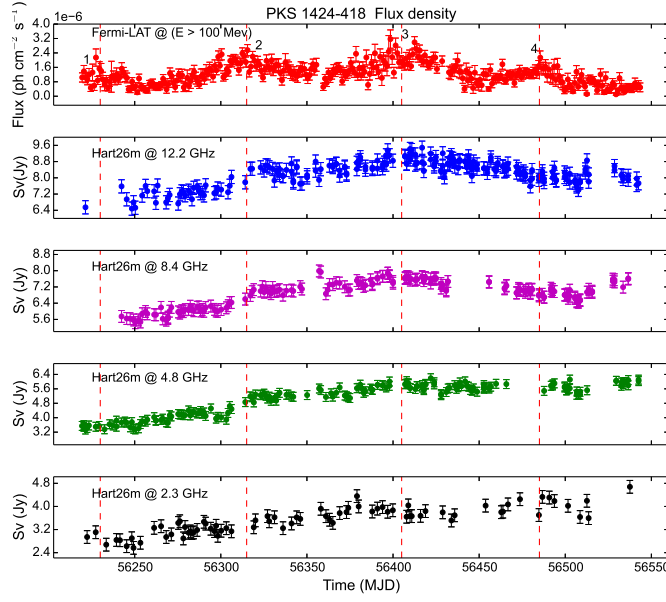


Figure 2. The light curves of observations from both the Fermi-LAT telescope and the HartRAO 26-m radio telescope. The top panel is data collected by Fermi for $E > 100$ MeV with the flares marked by numbers 1 to 4. The subsequent panels are light curves of data collected by HartRAO 26-m at frequencies 12.2 GHz, 8.4 GHz, 4.8 GHz and 2.3 GHz respectively. The dashed lines indicate the peak flux density epochs of the corresponding flares.

of the γ -ray/12.2-GHz correlation where we found a 23.2 ± 1.1 day correlation. In this case the γ -rays were leading the radio frequencies as expected. However, figure 4 shows the opposite case with the lower radio frequencies leading the higher radio frequencies for the 12.2-GHz/8.4-GHz correlations, where we found a time-lag of -4.9 ± 0.6 days. This connection might imply that the HE photons might be relativistically boosted Synchrotron photons, but we need more observations to draw any firm conclusions.

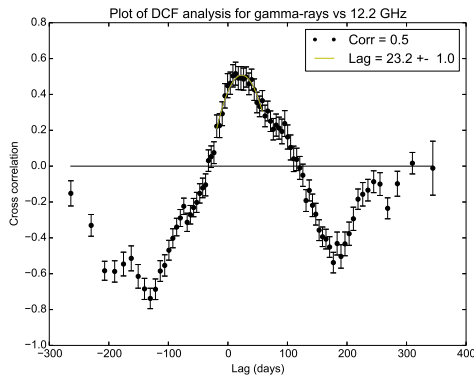


Figure 3. Gamma-rays vs 12.2-GHz light-curve correlation result.

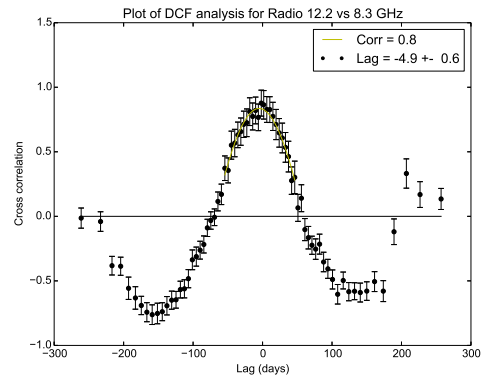


Figure 4. Radio 12.2-GHz vs 8.4-GHz light-curve correlation result.

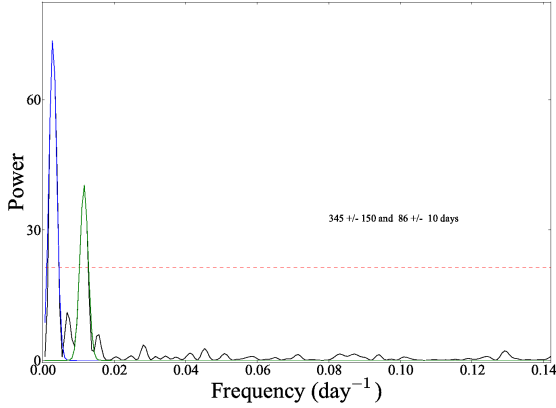


Figure 5. Results for γ -rays/radio and radio/radio DCF correlations. Horizontal line is the 3σ false detection limit.

Light-curve 1	Light-curve 2	Time-lag (days)
γ -rays	12.2 GHz	23.2 ± 1.0
γ -rays	8.4 GHz	27.9 ± 1.8
γ -rays	4.8 GHz	68.8 ± 1.4
γ -rays	2.3 GHz	70.2 ± 2.4
12.2 GHz	8.4 GHz	-4.9 ± 0.6
12.2 GHz	4.8 GHz	4.8 ± 1.1
12.2 GHz	2.3 GHz	13.8 ± 3.1
8.4 GHz	4.8 GHz	3.9 ± 1.7
8.4 GHz	2.3 GHz	10.4 ± 2.3
4.8 GHz	2.3 GHz	4.7 ± 2.9

Table 1. Results for γ -rays/radio and radio/radio DCF correlations.

4.2. Spectral index and periodicity

From the data collected, we also wanted to investigate whether there was a relationship between the variability observed and the spectral index (α), where $S_\nu \sim \nu^\alpha$. We used the two-point spectral index relation $\alpha = \frac{\log(S_1/S_2)}{\log(\nu_1/\nu_2)}$ where S_1, S_2 and ν_1, ν_2 are the flux density and frequency at two independent radio frequencies respectively. Results obtained showed that there was indeed a relationship between the spectral index and the variability. The spectral index showed increased variability with time between radio light curves that appeared out of phase than those that were in phase during flaring events [10], showing a general trend moving from inverted towards flat. Light curves that appeared in phase e.g. 12.2 and 8.4 GHz, however showed exceptions to the trend by having $\alpha_{12.2-8.4}$ remaining almost constantly inverted over time. The periodicity search produced an 86 ± 10 day period for the varying flares, the time between two successive local minima, consistent with observations (see figure 4.1).

4.3. Conclusion

From the results it can be seen that in order to investigate the jets in higher detail we would require more observations to provide high cadence light curves. The data suggests that there might be a sharp upturn in the radio at higher frequencies, which could mean that the source was still active in radio long after the γ -rays had reduced to the blazars quiescent stage. From the observations we discovered the source spectrum was inverted due to the flaring, but more importantly we found that there was a γ -ray to radio correlation with the time-lags increasing with wavelength. Results from the spectral index study however suggest that the spectral index hardened after each flare event for light curves in phase, suggesting that each flare event is unique and has its own distinct characteristics. This could be a good indication that splitting each flare and evaluating them separately could prove useful for future studies into multi-variability, see for example [10]. Lastly, the γ -rays produced an 86 ± 10 day period between flares and a longer 345 ± 150 day trend for the observing period.

Acknowledgments

This work was jointly supported by the *National Astrophysics and Space Science program* (NASSP) and HartRAO through the *National Research Foundation* (NRF).

References

- [1] Urry, C., M., and Padovani, P., 1995, *Publ. Astron. Soc. Pac.*, **107**, 803
- [2] Raiteri, C. M., Villata, M., Kadler, M., et al. 2006, *aap*, **459**, 731
- [3] Raiteri, C. M., Villata, M., Smith, P. S., et al. 2012, *aap*, **545**, AA48
- [4] Hovatta, T. H., Tornikoski, M., Lainela, M., et al. 2007, *A&A*, **469**, 899
- [5] Nishikawa, D., Hayashi, S., Chamoto, N., et al. 1999 *in Proc. 26th ICRC (Salt Lake City) OG 2.1.17*, **3**, 354
- [6] Lomb, N., R., 1976, *Ap&SS*, **39**, 447
- [7] Scargle, J., D., 1982, *ApJ*, **263**, 835
- [8] Ott, M., Witzel, A., Quirrenbach, A., et al. 1994, *A&A*, **284**, 331
- [9] Edelson, R., A., and Krolik, J., H., 1988, *ApJ*, **333**, 646
- [10] Shabala, S., S., Rogers, J., G., McCallum, J., N., et al. 2014, *Journal of Geodesy*, **88**, 575

Modelling the average spectrum expected from a population of gamma-ray globular clusters

C Venter¹ and A Kopp²

¹Centre for Space Research, North-West University, Potchefstroom Campus, Private Bag X6001, Potchefstroom 2520, South Africa

²Institut für Experimentelle und Angewandte Physik, Christian-Albrechts-Universität zu Kiel, Leibnizstrasse 11, 24118 Kiel, Germany

Abstract. Millisecond pulsars occur abundantly in globular clusters. They are expected to be responsible for several spectral components in the radio through γ -ray waveband (e.g., involving synchrotron and inverse Compton emission), as have been seen by Radio Telescope Effelsberg, *Chandra* X-ray Observatory, *Fermi* Large Area Telescope, and the High Energy Stereoscopic System (H.E.S.S.) in the case of Terzan 5 (with fewer spectral components seen for other globular clusters). H.E.S.S. has recently performed a stacking analysis involving 15 non-detected globular clusters and obtained quite constraining average flux upper limits above 230 GeV. We present a model that assumes millisecond pulsars as sources of relativistic particles and predicts multi-wavelength emission from globular clusters. We apply this model to the population of clusters mentioned above to predict the average spectrum and compare this to the H.E.S.S. upper limits. Such comparison allows us to test whether the model is viable, leading to possible constraints on various average cluster parameters within this framework.

1. Introduction

There are nearly 160 Galactic globular clusters (GCs) known [14]. Each consists of hundreds of thousands of stars held together by their mutual gravity, and they orbit the Galactic Centre in a spherical distribution. GCs are typically about ten gigayears old and are therefore expected to harbour many evolved stellar objects, since the latter should have had ample time to complete their evolutionary processes. The high stellar densities in the cores of GCs also enhance stellar encounter rates, facilitating the formation of objects such as low-mass X-ray binaries (LMXRBS), cataclysmic variables, white dwarfs, and pulsars [18]. LMXRBs are believed to be the progenitors of millisecond pulsars (MSPs; [5]), and since they occur abundantly in GCs, the same should hold true for MSPs. This is indeed found to be the case: 28 of the Galactic GCs contain more than 144 confirmed radio pulsars¹, the bulk of these being MSPs. It was furthermore estimated that there should be 2 600 – 4 700 Galactic GC MSPs observable in γ -rays [1].

GCs are multi-wavelength objects, being visible from radio to the highest energies. For example, several radio structures are visible in the direction of Terzan 5 at 11 cm and 21 cm [9], while diffuse X-ray emission has also been detected from this GC [9, 11]. The *Fermi* Large Area Telescope (LAT) plausibly detected about a dozen GCs in the GeV energy band [17], and their spectral characteristics point to the cumulative emission from a population of GC MSPs.

¹ <http://www.naic.edu/~pfreire/GCpsr.html>

In the TeV domain, H.E.S.S. has published upper limits for 47 Tucanae [4], and detected a very-high-energy (VHE) excess in the direction of Terzan 5 [2].

Several models have been proposed to explain the observed GC spectra. The total GeV contribution from GC MSPs was estimated by summing up individual predicted pulsed curvature radiation (CR) spectra from a population of MSPs [13, 19]. An alternative scenario [8] assumed that the GeV emission was due to inverse Compton (IC) radiation by leptons escaping from the MSP magnetospheres, upscattering Cosmic Microwave Background (CMB), stellar, as well as Galactic background (infrared and optical) photons. Another model [6] considered MSPs that accelerate leptons either at the shocks originating during collisions of neighbouring pulsar winds or inside the pulsar magnetospheres. These leptons escape from the magnetospheres and diffuse through the GC, encountering target photon fields such as optical and CMB emission. The latter are upscattered via the IC process, leading to GeV / TeV emission. This model was extended [20, 23] using alternative particle injection spectra, and also calculating the expected synchrotron radiation (SR) from leptons moving in a homogeneous cluster magnetic field B . Recently, the model was significantly refined [15] and now includes a line-of-sight calculation of the X-ray surface brightness (which is used to constrain the diffusion coefficient) as well as full particle transport, assuming spherical symmetry and a steady-state regime. Apart from leptonic models, there two other ideas. Non-accreting white dwarfs may make a considerable contribution to the cumulative γ -ray flux seen from GCs [7], depending on their abundance. Lastly, a model invoking γ -ray burst remnants as sources of energetic leptons and hadrons was also put forward [10]. A short burst may accelerate hadrons, which may in turn collide with ambient target nuclei, leading to π^0 particles that eventually decay into γ -rays. In this model, X-rays may result from IC on optical stellar photons by primary electrons accelerated by the relativistic blast wave.

We note that a search for diffuse X-ray emission from several *Fermi*-detected GCs [12] failed to detect any such emission above the Galactic background level, although a recent reanalysis of the archival *Chandra* data detected a new diffuse X-ray emission feature within the half-mass radius of 47 Tucanae [22]. In the context of the leptonic MSP models, these upper limits and detection place constraints on parameters such as the number of embedded MSPs N_{MSP} as well as the typical cluster field strength B . One way to help discriminate between the various models would be to follow a population approach. H.E.S.S. has recently obtained an upper limit to the average TeV flux of 15 non-detected GCs [3]. They noted that their upper limit was lower than the flux predicted by a simple leptonic scaling model by a factor of $\sim 3 - 30$, depending on model assumptions. This paper represents a first attempt to model the average TeV flux from the population of 15 GCs using our refined leptonic GC model. Our motivation is to compare our results with those of the scaling model, and second, to assess the plausibility of the MSPs scenario by testing whether our average spectrum satisfies the TeV upper limits. In Section 2, we briefly discuss our model, after which follow a description of the assumed parameters and calculation method (Section 3), as well as our conclusions (Section 4).

2. The basic GC model

We solve the following transport equation numerically

$$\frac{\partial n_e}{\partial t} = \vec{\nabla} \cdot (\mathcal{K} \cdot \vec{\nabla} n_e) - \frac{\partial}{\partial E_e} (\dot{E}_e n_e) + Q, \quad (1)$$

with n_e the electron² density, which is a function of the radius vector \vec{r}_s , E_e the electron energy, \mathcal{K} the diffusion tensor, $\dot{E}_e > 0$ the particle energy losses, and Q the source term. We assume a

² We use the word ‘electron’ in a wider sense. Pair production may take place, additionally providing the possibility of positrons as efficient radiators.

source term located at $\vec{r}_s = \vec{r}_C$ that is of the form

$$Q = Q_0 E_e^{-\Gamma} \delta(\vec{r}_s - \vec{r}_C), \quad (2)$$

with the spectral index Γ . We assume stationarity and spherical symmetry, leading to

$$0 = \frac{1}{r_s^2} \frac{\partial}{\partial r_s} \left(r_s^2 \kappa \frac{\partial n_e}{\partial r_s} \right) - \frac{\partial}{\partial E_e} (\dot{E}_e n_e) + Q, \quad (3)$$

with κ the scalar diffusion coefficient. This may be rewritten as

$$\frac{\partial n_e}{\partial E_e} = \frac{1}{\dot{E}_e} \left(\frac{1}{r_s^2} \frac{\partial}{\partial r_s} \left(r_s^2 \kappa \frac{\partial n_e}{\partial r_s} \right) - n_e \frac{\partial \dot{E}_e}{\partial E_e} + Q \right), \quad (4)$$

and is solved numerically using a Crank-Nicolson algorithm. Our particle injection spectrum is normalised as follows:

$$\int_{E_{e,\min}}^{E_{e,\max}} E_e Q dE_e = N_{\text{MSP}} \eta \langle L_{\text{sd}} \rangle, \quad (5)$$

with η the particle conversion efficiency, and $\langle L_{\text{sd}} \rangle$ the average MSP spin-down luminosity. After solving for n_e , we calculate the SR and IC fluxes as detailed in [15]. We assume Bohm diffusion

$$\kappa(r_s, E_e) = \frac{c E_e}{3eB(r_s)}, \quad (6)$$

where c and e denote speed of light and electron charge. As an alternative, we also investigated a coefficient of the form

$$\kappa(r_s, E_e) = \kappa_0(r_s) \left(\frac{E_e}{E_{e,0}} \right)^\alpha, \quad (7)$$

with $E_{e,0} = 1$ TeV and $\alpha = 0.6$. For simplicity, we only considered spatially constant κ_0 and B .

3. Estimating the average GC spectrum - a first approach

The model described in Section 2 has been successfully applied to the case of Terzan 5 [15], where we performed a line-of-sight integration of the X-ray flux in order to constrain the lepton diffusion coefficient. We found that values of $\kappa_0 \approx 10^{-4} \text{ kpc}^2 \text{ Myr}^{-1}$ gave good fits. In this paper, we apply the same model to 15 non-detected GCs [3], using fixed parameters as noted in table 1. We have used the values of [1] for N_{MSP} where possible, and N_* values from [16], and obtained distances d and structural parameters³ r_c (core radius), r_{hm} (half-light radius, used as a proxy for half-mass radius), and r_t (tidal radius) from [14]. Typical values of $\eta \sim 0.01$, $N_{\text{MSP}} \sim 25$, and $\langle L_{\text{sd}} \rangle \sim 2 \times 10^{34} \text{ erg s}^{-1}$ led to values for the source strength Q_0 . We have used $B = 5 \mu\text{G}$ and $\Gamma = 2.0$ throughout, and used optical GC plus CMB photons as background fields for our IC calculation (assuming an average stellar temperature of $T = 4 \text{ 500 K}$).

We first produced SR and IC spectra for each individual cluster. We next added these spectra and divided by the number of GCs to obtain the average predicted spectrum. This is indicated in figure 1 (differential flux) and figure 2 (integral flux). None of the single-cluster spectra violate the TeV upper limits. The stacked upper limits are $F(> 0.23 \text{ TeV}) = 3.3 \times 10^{-13} \text{ cm}^{-2} \text{ s}^{-1}$ for a point-like source analysis, and $F(> 0.23 \text{ TeV}) = 4.5 \times 10^{-13} \text{ cm}^{-2} \text{ s}^{-1}$ for an extended source analysis [3]. Our average spectrum gives $F(> 0.23 \text{ TeV}) \approx 2 \times 10^{-13} \text{ cm}^{-2} \text{ s}^{-1}$, which satisfies

³ See also <http://gclusters.altervista.org/index.php>

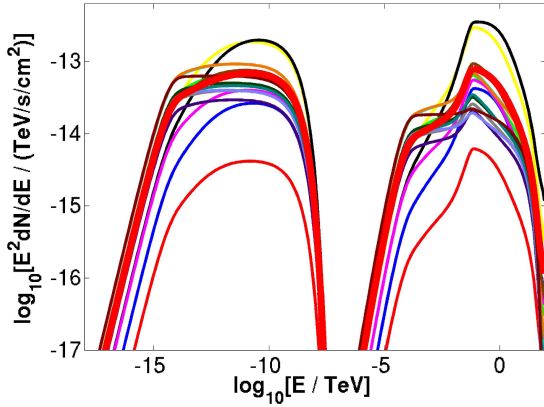


Figure 1. Predicted differential spectra $E^2 dN/dE$ for 15 single GCs (thin lines), as well as the average spectrum (thick red line). The two components represent the SR and IC spectra.

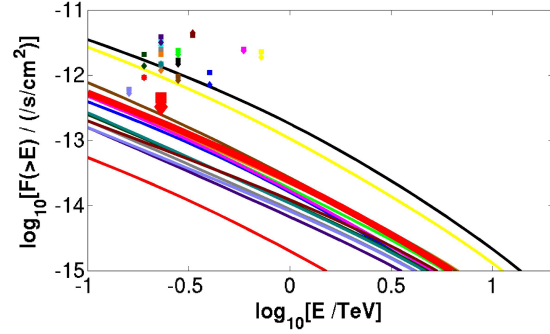


Figure 2. Same as figure 1, but for integral flux $F(>E)$. Upper limits for the point-like source analysis are indicated by diamonds, and those for the extended source analysis by squares [3]. Associated spectra and upper limits have the same colour. The larger (lowest) symbols are for the average spectrum.

these observational upper limits. However, we note that different choices of parameters will lead to average fluxes that could even exceed these limits in some cases.

We next tested the effect of changing the diffusion coefficient. Figure 3 indicates the average spectra assuming Bohm diffusion (solid red line), a coefficient $\kappa = \kappa_0 (E_e/1 \text{ TeV})^{0.6}$ with $\kappa_0 = 10^{-4} \text{ kpc}^2 \text{ Myr}^{-1}$ (dashed green line), and $\kappa_0 = 10^{-3} \text{ kpc}^2 \text{ Myr}^{-1}$ (dotted blue line). The change in spectral shape indicates the different energy dependencies ($\kappa \propto E_e^{1.0}$ for Bohm diffusion versus $\kappa \propto E_e^{0.6}$). The corresponding integral fluxes are shown in figure 4.

Table 1. Assumed GC parameters.

GC Name	d (kpc)	N_* (10^5)	N_{MSP}	Q_0 ($10^{32}/\text{erg/s}$)	r_c (')	r_{hm} (')	r_t (')
NGC 104 (47 Tuc)	4.5	4.57	33^{+15}_{-15}	9.55	0.36	3.17	42.86
NGC 6388	9.9	5.8	180^{+120}_{-100}	52.1	0.12	0.52	6.21
NGC 7078	10.4	4.13	25 (< 56)	7.24	0.14	1.00	21.5
Terzan 6	6.8	0.29	25	7.24	0.05	0.44	17.39
Terzan 10	5.8	0.38	25	7.24	0.9	1.55	5.06
NGC 6715	26.5	4.79	25	7.24	0.09	0.82	7.47
NGC 362	8.6	1.58	25	7.24	0.18	0.82	16.11
Pal 6	5.8	0.31	25	7.24	0.66	1.2	8.36
NGC 6256	10.3	0.21	25	7.24	0.02	0.86	7.59
Djorg 2	6.3	1.0	25	7.24	0.33	1.0	10.53
NGC 6749	7.9	0.24	25	7.24	0.62	1.1	5.21
NGC 6144	8.9	0.48	25	7.24	0.94	1.63	33.25
NGC 288	8.9	0.32	25	7.24	1.35	2.23	12.94
HP 1	8.2	0.48	25	7.24	0.03	3.1	8.22
Terzan 9	7.1	0.02	25	7.24	0.03	0.78	8.22

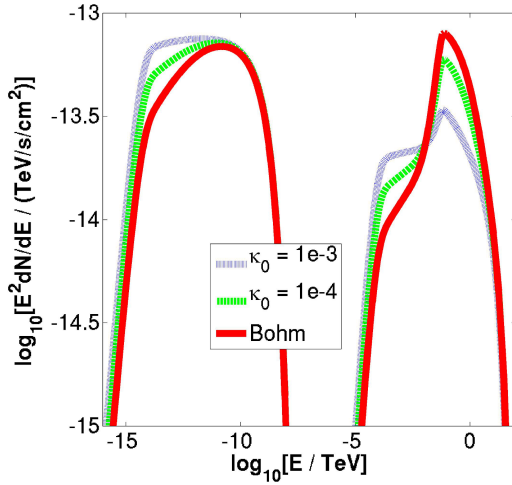


Figure 3. Predicted average differential spectra. The line types indicate different diffusion coefficient assumptions as noted in the legend.

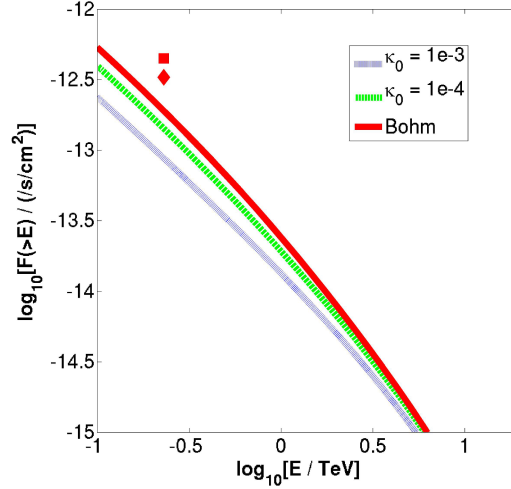


Figure 4. Same as figure 3, but for the integrated flux. The red square and diamond indicate the extended and point-like H.E.S.S. upper limits.

Our next task is a rigorous assessment of the error on the average integral spectrum, taking into account the range of values that each free parameter may assume, for each individual cluster. Preliminary investigations have shown that we can reduce the number of free parameters to five. Varying these parameters on a grid and calculating the average flux value and spread give quite large errors [21]. This process is very time-consuming, given the enormous number of combinations in which a cumulative (or average) spectrum may be obtained (15 GCs, with at least 5 free parameters each).

4. Conclusions

We have described the application of our leptonic GC model to a population of clusters that have been observed, but not detected, in VHE γ -rays. The fact that upper limits are available presented us with the opportunity to test our model using a sizable sample of clusters. The upper limit for the stacked (average) flux is stringent, and a simple scaling model violated this limit. In contrast, none of our single-cluster spectral predictions violated the respective upper limits. Our average model spectrum also satisfied the stacked upper limits for the given choice of parameters (table 1), which we regarded as reasonable. There is, however, considerable uncertainty in the single-GC parameters, so that the intrinsic error on the predicted average spectrum will be quite large. As an example, we showed that different assumptions for the diffusion coefficient lead to significant changes in spectral and flux. One should therefore attempt to reduce the number of free model parameters and also constrain their values so as to decrease the predicted flux error. This will allow one to more robustly test the viability of the MSP scenario for explaining the (non-)detection of TeV flux from GCs.

Acknowledgments

This research is based on work supported by the South African National Research Foundation.

References

- [1] Abdo A A *et al.* 2010 *Astron. Astrophys.* **524** A75
- [2] Abramowski A *et al.* 2011 *Astron. Astrophys.* **531** L18–22

- [3] Abramowski A *et al.* 2013 *Astron. Astrophys.* **551** A26
- [4] Aharonian F *et al.* 2009 *Astron. Astrophys* **499** 273–7
- [5] Alpar M A, Cheng A F, Ruderman M A and Shaham J 1982 *Nature* **300** 728–30
- [6] Bednarek W and Sitarek J 2007 *Mon. Not. Royal Astron. Soc.* **377** 920–30
- [7] Bednarek W 2012 *J. Phys. G: Nucl. Part. Phys.* **39** 065001
- [8] Cheng K S *et al.* 2010 *ApJ* **723** 1219–30
- [9] Clapson A-C, Domainko W F, Jamrozy M, Dyrda M and Eger P 2011 *Astron. Astrophys.* **532** A47
- [10] Domainko W F 2011 *Astron. Astrophys.* **533** L5–8
- [11] Eger P, Domainko W F and Clapson A-C 2010 *Astron. Astrophys.* **513** A66
- [12] Eger P and Domainko W F 2012 *Astron. Astrophys.* **540** A17
- [13] Harding A K, Usov V V and Muslimov A G 2005 *Astrophys. J.* **622** 531–43
- [14] Harris W E 1996 *Astron. J.* **112** 1487–8
- [15] Kopp A, Venter C, Büsching I and de Jager O C 2013 *Astrophys. J.* **779** 126–37
- [16] Lang K R 1993 *Astrophysical Data: Planets and Stars* (Heidelberg: Springer-Verlag) 257–77
- [17] Nolan P L *et al.* 2012 *ApJS* **199** 31–76
- [18] Pooley D *et al.* 2003 *Astrophys. J.* **591** L131–4
- [19] Venter C and de Jager O C 2008 *Astrophys. J.* **680** L125–8
- [20] Venter C *et al.* 2009 *Astrophys. J.* **696** L52–5
- [21] Venter C and Kopp A 2015 *Mem. S.A.It.* **86** 69–76
- [22] Wu E M H *et al.* 2014 *Astrophys. J.* **788** L40–4
- [23] Zajczyk A, Bednarek W and Rudak B 2013 *MNRAS* **432** 3462–73

Spectral comparison between AGN at $z = 0.1, 0.2$ and 0.3

Hartmut Winkler and Jessica Tsuen

Department of Physics, University of Johannesburg, South Africa

E-mail: hwinkler@uj.ac.za

Abstract. We identified three samples of ROSAT sources with Sloan Digital Sky Survey spectra, one at redshift $z = 0.1$, a second one at $z = 0.2$ and a third one at $z = 0.3$. 812 sources in total were examined. We determined the nature and spectral sub-types of the sources by visual inspection. The fraction of each sub-type at each of the three redshifts are then calculated. We consider selection biases caused by the luminosity cut-off threshold to determine whether any systematic trends in AGN type are evident with increasing redshift. We hence probe if an evolution effect is detected in our sample.

1. Introduction

The list of confirmed active galactic nuclei (AGN) has been dramatically boosted by a series of recent large scale spectroscopic surveys. In particular, the Sloan Digital Sky Survey (SDSS) has produced a vast number of optical spectra of extragalactic sources [1]. The resultant expanding database enables more detailed and rigorous study of the density and distribution of AGN in at least the comparatively nearby universe.

Despite the popularity and wide acceptance of the AGN unified model, which postulates that a common physical mechanism underlies all AGN types and that apparent differences can usually be ascribed to orientation effects, there is now also recognition that some AGN subclasses manifest real rather than apparent differences compared to other AGN. This then raises the question of whether the relative concentration of the different subclasses is a function of evolution, and hence effectively of distance, which is measured in terms of the redshift z .

Another aspect that becomes important in comparative studies of AGN subclasses is the relationship between their luminosities. While the range in AGN luminosities is vast, encompassing everything from the micro-AGN believed to exist in many otherwise ordinary galaxies to the most energetic distant quasars, some subclasses do on average appear to be intrinsically brighter than others.

In this work we attempt such a comparison of AGN subclass luminosities using a simple approach based entirely on visual spectral classification and population counts. We analyse spectra for sets of AGN identified in three ‘shells’ corresponding to redshifts of approximately $z = 0.1, 0.2$ and 0.3 . In order to circumvent selection biases caused by optical colour or spectral line effects, we defined our samples on the basis of x-ray detection. We recognise that x-ray detectability is in itself strongly dependent on redshift, but the x-ray luminosity bias should influence all subclasses equally, and thus not affect the ratio between subclass object counts.

2. Sample selection and classification of spectra

We started by making a positional cross-correlation of the x-ray sources detected in the ROSAT survey [2] with SDSS objects with spectra marking redshifts in one of the following three redshift ranges: $0.10 < z < 0.11$, $0.19 < z < 0.20$ and $0.29 < z < 0.30$. This resulted in a sample of 812

extragalactic objects, with at least 200 in each of the three ranges. Note that at greater z there is more contamination from the host galaxy spectrum, but this will not affect the AGN classification.

Spectra presented on the SDSS Data Release 10 website [1] were inspected and classified according to the scheme utilised in the ZORROASTER catalogue developed by one of us [3]. We point out that such visual inspections are able to distinguish between subtle spectral differences that are easily missed by automated line fitting and emission flux searches. A total of 379 of the initial 812 objects (= 47%) were thus confirmed to be AGN with Seyfert-like spectra. The remaining spectra either only exhibited weak starburst or weak LINER-like emission, or displayed no emission lines at all (meaning that the x-ray source was probably erroneously linked to them).

Having classified the AGN, we grouped them where appropriate into one of six sets listed in Table 1 below. Objects that could not be clearly assigned to one of these sets were not considered further.

Table 1. Distribution of AGN sub-types of ROSAT-detected galaxies in the three redshift ranges investigated.

Object class	$z =$ 0.10-0.11	$z =$ 0.19-0.20	$z =$ 0.29-0.30
Total ROSAT sources	338	269	205
Total ROSAT AGN	115	142	122
standard Seyfert 1	41	48	44
NLS1 (weak Fe II)	11	14	5
I Zw 1 type objects	8	17	18
strong Fe II spectrum	18	26	31
peculiar Balmer profile	2	9	4
very wide Balmer lines	3	9	9

The six sets may be described as follows. Two sets correspond to the so-called “narrow line Seyfert 1” subclass, or NLS1 [4]. We distinguish between such objects exhibiting a rich Fe II spectrum (often referred to as I Zw 1 objects after their prototype) and those NLS1 that do not have a similar iron line spectrum. Another set includes Seyferts with strong Fe II lines (but who are not NLS1). We then introduce sets for unusually wide Balmer lines and for Balmer lines with complex, asymmetric profiles often including red and/or blue bumps [5]. The final set is for ‘ordinary’ Seyfert 1 galaxies, i.e. objects with no distinct Fe II, He II or coronal line features, unremarkable Balmer broad line profiles (neither very wide nor very thin, reasonably symmetric with no bumps), and no hybrid spectrum typical of starburst or LINER galaxies.

3. Analysis of AGN spectral type distribution

In order to determine whether the fraction of specific AGN subtypes varies between the three samples, we employed a chi-squared statistical test. In each case, we compared the set tested with the set for the ‘ordinary’ Seyferts.

The chi-squared test yields a value that signifies that probability that the fraction of AGN belonging to a set in one of the three redshift ranges is significantly different to the fraction for that set in another redshift range. The values obtained are listed as percentages in Table 2.

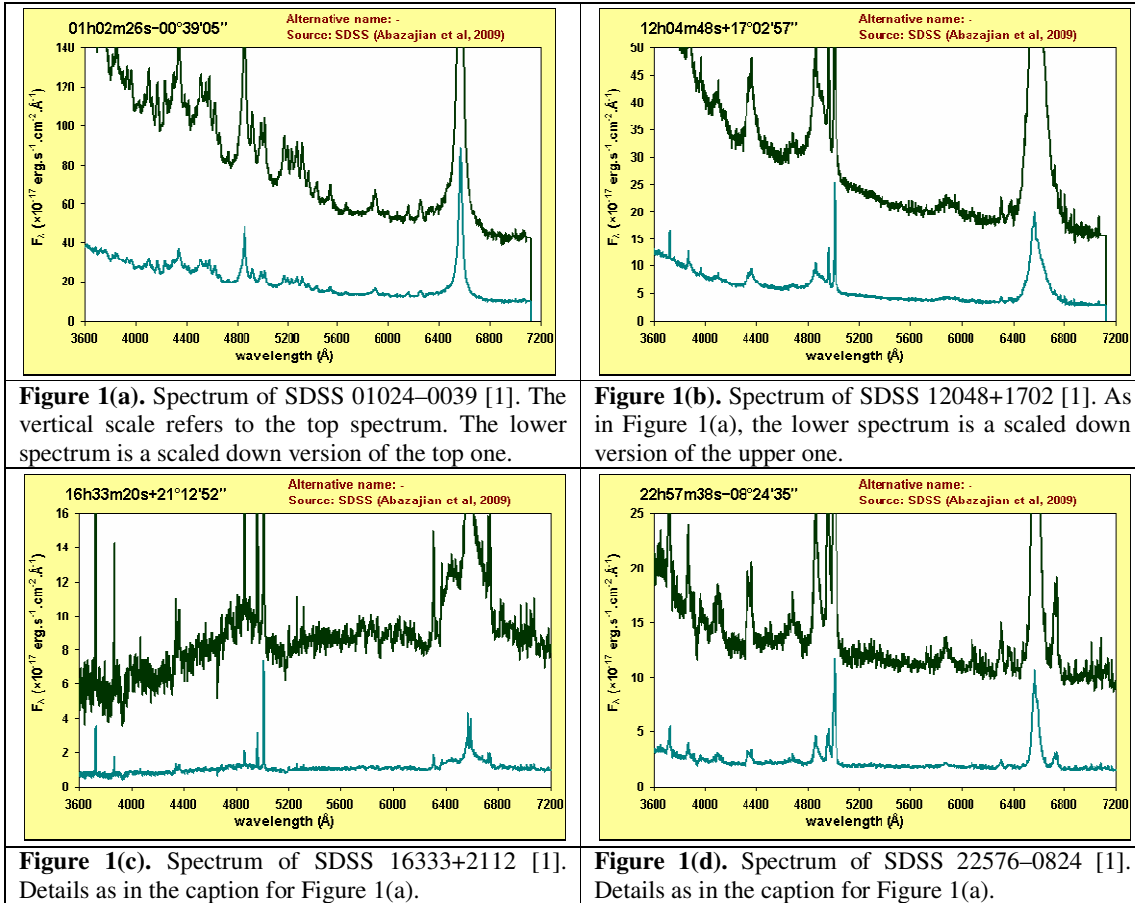
Table 2. Results of the chi-squared analysis of various AGN subtypes versus the ordinary Seyfert 1 samples. Values represent the probability that the samples are significantly different. Values above 90% are given bold.

	NLS1 (weak FeII)	I Zw 1 type objects	strong Fe II spectrum	peculiar Balmer profile	very wide Balmer lines
$z = 0.1$ vs $z = 0.2$	16%	84%	51%	94%	86%
$z = 0.2$ vs $z = 0.3$	94%	33%	68%	78%	15%
$z = 0.1$ vs $z = 0.3$	90%	93%	89%	53%	89%

None of the values exceeds the commonly employed 95% significance test, but several comparisons satisfy the less rigorous 90% significance mark. In particular, it is 90-95% likely that (i) AGN with peculiar Balmer profiles are less common at $z = 0.1$, (ii) NLS1 with weak Fe II are less common at $z = 0.3$ and (iii) I Zw 1 objects are less common at $z = 0.1$. There is also a hint that AGN with wide Balmer lines or strong Fe II occur more often at greater z , but this cannot be fully confirmed. Alternatively, the results may highlight intrinsic differences in the average luminosities of several of the sets. Resolving this requires an independent investigation. The present study also did not consider galaxy environments and clustering properties

Appendix. Description of interesting AGN identified

In Figure 1 we display four of the more noteworthy spectra in the investigated sample, which are briefly described in the following subsections.



3.1. SDSS 01024–0039

This is a well known nearby quasar, whose spectrum has been investigated in the past [6,7]. We note a very strong Fe II spectrum. No narrow lines are visible in Figure 1. The emission in the wavelength range in which the [O III] nebular lines are normally found are here in fact due to Fe II.

The spectrum in some degree resembles that of the nearby quasar Mkn 231 and the more recently discovered SDSS 12001–0204 [8], who also exhibit an exceptionally strong Fe II spectrum. Unlike these two objects, it does however not display a strong Na I absorption feature at 5892 Å.

3.2. SDSS 12048+1702

Figure 1(b) displays a Seyfert spectrum with very wide Balmer lines, which are in addition slightly asymmetric. In addition, the spectrum contains no detectable Fe II emission.

In contrast to other AGN featured in the ZORROASTER catalogue with a “w” descriptor for unusually wide lines, this object shows a distinctive He II 4686 Å broad emission feature.

3.3. SDSS 16333+2112

The spectrum of this object displays a complex Balmer line profile that is particularly striking in H-alpha (6562 Å). We note an H-alpha ‘blue bump’ clearly visible near 6400 Å, which would also be present in the remaining lines of the Balmer series. As the remaining Balmer lines are much weaker (both intrinsically, and also due to probable reddening), and due to greater contamination due to integrated starlight from the host galaxy, the ‘blue bump’ cannot be seen elsewhere.

The unusual profile is understood to be the result of the rotational dynamics of the AGN accretion disk and reverberation effects. This has been noted in a significant number of other AGN [5].

3.4. SDSS 22576–0824

This AGN exhibits comparatively narrow Balmer lines, only marginally wider than the forbidden lines. It is in that sense a typical example of a narrow line Seyfert 1 (NLS1), in this case without detectable Fe II emission.

What makes this object unusual is the strong He II 4686 Å emission, which requires exposure of the line emission region to abnormally powerful ultraviolet irradiance. The presence of this line is confirmed in some Seyfert 1’s (those characterised by the descriptor “x” in the ZORROASTER catalogue [3]), but in a NLS1 this feature is rare.

Acknowledgments – This paper utilized data from the Sloan Digital Sky Survey (SDSS). Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is <http://www.sdss.org/>.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

References

- [1] Ahn C P, Alexandroff R, Allende Prieto C, et al. 2014 *Astrophys. J. Suppl.* **211** A17
- [2] Voges W, Aschenbach B, Boller T, et al. 1999 *Astron. Astrophys.* **349** 389
- [3] Winkler H 2013 *ZORROASTER online catalogue*,
<http://www.uj.ac.za/EN/Faculties/science/departments/physics/Pages/ZORROASTER.aspx>
- [4] Osterbrock D E, Pogge R W 1985 *Astrophys. J.* **297** 166
- [5] Eracleous M, Halpern J P 2003 *Astrophys. J.* **559** 886
- [6] Dong X-B, Ho L C, Wang J-G, et al. 2010 *Astrophys. J. Lett.* **721** L143
- [7] Zamfir S, Sulentic J W, Marziani P, Dultzin D 2010 *Mon. Notices Roy. Astron. Soc.* **403** 1759
- [8] Rahman N, Winkler H 2014 Proceedings of the SAIP 2013 Annual Conference, Richards Bay, South Africa, 8-12 Jul 2013, 368

AUTOCAL: A software application for the calibration of stellar magnitudes

D J Wium¹ and B van Soelen²

¹ Department of Computer Science and Informatics, University of the Free State, Bloemfontein, South Africa

² Department of Physics, University of the Free State, Bloemfontein, South Africa

E-mail: WiumDJ@ufs.ac.za

Abstract. We present a software application that largely automates the process of calibrating stellar magnitudes. AUTOCAL uses a pattern-matching algorithm that Edward Groth and Peter Stetson developed around 1985 to match stars on a FITS frame (using a list of frame coordinates and magnitudes as generated by IRAF) to entries of stars in a catalogue covering the same area of sky. A weighted least squares fitting is performed to calculate a linear function that best describes the relationship between the observed and catalogue magnitudes for these stars. The automation of the matching saves time and decreases the possibility of errors. Additionally, the automated matching can be performed on thousands of stars in a frame, thereby reducing the error margins for the calibrated magnitudes. We present an overview of the software application as well as results obtained for SXP6.85, a Be/X-ray Binary system in the Small Magellanic Cloud. The data were captured with the IRSF telescope located at the South African Astronomical Observatory during three observing runs spanning December 2007 to December 2010. The results are compared to previously published results and show a good agreement.

1. Introduction

The software application discussed here, called AUTOCAL, automates the calibration of stellar magnitudes. The calibration is achieved by comparing the apparent magnitudes of numerous sources in the field of the target to their catalogue magnitudes, and then calculating the linear least squares solution that best describes the relationship between the two sets of magnitude values. The apparent magnitude of the target is then calibrated by applying this linear least squares equation to it.

In order to compare the apparent and catalogue magnitudes of an observed star, AUTOCAL must first identify which star in the catalogue, if any, corresponds to the star in question. The algorithm used to achieve this is a pattern-matching algorithm based on triangle-matching. In essence, the algorithm searches for pairs of similarly shaped triangles, formed by placing stars at their vertices, where one triangle is constructed from the observation data, and the other from the catalogue data. This algorithm was simultaneously developed by Edward Groth [1] and Peter Stetson [2] around 1985.

Although faster and more efficient algorithms to solve the same problem have been developed since, the simplicity and robustness of the algorithm made it preferable for this project. Only a few small adjustments to the algorithm were made during the development of AUTOCAL.

A typical calibration using AUTOCAL can be loosely divided into three sections.

1. Preparation, during which the user provides the input required by AUTOCAL and adjusts the user-defined parameters to suit the dataset that is used.

2. Calibration, which is automated. This includes matching stars from the frame to those listed in the catalogue and all additional tasks required to perform the magnitude calibration.
3. Analysis of the results, during which the user may store the results of the calibration and use the text-based and visual feedback provided by AUTOCAL to ascertain if the calibration went as expected. Furthermore, the user may manipulate the calibration plot (catalogue magnitude vs. instrumental magnitude) in order to improve the fit.

AUTOCAL was used to calculate apparent magnitudes for eight Be/X-ray binary systems in the SMC (Small Magellanic Cloud). The results for SXP6.85 are discussed as an example. Data from three different observation campaigns that searched for variability in a number of such systems in the SMC were used. These campaigns spanned December 2007 to December 2010.

The magnitudes that were calculated using AUTOCAL were compared to magnitudes published in [3] and [5] from the same data and there was good agreement between the results.

2. Calibration using AUTOCAL

The three phases of the calibration procedure, as outlined in the introduction, will be elucidated here.

2.1. Preparation

In this phase the user is required to provide the input necessary for the calibration. Additionally, the user may adjust a number of parameters that affect how the matching of stars is performed. Figure 1 shows the start screen of AUTOCAL, where these steps are performed. Two text files containing the coordinates and magnitudes of the stars that are to be matched, are required. The first is generated from the FITS frame using a standard astronomical software package like IRAF [4] that identifies the positions and calculates instrumental magnitudes for stars on the frame. The coordinates in this file are the pixel coordinates of the FITS frame that correspond to the stars' positions. The second text file is constructed from a catalogue that includes the stars in the field of the FITS frame. AUTOCAL provides an interface that allows star catalogues to be imported into a database that is linked to it. The text file is generated by querying the database through a provided user interface. The coordinates in the second text file are the catalogued RA and Dec values.

The user-provided parameters are used to adjust the automated calibration process for the specific FITS frame that is used. These parameters determine the number of stars from each list used for provisional matching, the minimum separation allowed between stars from the same list before they are discarded, the maximum difference between the coordinates of a star for it to be considered a match and the maximum ratio between the longest and shortest edges of a triangle for it to be used in the provisional matching phase.

Some settings may work better for sparse frames while others may be optimal for densely populated ones. During the testing of AUTOCAL we found that the default parameters worked well for all the frames. Although all the frames had a relatively high star density (as one would expect for the SMC), the number of stars identified per frame varied from 41 to 907, thus indicating that the matching algorithm and its parameters are robust with respect to the density of stars on the FITS frame.

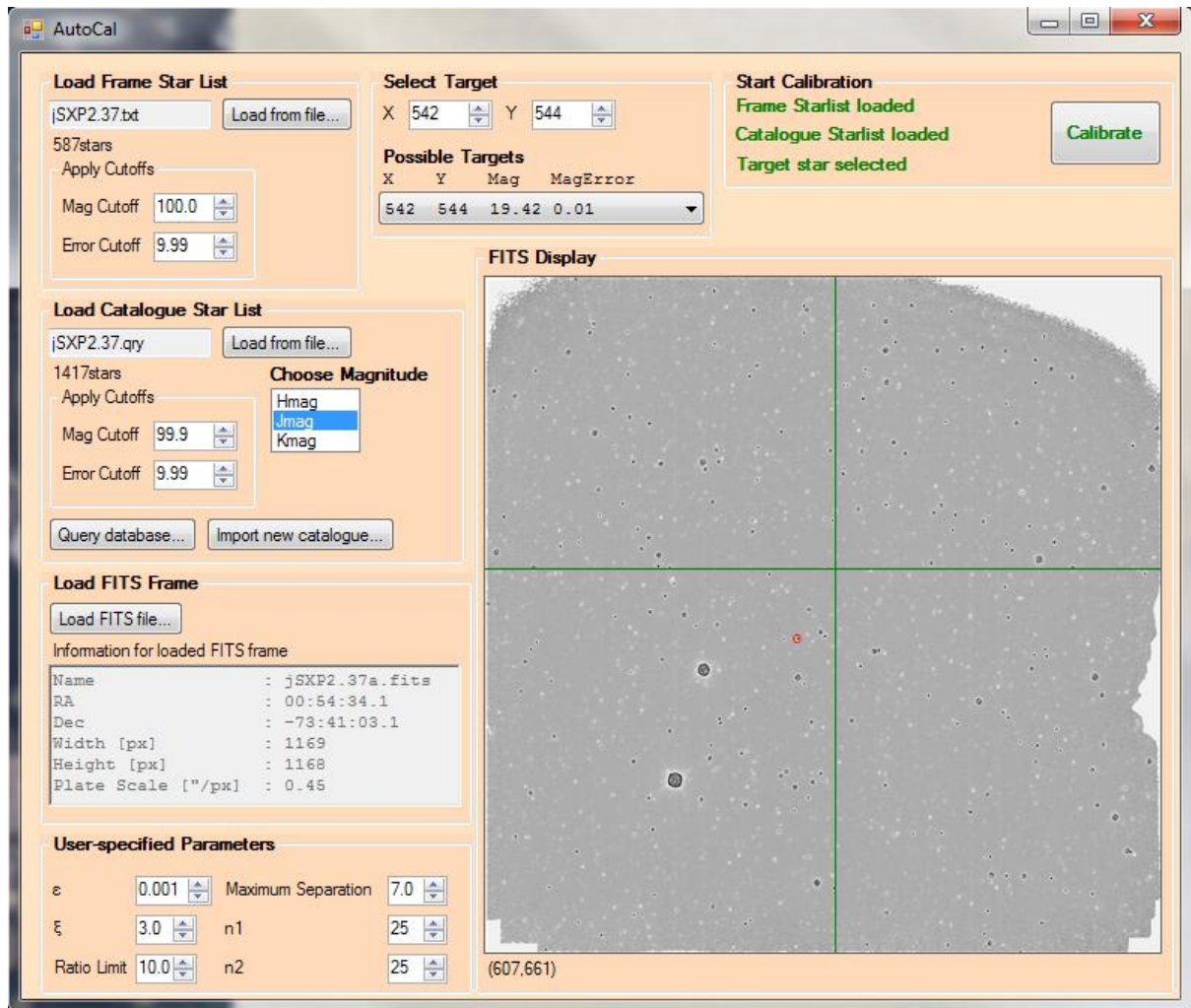


Figure 1. The start screen of AUTOCAL where the user provides input and adjusts the star matching parameters.

2.2. Automated calibration

The calibration procedure is completely automated in AUTOCAL and the software is generally able to match the vast majority of stars on the FITS frame to entries in the catalogue, and to calculate a linear relationship between the two sets of magnitudes, within a second or two. The steps that AUTOCAL follow are listed below. A full discussion of each step is beyond the scope of this document.

1. The target is removed from the list of FITS frame stars and all stars closer to one another than a user-specified threshold are eliminated from both the frame star list and the list of stars from the catalogue to reduce the possibility of incorrect matches being made.
2. The brightest stars from each of the two lists are selected for provisional matching. The number of stars selected from each list is specified by the user.
3. AUTOCAL matches triangles generated from the two lists using the procedure described in detail in [1] and [2]. Stars on corresponding vertices of matched triangles are matched provisionally.
4. The provisionally matched stars are used to transform the coordinate system of the catalogue to that of the FITS frame. The coordinate system of the catalogue is fixed, e.g. it uses equatorial coordinates for J2000.0. The pixel map of the FITS frame is transformed into the catalogue system.

5. With the two lists of stars now using the same coordinate system, entries from the two lists are matched based on their proximity to an entry from the other list.
6. A refined coordinate transformation for the catalogue is now calculated and applied using the larger number of matches found in the previous step.
7. Final matches between entries from the two lists are calculated as in step 5, but using the refined coordinate transformation calculated in step 6.
8. A weight is assigned to each of the matches from step 7, with matches with smaller uncertainties in their magnitudes receiving larger weights. The weights are calculated as $w = (\sigma_{\text{cat}}^2 + \sigma_{\text{ins}}^2)^{-1}$, with σ_{cat} and σ_{ins} denoting the uncertainties in the catalogued and instrumental magnitudes, respectively.
9. A weighted least squares approach, using the weights assigned in the previous step, is used to compute a linear relationship between the catalogue and instrumental (frame) magnitudes of the matched stars.
10. The linear equation calculated in step 9 is applied to the instrumental magnitude of the target, thus calibrating its magnitude to the catalogue. Additionally, the uncertainty in this calibrated magnitude is calculated.

2.3. Post-calibration Analysis

Upon completion of the calibration, the results of the process are displayed. The user can adjust the results by removing matches that are obvious outliers, caused by incorrect matching or faulty magnitude values, and also matches of stars with large magnitudes. Larger magnitudes generally have larger uncertainties and thus smaller weights in the calculation of the linear equation. Often a very good fit can be obtained by using a selection of the brightest matched stars only. An example of the above can be found in the calibration of the magnitudes for SXP6.85, which is discussed in the next section.

Additionally, AUTOCAL allows the user to save an automatically generated summary of the calibration to a text file for further analysis. This summary contains details of every step of the calibration, for instance the number of stars used from each list during the step and the transformation coefficients used for the coordinate transformations.

Once the user is satisfied with the results, the final results can be saved to a text file. This file includes entries for each star that was matched, indicates whether each of these entries were used in the final analysis, and records the equation of the final linear equation as well as the instrumental and calibrated magnitudes of the target. The graph of catalogue magnitudes vs. instrumental magnitudes for the matched stars, from which the linear equation is calculated, can be saved to a bitmap image file.

3. Calibration of magnitudes for SXP6.85 using AUTOCAL

We used AUTOCAL to calibrate magnitudes for SXP6.85 from data captured during three observing campaigns that searched for variability in Be/X-ray binary sources in the SMC. The first campaign took place during December 2007, the second during December 2009 and January 2010 and the third during December 2010. Variability had been reported for SXP6.85 in [3] and [5], both incorporating magnitudes calculated from the same data used here.

The data were captured with the IRSF (InfraRed Survey Facility) telescope at the SAAO's Sutherland observatory using the SIRIUS camera. SIRIUS allows simultaneous exposures to be taken at J (1.25 μm), H (1.65 μm) and K_s (2.15 μm) bands with a field of view of $\sim 7.7 \times 7.7$ arcmin² [6]. Science frames were constructed using the SIRIUS09 [7] pipeline package that is integrated into IRAF. Standard IRAF/NOAO packages [8] were used to perform PSF photometry on the science frames. The IRSF point source catalogue [9] was used to obtain published magnitude values for the stars on the science frames.

Figure 2 is a plot of catalogue magnitudes vs. instrumental magnitudes for stars that were matched during the calibration of SXP6.85 for an observation done on 18 December 2007 using a J-band filter.

The linear equation calculated by AUTOCAL is displayed in the top-left corner. For this observation the instrumental magnitude of SXP6.85 was 19.60 ± 0.01 and the calibrated magnitude was calculated to be 14.80 ± 0.01 . Blue dots indicate matches that were discarded due to being obvious outliers, representing faint stars or being more than 3σ removed from the final fit.

The results for all three observing campaigns are listed in table 1 along with the results published in [3] and [5]. There is good agreement between the results, with the H-band and K_s -band magnitudes for 18 December 2007 agreeing within the uncertainties specified. There is a slight discrepancy between the results obtained by AUTOCAL and those reported in [3]. The difference is introduced due to the removal of outlining data. If no outlier removal is performed, AUTOCAL calculates a J magnitude of 14.84 ± 0.03 , in agreement with [3]. It is clear from figure 2 that the inclusion of the outliers would skew the fit from the optimal solution.

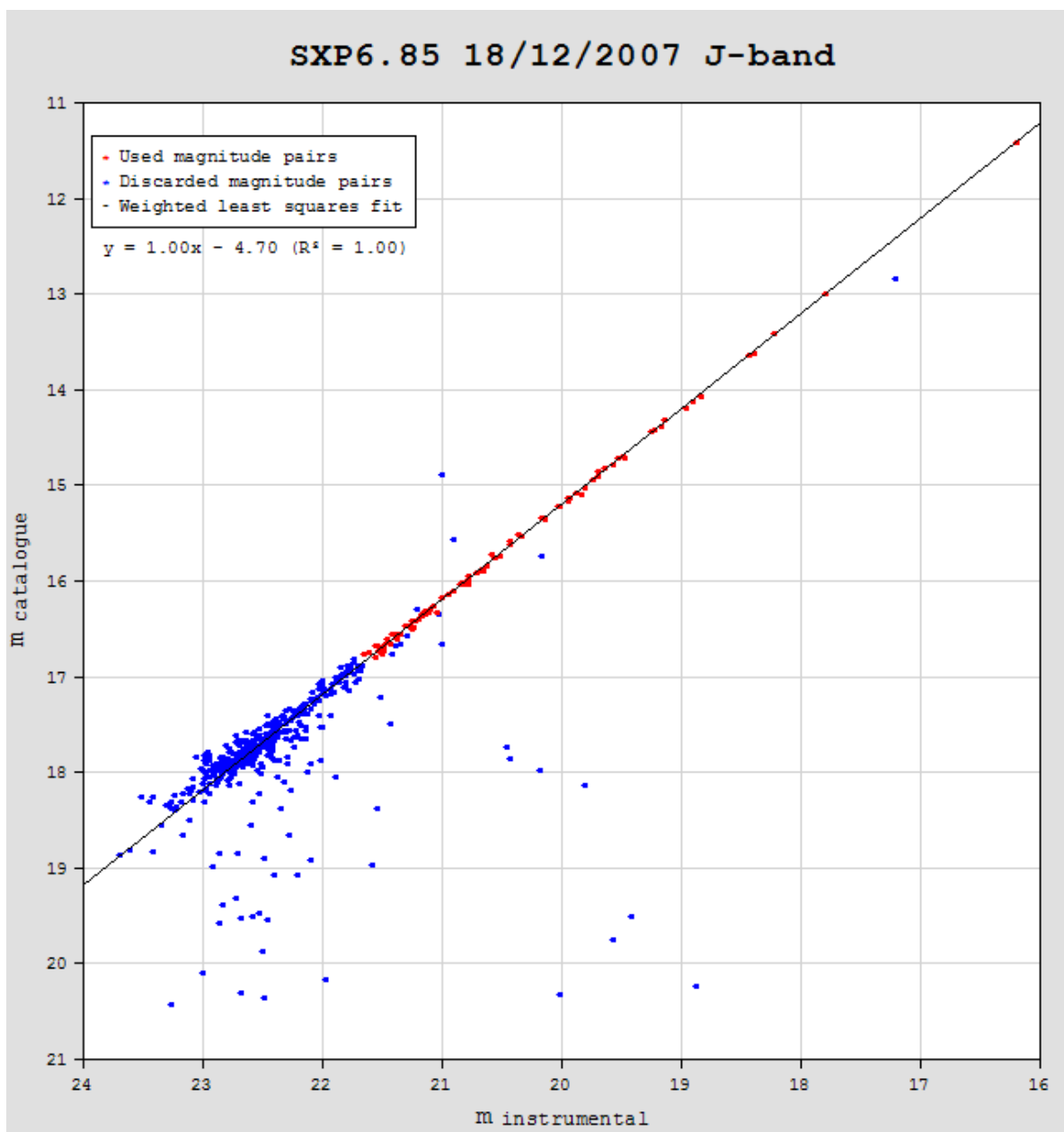


Figure 2. Plot of catalogue magnitudes vs. instrumental magnitudes for stars matched during calibration of SXP6.85, as displayed by AUTOCAL.

Table 1. Calibrated magnitudes calculated for SXP6.85, as calculated by AUTOCAL and published in [3] and [5].

Date	Filter	AUTOCAL	Townsend (2012)
18/12/2007	J	14.80±0.01	14.84±0.02
	H	14.76±0.01	14.77±0.02
	K _s	14.66±0.01	14.67±0.03
30/12/2009	J	14.85±0.01	
	H	14.83±0.01	
	K _s	14.81±0.02	
19/12/2010	J	14.60±0.01	
	H	14.50±0.02	
	K _s	14.37±0.05	

4. Conclusion

AUTOCAL was used to calculate calibrated magnitudes for a Be/X-ray binary system in the SMC and the results agree well with previously published magnitudes calculated from the same data. Generally, AUTOCAL matches a much larger number of stars to a catalogue than those used during a manual calibration. This results in smaller uncertainties in the calibrated magnitudes.

During its testing, which was performed on a total of 55 frames from 8 fields, AUTOCAL was able to calculate calibrated magnitudes for all 55 of the frames.

The advantages of using AUTOCAL over manual calibration include a reduction in the time spent on calibration, elimination of human error and the availability of reports and plots for further analysis. Perhaps the primary advantage of AUTOCAL is its interactive nature, specifically regarding outlier removal. This can lead to more accurate results, as illustrated in section 3.

References

- [1] Groth E J 1986 A pattern-matching algorithm for two-dimensional coordinate lists *AJ* **91**(5) 1244-8
- [2] Stetson P B 1990 The techniques of least squares and stellar photometry with CCDs *Available online at http://ned.ipac.caltech.edu/level5/Stetson/Stetson_contents.html* Accessed 26 June 2011
- [3] Townsend L J 2012 An X-ray, optical and infra-red study of high-mass X-ray binaries in the Small Magellanic Cloud *PhD thesis, University of Southampton*
- [4] IRAF version 2.16 *Available online at <http://iraf.noao.edu/>*
- [5] Townsend L J et al. 2010 Be/X-ray binary SXP6.85 undergoes large type II outburst in the Small Magellanic Cloud *MNRAS* **403**1239-45
- [6] Nagata T, Tanabé T and Glass I S 2003 *IRSF observing manual*
- [7] Nakajima Y 2011 *SIRIUS09 pipeline users manual*
- [8] Massey P and Davis L E 1992 *A user's guide to stellar CCD photometry with IRAF*
- [9] Kato D et al. 2007 The IRSF Magellanic Clouds point source catalog *PASJ* **59**(3) 615-41

Division D2 – Space Science

Influence of lightning on total electron content in the ionosphere using WWLLN lightning data and GPS data

M M Amin¹² M Inngs¹ P J Cilliers²

¹South African National Space Agency (SANSA) Space Science, P.O. Box 32, Hermanus 7200, South Africa

² Department of Electrical Engineering, University of Cape Town, Western Cape, Cape Town 7100, South Africa

E-mail: cyrusdv92@yahoo.com (M M Amin) mikings@gmail.com (M Inngs)
pjcilliers@sansa.org.za (P J Cilliers)

Abstract. An analysis of total electron content (TEC) variations detected over southern Africa has been performed. We included two days with high lightning activity, one day with low lightning activity and one day with a geomagnetic storm. An attempt to obtain a better understanding of lightning as a source mechanism is motivated by the fact that lightning couples energy directly to the mesosphere and lower ionosphere through quasi-electrostatic (QE) and electromagnetic pulsed (EMP) fields and through upward propagating gravity waves that transfer energy from the site of lightning into the ionosphere. Geomagnetic indices, GPS data from six dual frequency GPS reference stations and lightning data from the World-Wide Lightning Location Network (WWLLN) have been used to explore the origin of the TEC variations detected. The analysis reveals periods of TEC variations of ~ 1 TECU on geomagnetically quiet days which correspond to periods of intense lightning activity in the regions. The TEC variations on the days with high lightning activity appear to have more high-frequency TEC variation content than days with low lightning activity.

1. Introduction

Over the past few decades considerable attention has been given to the effects of thunderstorms and the lightning they produce on the middle and upper atmosphere. For example, a statistical study by Davis and Johnson [1] found a correlation between intensification in the sporadic E layer and lightning activity. In addition, modeling of atmospheric gravity waves (AGWs), originating from thunderstorms, has predicted variations in total electron content (TEC) associated with these AGWs of $\pm 7\%$ [2] and other observations [3, 4] indicate gravity wave effects on the ionosphere from tropical and mid-latitude storms. Recent studies by Lay and Shao [5] on the D layer ionosphere (~ 65 – 90 km altitude) have shown that AGWs originating from large mesoscale thunderstorms clearly perturb the electron distribution at the lower boundary of the ionosphere. Studies [6, 7] have shown that the most active regions of thunderstorms are over Africa, the Americas and South-East Asia, with the major proportion of lightning occurring over land. This work is motivated by the potential links between thunderstorms and variations in the ionosphere over South Africa as observed through the background TEC variations. The TEC

variation is an indicator of ionospheric variability and can be derived from GPS signals. The TEC variation is also the parameter of the ionosphere that produces most of the effects on GPS signals. Lightning measurement by the World-Wide Lightning Location Network (WWLLN) was used as a proxy to identify regions of intense thunderstorms and the observations through the ionosphere was conducted by analyzing TEC measurements from ground-based dual frequency GPS receivers in regions with different thunderstorm activity levels within South Africa.

2. Data and Method of Analysis

The automatic lightning detector at the South African National Space Agency (SANSA) Space Science in Hermanus and another in Durban, both in South Africa, and several other such detectors distributed across the globe constitute the WWLLN, which logs the incidences of lightning strikes all over the world (<http://webflash.ess.washington.edu/>). The WWLLN exploits the VLF (3-30 kHz) band term as “sferics” emitted by lightning during the return stroke and employs the Time of Group Arrival (TOGA) described by Dowden et al. [8] to determine the location of each stroke. Each lightning stroke location requires the TOGA from at least four stations with time residual of less than 30 microseconds [9, 10]. The WWLLN currently has over 50 sensors distributed across the globe and provides real-time lightning locations globally [11] with less than 10 km and 10 microseconds location and timing errors respectively [12].

The ionospheric data (TEC) presented in this study were derived from RINEX file which contains data obtained from six GPS receiver stations belonging to the Trignet (<http://www.trignet.co.za/>) network. The GPS-TEC software developed at Boston College [13] was used to process the RINEX files to derive the integrated electron density along the signal path between the receiver on the ground and the GPS satellite in space, commonly referred to as slant TEC (STEC). The STEC is calculated at an Ionospheric Pierce Point (IPP) altitude of 350 km with an elevation mask of 15° over the mid-latitude stations. The GPS-TEC software uses the phase and code values for both L1 and L2 GPS frequencies to eliminate the effect of clock errors and tropospheric water vapor to calculate relative values of STEC [14]. To obtain the absolute TEC values, the differential satellite biases (published by the University of Bern) are included along with the receiver bias values that are calculated by minimizing the TEC variability between 02:00 and 06:00 local time [15]. The biases in the observed STEC are accounted for by computation via the expression below.

$$(STEC)_{desired} = STEC + B_{Rx} + B_{Rich} + B_{sat} \quad (1)$$

where B_{Rx} , B_{Rich} and B_{sat} are the receiver bias, receiver inter-channel bias and satellite biases respectively. In order to minimize the multipath effects on GPS data, an elevation cutoff of 20° was applied in selecting data for the analysis. In the method adopted here, the control curves or “undisturbed” conditions (TEC_{fit}) used are modeled by fitting a 6th order polynomial to the STEC measurements by a GPS pseudo-random number (PRN) from each station, similar to the method described in Lay et al. [16]. The degree of variations in the STEC was estimated by subtracting the TEC_{fit} values from the STEC. The removal of the TEC_{fit} tends to filter out lower-frequency variations, hence this method is only sensitive to higher-frequency variation. Geomagnetic activity was expressed in terms of the global geomagnetic index (Kp).

3. Results

The normalized global lightning data from the WWLLN are used to map intense thunderstorm regions over Southern Africa with high time and spatial resolution. The lightning data was recorded in universal time (UT) system per event which is equivalent to South Africa standard time (SAST) (UT + 2 hours).

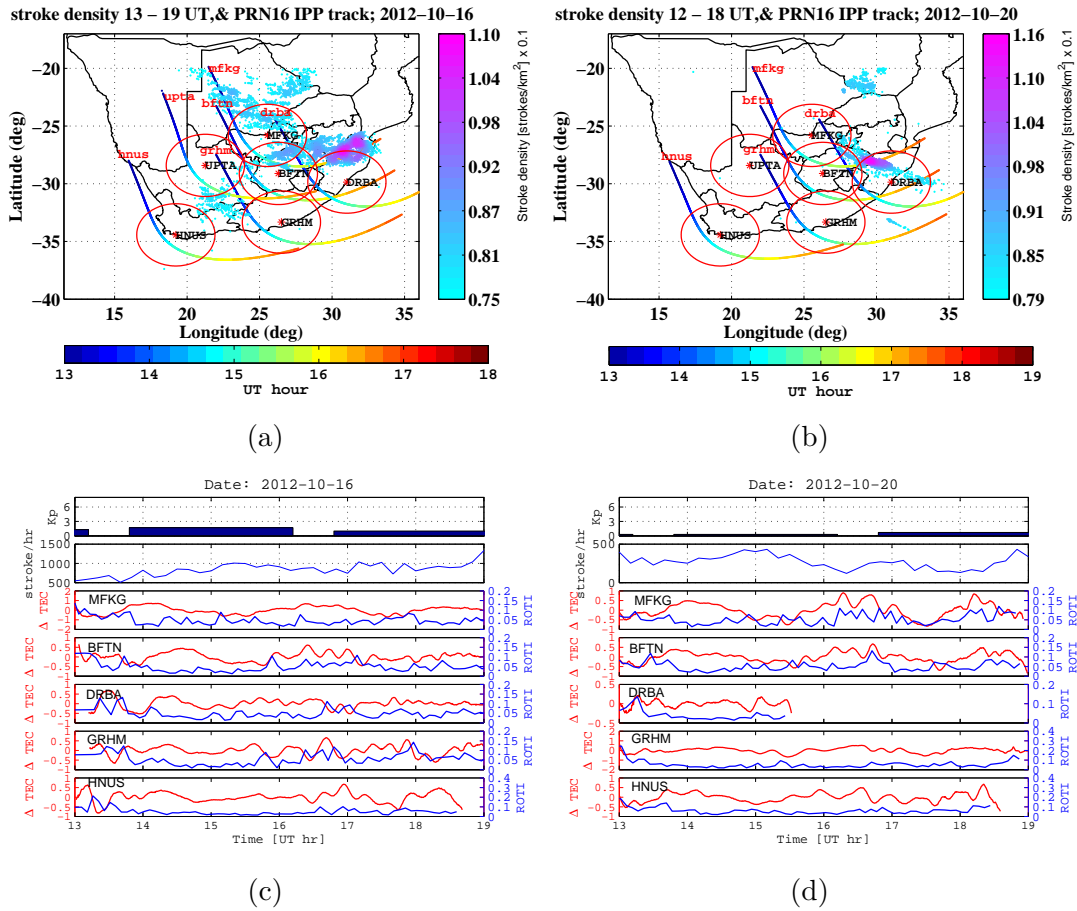


Figure 1: Top row: corresponding TEC measurement paths (Ionospheric Pierce Point (IPP) at 350 km) from PRN 16 between 13:00 and 19:00 UT visible at individual reference receiver stations for high lightning activity and geomagnetically quiet days (a) 16 October, 2012 and (b) 20 October, 2012. The rainbow colored curves in (a) and (b) indicate the measuring path of the GPS satellite (PRN-16). The color of the paths is the corresponding time in UT as shown in the colorbar below the figure. The red ellipses indicate 300 km radius around each GPS station. The distribution of stroke/km² is also presented in color scale (from light blue to violet): The light blue color indicates regions of relative low stroke/km² and the violet color indicates regions of relative high stroke/km². Bottom rows: Kp index, stroke rate, STEC variations (Δ TEC) and ROTI for PRN 16, (c) on 16 October, 2012 and (d) on 20 October, 2012.

GPS data from six reference GPS receiver stations between 13:00 and 19:00 UT on 5, 8, 16 and 20 October, 2012 has been selected, processed and analyzed and the results to be presented in Figures 1 and 2. The STEC obtained from the GPS-TEC algorithm is used to derived the rate of change of TEC (ROT) and ROT index (ROTI). The ROTI, which can be used as a proxy for the S_4 amplitude scintillation index (S_{4p}) is defined as the standard deviation over a 5-minute period of ROT calculated from 30-second sampled GPS data [17].

Figure 1, shows that the significant thunderstorm activity which occurred on 16 October, and 20 October 2012 correspond to enhanced TEC variations along the ray path for satellite PRN 16. Although 16 October day has relatively low geomagnetic activity ($K_p = 2.3$), it has a very active thunderstorm that sustained lightning rates of > 1000 per hour from 13:00 UT to 19:00 UT. The maximum lightning density was $\sim 0.11/\text{km}^2$ located at about 300 km to the north of DRBA Figure 1(a). The ionosphere over South Africa is significantly disturbed during this period as measured by the GPS stations BFTN, DRBA, GRHM, MFKG and HNUS from PRN 16 Figure 1(c).

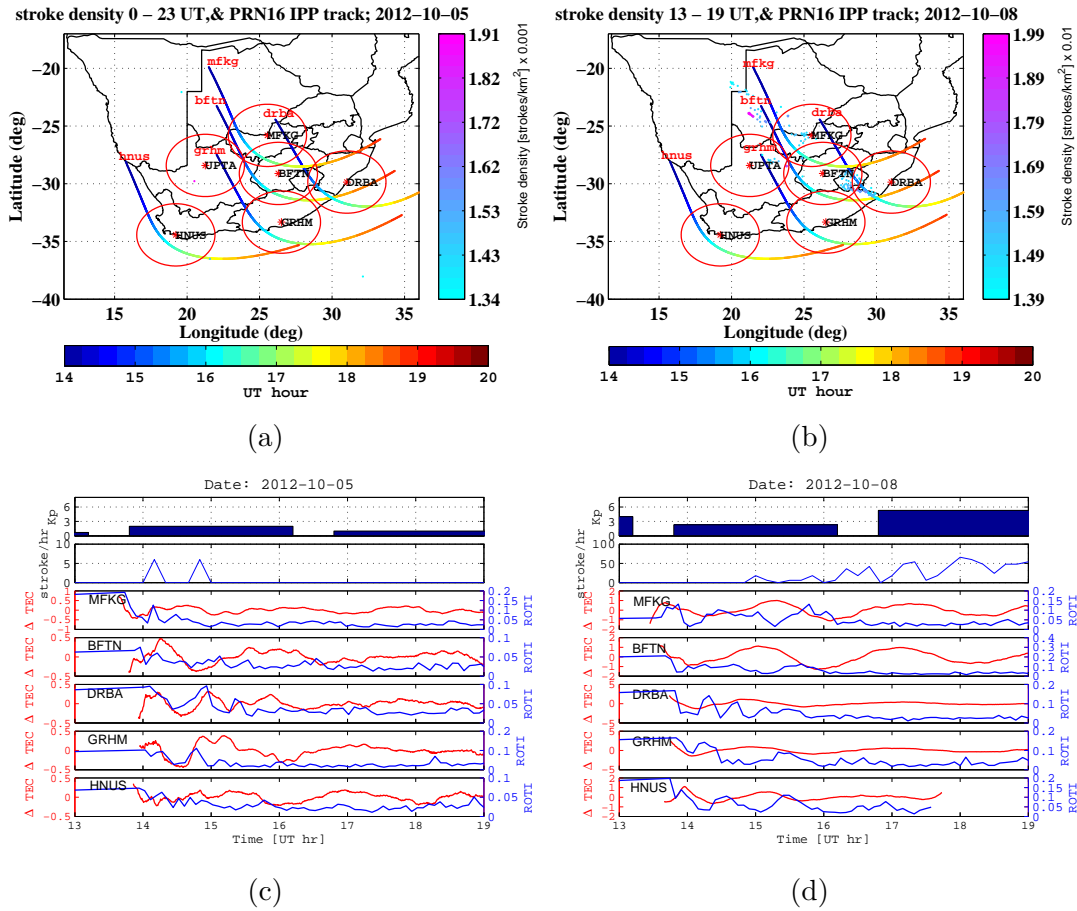


Figure 2: Top rows:corresponding TEC measurement paths (IPP at 350 km) from PRN 16 between 13:00 and 19:00 UT on (a) 5 October, 2012 and (b) 8 October, 2012. Bottom row: Bottom rows: Kp index, stroke rate, STEC variations (Δ TEC) and ROTI for PRN 16 on (c) 5 October, 2012 and (d) 8 October, 2012

October 20 has much less geomagnetic activity ($K_p = 0.7$) and a thunderstorm with a relatively higher lightning activity ($\sim 0.116/\text{km}^2$) in a fairly concentrated area within 300 km radius of DRBA. The lightning rate is relatively low (< 500 per hour) between 13:00 UT and 21:00 UT. Enhanced TEC variations also occurred as measured by the above mentioned GPS stations from PRN 16 during the same period when the thunderstorm is active. On 08 October, due to enhanced geomagnetic activity ($K_p = 6.3$), the TEC variation is larger as measured by GPS stations BFTN, DRBA, GRHM, MFKG, HNUS from PRN 16. Again no data is available from UPTA from PRN 16. On 05 October 2012 there was much less lightning activity ($\sim 0.0019/\text{km}^2$) as shown in Figure 2(b) than on the other three days, and there are no clear enhanced TEC variations as observed by the GPS stations from PRN 16 (Figure 2(d)).

It is interesting to note that the Δ TEC variations on the thunderstorm-active days (Figure 1(c) and Figure 1(d)) have more high-frequency TEC variation contents superimposed on low-frequency background TEC variation than the thunderstorm-quiet (Figure 2(c)) and geomagnetic-active (Figure 2(d)) days.

4. Discussions and conclusions

There is evidence [18] for multiple mechanisms by which lightning can enhance the ionospheric ionization. Our analysis show evidence of lightning related perturbations in ionospheric TEC in the range of 0.5 to 1.2 TECU on selected days when there was high lightning activity and low geomagnetic activity. The mechanism responsible for the perturbations detected

on the thunderstorm-active days is not clear at this stage. Some plausible mechanisms of the perturbations could be: propagating atmospheric gravity waves [1], vertical electrical discharge (such as sprites) [19, 20, 21], AGWs produced when convective thunderstorm activity overshoots the tropopause [22, 2], thunderstorm-produced atmospheric density bubbles [23], or thunderstorm-triggered Perkins instabilities [24] or by a combination of these mechanisms. Wilson [25] also speculated that the ionospheric electric field generated by a large rain cloud may be sufficient to induce ionization at ionospheric altitudes without lightning. Mechanisms involving electrical discharge and infrasonic waves would enhance the ionosphere directly above while atmospheric gravity waves launched by lightning [26, 27] have a very small vertical component and would need to propagate several hundred kilometers horizontally before reaching ionospheric altitudes.

The peak-to-peak magnitude of the ΔTEC variation along the measurement paths (i.e., *bftn* and *mfgk*) in Figure 1(d) is about 1 TECU on the thunderstorm-active and low geomagnetic-active (i.e. $K_p = 0.7$) day. In comparison, numerical simulations studies have suggested that density perturbations of the order of 5% or even less, attributed to atmospheric gravity waves, can trigger instability growth in the ionosphere leading to vertical development of bubbles to the topside ionosphere under the typical dynamic state of the background ionosphere during the sunset period [28]. In addition Vadas and Liu [2] reported 5% TEC variation due to AGWs generated by a convective plume 1-1.5 h before sunset. Based on the International Reference Ionosphere [29], a typical pre-sunset background of ~ 40 TECU is inferred for the general geographical location. Therefore, a 5% change implies a 2 TECU variation. The result obtained in this paper (1 TECU) and that inferred from Vadas and Liu [2] (2 TECU) could be considered comparable. This discussion presents some individual possibilities for the cause of these TEC perturbations, but it is also possible that multiple mechanisms simultaneously affect the ionosphere above thunderstorms, leading to even more complex coupling dynamics. The fact that there is no noticeable TEC perturbation for geomagnetic-quiet (05 October, 2012) day without lightning is evidence that lightning is necessary for the enhanced TEC perturbation on the thunderstorm-active days (16 and 20 October, 2012). We suggest that because the TEC variations typically occur some distance from the region where the lightning activity is the largest confirms the hypothesis that the enhanced TEC variations is likely due to propagating atmospheric gravity waves that transfer energy from the site of lightning into the ionosphere, or vertical electrical discharge, or by a combination of these two mechanisms. Future data analysis and modeling efforts are needed to understand the possible mechanisms associated with the TEC perturbation.

Acknowledgments

The lightning data used in this work is obtained from the World Wide Lightning Location Network through collaborations with research institutions across the globe including South African National Space Agency. The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged.

References

- [1] Davis C J and Lo K-H 2005 Lightning-induced intensification of the ionospheric sporadic-E layer, *Nature* **435** 799–801.
- [2] Vadas S L and Liu H -l 2009 Generation of large-scale gravity waves and neutral winds in the thermosphere from the dissipation of convectively generated gravity waves, *J. Geophys. Res.: Space Phys.* **114** A10 doi:10.1029/2009JA014108.
- [3] Bishop R L Aponte N Earle G D Sulzer M Larsen M F and Peng G S 2006 Arecibo observations of ionospheric perturbations associated with the passage of Tropical Storm Odette *J. Geophys. Res.: Space Phys.* **111** 1978–2012.

- [4] Kelley M C 1997 In situ ionospheric observations of severe weather-related gravity waves and associated small-scale plasma structure, *J. Geophys. Res.: Space Phys.* **102** 329–335.
- [5] Lay E H Shao X -M 2011 High temporal and spatial-resolution detection of D-layer variations by using time-domain lightning waveforms, *J. Geophys. Res.: Space Phys.* **116** A1 10.1029/2010JA016018.
- [6] Collier A B Hughes A R W Lichtenberger J Steinbach P 2006 Seasonal and diurnal variation of lightning activity over southern Africa and correlation with European whistler observations, *Ann. Geophys.* **24** 529–542.
- [7] Lay E H Jacobson A R Holzworth R H Rodger C J Dowden R L 2007 Local time variation in land/ocean lightning flash density as measured by the world wide lightning location network, *J. Geophys. Res.: Atmos.* **112** D13.
- [8] Dowden R L Brundell J B Rodger C J 2002 VLF lightning location by time of group arrival (TOGA) at multiple sites *J. Atmos. and Solar-Terres. Phys.* **64** 817–830.
- [9] Rodger C J Brundell J B Dowden R L 2005 Location accuracy of VLF World Wide Lightning Location (WWLL) network: Post-algorithm upgrade, *Ann. Geophys.* **23** 277–290.
- [10] Rodger C J Brundell J B Dowden R L Thomson N R 2004 Location accuracy of long distance VLF lightning location network *Anna Geophys.* **22** 747–758.
- [11] Rodger C J Brundell J B Holzworth R H Lay E H April 2009 Growing Detection Efficiency of the World Wide Lightning Location Network, In *American Inst. Phys. Conf. Series* **1118** 15–20 doi:10.1063/1.3137706.
- [12] Abarca S F Corbosiero K L Galarneau T J 2010 An evaluation of the Worldwide Lightning Location Network (WWLLN) using the National Lightning Detection Network (NLDN) as ground truth *J. Geophys. Res.: Atmos.* **115** doi:10.1029/2009JD013411.
- [13] Seemala G K Valladares C E 2011 Statistics of total electron content depletions observed over the south american continent for the year 2008, *Radio Sci.* **46** doi:10.1029/2011RS004722.
- [14] Sardón E Zarraoa N 1997 Estimation of total electron content using GPS data:How stable are the differential satellite and receiver instrumental biases? *Radio Sci.* **32** 1899–1910.
- [15] Doherty P Coster A Murtagh W 2004 Eye on the Ionosphere: space weather effects of October-November 2003 *GPS Solutions* **8** 267–271.
- [16] Lay E H Shao X -M Carrano C S 2013 Variation in total electron content above large thunderstorms *Geophys. Res. Lett.* **40** 1945–1949.
- [17] Pi X Mannucci A J Lindqwister U J Ho C M 1997 Monitoring of global ionospheric irregularities using the Worldwide GPS Network. *Geophys. Res. Lett.* **24** 2283–2286.
- [18] Johnson C G Davis C J 2006 The location of lightning affecting the ionospheric sporadic-E layer as evidence for multiple enhancement mechanisms *Geophys. Res. Lett.* **33** doi:10.1029/2005GL025294.
- [19] Neubert, T Rycroft M Farges T Blanc E Chanrion O Arnone E Odzimek A Arnold N Enell C -F Turunen E Bösinger T Mika A Haldoupis C Steiner R Velde O Soula S Berg P Boberg F Thejll P Christiansen B Ignaccolo M Fflekrug M Verronen P Montanya J Crosby N 2008 Recent results from studies of electric discharges in the mesosphere *Surveys in Geophys.* **29** 71–137.
- [20] Laštovička J 2006 Forcing of the ionosphere by waves from below, *J. Atmos. and Solar-Terr Phys.* **68** 479–497.
- [21] Sentman D D Wescott E M Osborne D L Hampton D L Heavner M J 1995 Preliminary results from the Sprites 94 Aircraft Campaign: 1. red sprites *Geophys. Res. Lett.* **22** 1205–1208.
- [22] Alexander M J Holton J R Durran D R 1995 The gravity wave response above deep convection in a squall line simulation *J. Atmos. Sci.* **52** 2212–2226.
- [23] Kuo C -L Lee L C Hsu R -R Huba J D 2012 Ionosphere density variations and formation of plasma bubbles caused by thunderstorm currents *Eos Trans. AGU; AGU Fall Meet. Suppl. Abstract* AE41A-04.
- [24] Perkins F 1973 Spread F and ionospheric currents, *J. Geophys. Res.* **78** 218–226.
- [25] Wilson C T R 1925 The electric field of a thundercloud and some of its effects *Proc. Phys. Soc. Lond.* **37** 32–37.
- [26] Chimonas G 1971 Enhancement of sporadic E by horizontal transport within the layer *J. Geophys. Res.* **76** 4578–4586.
- [27] Shrrstha K 1971 Sporadic-E and atmospheric pressure waves *J. Atmos. and Terr. Phys.* **33** 205–211.
- [28] Abdu M A Kherani E A Batista I S de Paula E R Fritts D C Sobral J H A 2009 Gravity wave initiation of equatorial spread F/plasma bubble irregularities based on observational data from the SpreadFEx campaign *Ann. Geophys.* **27** 2607–2622.
- [29] Bilitza D McKinnel L -A Reinisch B and Fuller-Rowell T 2011 International reference, ionosphere today and in the future, *J. Geod.* **85** 909–920 doi:10.1007/s00190-010-0427-x.

Division E – Physics Education

Relationship between grade 10 learners' views about nature of science and contextual factors

V M Baloyi¹, H I Nordhoff², W E Meyer¹, E Gaigher³ and M H W Braun³

¹ Department of Physics, University of Pretoria, Private bag X20, Hatfield, Pretoria, South Africa, 0028

² Sci-Enza, University of Pretoria, Private bag X20, Hatfield, Pretoria, South Africa, 0028

³ Department of Science, Mathematics and Technology Education, University of Pretoria, Private bag X20, Hatfield, Pretoria, South Africa, 0028

E-mail: vonani.baloyi@up.ac.za

Abstract: This study investigated relationships between learners' views about the nature of science (NOS) and contextual factors. The participants were a group of 100 grade 10 learners who applied to participate in a science enrichment programme offered by a South African university. A modified 'Views on the Nature of Science' (VNOS) questionnaire consisting of eleven open-ended questions was used to examine learners' views on seven aspects of the NOS. We investigated relationships between students' VNOS performance and parents' level of education, learners' gender, performance, home language as well as their school's performance in national final examinations. Findings showed no significant difference in the test scores between genders. The students' scores on the VNOS test were strongly correlated to the educational background of the parents, and learners with English as a home language performed better. However, there was poor correlation between the VNOS scores and learners' performance in Mathematics, Science and English. Also, the matric pass rate of the school did not correlate strongly with the test scores. From the results, it is concluded that the strongest factor influencing learners' performance on the VNOS test are parents' educational background and the use of English as a home language.

1. Introduction

One of the essential goals of science education reforms throughout the world is to prepare scientifically literate individuals [1, 2]. It is argued that an adequate understanding of NOS and scientific inquiry is essential for the development of scientific literacy [3]. Generally, scientific literacy includes an understanding of science content, scientific methods of inquiry and NOS [4]. Many learners in different parts of the world have poor conceptions of NOS [5]. This may be related to teacher-centred instruction methods which teachers have developed from their own experiences as students [6]. Teacher-centred instruction which does not promote the development of a better understanding of different features of NOS [7, 8] is still used in many South African schools. The science education reforms which supported scientific literacy in South Africa started after the introduction of Curriculum 2005 (C2005) policy in 1998. Many curriculum changes occurred after 1998 that eventually gave birth to National Curriculum Assessment Policy Statement (NCAPS) [9]. NCAPS encourages the development of science process skills, application of scientific knowledge in real-life problems and acknowledging the relationship between science, society and technology [9].

The current study was undertaken at a South African university that hosts a science enrichment programme for senior secondary school pupils. The name of the university is not mentioned for the sake of anonymity as required by ethical guidelines. The enrichment programme, presented by the Faculty of Natural and Agricultural Sciences, uses an inquiry approach in preparing learners in a three year programme from grade 10 to 12 in developing better understanding of science. The programme is focussed on physics and chemistry. Learners who successfully participate throughout the programme receive a full bursary if they enrol for a Bachelor of Science degree at this university, after completing grade 12. This paper reports on the first stage of a three year research project which tracks the development of NOS understanding amongst a cohort of participants in the enrichment programme. The focus of this paper is on exploring the views of NOS amongst learners upon entering the programme when in grade 10. The research question addressed in this paper is: *How do grade 10 learners' views about nature of science relate to contextual factors?* In this study the contextual factors include: gender, parents' level of education, home language, learner performance, and school performance.

2. Literature review

NOS is a dynamic scientific construct without a generally accepted definition. There are still disagreements amongst science educators, scientists, philosophers of science, and historians of science, regarding the precise meaning of NOS [1, 10]. This study was informed by Lederman's [11] perception of NOS, which he described as "epistemology of science, science as way of knowing, or the values and beliefs inherent to the development of scientific knowledge". Irrespective of the differences among aforementioned scholars with regard to the definition of NOS [19], there is a level of consensus on aspects of NOS that should be taught to high school learners [12, 13]. These aspects are discussed below.

The seven aspects of NOS include that scientific knowledge is tentative; empirical; theory-laden; subject to imagination and creativity; influenced by social and cultural values; based on observation and inferences; and described by scientific theories and laws and developed in the absence of a universal scientific method. Individual views of nature of science vary between well informed and naïve. The former views include understanding that scientific knowledge is tentative; empirically based; subjective; partly the product of human inference, imagination, and creativity; and socially and culturally embedded. Other aspects of NOS include the relationship between observations and inferences, the relationship between scientific theories and laws and use of variety of scientific methods [14]. The latter views include understanding that scientific knowledge is a discovery of an objective reality external to human beings; is generated through use of a single universal scientific method and that evidence collected through scientific processes will produce reliable knowledge [15] which is not affected by social and cultural values [16].

3. Methodology

The study is located in the interpretivist paradigm [17], using a mixed methods approach. Qualitative data were collected using the slightly modified open-ended Views of the Nature of Science (VNOS Form C) questionnaire containing eleven questions adapted from Lederman *et al* [12]. The modifications made to the test were mainly to adapt the test to the South African context. The VNOS questionnaire has been validated [13], and used regularly in research [18]. The changes made were considered minor with no significant effect on the validity of the test. The data were then analysed and quantified by scoring responses on a five point likert scale ranging from naïve to informed views (1 = naïve, 2 = partially naïve, 3 = moderate, 4 = partially informed, 5 = informed). These scores were then analysed using descriptive statistics. The scores were added to obtain an average score per individual and used in further analysis of the data. Descriptions of NOS aspects by [19] were used as a guide during the scoring process.

Purposeful sampling was used, utilizing the entire group of 100 applicants to the university's science enrichment program. From these, science enrichment programme annually selects a maximum

of 50 well performing grade 10 learners to enrol in the programme. For the current paper, the entire group of 100 grade 10 applicants acted as the sample. The ethical procedures were followed when selecting participants in this study. Permission was granted by the Ethics Committee of the University prior to the commencement of this study.

4. Results

The results from this study are represented in the three graphs shown in Figure 1, 2 and 3.

4.1 Average score per question

Figure 1 shows the performance of the students on different VNOS questions, using the 5 point scale. The majority of learners scored well on questions 2, 5 and 9 which were addressing VNOS aspects that science knowledge is empirically based, that scientific theories are tentative and that social and cultural values influence the development of scientific knowledge.

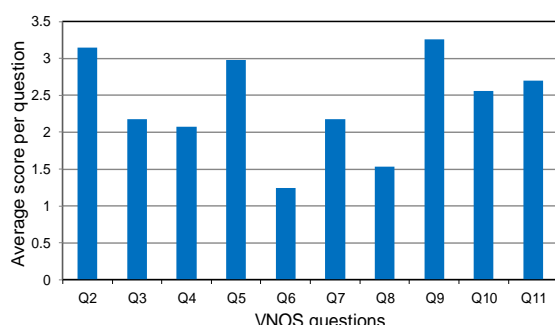


Figure 1. Learners' VNOS score averages on eleven open-ended questions.

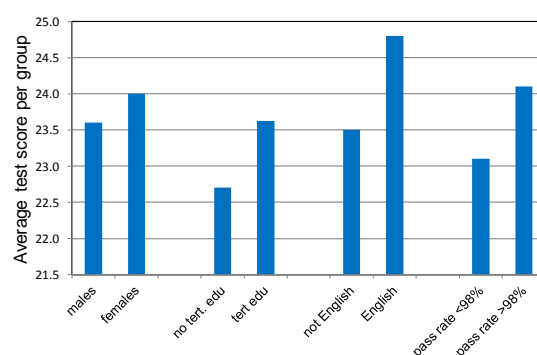


Figure 2. Learners' average test scores against contextual factors.

Learners obtained lower scores on questions 3, 4, 7, 10 and 11. These questions targeted the following VNOS aspects: learners' understanding of that scientific investigations are not always following a recipe-based scientific method, but involve observations and inferences to generate science knowledge (3, 4 and 11) and learners' understanding of scientists' application of imagination and creativity in the development of scientific knowledge (7 and 10). Finally, learners scored lowest on questions 6 and 8 focussing on learners' understanding of the difference between a scientific theory and a scientific law (6) and learners' understanding that science is subjective with opportunity for different interpretations (8).

4.2 Contextual factors

The average test scores for students from different backgrounds are shown in Figure 2. Students whose parents have a tertiary education obtained higher VNOS scores than those of parents without tertiary education ($p = 0.054$). In addition, students who had English as home language scored higher ($p = 0.074$) than other students using other languages. Also, the school background showed that VNOS scores were higher for the students from schools which had a matric pass rate over 98% ($p = 0.30$). There was no significant correlation ($r = -0.04$) found between VNOS test scores and Grade 10 marks for the group. However, there was a gender effect: Males showed a slight positive correlation between test scores and grade 10 marks ($r = 0.33$), whereas females showed a slight negative correlation ($r = -0.27$). After investigating correlation between genders in terms of understanding different aspects of the NOS, as determined by a t -test, there was no statistical difference between them ($p = 0.95$). Figure 3 shows the average score per gender that students obtained in the VNOS test. It can be seen that in different individual questions, there were only small differences between males and females. Statistically the most significant differences were observed when asked on their knowledge of the

difference between a scientific theory and a law, where males scored marginally better (question 6, $p = 0.16$), and the VNOS aspect demonstrating an understanding that scientists use their imagination and creativity in the generation of scientific knowledge, where females performed marginally better than males (question 7, $p = 0.13$).

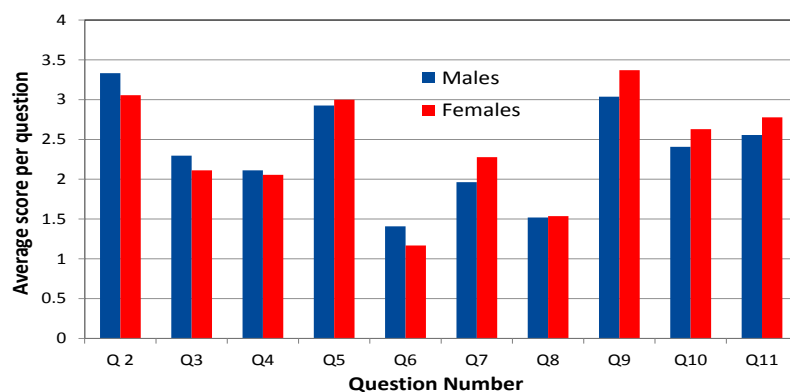


Figure 3. Gender differences in average scores per question

5. Discussion and conclusions

This study has demonstrated a noticeable variability in understanding different aspects of NOS. Students generally scored well on NOS aspects such as the tentative, empirical, and influence of social and cultural values on the generation of scientific knowledge. Previous research shows that many teachers and students perceive scientific knowledge as fixed and unchangeable. Additionally, they believe that scientific knowledge generated through experiments and technological means only is legitimate and reliable [20]. However, learners' performance in questions 2, 5 and 9, demonstrates that learners in this study are reasonably well informed in three VNOS aspects: empirical, tentative, and sociocultural embeddedness. This finding differs from the previous international studies.

Previous research shows that learners believe that scientific knowledge is tentative and modified in the light of new technological discoveries [12]. The history of science indicates the view that science knowledge could be changed through evolution and revolution of scientific ideas [21]. Students in this study struggled with differentiating between a theory and a law and understanding that different conclusions are possible given the same data due to the inherent subjectivity of scientists when interpreting data. Many teachers and students do not understand that observations and inferences are subjective and theory-laden as they are guided by scientists' prior knowledge and experiences of current science knowledge which could lead to multiple valid inferences [12]. Instead they believe that scientists would make objective observations and inferences of the same natural phenomenon [22]. Students also have misconceptions on the use of different scientific methods when generating scientific knowledge. This finding is similar to the results of studies conducted in the North of America and Australian contexts, where high school students experienced the same difficulties as South African learners [5].

Empirical research has shown that many teachers and students believe that laws are certain and proven whereas theories are unproven ideas [21]. Teachers and students also believe that there is a hierarchical relationship between theories and laws [23]. They also believe that theories will later change into laws depending on the accumulation of supporting evidence [12]. Students should know that science involves translation of data into evidence, invention of explanations and theoretical models which demand the use of creativity and imagination by scientists [12]. Furthermore, question 8 assessed learners' beliefs on different subjective factors such as theoretical commitments, and influence of social and cultural values that affect scientists' work. Learners should know that although science knowledge is considered to be general and universal, it is influenced by the societal and cultural values of the context in which it is generated. Many teachers and students consider science as a hunt for the universal truth not influenced by cultural and societal values [20]. Lederman *et al* [12]

posited that many learners have a misconception that there is a universal scientific method used by all scientists when they conduct scientific investigations. Kuhn [21] asserted that scientists use a mixture of different scientific approaches depending on their prior experiences, creativity and imagination and current paradigms in their fields of research.

Regarding contextual factors, it was found that the strongest correlation with learner's performance in the VNOS test was their parents' education, followed by home language and the matric pass rate of the school attended. These findings may be attributed to various factors such as the existence of many indigenous languages and use of English as a language of learning in the South African context. Additionally, majority of parents have low level of education due to poor schooling and segregated education system during the apartheid era. Similar findings were reported in a study conducted by Dogan and Abd-El-Khalick [24] in Beirut, Lebanon. They investigated the influence of social backgrounds on grade 10 Turkish students' and science teachers' conceptions of NOS. Results from their study showed that teachers' and students' views were similar and related to teacher graduate degree and geographical region, and student household socioeconomic status (SES), parent education, and SES of their city and geographical region. These relationships demonstrated the influence of western culture in their understanding about NOS. Similarly it can be argued that the effects observed in the current study can be attributed to the influence of western culture on development of scientific literacy in South Africa.

There was no statistical significant difference on average scores between genders. Lederman *et al* [12] posited that understanding of the different features of nature of science is not gender biased. Research studies in psychology have shown that there are differences with regard to the cognitive skills between boys and girls. One group of research has shown that boys outperform girls on tasks relying on spatial orientation and visualisation skills [25]. Many people believe that success in STEM (Science, Technology, Engineering and Mathematics) courses is linked to spatial capabilities, but research has shown that there is no relationship between STEM and spatial skills [26]. In contrast, girls outscore boys on tasks relying on verbal and writing skills, memory and perceptual speed [27]. This could be one reason why female learners outperform boys in question 7 which was addressing the role of human inference and creativity in science, and the notion that scientific models are not copies of reality.

The results of this study should not be generalized to the entire grade 10 population of South Africa. However, results could be generalized within the population of well performing students in the city from which the sample was drawn. The reproduction of results would be possible drawing a similar sample of learners from city schools and administering the questionnaire. The findings of this study especially the students' informed views on the tentative nature of science agree with the international research. These results may be attributed to the changes that occurred in the education system in South Africa after the introduction of C2005 in 1998. However, there are currently still many disadvantaged students with poorly educated parents in South Africa. This is a persisting consequence of the segregated education system during the apartheid era and which cannot be overcome in a short time. The language issue, that is the use of English as a medium of instruction in schools also disadvantages learners who speak indigenous languages at home. This language problem has no simple solution. Clearly, changes which occur at the political level have an influence at what is happening in teaching and learning process in a classroom situation [28]. A follow up study will again be conducted with the same group of students after the three year programme to examine the changes in their views of NOS due to their involvement in the science enrichment programme. In the next study, the relationships between parents' level of education and the use of English as a medium of instruction on students' views of the nature of science will also be investigated.

Acknowledgements

Funding was provided by the Sasol Inzalo Foundation and the National Research Foundation of South Africa.

References

- [1] Laugksch R C 2000 *Sci. Educ.* **84** 71
- [2] Millar R and Osborne J F (Eds.) 1998 *Beyond 2000: Science education for the future.* (London: King's College London)
- [3] De Boer G E 2000 *J. Res. Sci. Teach.* **37** 582
- [4] Anderson R D 2007 Inquiry as an organizing theme for science curricula *Handbook of Research on Science Education* ed KS Abell KS and NG Lederman (Mahwah, NJ: Lawrence Erlbaum Associates, Inc) pp.807-830
- [5] Abd-El-Khalick F 2006 Over and over again: College students' views of Nature of Science *Scientific Inquiry and Nature of Science: Implications for Teaching, Learning and Teacher Education* ed LB Flick and NG Lederman (Dordrecht, The Netherlands: Springer)
- [6] McDermott L C, Shaffer P S and Constantinou C P 2000 *Phys. Educ.* **35** 411
- [7] Dekkers P 2006 *Afr. J. Math. Sci. Tech. Educ.* **10** 81
- [8] Ogunniyi M B 2006 *Afr. J. Math. Sci. Tech. Educ.* **10** 93
- [9] Department of Education 2012 *Curriculum and Assessment Policy Statement: Physical Sciences (Grade 10 – 12)* (Pretoria: Government Printer)
- [10] National Research Council 2001 *Science education standards.* (Washington D.C.: National Academy Press)
- [11] Lederman N G 1992 *J. Res. Sci. Teach.* **29** 331
- [12] Lederman N G, Abd-El-Khalick F, Bell R L and Schwartz R S 2002 *J. Res. Sci. Teach.* **39** 497
- [13] Abd-El-Khalick F, Bell R L and Lederman N G 1998 *Sci. Educ.* **82** 417
- [14] Hanuscin D L, Akerson V L and Phillipson-Mower T 2006 *Sci. Educ.* **90** 912
- [15] Tsai C C 1998 *Curric. Teach.* **13** 31
- [16] Bencze L, Di Giuseppe M, Hodson D, Pedretti E, Serebrin L and Decoito I 2003 *Syst. Pract. Action. Res.* **16** 285
- [17] Creswell J W 2003 *Research design: Qualitative, quantitative, and mixed methods approach (2nd Ed.)* (Thousand Oaks, CA: Sage)
- [18] Abd-El-Khalick F and Lederman N G 2000 *J. Res. Sci. Teach.* **37** 1057
- [19] Schwartz R S, Lederman N G and Crawford B A 2004 *Sci. Educ.* **88** 610
- [20] McComas W 1998 The principal elements of the nature of science: Dispelling the myths. *The nature of science in science education: rationales and strategies* ed W.F. McComas (Dordrecht: Kluwer Academic Publishers) pp. 53–70
- [21] Kuhn T S 1970 *The structure of scientific revolutions (2nd Ed.)* (Chicago: University of Chicago Press)
- [22] Chen S 2006 *Sci. Educ.* **90** 803
- [23] Aikenhead G and Ryan A 1992 *Sci. Educ.* **76** 477
- [24] Dogan N and Abd-El-Khalick F 2008 *J. Res. Sci. Teach.* **45** 1083
- [25] Voyer D, Boyer S and Bryden M P 1995 *Psychol. Bull.* **117** 250
- [26] Ceci S, Williams W M and Barnett S M 2009 *Psychol. Bull.* **135** 218
- [27] Halpern D, Aronson J, Reimer N, Simpkins S, Star J and Wentzel K 2007 *Encouraging Girls in Math and Science* (NCER 2007-2003) (Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education)
Online: <http://ncer.ed.gov>.
- [28] Taylor S and Coetzee M 2013 *Estimating the impact of language of instruction in South African primary schools: A fixed effects approach.* Stellenbosch Working Paper Series No. WP21/2013
Online: <http://www.ekon.sun.ac.za/wpapers/2013/wp212013> accessed 2014/04/07

Learners' conceptual resources for learning kinematics graphs

Grace Djan¹, M Lemmer² and ON Morabe³

The Faculty of Natural Sciences, North-West University, Potchefstroom, South Africa

E-mail: gdjan@hotmail.com¹

Abstract. Learners' difficulty in the application of basic graphs concepts to solve kinematics graphs problems contributes to underperformance in physical sciences. Their ability to handle problems in kinematics graphs could be enhanced if they have effective conceptual resources on graphs. In South Africa, this underperformance could be due to a gap between the GET [General Education and Training] and FET [Further Education and Training] band's requirements on graphs. For this reason this study selected to investigate the conceptual resources acquired by grade 10 learners from grade 9 that can be used productively for the learning of kinematics graphs in grade 10. The use of mixed method approach was considered appropriate for this study. The mixed method depended on a quantitative method (questionnaire) to produce precise and measurable data, while a qualitative method (interviews) was to enhance the understanding of the data produced by the quantitative method. Patterns and trends in learners' knowledge and reasoning were probed. The results showed that many learners could answer mathematics questions but struggled with similar questions in kinematics. From the results it can be deduced that learners' conceptual resources can influence their understanding of kinematics graphs in physics. These resources are gained from previous learning in mathematics and natural sciences.

1. Introduction

Graphs can be found in many of the natural sciences and physical sciences textbooks used in South African schools. It is therefore apparent that learners have opportunities to develop graphical skills that include amongst others interpretation of data, drawing of graphs and tables, reading from graphs and interpretation of graphs. These skills are important for understanding sciences and analysis of data for research purposes and should be understood especially by learners who wish to continue with formal education.

Graphical skills are also essential for everyday life activities since they are commonly found in reports, periodicals and journals. In South Africa, grade 9 is the first exit level that serves as a gateway to the world. Learners can end their formal education at the end of grade 9 and they should be able to use the knowledge gained at this level to fend for themselves.

The research question is: What conceptual resources have Grade 10 physical sciences learners obtained in the GET band that can be productively used for the learning of kinematics graphs in the FET band?

¹ To whom any correspondence should be addressed.

1.1. Components of graph in the National Curriculum Statement for grade 10 to grade 12

Graphs are an inherent component of the natural sciences and physical sciences curricula in South Africa. The National Curriculum Statement (NCS) for physical sciences from grades 10 to 12 considers graphs as an important process skill that contributes to the construction and applications of science knowledge [1]. For example, grade 10 learners are expected to seek patterns and trends in data and represent them in different forms, including graphs.

1.2. Importance of graphs

Many teachers consider the use of graphs in the laboratory as important for developing understanding of topics in physics [2]. The ability to use graphs correctly is an important gate way to produce expertise in problem solving in science [3]. Reflecting on the importance of graphs [4] are quoted as follows, "Line graph construction and interpretation are very important" because they are "an integral part of experimentation, the heart of science" [5]. On the contrary, 38.3% of the learners that participated in the study agreed that graphs are important.

1.3. Difficulties learners experience in graphs

According to [3] learners' difficulty in application of basic concepts in graphs to solve problems leads to difficulty in understanding concepts such as the gradient of kinematics graphs. Their learning problems are enhanced if they do not have an effective knowledge base (or resources) on graphs. Most learners find graphs, especially line graphs, difficult to draw and interpret [6]. In a study of South African physical sciences learners' mathematics procedural and conceptual knowledge, [7] found that the section on interpretation and application of graphs had the lowest overall performance. This can be attributed to the fact that the comprehension of kinematics graphs is influenced by learners' prior knowledge and familiarity of graph concepts.

1.4 Learning

Learning is defined as an act of gaining knowledge [8]. From the constructivist perspective, learning involves activities that lead to learners constructing their own theories, by building on their prior knowledge [9], [10]. Learners relate well to learning when they are given the opportunity to construct their own theories and ideas through constructive scientific activities guided by their prior knowledge. That means activities developed from learners' existing knowledge can lead to effective learning.

2. Design and method

A mixed method was used where qualitative and quantitative research strategies were combined and applied in the investigation. A questionnaire was used to quantitatively determine the occurrence, and determine correlations between mathematics and science questions that deal with the same aspect (e.g. gradient understanding and use of scale to draw graphs), amongst 201 grade 10 learners from 7 local schools. The questionnaire used was compiled by the researchers. It contained relevant questions from existing research questionnaires [1],[11], South African Grade 9 and 10 textbooks and examination papers. The validity and reliability of the questionnaire was ensured by content review by researchers in the field, a pilot study and statistical calculation of Cronbach Alpha values. The descriptive statistics that were used to describe the data were the frequency distribution and mean of the data.

The questionnaire was followed by interviewing three grade 10 learners to determine the conceptual resources and understanding thereof, which was qualitatively analysed. The learner with the highest mark, lowest mark and one average were randomly selected. The varying abilities of the extremes give a wider range of coverage of learners' conceptual resources and understanding.

The mixed method (QUAN-qual) used is summarized in the diagram below:

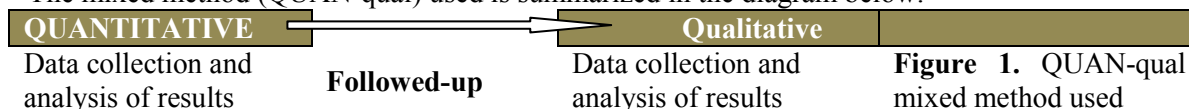


Figure 1. QUAN-qual mixed method used

3. Results of empirical study

3.1 Questionnaire results

The Cronbach's alpha coefficient of the questionnaire was above 0.7 after deletion of some items, which denotes consistency, reliability and validity of the questions used. These questions can therefore be understood in the same way by different populations.

Table 1 shows the learners' responses to the group of basic concepts items from the questionnaire.

Table 1. Learners' response to basic concepts questions

Grouping-Criteria	Description	Frequency (N=201)	Percentage %
Basic concepts	Distance	48	23.9
	Speed	22	10.9
	Unit for speed	46	22.9
	Unit for distance	90	44.8
	dependent variable	57	28.4
	independent variable	60	29.9
	Constant variable	23	11.4

The constant variable was correctly identified in the given problem situation by 23(11.4%) of the learners (table 1), while 57(28.4%) could identify the dependent variable and 60(29.9%) the independent variable. This result can probably be attributed to the lesser degree of application of these terms in grade 9 mathematics and natural sciences. The concepts of *distance* and *speed* were explained using everyday applications. For example, definitions such as "how far things are away from each other" and "how far you travelled" for distance. "How fast or slow an object moves" for speed was provided as response. Only 48(23.9 %) and 22(10.9%) of the learners were able to give scientifically acceptable descriptions of distance and speed respectively. The questions on units of measurement were answered much better than the definitions. The unit for speed was given correctly by 46(22.9%) of the learners while 90(44.8%) learners could provide the unit of distance.

The basic concepts needed to learn kinematics graphs seem to be lacking among the learners who participated in the study. A total average percentage of 23.9% responded positively by providing the expected answers to the questions.

Table 2 below shows the learners' responses to graph-related questions from the questionnaire.

Table 2. Learners' responses to graph-related questions from the questionnaire

Grouping-Criteria	Description	Frequency(N=201)	Percentage %
Graphs	Interpreting graphs	44	21.9
	Draw bar graph using correct label & scale.	53	26.4
	Draw line graph using correct label & scale.	44	21.9
	Use of scale	39	19.4
	Draw distance/time graph	32	15.9

A small number 44(21.9%) correctly interpreted the given graphs. The number 53(26.4%) of the respondents demonstrated acceptable skill of drawing bar graphs, however 44(21.9%) demonstrated all the skills of drawing line graphs, labelling axis, plotting points and using scale correctly. The use of scale was poorly answered, 39(19.4%) of learners from all the schools that participated in the study

were able to draw the axis to scale. A total of 32(15.9%) learners demonstrated the skill of drawing a distance/time graph from the given data. Drawing of line graphs appeared to be challenge for the learners.

Table 3. The mean score of the various groupings

Groupings-Criteria	Number of participants	Mean
Basic concepts	201	24.60
Graph interpretation–Mathematics	201	39.89
Graph interpretation-Physics	201	34.92
Gradient – Physics	201	40.73
Gradient – Mathematics	201	60.30

Table 3 refers to the mean score illustrating learners' performance. The learners performed better in gradient -mathematics than gradient physics with a mean score of 60.30 and 40.73 respectively than in graph interpretation-mathematics with a mean score of 39.89. The answers of graph interpretation-mathematics showed a large variability. A mean value of 39.89, for graph interpretation-mathematics, shows that there was a wide variety of responses to the interpretation of the graphs-related questions.

For given intervals learners were asked to state whether the gradient is positive, negative, or zero and over 45% were able to provide the accepted answer for each interval for the physics questions. However over 70% of the learners were able to interpret the given mathematics graphs correctly. It is perhaps due to the lack of exposure to such questions in grade 9 that contributed to fewer learners' 99(49.3%) ability to identify zero gradients.

3.2 Interview results

All the interviewees agreed that they were taught how to calculate the gradient from a given formula and to determine the gradient from graphs in grade 9 mathematics. That was why they were able to attempt the gradient related problems. This implies that if emphasis is placed on transfer and integration of graph skills from mathematics to physics by the grade 9 mathematics teachers, the gap between grade 9 and grade 10 could be narrowed. The interviewees confirmed that there is currently a huge gap between grade 9 natural sciences and grade 10 physical sciences.

4. Analysis and Discussions

4.1 Conceptual resources

The knowledge and skills obtained in graphs as learners proceed from grade 8 to grade 9 will be built upon in grade 10. The concepts and terms related to graphs that were acquired in grade 9 should serve as conceptual resources in learning kinematics graphs in grade 10.

4.1.1. Basic concepts: The findings revealed that the learners had difficulties with the basic kinematics concepts, for instance only 15.9% of the participants could draw the graph (distance/time) from given data. On the average, 33.2% could distinguish between speed and distance. The difficulty in differentiating between related concepts such as distance and speed remains a challenge. Molefe stated that one of the most common and a critical problem in physics education is the failure to differentiate between various kinematical quantities [12]. This was found to be true for this study. For example, only 26.4% and 22.9% of the respondents could write the correct unit for speed and distance respectively. They used everyday life descriptions for speed and distance instead of giving the scientific definitions.

4.1.2. Drawing and interpretation of graph: [2] found that most of the graphs in his study were not drawn to scale. Very few (19.4%) participants could draw the x and y axis to scale. This can be attributed to inadequate practice in grade 9. Interpretation of graphs can be a resource for learning kinematics graphs if the skills are properly mastered and correctly applied in grade 9. The educator should also ensure that the learners transfer their resources to new contexts such as kinematics graphs. If learners can interpret graphs and calculate the gradients then they have the needed resources to interpret and solve kinematics graphs problems.

The analysis of the quantitative and interview data indicate a lower percentage of learners' ability to interpret the kinematics graphs and this is in line with Beichner results where only 39% of the participants were able to provide the correct description for given kinematics graphs [1].

4.1.3. Gradient: The conceptual resource, gradient of graphs, was taught in grade 9 mathematics. The calculation of gradient, determination of gradient from a graph and drawing a graph are resources related to kinematics graphs that learners should have acquired in grade 9. These concepts are appropriately included in the grade 9 mathematics textbooks used by the participants of this study [14]. Calculation of gradient should have been taught intensively in grade 9 mathematics classes. This seems to be reflected in the results of the quantitative analysis, according to which more than 70% of the learners were able to distinguish between positive and negative gradients. The learners did not use the same concept in solving similar problems on gradients in physics. However, the interview results indicated that they do not have a good understanding of the concept of gradient. The reasoning given by the learners was not scientifically correct, e.g. "the line is longer therefore it's a positive gradient". With reference to the question asked they should have referred to the definition of gradient in terms of the ratio of changes in displacement and time or ($\Delta y/\Delta x$). Learners explained positive gradients as increase in gradient but did not apply their concept scientifically to solve problems related to gradient.

Analysis of the results revealed that learners struggled the most in the items concerning the following:

- Understanding and identifying variables.
- Identification of zero gradients.
- Interpretation of the point of intersection of line graphs.
- Differentiation between position and speed.
- Interpretation of kinematics graphs.
- Scaling.

5. Summary and discussion of findings

Based on the arguments from paragraph 1.4 on learning, one can deduce that having pre-knowledge about a topic, or possessing conceptual resources should contribute towards effective learning. Therefore the poor performance in kinematics graphs can be attributed to the fact that learners do not have all the necessary conceptual resources to learn kinematics graphs in the FET phase which include physics concepts and mathematics skills.

Conceptual resources

The conceptual resources that the grade 10 participants have for effective learning of kinematics graphs as obtained from the analysis of the results of the study are:

- Gradient: Calculating gradient.
- Deducing information from a graph: Identifying positive and negative slopes.
- Identifying trends: Decrease, increase, remains the same.
- Drawing graph: Plotting points. Table to graph conversion.
- Units: The units of basic kinematics physical quantities.

6. Conclusion

In conclusion the grade 10 learners who participated in the study do not have adequate conceptual resources (prior knowledge) to solve kinematics graph problems. Their resources are not coherent, i.e. learners do not apply knowledge consistently in different problems.

7. Recommendation

In grade 9 the science teachers should teach the basics of graphs required by the curriculum. Learners should be guided to draw graphs correctly and completely, e.g. with the appropriate scales, axes labels and graph headings.

They should also know how to calculate the gradient of a line graph and the area under a graph. They can then attach meaning to the gradients and areas of line graphs easily in grade 10 physical sciences. This can contribute to a smooth transition from grade 9 to grade 10.

The grade 10 science teachers must ensure the learners have the resources, the understanding, and the ability to transfer and connect grade 9 mathematics skills in grade 10.

On the other hand the grade 9 mathematics teachers must ensure proper understanding of gradient, variables, transfer and integration of mathematics skills in science.

References

- [1] Beichner RJ 1994 Testing students understanding of kinematics. *American J.Phys.***62** 750-761.
- [2] Robbins N B 2005 *Creating More Effective Graphs* (New Jersey: Wiley) p 377
- [3] Halloun I A and Hestenes D 1985 The initial college knowledge state in physics. *American Journal of Physics.* **33**
- [4] Hornby A S 2006 *Oxford Advanced Learner's Dictionary of Current English* [Oxford University] 7th ed. p 840
- [5] Driver R, Asoko H, Leach J, Mortimer E and Scott P 1994 Constructing scientific knowledge in the Classroom. *Educational Researcher* **23**
- [6] Bybee R W, Powell J C and Trowbridge L W 2008 *Teaching Secondary School Science: For Strategies Developing Scientific Literacy.* 9th ed. (Upper Saddle River: Pearson Prentice) p362
- [7] Bruner J S 1969 *The Process of Education* (London: Oxford University Press) p 97
- [8] Cothron J H, Giese R N and Rezba R J 2006 *Science Experiments and Projects for Students: Students' Version for Students and Research.* 4th ed. (Dubuque: Kendall/Hunt) p 266
- [9] Department of Education 2003 *National Curriculum Statement Grades10-12 (General) Physical Sciences.* (Pretoria: Government printer) p 76
- [10] Glazer N 2011 Challenges with graph interpretation: a review of the literature. *Studies in science education*, **47** 83-210
- [11] Hestenes D and Wells M 1992 A mechanics baseline test. *The physics teacher*, **30** 159-166
- [12] Molefe F K 2006 *Mathematical knowledge and skills needed in science education for grades 11 and 12.* Potchefstroom: NWU. (Dissertation - M.Ed.) p 98
- [13] Mckenzie D L and Padilla M J 1986 The construction and validation of the test of graphing in science: TOGS. *J of research in science teaching* **23** 571-57.
- [14] Djan G 2014 *Learner's conceptual resources for kinematic graphs.* Dissertation: *Magister Educationis* in Natural Science Education, Potchefstroom Campus of the North-West University p 209

Student understanding of DC circuits: fine-grained issues

Ignatius John¹ and Saalih Allie²

¹Department of Physics, Cape Peninsula University of Technology, South Africa-7530 & University of Cape Town. johni@cput.ac.za

²Department of Physics and ADP, University of Cape Town, South Africa – 8000. saalih.allie@uct.ac.za

Abstract. We report on a study with 60 first year physics students in which we made contextual changes to an “open circuit” in order to measure the effect of such changes to student’s responses. The 8 question instrument that we designed included representational, linguistic and (circuit) elemental variations. Our findings indicate that while the changes might appear trivial to an expert they significantly affect the way in which students respond.

1. Introduction

The general experience of teachers and lecturers is that students have difficulties in dealing with basic resistive DC circuits. Many studies have been carried out to probe the nature of these difficulties [1] and a number of interesting teaching approaches have been proposed in response to the research findings. However, it is clear that no single approach has been able to solve all the difficulties [2]. One reason for this may be that both the studies and the proposed solutions are not directed at a fine enough “cognitive grain size” to effect conceptual change [3]. For example, it is common to find (a) studies of student understanding and (b) teaching approaches in which the brightness of light bulb is used as a proxy for current. In doing so there is an underlying tacit assumption that is consistent with the classic “misconceptions” perspective, namely that students have unitary, static, prior conceptions in place. However, as described in [3] this is to be contrasted with the “Knowledge in Pieces” framework in which concepts arise more dynamically from smaller cognitive elements that come together in the moment. Key to this process is the triggering role of context. While the ambit of context is complex the present study confines itself to issues of contextual differences at the level of the circuit itself. In particular we probed the response of students to situations in which seemingly small representational, linguistic and circuit element changes are made, which from an expert point of view should have no bearing on the physics at hand.

2. Methodology

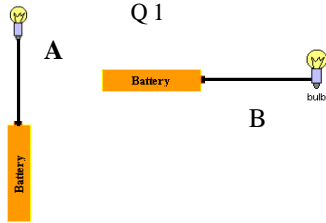
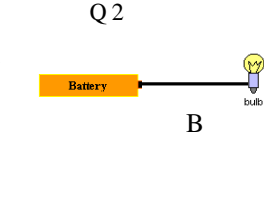
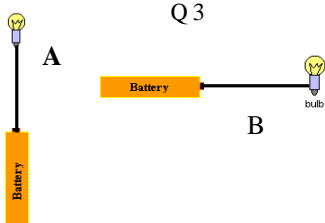
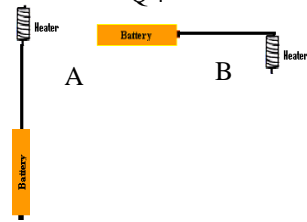
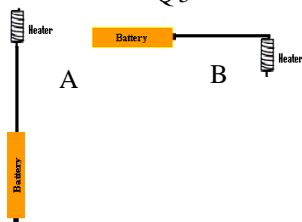
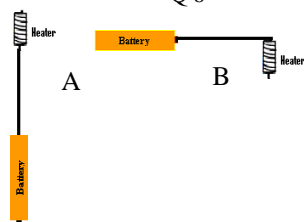
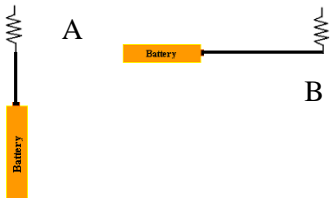
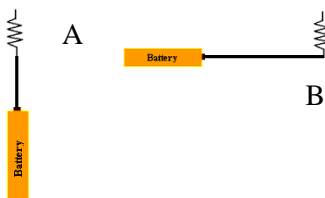
From a broad range of possibilities we honed in on the areas of question listed below and used these to construct a research instrument that required students’ written responses. (a) To what extent does changing a resistor for a light bulb or a heating element affect student response? (b) To what extent does drawing the same circuit vertically or horizontally affects student response? (c) To what extent does the phrase “charge flow” elicit the same response as the term “current”?

2.1. Research Instrument

A series of questions based on an open circuit was developed in which three circuit elements were interchanged with each other, namely, a resistor, a light bulb and a heater. These circuit elements were

connected with a single wire to one end of a battery in either a horizontal or vertical configuration. The terms “charge flow”, “current” and “heat up” (or “light up”) were interchanged in the text for identical circuits. Thus, we developed a question bank of 120 questions that spanned the “question space” in a systematic manner. For the purposes of the present study 8 questions were selected as shown in the table 1 below. (Sample question is given in figure 1). The table is arranged so as to facilitate comparison both across the rows and down the columns.

Table 1: The eight questions used in the probe. In the horizontal rows the words are being changed and in the vertical columns the elements are being changed.

<p>Q 1</p>  <p>“will the bulb light up”</p>	<p>Q 2</p>  <p>“will charge flow in the bulb”</p>	<p>Q 3</p>  <p>“will there be current in the Bulb”</p>
<p>Q 4</p>  <p>“will the heater heatup”</p>	<p>Q 5</p>  <p>“will charge flow in the heater”</p>	<p>Q 6</p>  <p>“will there be current in the heater”</p>
<p>Q 7</p>  <p>“will charge flow in the resistor”</p>	<p>Q 8</p>  <p>“will there be current in the resistor”</p>	

The pattern of contextual differences between questions across the rows and the columns can be summarized as follows. With regard to the rows, Row 1 shows 3 questions each consisting of identical scenarios. Each scenario shows two variations of the same circuit, one orientated vertically the other horizontally, comprising one end of a battery connected to a light bulb with a single wire. In the vertical cases the battery is connected to the bottom of the bulb while in the horizontal cases the battery is connected to the side of the bulb. The key variation in each question is the wording of the accompanying text as indicated. Row 2 shows what are essentially the same circuits as in Row 1 but with the bulbs replaced by heaters. In the vertical cases the battery is connected to the bottom of the heater while in the horizontal cases the battery is connected to the top of the heater. The key variation across the row again lies in the text as indicated. Row 3 shows 2 questions each consisting of identical scenarios. Each scenario shows the same circuit configuration but orientated either vertically or horizontally in which one end of a battery is connected to a resistor with a single wire. In both vertical and horizontal cases the battery is connected to the bottom of the resistor. The key variation in each

question is the text as indicated. The reason for no scenario being depicted in column 1 is that there is no text that is equivalent to “light up” or “heat up” in the case of the resistor.

If we consider the columns, Column 1 shows 2 scenarios in which the bulb is changed to a heater and the text “light up” is changed to “heat up”, respectively. In addition the point of connection (in the horizontal circuit) to the circuit element is varied from middle (bulb) to top (heater). Column 2 shows 3 scenarios in which the text “is there charge flow” is the same in each case but the circuit element (bulb, heater, resistor) changes. In addition the point of connection to the circuit element is varied (middle-bulb, top-heater, bottom-resistor). Column 3 shows 3 scenarios in which the text “is there current” is the same in each case but the circuit element changes (bulb, heater, resistor). In addition the point of connection to the circuit element is varied (bulb-middle, heater-top, resistor-bottom).

Each question was presented as a situation involving a discussion among a group of students who are posited to be setting up a circuit. A number of different points of view are articulated by the students and offered as options for which (a) a particular option has to be made and (b) the reason for the option has to be provided in detail. In total five options were presented of which the fifth allowed for ideas that were not offered. The response options were presented in a manner that the correct answer was not the same number (1-5) for each question. Figure 1 shows an example of a question in detail.

2.2 Sample

The sample consisted of 60 first year degree students in the extended programs in various courses in the science faculty and was doing Physics1. All these students passed physical science in Grade 12 but they had not received any formal instruction on DC circuits at the university. The majority of the students were from rural schools, aged about 18 years, and for whom English is a second or third language. These students are the first group of students who passed matric in 2008 having followed the new Outcomes Based Education curriculum.

One student connects a light bulb to a battery as shown in figure A. Another student connects the light bulb to a battery as shown in the figure B. The following discussion takes place among the students.

Student 1 says “The bulb in figure A will light up but not the bulb in figure B!”

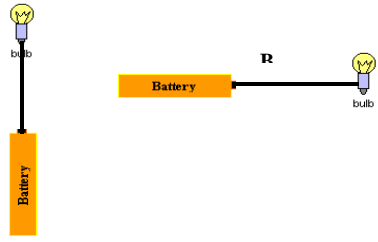
Student 2 says “No! The bulb in figure B will light up but not the bulb in figure A”

Student 3 says “I disagree! Both bulbs will light up!”

Student 4 says “No! None of the bulbs will light up!”

Student 5 says “I have another idea which I will explain to you!”

With whom do you most closely agree? Circle only one of 1, 2, 3, 4 or 5.



1
2
3
4
5

Explain the reasons for your choice in detail below.

.....

Figure 1: Question format.

2.3 Presentation of the instrument

The questionnaire was presented to the students during a physics practical afternoon session. Students were informed that “we are developing a new curriculum for the 1st years; hence we want to know your basic knowledge in electricity, so that we can develop a suitable curriculum for you. So please answer these questions as sincerely as possible so that we can help you to prepare for your exam”. Of the 60 students, Six completed the questionnaire in five minutes while others took more than twenty minutes to complete the questionnaire.

3. Results

Analysis of the data comprised three phases. In the first phase only the numerical choices were analysed while the second phase considered the free response writing and the third is interview. Results from the data in which only the numerical choice was analysed are summarised in table 2. The four different options given in the questions are symbolically represented in the table. The square \square represents option 1: (element in circuit A will activate), the rhombus \diamond represents option 2: (element in circuit B will activate), the triangle \triangle represents option 3: (elements in circuits A & B will activate) and the closed circle \bullet represents option 4: the correct answer in that “none of the circuits will activate”, since all the circuits are open. Each of the 60 rows in table 2 represents the 8 answers of a particular student.

From the table 2 it can be seen that only 15% of the sample (the first 9 rows of filled circles) chose the correct answers for all 8 questions. In the remaining cases answers can be seen to be changing from one question to another. We then analysed the different types of reasoning used by each student in answering the 8 questions. A number of different lines of enquiry ensued from these results of which we present the following one. Respondents were grouped on the basis of the number of reasons used to explain their answers to the 8 questions. The results were as follows: 12 students used one reason to answer all the 8 questions, 16 students used two reasons, 13 used 3 reasons, 10 used 4 reasons, 6 used 5 reasons and 3 used 6 reasons. The situation is summarised in the bar graph shown in figure 2 where each bar represents the number of students in each group with the number of different reasons cited per individual student.

4. Discussion

In general students who used more than one explanation did not get all the answers correct. Of the 12 students who used only one reason, 9 of them answered all the 8 questions correctly. Eight students of these 9 students who based their reasoning on only one concept appeared to consider the circuit as a whole as the unit of reasoning, using the following words (verbatim)

1. *Incomplete circuit.*
2. *Both positive and negative should be connected.*
3. *Only one end of the battery is connected.*
4. *No circuit.*
5. *No closed circuit.*
6. *Open circuit.*

Thus, only students who explicitly or implicitly used “circuit” reasoning got all 8 questions correct i.e. circuit reasoning provided a productive foothold for dealing with all the situations (light bulb, heating element, resistor, charge flow, current, light up and heat up). In all the other cases where specific features of the context were used as the primary foothold the resulting reasoning ended up with an incorrect conclusion.

Table 2: The eight columns represents the eight answers and each row represents the answers of each student.

St. No.	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
1	●	●	●	●	●	●	●	●
2	●	●	●	●	●	●	●	●
3	●	●	●	●	●	●	●	●
4	●	●	●	●	●	●	●	●
5	●	●	●	●	●	●	●	●
6	●	●	●	●	●	●	●	●
7	●	●	●	●	●	●	●	●
8	●	●	●	●	●	●	●	●
9	●	●	●	●	●	●	●	●
10	□	●	●	●	●	●	●	●
11	□	●	●	●	●	●	●	●
12	△	●	●	●	●	●	●	△
13	●	●	●	●	◇	●	●	●
14	●	●	●	●	◇	●	●	●
15	●	●	●	●	●	◇	●	●
16	●	●	●	●	△	●	●	□
17	●	●	●	●	●	◇	●	●
18	●	●	●	●	●	△	●	△
19	△	●	●	●	●	◇	●	△
20	●	□	△	●	△	●	●	●
21	●	●	△	△	●	●	△	●
22	●	●	●	●	●	●	●	△
23	●	●	□	□	□	●	●	●
24	□	●	◇	●	□	●	●	●
25	△	△	△	●	●	●	●	●
26	△	●	●	□	△	●	□	△
27	●	●	△	◇	△	△	△	●
28	●	□	●	●	△	●	◇	△
29	△	◇	●	△	●	●	◇	●
30	□	◇	△	□	△	●	●	●
31	□	●	△	□	△	●	●	□
32	●	●	●	●	△	△	●	△
33	□	◇	●	□	●	△	●	□
34	△	△	△	●	△	△	●	△
35	●	□	△	△	△	●	◇	●
36	●	□	□	△	△	△	●	△
37	△	△	△	●	△	□	●	△
38	●	●	□	□	△	◇	◇	□
39	△	△	●	□	◇	△	◇	◇
40	△	△	△	△	△	△	△	●
41	□	△	△	□	△	△	●	△
42	□	△	◇	□	△	□	●	●
43	△	◇	△	△	△	△	●	□
44	□	●	△	□	△	◇	●	□
45	△	△	△	△	△	△	△	△
46	□	△	△	□	△	△	△	□
47	●	●	△	□	△	△	◇	△
48	□	△	△	□	●	△	△	□
49	△	◇	△	△	△	◇	◇	△
50	□	△	△	□	△	△	△	□
51	△	◇	△	□	△	◇	◇	□
52	△	△	△	●	△	△	△	●
53	△	△	△	●	△	△	△	●
54	△	△	△	△	△	△	△	△
55	□	△	●	□	△	△	□	□
56	□	◇	△	△	△	◇	◇	△
57	□	◇	◇	□	□	◇	◇	□
58	□	△	△	□	△	△	△	□
59	△	△	△	△	●	△	△	△
60	□	◇	△	□	△	◇	△	□

In keeping with the “Knowledge in Pieces” framework, individual student reasoning about circuits depended strongly on contextual cues provided by the type of circuit elements, the words used and point of connection to the light bulb i.e. bottom or side. A number of curious conceptions also surfaced amongst the reasons provided. For example, regarding resistors, ideas included that a resistor is a device that is used to connect different elements in the circuit, to notions that the resistor stores electricity or is used to stop current. It is possible that students are engaged in local sense-making (and there is anecdotal evidence to suggest this) but direct interviews will be necessary to probe these aspects further. A key point of interest that also requires further elaboration is to try and understand which contextual features trigger starting points for reasoning. This will almost certainly influence the design of future teaching approaches.

However, it is already suggestive from the present results that any curriculum for teaching circuits will need to address prior experience in a manner that turns it into a productive cognitive resource. What seems to be clear at this stage is that the use of light bulbs in both the teaching of DC circuits and probing student understanding requires closer scrutiny.

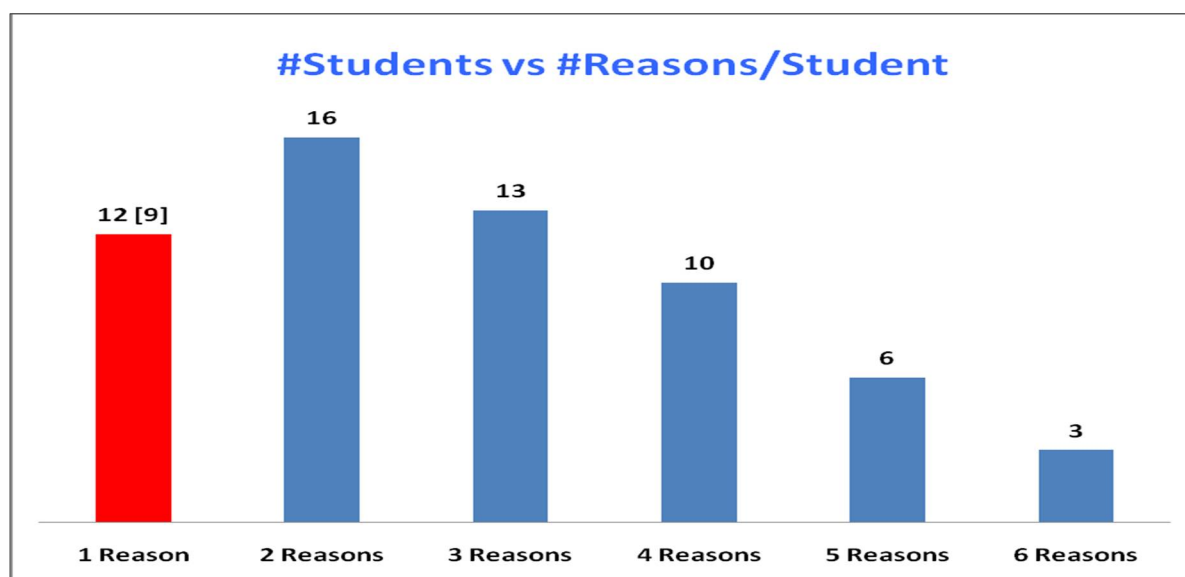


Figure 2: The number of students in six groups with different number of reasons.

References

- [1] McDermott L.C and Shaffer P. S 1992 Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding Am. J. Phys. **60** (11), 994-1000
- [2] MacKenzie R Stetzer, Paul Kampen, Peter S. Shaffer, and LC McDermott 2013 New insights into students understanding of complete circuits and the conservation of current. Am J Phys. **81** (2), 134-143
- [3] diSessa, A A 2006. A History of Conceptual Change Research: Threads and Fault Lines. Sawyer, R. Keith (Ed), The Cambridge handbook of: The learning sciences. , (pp. 265-281). New York, NY, US: Cambridge University Press, xix, 627 pp.

A CHAT perspective on the tensions and dynamics in the professional development of Physical Sciences teachers in a mentoring relationship

¹Sam Ramaila and ²Umesh Ramnarain

¹Department of Applied Physics and Engineering Mathematics, University of Johannesburg, P.O. Box 17011, Doornfontein, 2025, Johannesburg, South Africa

²Department of Science and Technology Education, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, Johannesburg, South Africa

E-mail: samr@uj.ac.za

Abstract. This research explored mentoring in terms of the dynamics and tensions associated with the interaction between a ‘keystone species’ and a novice teacher within communities of practice using a case study method underpinned by the Cultural Historical Activity Theory (CHAT). A primary constraint in the implementation of curriculum reform has been the lack of professional development for teachers. Insights into the nature of the mentoring relationship between Physical Sciences teachers revealed distinctive tensions and contradictions in terms of the activity system. The benefits of mentoring relationship suggest that teacher professional development could receive a major boost if the Department of Basic Education more overtly encouraged mentorship relationships between teachers.

1. Introduction

Large-scale research in South Africa reveals that despite significant reform in science education in this country there is little to suggest that the quality of science education has improved. For instance, the Trends in International Mathematics and Science Studies (TIMSS) repeated over the years revealed that the performance of South African learners in mathematics and science was very poor compared to other developing countries [1,2,3,4,5]. A primary constraint in the implementation of the National Curriculum Statement (NCS) has arguably been the lack of professional development for teachers [6] in response to curriculum reforms. Teachers in South Africa feel overwhelmed by the challenges presented by the reforms in the Physical Sciences curriculum. Research has pointed to teachers lacking confidence in teaching topics in the new curriculum [7,8,9] and teachers lacking in competence in implementing meaningful learner-centred scientific investigations [10,11]. Furthermore, teachers’ lack of pedagogical content knowledge (PCK) may limit their ability to facilitate meaningful learning [12,13].

It is against this background that our study sought to explore how mentoring can be exploited as a means by which a teacher facing challenges in teaching Physical Sciences in secondary schools in South Africa could be support from a skilled and experienced colleague (“keystone species”). Mentoring has increased in popularity as a way by which a teacher experiencing some weakness in his/her practice receives support from a skilled and experienced colleague [14]. A review of literature on mentoring in the field of teacher education reveals that it involves complex personal interactions conducted under different circumstances in different schools in which it cannot be rigidly defined [15]. In this research we invoked the theoretical lens of Cultural-Historical Activity Theory (CHAT) in trying to unpack the tensions and dynamics in the experiences of two physical sciences teachers involved in a mentor-mentee relationship.

2. CHAT as a theoretical framework

Activity Theory is a socio-cultural and historical lens through which human activity systems can be

holistically analyzed [16,17]. Activity Theory is thus a framework for understanding the totality of human activity in context [18]. This theory invokes premises of Vygotskian socio-cultural theory [19,20,21], namely, that learning takes place in a social and cognitive space known in Vygotskian parlance as the “zone of proximal development.” Such a perspective sees learning as a mediated process in which not only subject content is learned, but in which the relationships of mediation are crucial [22,23,24]. Suffice to indicate that CHAT has extensively been utilized as a key framework in computer science for analysing human computer interfaces [25,26]. We use this theory to explore the social dynamics and tensions of teachers engaged in a mentoring relationship within a community of practice. Accordingly, the following research question was formulated:

How can CHAT reveal the tensions and dynamics of the mentoring experiences of physical sciences teachers within a school educational setting?

The basic elements of an activity include subject, object, tools, community, rules, and division of labor [27]. All of the elements influence the others and are influenced by social, cultural, and historical factors, such as background knowledge, personal bias, availability of tools, and other factors. **Tools** in this research included the national curriculum policy document, namely the National Curriculum Statement (NCS). The policies of a district (cluster) and the relevant school on professional development are also tools since the activity system in this research study focuses on professional development through mentoring. The **subject** is the Physical Sciences teacher and the concomitant professional development. The **object** refers to the effective development of the teacher by means of mentoring support within the community of practice.

Rules refer to the guidelines of the NCS, but also the rules of the school in which the teacher operates. In particular the rules can be the “rules of engagement” of teachers with their line managers (subject heads or HODs). The **community** refers to the school (or ecology of practice) in which the teacher works, and the symbiosis between teachers, learners, parents and institutions. **Division of labour** in the context of this study refers to the roles of the teachers who participate in this study. Figure 1 below shows an activity system adapted for this research study.

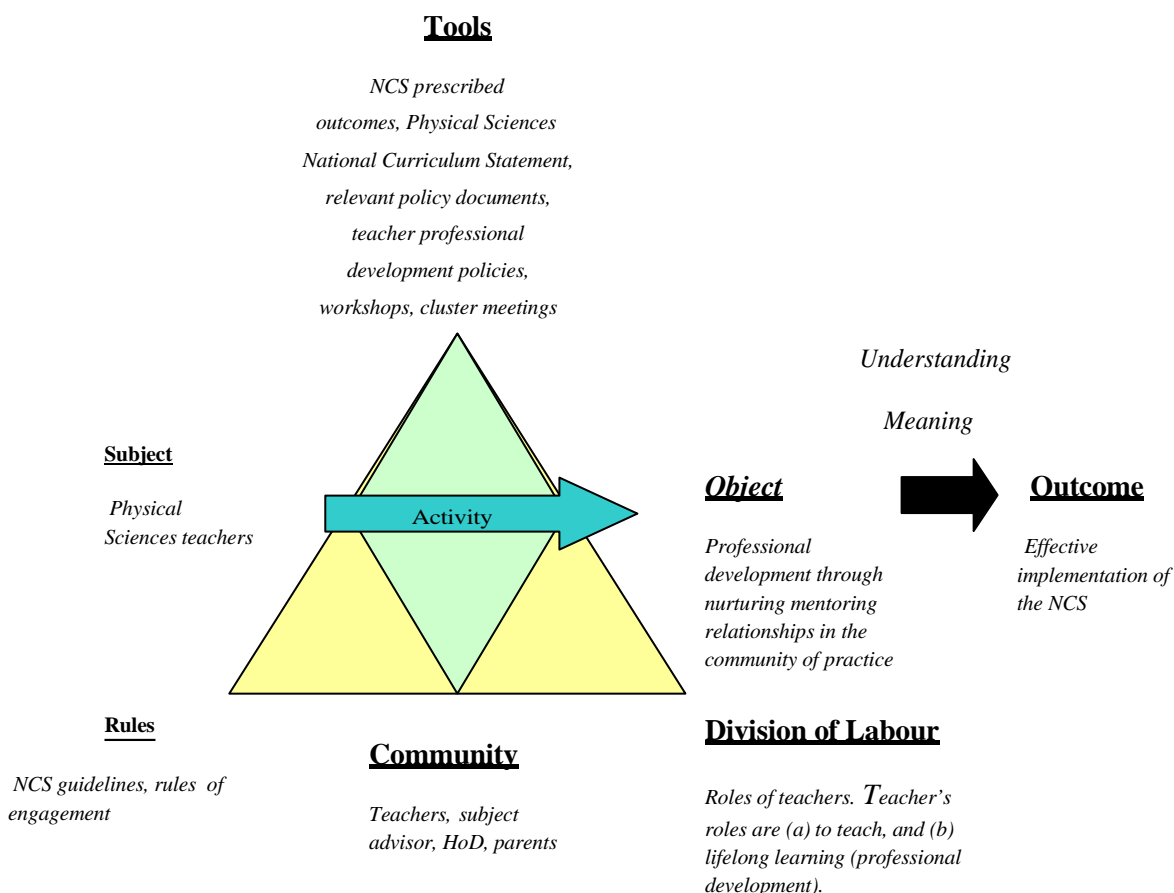


Figure 1: An activity system adapted for the research study

3. Research design and methodology

The research adopted a case study design that was used to gain an in-depth understanding of the situation. The research was conducted at a city school. The school was purposefully selected as it depicted a case where Physical Sciences teachers were involved in a mentoring relationship. The location of the school was convenient as it was accessible to us in terms of travelling distance. The research extended over a period of three months. Data collected from the interviews and classroom observations were analysed qualitatively. The interviews with teachers were transcribed verbatim. Written notes were also taken for lessons observed. The other data for qualitative analysis were the comments made by teachers on the questionnaires. Data were coded using the Atlas.ti computer programme. With regard to this study, we were particularly guided in this process by the Cultural Historical Activity Theory grounded in the Vygotskian tradition as the underlying theoretical framework. The related codes were then grouped into categories. The ensuing categories were then named inductively using the data as a guide in deciding what a category should be called. The patterns within the categories then led to the formulation of themes. These themes were then related to the elements within the activity system.

We sought to establish reliability in this process of coding and grouping codes into families by conferring with two researchers in science education. Due to the large volume of interview transcripts it was not feasible to ask them to go through the entire process as I had done so. I therefore randomly chose a transcription of one interview, which was analyzed using the same software. There was an 82% agreement amongst the three of us in this process of data analysis.

3.1 The school

Whale High School is a city school that is located in central Johannesburg, South Africa, and is described as a former model C school. In the Apartheid education system a model C school was designated for white children. The school is now racially integrated, with many Black children travelling from a neighbouring Black township. The school now has approximately 80% Black learners with the rest of the school population comprised of Coloured, White and Indian children. The school is adequately resourced for science with two laboratories that are being used. The school has 995 students. The pass rate for the Grade 12 national exit examination in the previous year was 83%. The school fee was R5000, with a 65% collection rate. The teachers were all employed by the state. The average class size is 35.

3.2 The teachers

Mr Ndlovu has 15 years' experience teaching Physical Sciences. He is very well qualified and has a Bachelor of Science degree with mathematics and physics as his majors. He also has a Higher Diploma in Education. He is the subject head in Physical Sciences at the school. He teaches grade 11 and grade 12 Physical Sciences. The principal of the school described him as "a master educator who is an inspiration to all teachers". Mr Ndlovu appears to hold a social constructivist view of learning as he believes that "learners develop their knowledge by working socially with the teacher guiding them". The learners in his class have consistently produced excellent results. In the previous matric examination, five learners had a distinction in Physical Sciences, and all learners had achieved a mark in excess of 50%. His expertise in the subject has also been recognised by his subject advisor who has asked him to act as an examiner for a district examination in Physical Sciences. He stated that despite the enormous challenges posed by the new curriculum he had adapted to it with ease. He explained that he had always subscribed to a teaching approach where learners were "centre stage" and he would design his activities around them. Mr Ndlovu therefore has all the credentials to be considered a "keystone species" in his profession. Mr Ndlovu acted as a mentor to a novice teacher, Mr Ngidi. Mr Ngidi had graduated the previous year from a university with a Bachelor of Education degree. His specialist teaching subjects are Physical Sciences and Mathematics. He teaches grade 9 Natural Sciences and grade 10 Physical Sciences.

4. Discussion of findings

Insights into the nature of the mentoring relationship between Physical Sciences teachers in this inquiry revealed distinctive tensions and contradictions in terms of the activity system. In the activity system the object is the professional development of the novice teacher and the subject is the vastly experienced teacher providing mentorship. The achievement of the outcome, which is the effective implementation of the National Curriculum Statement was complicated by several factors which can be described in terms of the elements of the activity system. Workshops and cluster meetings organised by the Department of Basic Education were largely ineffective in addressing the professional development needs of the novice teacher. However, the mentor provided much needed scaffolding which benefited the novice teacher to a considerable extent. The critical tension that emerged in this regard is that crucial elements such as workshops and cluster meetings were largely not responsive to the effective implementation of the National Curriculum Statement. In essence, workshops and cluster meetings as vital tools served a minimal purpose in providing meaningful professional development opportunities for teachers in order to enhance their practice. The concomitant consequence was that teachers' creativity was essentially stifled thus limiting their capacity to become reflective practitioners in their profession.

The absence of the subject advisor in the professional life of the novice teacher appeared to be detrimental to his professional development aspirations. The subject advisor appeared to be the weak or missing link in this professional development process and this renders the existing community, a pseudo-community of practice. This is a key area of concern that must be addressed by the Department of Basic Education in order to facilitate meaningful professional development of novice Physical Sciences teachers in particular. This crucial step has the potential to eventually engender dynamic, innovative and well-functioning communities of practice. The onus is on the Department of Basic Education to provide meaningful professional development opportunities through subject facilitators. The study also recognised the potential of CHAT in affording a holistic and analytical description of the mentor-mentee relationship in the professional development of teachers. This unharnessed potential ought to be brought to full fruition through critical understanding and interpretation of complex dynamic activity systems affecting meaningful professional development of teachers guided by firmly established and coherent analytical frameworks.

5. Conclusion

The significant reforms associated with the implementation of the South African Physical Sciences curriculum frustrated teachers' efforts in implementing the curriculum effectively. The two critical areas of reform have been in the content of the curriculum and an investigative approach to practical work. Teachers largely endorsed these reforms but their efforts at implementing such reforms have been compromised due to their lack of competence and experience in this regard. Clearly, the expectations of teachers have grown more complex and demanding. The In Service Educational Training (INSET) from the department of education by way of the one-shot workshops has been ineffective in developing teachers. The mentoring relationship explored in this inquiry showed that mentoring does hold promise as a form of professional support for developing teachers by keystone species in the profession.

References

- [1] *Trends in International Mathematics and Science Study (TIMSS)*. (2011). Assessing the mathematics and science achievement of students in the third, fourth, seventh, and eighth grades, and in the final year of secondary school.
- [2] *Trends in International Mathematics and Science Study (TIMSS)*. (2007). Assessing the mathematics and science achievement of students in the third, fourth, seventh, and eighth grades, and in the final year of secondary school.
- [3] *Trends in International Mathematics and Science Study (TIMSS)*. (2003). Assessing the mathematics and science achievement of students in the third, fourth, seventh, and eighth grades, and in the final year of secondary school.
- [4] *Trends in International Mathematics and Science Study (TIMSS)*. Assessing the mathematics and science achievement of students in the third, fourth, seventh, and eighth grades, and in the final year of secondary school.
- [5] *Trends in International Mathematics and Science Study (TIMSS)*. (1995). Assessing the mathematics and science achievement of students in the third, fourth, seventh, and eighth grades, and in the final year of secondary school.
- [6] Organisation for Economic Co-operation and Development (OECD). (2009). *Creating Effective Teaching and Learning Environments: First Results from TALIS*. Paris: OECD.
- [7] Kriek, J. (2005). Construction and evaluation of holistic development model for the professional development of physics teachers via distance education. Unpublished dissertation, University of South Africa.
- [8] Muwanga-Zake, J.W.F. (2004). Is science education in a crisis? Some of the problems in South Africa. *In: Science in Africa*, (Science Education), Issue 2, November 2004, pp. 17.
- [9] Pandey, D. & Braun, M. (2003). Investigating conceptual development of teachers doing further diploma in education course in high school current electricity. In *Proceedings of 11th Annual Southern African Association for Research in Mathematics, Science and Technology Education (SAARMSTE) Conference*, ed. B. Putsoa, M. Dlamini, B. Dlamini, and V. Kelly, 59-66. Mbabane: University of Swaziland.
- [10] Onwu, G. & Stoffels, N. (2005). Instructional functions in large, under-resourced science classes: Perspectives of South African teachers. *Perspectives in Education*, **23**(3), 79-91.
- [11] Ramnarain, U.D. (2007). A study of the implementation of scientific investigations at Grade 9 with particular reference to the relationship between learner autonomy and teacher support. Unpublished doctoral dissertation, University of KwaZulu-Natal.
- [12] Adams, P.E., & Krockover, G.H. (1997). Concerns and perceptions of beginning secondary science and mathematics teachers. *Science Education*, **81**, 29–50.
- [13] Lockheed, M.E. & Verspoor, A.M. (1991). *Improving Primary Education in Developing Countries*, ed. Associates, Oxford: Oxford University Press for the World Bank.
- [14] Bradbury, L.U. (2010). Educative mentoring: Promoting reform-based science teaching through mentoring relationships. *Science Education*, **94**(6), 1049-1071.
- [15] Wildman, T.M., Magglio, S.G., Niles, R.A., & Niles, J.A. (1992). Teacher mentoring: An analysis of roles, activities and conditions. *Journal of Teacher Education*, **43**(3), 205-213.

- [16] Engeström, Y. (1999). Innovative learning in work teams: Analysing cycles of knowledge creation in practice. In: Y. Engeström et al (Eds.). *Perspectives on Activity Theory* (pp. 377-406). Cambridge: Cambridge University Press.
- [17] Jonassen, D. & Rohrer-Murphy, L. (1999). Activity theory as a framework for designing constructivist learning environments. *Educational Technology, Research & Development*, **47** (1), 61-79.
- [18] Bodker, S. (1991). *Through the Interface – A Human Activity Approach to User Interface Design*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- [19] Vygotsky, L. (1978). Interaction between learning and development. *Mind and Society* (pp. 79-91). Cambridge, MA: Harvard University Press.
- [20] Vygotsky, L. (1986). *Thought and Language*. Cambridge, MA: The MIT Press.
- [21] Vygotsky, L.S. (1983). *Sobraniye Sochinenii* [Collected Works], Vol. 5, Moscow, Pedagogika Publishers.
- [22] Veresov, N. (2004). “Zone of proximal development (ZPD): The hidden dimension?”. In Ostern, A., Heila-Ylikallio, R. & Heila-Ylikallio, R. (Eds.), *Sprak som kultur – brytningar I tid Sprak son kultur – brytningar I tid och rum. Och rum (Language and Culture – Tensions in Time and Space)*. Vol. 1(1), pp. 13-30.
- [23] Veresov, N. (2007). Leading activity in developmental psychology. *Journal of Russian and East European Psychology*, **44**(5), 7-26.
- [24] Veresov, N. (2008). Generation “as if”: Some considerations from the point of view of cultural-historical psychology. *Bulletin of Academy of Pedagogical and Social Sciences*, **7**, 341-350.
- [25] Nardi, B.A. & O’ Day, V.L. (1999). *Information Ecologies: Using Technology with Heart*. Cambridge, MA: The MIT Press.
- [26] Nardi, B. (Ed.). (1996). *Context and Consciousness: Activity Theory and Human-computer Interaction*. Cambridge, MA: MIT Press.
- [27] Engeström, Y. (1987). *Learning by Expanding: An Activity-Theoretical Approach to Developmental Research*. Helsinki: Orienta Konsultit Oy.

First year university physics students' perceptions of the teaching-learning environment: In search of a coherent pedagogic learning orientation

¹Sam Ramaila and ²Umesh Ramnarain

¹Department of Applied Physics and Engineering Mathematics, University of Johannesburg, P.O. Box 17011, Doornfontein, 2025, Johannesburg, South Africa

²Department of Science and Technology Education, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, Johannesburg, South Africa

E-mail: samr@uj.ac.za

Abstract. Classroom inquiry and reflective activities are central to the improvement of the effectiveness of teaching practices. To this end, the investigation of aspects of teaching and learning in various educational settings is of paramount significance in order to provide insightful elucidation into the nature of teaching-learning environments. As part of this inquiry, the Experiences of Teaching and Learning Questionnaire was administered to first year university physics students at the University of Johannesburg in a bid to unearth their perceptions of the teaching-learning environment using the Teaching and Learning Research Programme's framework for analysis as the underlying conceptual framework. The Experiences of Teaching and Learning Questionnaire has been produced under the auspices of the Enhancing Teaching-Learning Environments in Undergraduate Courses Project in the United Kingdom. Critical interrogation of responses provided valuable insights into students' perceived nature of the teaching-learning environment forming an integral part of their academic training. In particular, students' pedagogic learning orientation appeared to hinge on the approaches they use as well as the extent to which they are well-organised in their study methods, use their time efficiently, and put concentrated effort into their work.

1. Introduction

A critical reflection on ways to enhance the teaching-learning environments is a crucial imperative for higher education institutions in their quest for human capital development. This imperative has direct bearing on students' existing knowledge and characteristics, students' approaches to learning and studying, students' perceptions of the teaching-learning environments as well as the quality of learning achieved. Students' approaches to learning and studying are premised on five distinct aspects in the form of deep approach, surface approach, monitoring studying, organised studying and effort management [1].

In terms of theoretical clarity, a deep approach indicates the intention to understand for oneself, backed up by processes of learning that involve relating ideas and use of evidence while surface approach indicates the intention to cope minimally with course requirements relying mainly on routine memorisation and procedural working [1]. Monitoring studying refers to the ability to stand back from one's learning and studying and reflect on how it affects academic progress [1]. Organised studying indicates how well students systematise their work on required tasks and manage their time effectively with effort management referring to the extent to which effort is well directed and concentration is maintained [1]. Key aspects which serve to characterise students' perceptions of the teaching-learning environments include aims, organisation and alignment; encouraging learning; assessment, assignments and feedback; supportive climate as well as evoking and showing evidence [1].

As part of this inquiry, the Experiences of Teaching and Learning Questionnaire (ETLQ) was administered to first year university physics students in a bid to unearth their perceptions of the teaching-learning environment using the Teaching and Learning Research Programme's framework for analysis as the underlying conceptual framework. The Experiences of Teaching and Learning Questionnaire has been produced under the auspices of the Enhancing Teaching-Learning Environments in Undergraduate Courses Project in the United Kingdom [2].

2. Teaching-learning environments

The manner in which teaching-learning environments are structured should ideally provide meaningful opportunities for the maximisation of students' academic experience. The influence of factors such as the nature of teaching, assessment and powerful learning environments [3] has widely been explored in various studies. It has been established that constructivist theories of learning provide a rich array of experiences which could arguably facilitate the development of students' conceptual understanding [4]. Emphasis has also been placed on the provision of authentic learning tasks which mirror experiences in the real world [5].

At another pragmatic level, there has been discourse about the nature of teaching-learning environments which encourage and support high-quality learning [6]. In this regard, it has been established that differing approaches to teaching affect approaches to studying [7] and good teaching encourages engagement and reflection through which conceptual understanding develops [8]. Research efforts in this regard have been augmented by the development of various conceptual frameworks [9]. In essence, these conceptual frameworks primarily served to indicate how students' pre-existing abilities and skills, in interaction with the teaching and departmental context they experience, influence their ways of learning and studying as well as the quality of their learning outcomes. In addition, the ways in which particular groups of academics organise their professional lives are intimately related to the intellectual tasks on which they are engaged [10] and teaching practices are mediated by the educational ideologies and the conceptions of teaching held by individual academics [11].

3. Conceptual framework

This inquiry is underpinned by the Teaching and Learning Research Programme's framework for analysis [12] shown in Figure 1 below as the underlying conceptual framework. This framework serves as a conceptual lens through which the interconnectedness of the factors characterising the nature of teaching-learning environments is explicated.

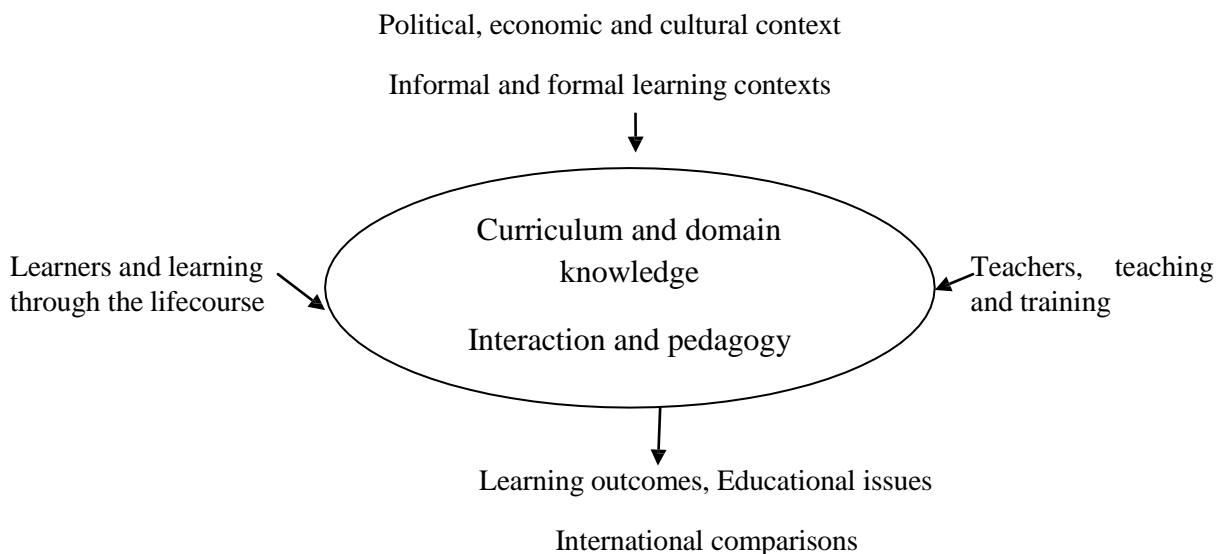


Figure 1: Teaching and Learning Research Programme's framework for analysis

4. Research design and methodology

This inquiry seeks to establish first year university physics students' perceptions of the teaching-learning environment as its primary objective. This was accomplished by administering the Experiences of Teaching and Learning Questionnaire (ETLQ) to first year mainstream Electrical Engineering (National Diploma Programme) students at the University of Johannesburg ($N = 114$) as part of a survey. For most of the items in the questionnaire, students responded on a 1-5 (5 = high) Likert Scale. Subscales were formed by adding together the responses on the items in that subscale. Each item was set as a variable and then a subscale total was produced by creating a new variable by summing the items. The average score was subsequently calculated for each category. The nature of the data obtained necessitated analysis by means of statistical tools. This inquiry is underpinned by the Teaching and Learning Research Programme's framework for analysis as the underlying conceptual framework with a view to ensure appropriate epistemological coherence and premised on the following research question:

What are first year university physics students' perceptions of the teaching-learning environment?

5. Findings

Table 1 below provides average values reflecting students' perceptions in terms of the various constructs which underpin learning and studying.

Table 1: Average values reflecting students' perceptions in terms of the various constructs which underpin learning and studying

Constructs	Key underlying aspects	Score
Deep approach	Intention to understand for oneself	6.7
	Relating ideas (including constructivist learning)	7.6
	Use of evidence	7.4
Surface approach	Memorising without understanding	3.0
	Unreflective studying	2.5
	Fragmented knowledge	3.1
	Unthinking acceptance	3.0
Monitoring studying	Monitoring understanding	7.5
	Monitoring generic skills	6.5
Organised studying	Organised studying	3.4
and effort management	Time management	3.2
	Effort	4.1
	Concentration	3.3

In terms of the scores, students appeared to experience difficulties with organised studying, time management, effort as well as concentration which are vital ingredients for academic success. Table 2 below depicts average values in terms of students' perceptions of the teaching-learning environment. While students' perceptions of the teaching-learning environment appeared to be largely positive in terms of the constructs specified, it was felt that the teaching-learning environment provided limited opportunities for autonomy (choice) and indulgence in broad general discussions (doubtful generality). In addition, the students were particularly pleased with aspects such as encouragement of high-quality learning, clarity and feedback about assessment, assessment for understanding, staff enthusiasm and support, as well as interest, enjoyment and relevance associated with the teaching-learning environment.

Table 2: Average values in terms of students' perceptions of the teaching-learning environment

Construct	Score
Organisation, structure and content	10.3
Alignment	10.9
Integration of teaching and learning materials	10.4
Choice	6.1
Encouraging high-quality learning	16.5
Clarity and feedback about assessment	18.3
Assessment for understanding	14.6
Staff enthusiasm and support	12.8
Support from other students	10.1
Interest, enjoyment and relevance	17.2
Doubtful generality or not easy to place	8.8

As reflected in Table 3 below, students' perceptions of the teaching-learning environment in terms of the cognitive demands made by the learning unit appeared to paint an optimistic picture as it was felt that meaningful opportunities for the acquisition of knowledge and subject-based skills were provided.

Table 3: Average values in terms of the demands made by the learning unit

Construct	Score
Acquiring knowledge and subject-based skills	17.2
Organising and communicating	18.5

While students' perceptions of the teaching-learning environment appeared to express a sense of appreciation for the opportunities provided for the acquisition of skills, more still needs to be done in this regard (Table 4).

Table 4: Average values in terms of the learning achieved

Construct	Score
Acquiring knowledge and subject-based skills	9.8
Organising and communicating	17.3

6. Discussion

First year at university can have many overly complex ramifications as it may in some instances turn out to be a daunting experience for students. It has been established that first year is in several respects a more challenging year for students [13]. Within the context of this inquiry, the key difference in the way students learn and study appeared to hinge on the approaches they use. The other key difference manifested itself in terms of the extent to which students are well-organised in their study methods, use their time efficiently, and put concentrated effort into their work. Students' perceptions of the teaching-learning environment in this inquiry have been considered in terms of various constructs which are dependent on curriculum dynamics. Awareness of disciplinary norm versus the diversity of student intake and the need to bridge the theory-real world divided are strongly advocated [13]. In terms of the conceptual framework outlined for this inquiry, the impact of contextual factors associated with curriculum change such as departmental context, institutional culture and history and policy environment characterising the teaching-learning environment ought to be quantified given the fact that the academic personnel may be operating under conditions of considerable pressure and increasing workloads [13]. The teaching-learning environment in this regard is essentially viewed as occurring within a formal and cultural learning context. The dynamics associated with the teaching-learning environment under consideration are encapsulated in the adapted conceptual framework as illustrated in Figure 2 below.

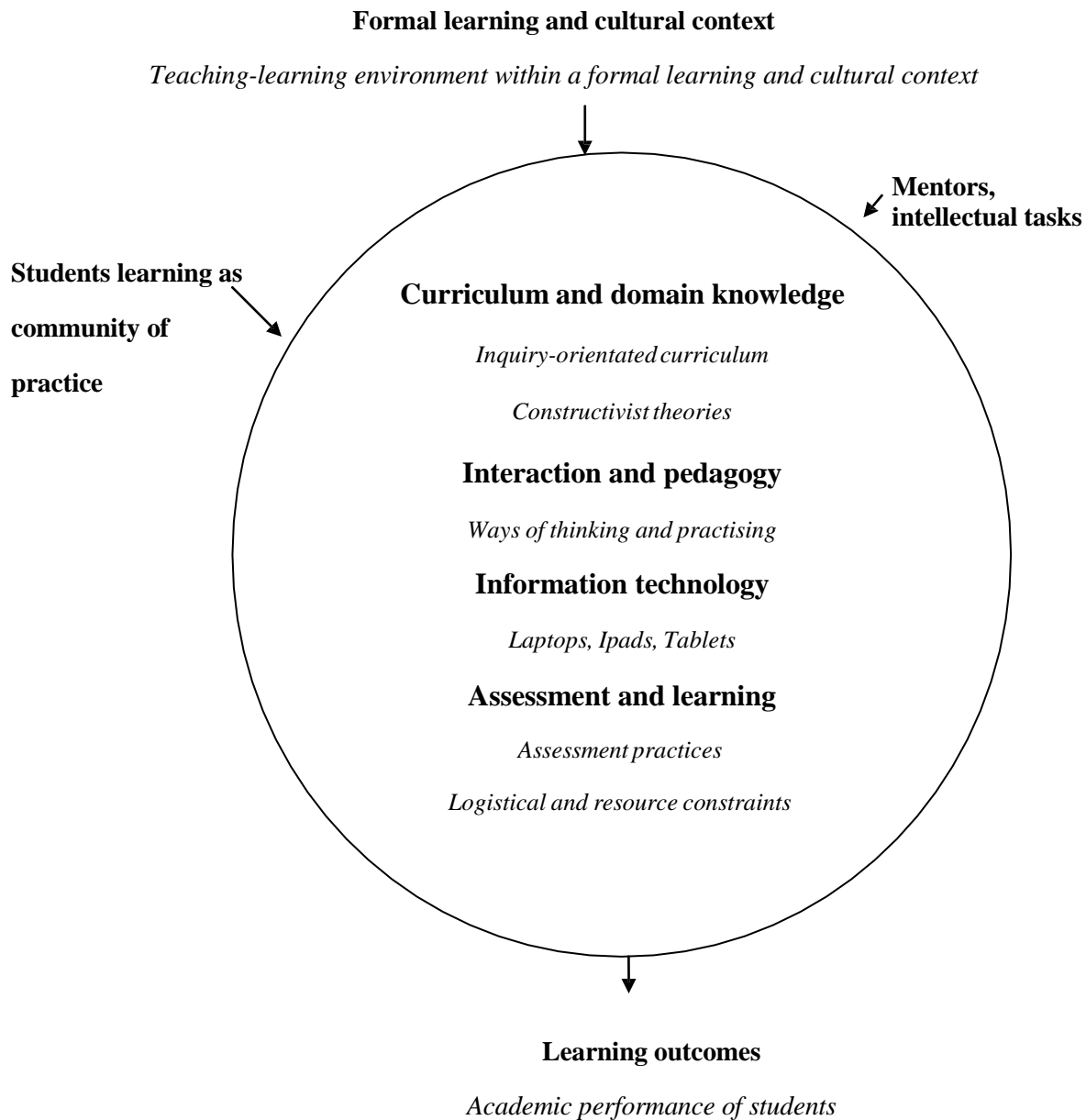


Figure 2: Adapted conceptual framework reflecting dynamics associated with the teaching-learning environment

7. Conclusion

There is an insatiable need to explore the nature of teaching-learning environments in more creative and meaningful ways in order to engender rational discourse about the interwoven factors afflicting the teaching-learning process. Students' perceptions of teaching-learning environments would hopefully serve to harness and enrich the relevant rational discourse in view of the fact that students' pedagogic learning orientation appeared to hinge on the approaches they use within the context of this inquiry. Within the context of this inquiry, students' pedagogic learning orientation appeared to hinge on the approaches they use as well as the extent to which they are well-organised in their study methods, use their time efficiently, and put concentrated effort into their work.

References

- [1] Entwistle, N., McCune, V., & Hounsell, J. (2003). Investigating ways of enhancing university teaching-learning environments: Measuring students' approaches to studying and perceptions of teaching. In E de Corte, L. Verschaffel, N. Entwistle, & J. Van Merriënboer, (Eds.), *Unravelling Basic Components and Dimensions of Powerful Learning Environments*: Oxford: Elsevier Science.
- [2] Experiences of Teaching and Learning (ETL) Project. (2001-2004). Higher and Community Education, The School of education, University of Edinburgh, Paterson's Land, Holyrood Road, Edinburgh EH 8 8AQ.
- [3] De Corte, E. (1995). Fostering cognitive growth: A perspective from research on mathematics. *Educational Psychologist*, **30**, 37-46.
- [4] Phillips, J.A. (1995). *The Theory of Multiple Intelligences: A Framework for Assessing the Competencies and Potentials of Malaysian Students*. Abstracts of the Second Asia-Pacific Conference in Educational Assessment and Research. Chiang Mai, Thailand.
- [5] Schank, R. C. (1999). *Dynamic Memory Revisited* (2nd ed.). Cambridge: Cambridge University Press.
- [6] Ramsden, P. (1981). *A Study of the Relationship between Student Learning and its Academic Context*. Unpublished PhD thesis, University of Lancaster.
- [7] Trigwell, K., Prosser, M., & Waterhouse, F. (1999). Relations between teachers' approaches to teaching and student learning. *Higher Education*, **37**, 57-70.
- [8] Anderson, C.D.B. (1997). Enabling and shaping understanding through tutorials. In F. Marton, D.J. Hounsell, & N.J. Entwistle (Eds.), *The Experience of Learning* (2 nd ed.) (pp. 72-88). Edinburgh: Scottish Academic Press.
- [9] Entwistle, N.J. (1987). A model of the teaching-learning process. In Richardson, J.T.E., Eysenck, M.W., and Warren Piper, D. (eds.), *Student Learning: Research In Education and Cognitive Psychology*. London: S.R.H.E. / Open University Press, pp. 13-28.
- [10] Buchner, T. (2009). Deinstitutionalisation and community living for people with intellectual disabilities in Austria: History, policy, implementation and research. *Tizard Learning Review*, **14**(1), 4 -13.
- [11] Trowler, P. (1998). *Academics Responding to Change: New Higher Education Frameworks and Academic Cultures*. Buckingham: Open University Press
- [12] Teaching and Learning Research Programme (T.L.R.P). (2006). *Improving Teaching and Learning in Schools – A Commentary by the Teaching and Learning Research Programme*. Institute of Education. University of London.
- [13] Teaching and Learning Research Programme (T.L.R.P). (2006). Enhancing learning and teaching in Economics: A digest of research findings and their implications. *Enhancing Teaching-Learning Environments in Undergraduate Courses Project*, Higher and Community Education, School of Education, University of Edinburgh, Edinburgh.

Lesson planning perceptions and experiences of South African Physical Sciences teachers in a new curriculum

¹Sam Ramaila and ²Umesh Ramnarain

¹Department of Applied Physics and Engineering Mathematics, University of Johannesburg, P.O. Box 17011, Doornfontein, 2025, Johannesburg, South Africa

²Department of Science and Technology Education, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, Johannesburg, South Africa E-mail:

samr@uj.ac.za

Abstract. This paper reports findings on the perceptions of South African Physical Sciences teachers on lesson planning. Significant changes to the school science curriculum appear to put considerable pressure on teachers to revisit how they plan lessons. We adopted a mixed method approach in collecting and analysing data from a large-scale survey of teachers through a structured questionnaire, and followed this with interviews with 10 teachers in seeking more in-depth explanations of the findings which emerged from the survey. The study revealed that when feasible, teachers work collaboratively in a community of practice when planning lessons. Apart from reducing the planning time, this strategy also leads to creative and innovative ideas that are shared. This is especially the case when teaching topics that are new in the curriculum and also lessons that are inquiry-based. Teachers also believe that writing a lesson plan does have pedagogical value because it serves as a support mechanism in planning deliberately for difficulties they encounter in addressing curriculum implementation challenges. The lesson plan therefore supports teachers in their role as reflective practitioners.

1. Introduction

The democratic breakthrough in 1994 in South Africa resulted in the need to transform the education landscape. A key goal in transforming the Apartheid education system was the need to provide quality education for all racial groups [1]. The new democratic South African government sought to develop and enact policies that sought to enhance the quality of school science education that was traditionally characterised by inequity. In particular, curriculum reform in Physical Sciences education was driven by the view that the previous curriculum during the Apartheid era was both inaccessible and irrelevant to historically disadvantaged black learners. In an attempt to address the above concerns and also the imbalances in the education system, an outcomes-based curriculum was introduced in 1998. This curriculum advocates a learner-centred and activity-based approach to education [2]. This development signalled a significant paradigm shift in the manner in which the learner was portrayed as the previous curriculum was characterised by learner passivity and teacher-directedness.

In relation to Physical Sciences, curriculum planners were tasked with developing a curriculum with content that would be meaningful, accessible and relevant to all learners. This resulted in restructuring of the existing content and the addition of new topics. The planners were also faced with the demand of producing a curriculum which is more representative of the basic tenets of the nature of science such as the tentativeness of scientific knowledge and learners appreciating other systems of knowledge such as indigenous knowledge. Practical work in the school science curriculum is an area which received considerable attention in the curriculum reform initiatives which took place in South Africa. Traditionally, practical work took the form of teacher demonstrations or embodied a “cookbook approach” in which learners follow recipes for the execution of procedures handed down by the teacher.

The National Curriculum Statement advocated an inquiry-based approach to practical work that encourages learners on their own to explore objects, situations and events in their immediate environment, to collect data and record information and draw conclusions accurately [2]. These significant changes to the school science curriculum require teachers to revisit how they plan lessons.

2. Conceptualising lesson planning in science teaching

Lesson planning is widely regarded as being pivotal to effective curriculum delivery. Accordingly, it has been suggested that lesson plans serve as frameworks for conceptualizing, designing and delivering instruction [3]. A review of literature provides a compelling argument for teachers to engage in lesson planning. Lesson planning reminds the teacher of the specific goals to be achieved through classroom teaching and provides a picture of what is to be pursued and what is to be avoided [4]. In terms of the South African Physical Sciences Curriculum and Assessment Policy Statement (CAPS), such goals are expressed through specific aims, content specification and requirements for practical work [5]. Furthermore, a lesson plan is regarded as being essential to economize time and to make the teaching systematic and coherent [4].

The format of a lesson plan can vary. The dominant model for lesson planning is linear. It begins with the specification of objectives and ends with a lesson assessment. This rational approach to planning has been strongly influenced by instrumental interpretations of Tyler's Basic Principles of Curriculum and Instruction [6] that regards the curriculum as comprising of planned learning experiences that are guided by intended learning outcomes and theorists such as Benjamin Bloom [7] who have developed taxonomies of learning [8]. Such a simplistic conception of lesson planning may reinforce views especially those of student teachers that teaching is the manifestation of classroom procedures [9]. Typically such a lesson plan grid will comprise of the following sections: topic; lesson outcome; specific aim; resources; teaching-learning activities; assessment opportunities and reflection. Such a format does not make explicit the internal logic that should guide prospective teachers in the way they think about conceptual progression, the conceptual relations within the knowledge field and what constitutes appropriate inquiry processes within a subject [9]. In view of the significant changes to the school science curriculum, this research investigated the perceptions and experiences of South African Physical sciences teachers on lesson planning.

3. Methodology

The study employed a triangulation mixed methods design in which investigators collect both quantitative and qualitative data, merge the data, and use the results to best understand a research problem [10]. In this design, both qualitative and quantitative methods have equal status. Quantitative data was collected by means of a questionnaire that was administered to Physical Sciences teachers at various schools in the Gauteng Province of South Africa. The findings that emerged from the questionnaire analysis were then explored further by means of interviews with a smaller sample of teachers. Figure 1 below depicts a visual representation of the research design used in this inquiry.

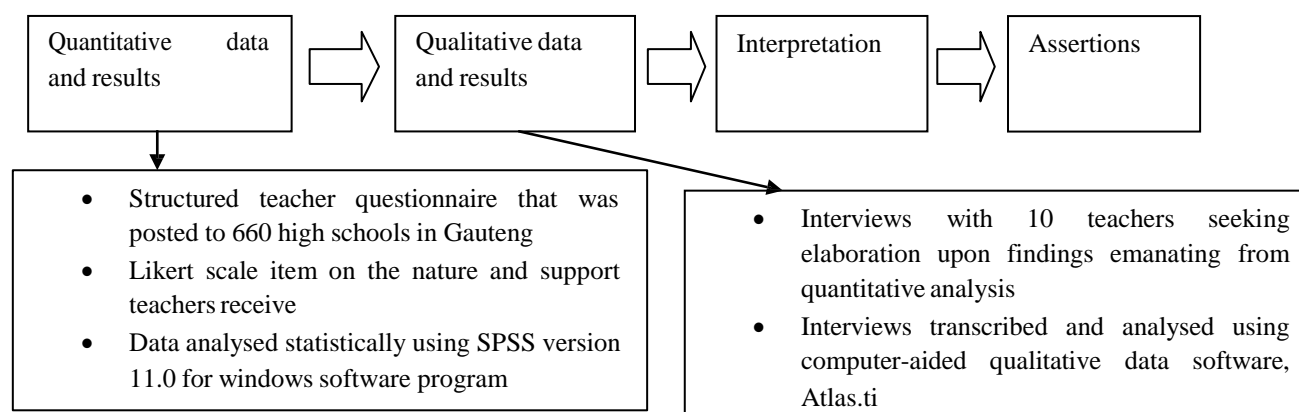


Figure 1: Visual representation of the research design used in this study

Quantitative data was collected by means of a structured questionnaire (Appendix A) that was posted to 660 high schools in the Gauteng Province of South Africa. Of the 660 questionnaires that were sent out, 263 questionnaires were returned. There was a reasonably large representation of respondents from suburban, township and city schools. In comparison, only a few teachers from rural schools responded. In developing the questionnaire items, we firstly conducted a focus group interview with 6 teachers. We were able to identify the core issues surrounding the implementation of the curriculum through this interview. The content validity of this questionnaire in terms of which items related to a particular aspect of curriculum implementation was established by having it reviewed by researchers in science education at three South African universities. Face-to-face interviews were conducted individually with 10 teachers from the sample of teachers who had filled in the questionnaire. These were teachers who had indicated on the questionnaire their willingness to be interviewed. In these interviews we sought more in-depth explanation of some of the findings which emerged from the questionnaire survey. Table 1 below provides a demographic description of each teacher in this sample.

Table 1: Demographic description of teachers interviewed

Teacher	Gender	Age	Qualifications	Type of school	School resources
Teacher 1	Male	35	Teaching Diploma	Township school	Poorly resourced
Teacher 2	Female	44	Teaching Diploma	Township school	Poorly resourced
Teacher 3	Male	28	Education degree	Township school	Poorly resourced
Teacher 4	Male	37	Science degree	Suburban school	Poorly resourced
Teacher 5	Female	45	Education degree	Suburban school	Well resourced
Teacher 6	Male	34	Science degree	Suburban school	Adequately resourced
Teacher 7	Male	43	Teaching Diploma	Suburban school	Adequately resourced
Teacher 8	Female	45	Education degree	Suburban school	Well resourced
Teacher 9	Female	34	Science degree	Suburban school	Adequately resourced
Teacher 10	Female	43	Teaching Diploma	Suburban school	Adequately Resourced

The responses to the questionnaire were analyzed statistically using the SPSS version 11.0 for Windows software program. Frequency tables of data collected for Section A were generated. This enabled us to identify trends on teacher perceptions and experiences of lesson planning in teaching a revised Physical Sciences curriculum. The interviews were transcribed and analyzed using computer-aided qualitative data software, Atlas.ti. Data was then coded and classified, leading to the formulation of themes. This process was largely guided by the trends and patterns which had emerged from the analysis of the questionnaire data in relation to the teachers' perceptions and experiences of lesson planning.

4. Findings

The findings from the analysis of the survey were integrated with the findings from the teacher interviews into a coherent whole. This integration of quantitative and qualitative data supported the production of assertions in relation to teachers' perceptions and experiences of lesson planning in the implementation of a revised Physical Sciences curriculum.

Assertion 1:

Teachers view the lesson plan as a scaffold in addressing implementation challenges in teaching a new curriculum.

The great majority of teachers believe that lesson planning is important, especially in view of the new curriculum, with 76.2% of all teachers either agreeing or strongly agreeing that "Due to the newness of the curriculum all teachers should write lesson plans" (Item 7, Section C). In the interviews, teachers explained that a lesson plan served as a conceptual framework that enabled them to deliberately articulate their ideas on how they prepare their lessons to address changes to the curriculum. They mentioned in particular how the curriculum imperative for inquiry-based education meant that they had to plan activities that were directed at developing science process skills. This is evident in the following excerpts from the interviews:

I am starting to see some value in the lesson plan. Whereas before it was a monotonous thing to do because I was teaching the same topics over and over, now we are being asked to apply our minds. I can put all my ideas on the plan and then start to imagine how it unfolds.

Now the practical work is challenging to organize. The CAPS says let them investigate, and now we must try and organize it for them. I must think about a whole lot of issues, like what skills are being targeted and how to achieve it. I have a lesson plan grid I adapted that shows in time segments what they do and how I must help them.

Assertion 2:

In planning their lessons, teachers find it useful to work within a community of practice so that ideas can be shared, especially when teaching new content.

To a large extent, teachers find much value in working collaboratively when planning lessons. This was underlined in their responses to item 3, "When I am planning my lessons I find it useful to have a discussion with a colleague", in which 82% of teachers either agreed or strongly agreed with this statement. In these discussions, teachers exchange ideas on the most effective pedagogy to teach a particular topic. This was reflected in their responses to item 33 where 84% of all teachers either agreed or strongly agreed with "I exchange ideas with colleagues on how best to teach a topic". During interviews teachers elaborated as follows on this practice in lesson planning.

It is always useful to find a colleague who is teaching the same section and then we exchange our notes on it. I learn a lot from these meetings. Sometimes we are lost on the new topics and then I know a colleague can support me with something.

Teachers also explained that working within a community of practice enabled them to develop pedagogical strategies on how best to support learners when performing inquiry-based activities. A teacher mentioned that in inquiry-based learning, the teacher needed to redefine his role from that of a transmitter to a facilitator of learning. He explains as follows how interaction with a colleague from another school helped him acquire some understanding of what this role entailed.

The curriculum speaks of us facilitating learning when learners are busy with inquiry. I struggled with this and had no clue on what this meant. I then met another science teacher at a district and he sat with me and told me using some examples how I needed to plan my lessons. I found this to be enlightening.

Assertion 3:

Although teachers recognize the value of lesson planning, they bemoan the lack of time for it.

Due to the overwhelming demands of assessment requirements and the related marking associated with this, teachers struggle to find time to plan their lessons. In responding to item 13 “I wish I had more time available to plan my lessons”, 89% of teachers expressed the need for more time in planning. The teachers also highlighted that the demands of the new curriculum necessitated that they had to spend more time in planning their lessons. The following excerpts attest to this:

Where is the time for writing out a lesson plan? I spend more time both at school and at home sitting with portfolio files and assessment.

I wish I had more time in planning out my lessons. I try, but I get bogged down with other stuff like marking and so on. Also, planning a lesson is now no longer a 10 minute job. We have to be very direct on what is being taught and our approach in doing it. We must also think about assessment both formative and summative.

Assertion 4:

The teachers use the internet quite extensively in lesson planning.

Teachers maintain that the current set of school textbooks are inadequate in lesson planning, and they therefore sourced material from the internet. A large number of teachers made extensive use of the internet in seeking learner support materials. In fact, 65% of the teachers either agree or strongly agree with the statement “I make use of the internet to get ideas on teaching a new topic” (Item 15). They regard textbooks as lacking in learner support material that is sensitive to the cultural background of the learner, and they also regarded some of the activities presented as being far removed from the experiences of learners. Furthermore, they expressed the view that the current set of textbooks still promoted a didactic, teacher-centred approach to the curriculum imperative for inquiry-based learning. As a result, they had to search the internet websites such as Biological Sciences Curriculum Study (BSCS) for material that promoted an inquiry-based approach to teaching. The following excerpts reflect some of these views:

The textbooks we were asked to order do not have enough activities for learners to do for practical work according to inquiry-based approach. How should we cope now when the curriculum says teach inquiry? When I have time I try and make use of the internet education sites I can find.

All the learners have textbooks we handed out, but these books do not have what we want learners to do. They are still written in the old traditional style where we must be transmitting knowledge to them. How do I get them to explore something if there is nothing in the textbook? Fortunately with the internet I can google search some things that are relevant and current.

5. Discussion and conclusion

The research has revealed that teachers place much importance on lesson planning, and this is especially pronounced against the backdrop of significant curriculum reform that took place. Teachers believe that the process of engaging writing a lesson plan enables them to explicate their ideas on how the lesson should unfold. The lesson plan serves as a support mechanism in planning deliberately for difficulties they encounter in addressing curriculum implementation challenges. The teachers also indicated that if given the opportunity, they worked collaboratively with colleagues within a community of practice when planning lessons. In this community of practice they shared ideas on how to teach certain topics, especially those that were new in the curriculum, and this engagement enabled them to teach topics more effectively.

The teachers also highlighted the challenges they experience in providing inquiry-based lessons due to their lack of expertise and experience in this regard. Furthermore, they were largely reliant on colleagues for support and guidance when planning lessons. The teachers regarded the current set of textbooks as being inadequate in facilitating a learner-centred, constructivist approach which provides opportunities for learners to investigate their ideas practically. It was evident from the perceptions and experiences of teachers in lesson planning that the lesson plan is a mechanism through which ideas can be articulated on how learners can be supported in the construction of scientific knowledge. The conception of lesson planning that emerges from this research is one that goes beyond the dominant simplistic notion of a lesson plan being a linear and procedural manifestation of classroom procedures.

References

- [1] Department of Education (DoE). (2001). *Education in South Africa: Achievements Since 1994*. Pretoria: Government Printer.
- [2] Department of Education. (2003). *National Curriculum Statement Grades 10-12: Physical Sciences*. Pretoria: Government Printer.
- [3] Ozogul, G. & Sullivan, H. (2009). Student performance and attitudes under formative evaluation by teacher, self and peer evaluators. *Educational Technology Research and Development*, **57**, 393-410.
- [4] Stella, O.M. (2012). Instilling right attitudes towards the use of lesson plans in chemistry instruction in Mosoch Division of Kisii District, Kenya. *Journal of Emerging Trends in Educational Research and Policy Studies*, **3**(2), 143-146
- [5] Department of Basic Education (2011). *Curriculum and Assessment Policy Statement: Grades 10-12 Physical Sciences*. Pretoria: Government Printer.
- [6] Tyler, R.W. (1949). *Basic Principles of Curriculum and Instruction*. Chicago: University of Chicago Press.
- [7] Bloom, B. S. (ed.) (1956) *Taxonomy of Educational Objectives: Handbook I: Cognitive Domain* (New York: McKay).
- [8] John, P. (2006). Lesson planning and the student teacher: re-thinking the dominant model. *Journal of Curriculum Studies*, **38**(4), 483-498.
- [9] Rusznyak, L. & Walton, E. (2011) Lesson planning guidelines for student teachers: A scaffold for the development of pedagogical content knowledge. *Education as Change*, **15**(2), 271-285.
- [10] Creswell, J. W. (2002). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Approaches to Research*. Upper Saddle River, NJ: Merrill/Pearson Education.

Students' perceptions of the study process

Sam Ramaila, Leelakrishna Reddy, Padmanabhan Nair and Suzan Bvumbi

Department of Applied Physics and Engineering Mathematics, University of Johannesburg, P.O. Box 17011, Doornfontein, 2025, Johannesburg, South Africa

E-mail: samr@uj.ac.za

Abstract. The inherent structure of the study process is central to the provision of meaningful opportunities geared towards the achievement of success in any academic programme. Given this scenario, Biggs' Study Process Questionnaire (SPQ) was administered to first year university physics students in order to establish their perceptions about the nature of the study process which underscores the academic achievement. Analysis of SPQ responses provided valuable insights into students' perceptions about the nature of the study process. While the use of deep approach appeared to be a predominant learning tool within the context of this inquiry, a substantial number of students still appeared to employ surface approach to navigate their studies.

1. Introduction

There has been considerable focused attention on students' approaches to learning with a view to search for the fundamental differences in relation to students' engagement in learning tasks [1,2]. Deep and surface approaches have been extensively utilised in educational research as key constructs [1]. Deep approach is mainly concerned with students' intentions to understand and construct the meaning of content to be learned while surface approach specifically refers to students' intentions to learn by memorizing and reproducing factual contents of the study materials.

Both quantitative and qualitative inquiries have been undertaken into students' approaches to learning [2,3]. These inquiries revealed a complex dichotomy between deep approach and surface approach in students' learning [4]. Over and above the two approaches, a mixed approach to learning called the achieving approach has also been identified [5,6]. The achieving approach manifests itself either in the form of deep or surface processing depending on the key critical contextual demands [7]. This inquiry is largely inspired by the intellectual curiosity into the relationship between students' approaches to learning and students' learning outcomes. Congruent with this endeavour, this article provides insights into first year university students' perceptions about the nature of the study process.

2. Student approaches to learning and studying

Student approaches to studying comprise two elements: motive and strategy [2]. Three approaches to learning each consisting of motive and strategy [2] are illustrated in Figure 1 below.

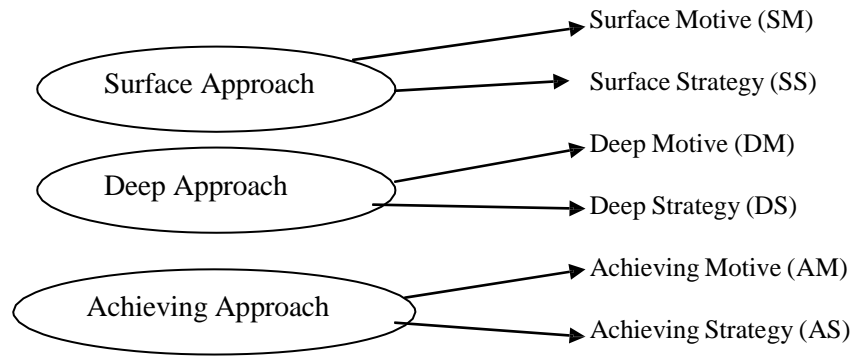


Figure1: Biggs' conception of a 6-factor structure in students' approaches to learning (Biggs, 1987)

Surface approach is premised on an extrinsic motive which seeks to carry out the task because of either positively or negatively reinforcing consequences. Deep approach is underpinned by an intrinsic motive which arises from the curiosity to seek meaning. The achieving approach is inspired by an achieving motive which primarily focuses on a product. Students' development of a specific learning approach is largely influenced by personal factors and the teaching context [8]. Personal factors such as background and personality are associated with a surface approach [8] and others with a deep approach [9]. Teaching factors such as time pressures, examination stress, and using test items emphasising low level cognitive outcomes encourage surface approach while learner activity, student-student interaction, and interactive teaching, particularly problem-based teaching encourage a deep approach [10]. In order to provide further elucidation, a summary of the differences in motivation and study process of surface, deep, and achieving approaches to study [2] is provided in Table 1 below.

Table 1: Summary of the differences in motivation and study process of surface, deep, and achieving approaches to study (Biggs, 1987)

Approach	Motivation	Process (strategy)
Surface	Fear of failure	Rote learning of facts and ideas
	Desire to complete their course of study	Focusing on task components in isolation
		Little real interest in content
Deep	Interest in the subject	Relate ideas to evidence
	Vocational relevance	Integration of materials across courses
	Personal understanding	Identifying general principles
Achieving	Achieving high grades	Use any technique that achieves highest grades
	Competing with others to be successful	Level of understanding patchy and variable

A coherent inquiry on student approaches to learning in educational research is underpinned by a systems model of learning called the presage-process-product model [11]. According to the model, learning is characterised by three inter-related components in the form of presage (student-based factors and the learning environment), process (how students engage in the task) and product (learning outcome) as depicted in Figure 2 below. Presage factors are independent of the learning situation and include personal factors (intelligence quotient, background, personality characteristics) and situational factors (subject content, methods of teaching and evaluation, course structures) [2]. It is important to point out that presage factors may affect student's performance directly or indirectly through their influence on process factors [2]. By their very nature, process factors determine the way the student goes about learning [2]. This implies that process factors are essentially anchored on students' motives for learning and their concomitant strategies.

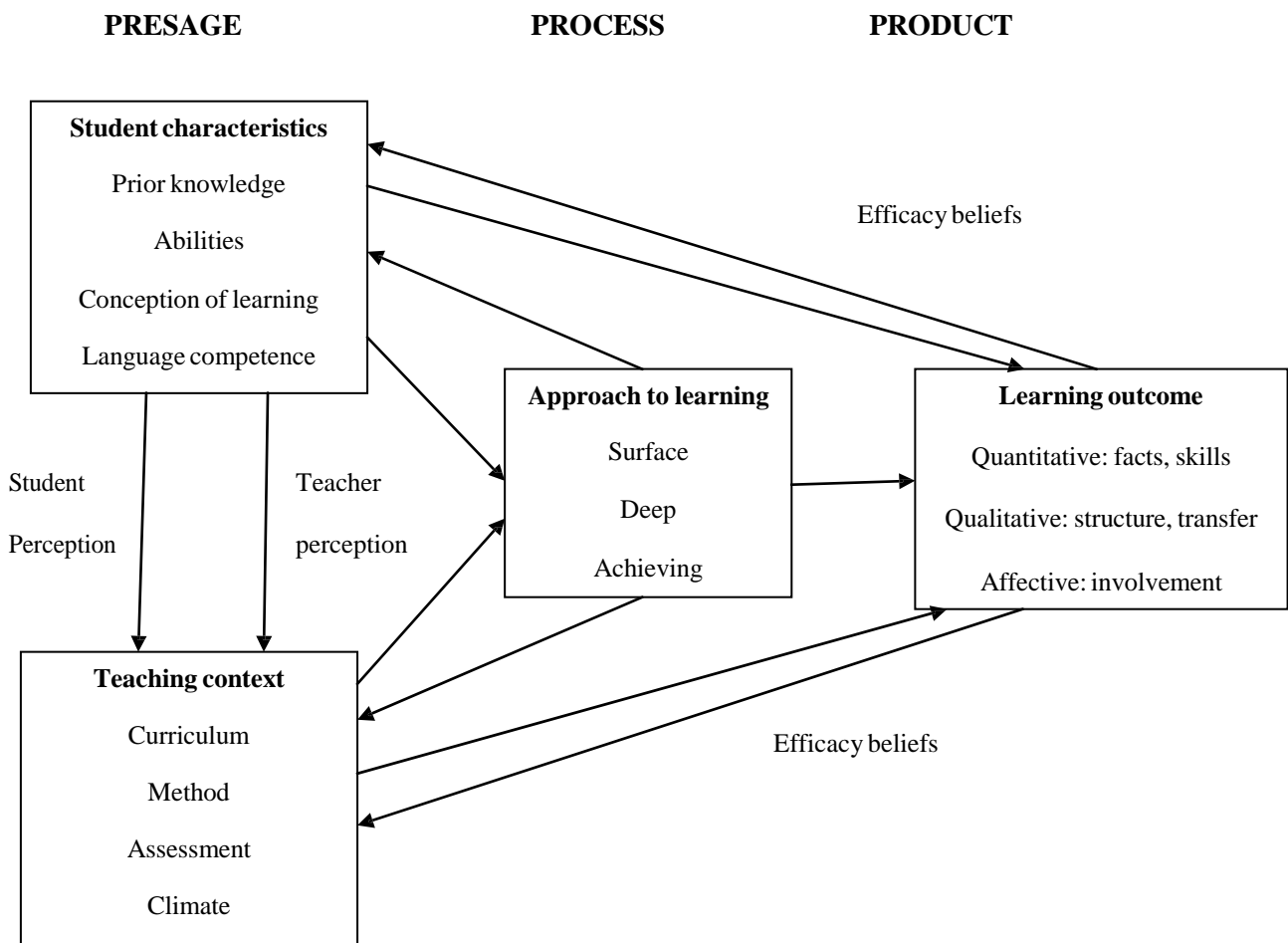


Figure 2: Systems Model of Study Process (Biggs, 1999)

Meta-learning is an important key parameter in educational research on student approaches to learning. Meta-learning refers to students' awareness of and control over their own learning [2]. It has been established that students show lack of meta-learning capability when they choose strategies that are incongruent with motives such as rote learning (surface strategy) to satisfy intrinsic curiosity (deep motive) [2]. Thus, the relationship between personal and situational factors associated with approaches to learning and performance can be meaningfully assessed in terms of a general model of student learning [2] illustrated in Figure 3 below.

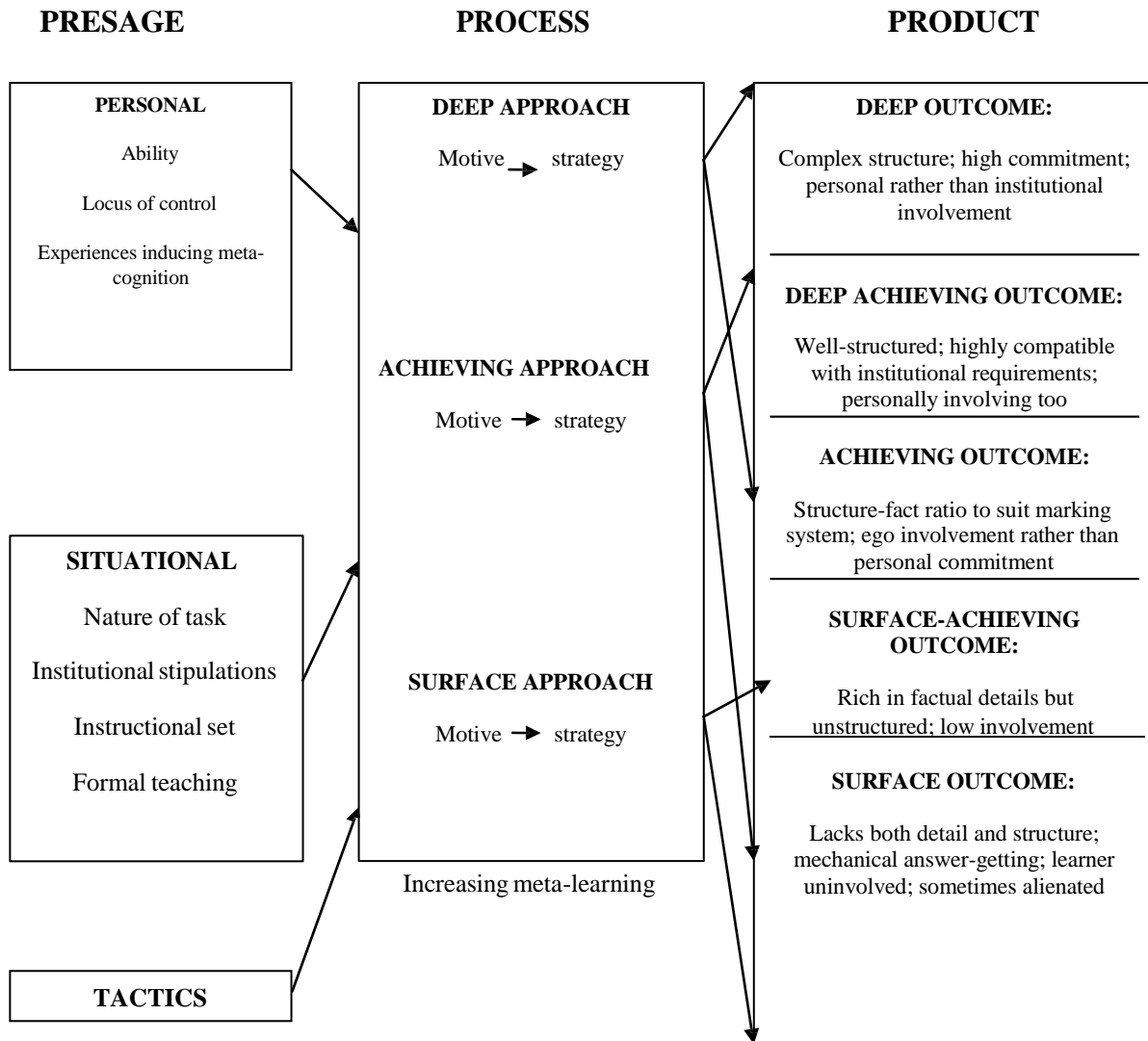


Figure 3: A General Model of Student Learning (Biggs, 1987)

3. Research design and methodology

This inquiry seeks to establish first year university students' perceptions about the nature of the study process as its primary objective. This was accomplished by administering the Biggs' Study Process Questionnaire [2] to two groups of first year physics students enrolled for the Optometry Degree Programme (main stream) ($N = 60$) and National Diploma in Mechanical Engineering Programme (Extended) ($N = 50$) as part of a survey. The Biggs' Study Process Questionnaire is a 20 item inventory with a 5 point Likert scale. The nature of the data obtained necessitated analysis through quantitative means. This inquiry is underpinned by the Systems Model of Study Process [2] and a General Model of Student Learning [2] as theoretical frameworks with a view to ensure appropriate epistemological coherence.

4. Results

The questionnaire data provided predominant students' perceptions about the nature of the study process. Table 2 below provides students' predominant perceptions about the nature of the study process. These predominant perceptions are based on the SPQ items which reflected higher scores.

Table 2: Predominant students' perceptions about the nature of the study process

I find that at times studying gives me a feeling of deep personal satisfaction.
I find that I have to do enough work on a topic so that I can form my own conclusions before I am satisfied.
I only study seriously what's given out in class or in the course outlines.
I feel that virtually any topic can be highly interesting once I get into it.
I find most new topics interesting and often spend extra time trying to obtain more information about them.
I learn some things by rote, going over and over them until I know them by heart even if I do not understand them.
I find that studying academic topics can at times be as exciting as a good novel or movie.
I test myself on important topics until I understand them completely.
I work hard at my studies because I find the material interesting.
I make a point of looking at most of the suggested readings that go with the lectures.
I find the best way to pass examinations is to try to remember answers to likely questions.

The scores for deep and surface approaches were 1733 (59%) and 1181 (41%), respectively. The scores suggest that a significant number of students (41%) enrolled for the National Diploma in Mechanical Engineering Programme (Extended) appeared to lean towards surface approach. This practical consideration appears to be commensurate with the nature of Grade 12 achievement levels for students admitted to the National Diploma in Mechanical Engineering Programme (Extended) which are generally lower in terms of the admission point score. As a key requirement, students in this programme are subjected to intensive tuition characterised by several compulsory academic support interventions in order to address knowledge gaps associated with their academic background. The scores for deep and surface approaches were 1647 (56%) and 1317 (44%), respectively. The scores suggest that a significant number of students (44%) enrolled for the Optometry Degree Programme (main stream) also appeared to lean towards surface approach. Within the contexts of the two academic programmes, Optometry students were expected to demonstrate deep approach in view of their admission requirements while the extended mechanical engineering students were expected to demonstrate surface approach. Table 4 below provides a summary of the scores for both surface and deep approaches for the two groups.

Table 4: A summary of the scores for both surface and deep approaches for the two groups

Group	Deep Approach Score	Surface Approach Score
Optometry (Degree Programme)	1647 (56%)	1317 (44%)
Mechanical Engineering (National Diploma Programme)	1733 (59%)	1181 (41%)

5. Discussion

The scores obtained seem to suggest that a significant number of students enrolled for the National Diploma in Mechanical Engineering Programme (Extended) and the Optometry Degree Programme (main stream) appeared to lean towards surface approach. In terms of the Systems Model of Study Process [2], this prevailing scenario may be attributed to the complexity of factors such as student characteristics and teaching context. Student characteristics include key parameters such as prior knowledge, abilities, conception of learning and language competence while teaching context include critical aspects such as curriculum, method, assessment and climate. As a matter of fact, the use of a

deep learning approach is generally associated with higher quality learning outcomes and a surface approach with lower quality learning outcomes [12,13].

In a similar vein, students are admitted to the National Diploma in Mechanical Engineering Programme (Extended) on the basis of relatively weaker Grade 12 achievement levels as required by the stipulated admission point score in recognition of various deficiencies associated with their schooling background. Key parameters associated with critical factors such as student characteristics and teaching context may to some degree create a dissonance with regard to the adoption of surface and deep approaches by the students. In fact, the differences in quantitative learning outcomes between students using surface or deep approach is only significant for questions measuring insights and not for questions measuring reproduction of knowledge [14]. In addition, it has been established that there is no evidence that a deep approach to learning would be more effective for questions assessing more complex components of problem-solving [15]. However, problem-solving as an instructional approach has the potential to facilitate deep approaches to learning [16].

6. Conclusion

While the use of deep approach appeared to be a predominant learning tool within the context of this inquiry, a substantial number of students still appeared to employ surface approach to navigate their studies. There is a need to develop a critical understanding of students' approaches to learning in order to provide a coherent interpretation of their prevalent pedagogic learning orientation.

References

- [1] Marton, F., & Säljö, R. (1976). On qualitative differences in learning – I: Outcome and process. *British Journal of Educational Psychology*, **46**, 4-11.
- [2] Biggs, J.B. (1987). *Student Approaches to Learning and Studying*. Melbourne: Australian Council for Educational Research.
- [3] Marton, F. (1981). Phenomenography: Describing conceptions of learning. *International Journal of Educational Research*, **19**, 277-300.
- [4] Prosser, M., & Trigwell, K. (1999). *Understanding Learning and Teaching: The Experience in Higher Education*. Buckingham: The Society for Research into Higher Education.
- [5] Biggs, J.B. (1993). What do inventories of students' teaching processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology*, **63**, 1-17.
- [6] Entwistle, N.J. (1991). Approaches to learning and perceptions of the learning environment. *Higher Education*, **22**, 201-204.
- [7] Mäkinen, J. (2003). *University students' General Study Orientations – Theoretical Background, Measurements and Practical Implications* (Dissertation). Turku: Turun Yliopisto.
- [8] Biggs, J.B. (1989). Approaches to the enhancement of tertiary teaching. *Higher Education Research and Development*, **8**, 7-25.
- [9] Biggs, J.B. (1987b). *The Study Process Questionnaire (SPQ): Manual*. Hawthorn, Vic.: Australian Council for Educational Research.
- [10] Biggs, J.B., & Tefler, R. (1987). *The Process of Learning* (2nd ed.). Sydney: Prentice Hall of Australia.
- [11] Biggs, J.B. (1999). *Teaching for Quality Learning at University*. Buckingham: The Open University Press.
- [12] Snelgrove, S. & Slater, J. (2003). Approaches to learning: Psychometric testing of a study process questionnaire. *Journal of Advanced Nursing*, **43**(5), 496-505.
- [13] Zeegers, P. (2001). Student learning in science: A longitudinal study. *British Journal of Educational Psychology*, **71**, 115-132.
- [14] Biggs, J.B. & Collis, K. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy*. New York: Academic Press
- [15] Minbashian, A., Huon, G.F., & Bird, K.D. (2004). Approaches to studying and academic performance in short-essay exams. *Higher Education*, **47**(2), 161-176.
- [16] Biggs, J.B. (2003). *Teaching for Quality Learning at University* (2nd ed.). Buckingham: The Open University Press.

A critical assessment of first year entering university science students' conceptual understanding

Leelakrishna Reddy, Padmanabhan Nair and Sam Ramaila

Department of Applied Physics and Engineering Mathematics, University of Johannesburg, P.O. Box 17011, Doornfontein, 2025, Johannesburg, South Africa

E-mail: samr@uj.ac.za

Abstract. Meaningful conceptual understanding is a key requirement for the acquisition of applied knowledge for the navigation of sophisticated studies in science and engineering. It is a known fact that the academic migration of students from the school sector to institutions of higher learning is essentially accompanied by the existence of conceptual knowledge gaps in various content domains. In response to this difficulty and as a critical component of the First-Time Entering Students' Orientation programme, a diagnostic questionnaire was administered to 222 first year science and engineering students at the University of Johannesburg prior to the commencement of the academic programme in order to establish the nature and extent, if any, of their conceptual knowledge gaps. In terms of its design, the diagnostic questionnaire encapsulated items based on various conceptual knowledge areas pertaining to Grade 8 to 12 Physical Science content domains. Analysis of students' responses revealed the existence of conceptual knowledge gaps which may impede meaningful learning.

1. Introduction

It has been observed that school leavers are inadequately prepared for higher education studies in science within the South African context due partly to lack of adequate mathematical skills (analytical, problem solving, manipulative, visualization and algebraic skills) and fundamental changes associated with curriculum reform as major impediments afflicting student learning [1]. In the main, curriculum reform in South Africa entailed phasing out the differentiated Higher Grade/Standard Grade system in all the learning domains and the inclusion of Mathematics Literacy as a viable alternative for learners not taking core Mathematics.

Local and international benchmark tests largely painted a gloomy picture about the quality of the South African education system. For instance, Grade 8 and 9 learners demonstrated dismally inadequate performance in mathematics and science [2]. This performance was commensurate with South Africa's global ranking in terms of the quality of mathematics and science education as well as the quality of the overall education system. South Africa was placed 143rd out of 144 countries surveyed during the 2012-2013 period and subsequently 148th out of 148 countries during the period 2013-2014 in terms of the quality of mathematics and science education [3,4].

Yet higher achievement levels are required for students to be admitted for higher studies in science. Results from the National Senior Certificate Examination written during 2013 indicated that a meagre 25.6% of the learners who passed were eligible to enrol for degree studies while 36.8% were eligible for diploma studies [5]. The breakdown in terms of the number of students eligible for higher studies during the period 2010-2013 is as follows [5]: 205 364 (2010), 180 585 (2011), 179 194 (2012) and 184 384 (2013). The breakdown itself presents a grim reality in relation to key aspirations pertaining to the attainment of meaningful human capital development within the broader South African context.

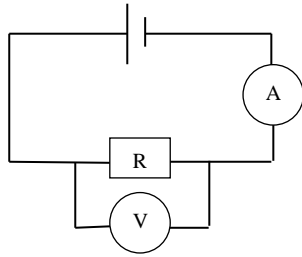
The prevailing articulation gap has a potentially detrimental impact on students' academic success at higher education institutions as a large majority of the students find it increasingly difficult to complete their studies in regulation time. Academic personnel are consequently grappling with various challenges such as large classes and cultural diversity coupled with serious knowledge gaps and conceptual inadequacies exhibited by the students. The conflation of various factors afflicting student academic performance prompted universities to introduce extended programmes underpinned by student-centred approaches making provision for remedial intervention and early identification of at risk students. It is against this background that this article provides a critical assessment of first year entering university science students' conceptual understanding.

2. Methodology

A questionnaire covering Physical Science topics from Grades 8 to 12 was administered to first year entering university students ($n = 222$) in Engineering, Optometry, Food Technology and Biotechnology. The topics covered ranged from general science in the General Education and Training (GET) Phase to Physical Science in the Further Education and Training (FET) Phase. More specifically, these topics included density, work, energy and electricity. Substantial effort was made to ensure adherence to the school syllabus. Suffice to indicate that the questionnaire itself was administered during the orientation week when students were formally introduced to the university setting. A summary of the type of questions, excluding the options (for most questions) and the relevant syllabus coverage is given in Table 1 below.

Table 1: Type of questions used

QUESTION	SYLLABUS IDENTIFICATION
1. A rectangular wooden block is 5 cm long, 3 cm high and 2.5 cm wide. The mass of the block is 26 g. Calculate the density of the wooden block.	Basic general science question from the grade 8 science syllabus.
2. "Frosted glass" is an example of A. Transparent material B. Translucent material C. Opaque material D. None of the above E. I cannot remember this one	A definition question from the grade 9 syllabus
3. Consider a closed circuit as shown below How will the voltmeter (V), ammeter (A) reading change if bulb of resistance R is burnt out during an experiment?	An electricity question from the grade 10 syllabus



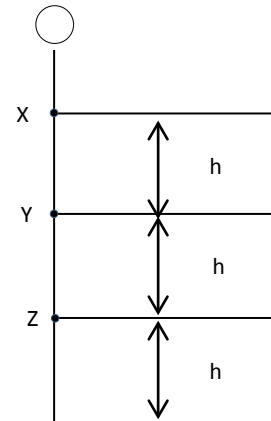
Options	Ammeter reading	Voltmeter reading
A	increases	Increases
B	Becomes 0	Becomes 0
C	Becomes 0	Does not change
D	Decreases	Becomes 0
E	Does not change	Decreases

4. A person stands on a bathroom scale which is placed inside a lift. When the lift is stationary, the reading on the scale is X newton's. What will be the reading on the scale if the lift is moving at constant velocity downwards?

A grade 11 question on Mechanics

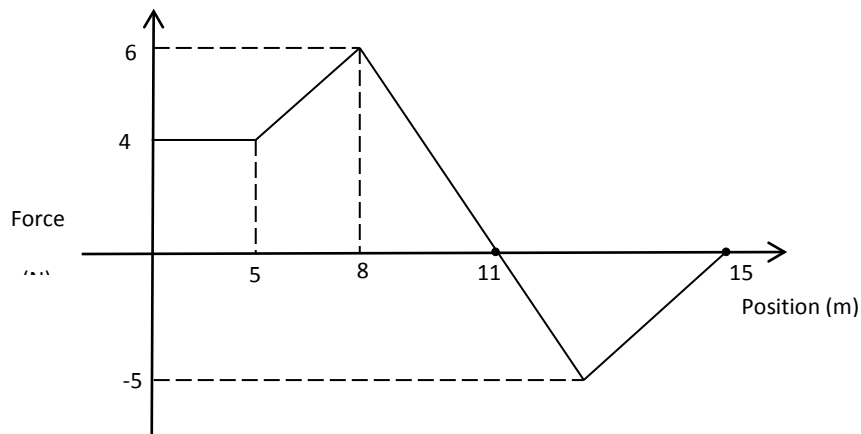
5. An object is in free fall. The picture below shows its motion. Find the ratio of the kinetic energy at point Z to the total mechanical energy of the object

A grade 11 question on Energy



6. A graph representing applied force versus position of an object is given below. Find the kinetic energy gained by the object as it travels a distance of 15 m.

A grade 12 question on Work and Energy



7. Lithium, beryllium and mercury have work functions of 3.68×10^{-19} J, 3.90×10^{-19} J and 4.5×10^{-19} J, respectively. Light with a wavelength of 400 nm is incident on each of these metals. Which one of these metals will exhibit photoelectric effect?	Grade 12 question on Photoelectric Effect
--	---

3. Findings

Table 2 below provides student performance across the seven questions which formed an integral part of the questionnaire administered.

Table 2: Summary of performance

QUESTION NUMBER	GRADE	CORRECT OPTIONS
1	8	35%
2	9	26%
3	10	15%
4	11	45%
5	11/12	21%
6	11/12	11%
7	11/12	16%

Question 1 involved the calculation of volume as one of the essential concepts in science. Students' performance on Question 1 (35%) suggests lack of understanding of basic concepts in Physics. Question 2 primarily focused on the recall of basic laws, principles and definitions taught in the earlier grades. Students' performance on Question 2 presents a grim reality in terms of their understanding of electrical circuits. In fact, it has been observed that learners have consistently demonstrated inadequate performance on questions involving electrical circuits in the Senior Certificate Examination [4]. Considerable amount of time appears to be devoted to the teaching of mechanics at school as the performance on Question 4 seems to suggest. Question 5 involved the determination of work done based on a graphical representation. Graphical interpretation further appeared to be a critical challenge for the students as the performance on Question 5 (21%) indicates. Question 6 involved the application of the principle of conservation of energy and the performance in this regard (11%) leaves much to be desired. As the performance on Question 7 (16%) illustrates, students appeared to confuse work function of a metal and the associated threshold frequency. In addition, this performance points to insufficient meaningful coverage of the photoelectric effect as a crucially important concept in Physics. A summary of the performance of top students is given in Table 3 below.

Table 3: Summary of the performance of the top students

QUESTION NUMBER	GRADE LEVEL OF THE QUESTION	CORRECT OPTIONS SELECTED (%)
1	8	55%
2	9	27%
3	10	27%

4	11	27%
5	11/12	36%
6	11/12	18%
7	11/12	18%

Some of the items in the questionnaire probed the availability of laboratory facilities and the concomitant provision of adequate practical exposure at schools. A vast majority of the students indicated that laboratory facilities were nonexistent at their schools resulting in limited opportunities for practical exposure. In essence, lack of adequate practical exposure and inadequately developed conceptual understanding appear to be the root causes of pronounced knowledge gaps and conceptual inadequacies within the context of this inquiry.

4. Recommendations

South Africa is duty bound to provide a globally competitive curriculum that responds to the economic and societal needs of the country. In terms of curriculum reform, there is a critical need for Physics and Chemistry to be treated as two separate entities in order to provide meaningful opportunities to assess learner performance in these content domains. Particular emphasis has to be put on the development of pedagogically valid instructional approaches that serve to enhance meaningful conceptual understanding coupled with a coherent infusion of appropriate practical activities. However, these ideals cannot be brought to full fruition without the provision of concrete and relevant teacher professional development opportunities within firmly established communities of practice.

5. Conclusion

Inadequate mathematical skills can be a major impediment to meaningful learning of science at institutions of higher learning. Within the context of this inquiry, students demonstrated knowledge gaps and conceptual inadequacies which may potentially serve as conceptual hurdles to be overcome in order to make appropriate progress in their studies. Institutions of higher learning are faced with an added imperative to put systems in place that adequately respond to the critical needs of students in order to give practical expression to human capital development priorities.

References

- [1]. South African Institute of Physics and Council on Higher Education. (2013). Report on the Review of Undergraduate Physics Education in Public Higher Education Institutions. Pretoria.
- [2]. Trends in International Mathematics and Science Study (TIMSS). (2011). International Results in Mathematics, PIRLS International Study Centre, Boston College, USA.
- [3]. Work Economic Forum – Global Competitiveness Report 2012-2013
- [4]. Work Economic Forum – Global Competitiveness Report 2013-2014
- [5]. Department of Basic Education. (2013). National Senior Certificate Examination- National Diagnostic Report. Pretoria.

A triggering strategy for improved pass rate in software-managed evaluations of Physics practicals for the Engineering Programmes at the University of Johannesburg

Leelakrishna Reddy, Padmanabhan Nair, Sam Ramaila and Jan Oelofse

Department of Applied Physics and Engineering Mathematics, University of Johannesburg,
Doornfontein Campus, P.O. Box 17011, Doornfontein, 2028

E-mail: samr@uj.ac.za

Abstract. The non-traditional approach to the evaluation of Physics practicals through an innovative software embedded system, first of its kind in South Africa, allows students to execute and report the results of an experiment independently. This system evaluates the individual performance in an experiment in terms of accuracy, analysis of data and report of the results obtained. As part of the evaluation of the experimental report, the student's data is loaded into a software system and checked against the pre-loaded data for a particular experimental work station. Thus, students are unaware of the exact requirements for securing marks at each step and process of the report. To overcome this difficulty, the experimental group of students were given support through rigorous tutorials and consultations in contrast with the control group. Results reveal that the support system appears to have huge potential in terms of the improvement of student performance with regard to practical work. The observed improvement of student performance is viewed in a positive light as a potentially valuable outcome that can meaningfully inform pedagogic innovation in recognition of the fact that Physics practicals carry a full modular credit in the diploma programmes offered by the Engineering Faculty at the Doornfontein Campus of the University of Johannesburg.

Keywords: experimental skills, strategy, practicals, engagement, mentorship

1. Introduction

It has repeatedly been observed that the first-time entering students enrolled for Diploma programmes at the University of Johannesburg lack the basic experimental skills needed in Physics such as observations and manipulation of equipment as well as communication skills such as reporting of observations and results [1]. This skills deficiency is most prevalent in many students coming from under-resourced public schools in South Africa. These students come to tertiary institutions with an inherent fear for practical work. The Department of Applied Physics and Engineering Mathematics (APEM) of the University of Johannesburg responds innovatively to the critical challenges associated with lack of practical skills. The students doing Physics as a service module in the Engineering Faculty are introduced to an array of uniquely designed student friendly and interactive fundamental practicals in Physics.

Contrary to the traditional assessment approach to practical work, APEM developed a software embedded rubric marking system that is linked to a database of 350 practicals housed in 7 laboratories. This approach seeks to ensure that students perform meaningful practical work through the provision of opportunities for meaningful interaction with the materials, data collection and analysis as well as making sense of the natural world [2]. Such an innovative assessment marking system has many parameters. This article reflects on a triggering strategy taking the form of support and controlled mentorship through engagement. It is envisaged that this strategy would serve to demystify the fear associated with practical work and enhance laboratory skills.

2. The purpose of practical work in science

Various arguments have been advanced in an attempt to make sense of the aims of practical work in science, namely, cognitive arguments, affective arguments and skills arguments [3]. Cognitive arguments advocate that practical work can improve understanding of science and promote conceptual development through visualisation of laws and theories of science thereby providing opportunities for illustrating, verifying or affirming theory [3]. Affective arguments advocate that practical work provides motivation and excitement as well as the generation of interest and enthusiasm [3]. Skills arguments advocate that practical work promotes development of higher-level transferable skills such as observation, measurement, prediction and inference [3].

At another level, several counter arguments to all these claims for practical work have been advanced. These arguments are premised on the notions that doing science and understanding science theories are two different entities [4], and that there is evidence that many students are not very positive about doing experiments [5] as well as the existence of the evidence that the transferability of skills is limited [6]. Clearly, the discourse pertaining to the aims and purpose of practical work is a contested terrain in many ways. In fact, the plurality of espoused aims for practical work in science make the task of assessment increasingly difficult [7].

3. The place of practical work within the science curriculum

It has been argued that the centrality of the laboratory to the teaching of science has become like the addicts' relationship to their drug – an unquestioned dependency which needs to be re-examined and weakened if not broken altogether [8]. It has also been noted that despite a shift of emphasis towards learning outcomes, the evidence suggests that there is a chasm between what teachers identify as their outcomes before lessons and the outcomes that their students perceive [9]. A major hurdle that has been identified earlier is that despite the aim of curriculum reform at improving the quality of practical work, students spend too much time following 'recipes' and, consequently, practising lower level skills [10] leading to failure to perceive the conceptual and procedural understandings as intended goals for the laboratory activities [11].

4. The role of information technology in supporting teaching and learning in practical work

Significant technological innovations have offered new resources for teaching and learning, but insufficient attention has been directed to examine critically how these new technologies can enhance or confound experiences in the laboratory [11]. In fact, it has been claimed that rapid advances in technology offer a wide range of new opportunities for innovative science education [12] which include the use of sensors, simulations and the internet [13]. Vitals tools such as computer-based simulations may also help to reduce the noise associated with the laboratory bench and focus attention on important aspects of experimental planning and data interpretation [14]. In addition, computers and their peripherals can be used to aid long-term investigations, for example, in data-logging experiments [15] and can also be used in visualizing data as well as modelling scientific phenomena [16]. Given the potential of the role of information technology in supporting practical work, a software-embedded system to evaluate physics practicals was developed at the University of Johannesburg.

5. Methodology

Students enrolled for Electrical Engineering Diploma were participants in the implementation of the engaged triggering strategy mechanism because of their comparatively lower achievement levels in Grade 12 Physical Science. A total of 178 students were randomly divided among 4 staff members resulting in two groups per staff member. These groups of students were accordingly allocated labels A1, A2, B1, B2, C1, C2, D1 and D2 and each group performed a total of 17 practicals during the

semester. The groups of students with labels D1 and D2 were regarded as the experimental group due to the specific implementation of the engaged monitoring strategy and the other groups with labels A1 to C2 were called the control groups.

Each student is required to purchase a Laboratory Manual and a Laboratory Result Book [17] as key instructional resources. The Laboratory Manual contains theories, methods and procedures while the Result Book provides templates for data collection, calculation of unknown parameters, plotting of graphs and submission of the final report. At the beginning of each practical session, students are briefed on the practicals to be performed before being afforded the opportunity to perform the practicals individually. A typical laboratory cubicle and excerpt from the Result Book are shown Figures 1 and 2 below. On completion of the experiment, students are afforded an opportunity to discuss related laws and theories coupled with a comprehensive explanation of the pitfalls and expectations of the excel rubric marking programme (ERMP).



Figure 1: Laboratory cubicle

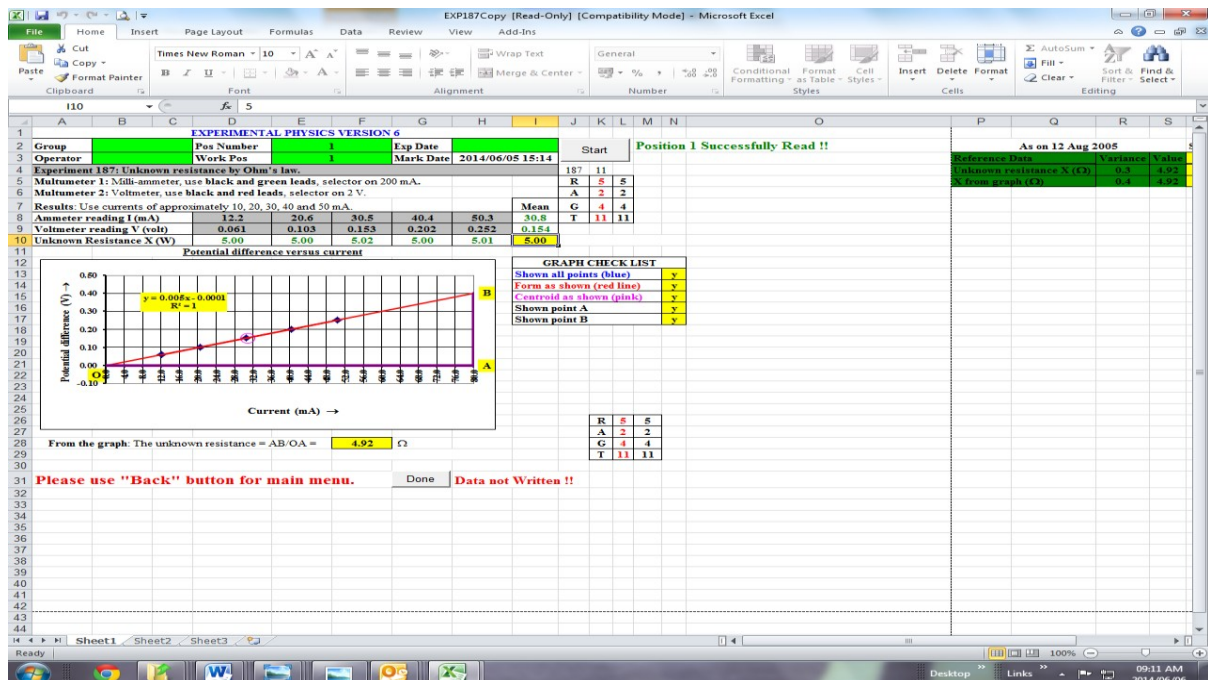


Figure 2: Excerpt from the Result Book

A set-up for a typical electricity experiment (EXPERIMENT 187) is depicted in Figure 3 below.

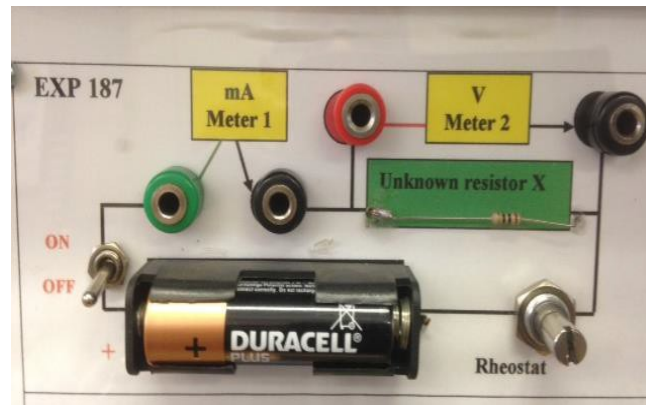


Figure 3: Illustration of Experiment 187

In addition, the students were given few days to perform necessary calculations and plot graphs as part of laboratory report before presenting it to the lecturer for critical scrutiny before final submission. This takes the form of a compulsory consultation. The rules are gradually relaxed as the students become familiar with the pitfalls associated with the software embedded assessment system. Suffice to indicate that students who fail to present their work for critical scrutiny incur penalties.

6. Discussion

Analysis of the marks of the six experiments numbered EXP 187-190 and 193-194 from the various groups A1 to D2 is shown in Table1 below.

Table 1: Mark analysis for the six experiments performed

	Group allocation						
Experiment number		187	188	189	190	193	194
Control group	A1	65%	84%	81%	40%	73%	33%
	A2	72%	85%	79%	61%	81%	74%
	B1	76%	83%	78%	51%	83%	54%
	B2	Did not do these experiments					
	C1	79%	85%	78%	67%	88%	59%
	C2	79%	85%	78%	67%	88%	59%
Average		75%	85%	80%	57%	80%	58%
Experimental group	D1	85%	88%	83%	66%	75%	72%
	D2	79%	85%	78%	67%	88%	59%
Average		86%	94%	89%	76%	86%	74%
Shift in % mark		+11%	+9%	+9%	+19%	+6%	+16%

The experimental group demonstrated improved performance as compared to the control group. Percentage shifts suggest benefits that accrue from the engaged strategic intervention programme. A comparison of the semester overall performance is shown in Table 2 below.

Table 2: Comparison of the semester overall performance

	Control group						Experimental group	
Groups	A1	A2	B1	B2	C1	C2	D1	D2
Overall percentage	66%	70%	68%	67%	65%	65%	76%	79%
							Average = D1 + D2 = 78%	
Shift in % marks	-12%	-8%	-10%	-11%	-13%	-13%		

The improved overall performance demonstrated by the experimental group points to the potential efficacy of the intervention. It has to be accentuated that the majority of students in South Africa have limited practical exposure and this is commensurate with South Africa's global ranking in terms of the quality of mathematics and science education as well as the quality of the education system [18].

Appropriate and concrete steps geared towards the development and improvement of instruction with the potential to enhance pedagogic innovation within a broader South African context are required in the light of the key findings in this inquiry. The prevailing articulation gap which appears to be a result of the conflation of various factors ought to be adequately addressed in order to significantly improve the overall quality of the education system within the South African context and the development of innovative strategies that can serve to enhance practical skills is no exception.

7. Conclusion

The heterogeneous sample of 178 students comprising experimental and control groups is a reasonably good representative sample of students performing practicals in the Physics laboratories as the highest number of students come from the Engineering Faculty. The engaged mentoring system appears to hold promise in terms of the development of a repertoire of essential practical skills and the reinforcement of positive attitudes. This intervention may potentially serve as a sustainable and viable mechanism through which the development of practical skills can be meaningfully fostered.

8. References.

- [1] www.examfear.com/files/00/45/2010-12-09-20-08-05.pdf, retrieved 30/5/2014.
- [2] Lunetta, V. N., Hofstein, A., & Clough, M. P. (2007). Learning and teaching in the school science laboratory: An analysis of research, theory, and practice. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of Research on Science Education* (pp. 393-441). NJ: Lawrence Erlbaum Associates.
- [3] Wellington, J. (Ed.) (1998). *Practical Work in School Science. Which Way Now?* London: Routledge.
- [4] Leach, J. and Scott, P. (1995). The demands of learning science concepts – Issues of theory and practice, *School Science Review*, 76(277), 47–51.

- [5] Murphy, P. and Qualter, A. (1990). Constructing gender - Fair science activities and assessments. Contributions- GASAT, Sweden.
- [6] Lave, J. (1988). *Cognition in Practice*. New York: Cambridge University Press.
- [7] Bennett, J. and Kennedy, D. (2001). Practical work at the upper high school level: The evaluation of a new model of assessment, *International Journal of Science Education*, 23, 97–110.
- [8] Osborne, J. (1998). Science education without a laboratory? In, J. Wellington (Ed.), *Practical Work in School Science Which Way Now?* (156–173). London: Routledge.
- [9] Hodson, D. (2001). *Research on practical work in school and universities: In pursuit of better questions and better methods*. Proceedings of the 6th European conference on research in chemical education. University of Aveiro; Aveiro, Portugal.
- [10] Tamir, P. and Lunetta, V. N. (1981). Inquiry related tasks in high school science laboratory handbooks, *Science Education*, 65, 477–484.
- [11] Lunetta, V. N., Hofstein, A. and Clough, M. P. (2007). Teaching and learning in the school science laboratory. An analysis of research, theory, and practice. In, S. K. Abell and N. G. Lederman (Eds), *Handbook of Research on Science Education* (pp. 393–431). Mahwah, NJ: Lawrence Erlbaum Associates.
- [12] Lunetta, V. N. (1998). The school science laboratory: Historical perspectives and centers for contemporary teaching. In, P. Fensham (Ed.). *Developments and Dilemmas in Science Education* (pp. 169–188). London: Falmer Press.
- [13] Braund, M. and Reiss, M. (2006). Towards a more authentic science curriculum: The contribution of out-of-school learning, *International Journal of Science Education*, 28, 1373–1388.
- [14] Millar, R. (1999). *Understanding how to deal with experimental uncertainty: A ‘missing link’ in our model of scientific reasoning?* Paper presented at the conference of the European Science Education Research Association (ESERA), Kiel, Germany, August 31 – September 4.
- [15] Dori, Y. J., Sasson, I., Kaberman, Z. and Herscovitz, O. (2004). Integrating case-based computerized laboratories into high school chemistry, *The Chemical Educator*, 9, 1–5.
- [16] Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B. L., Steinmuller, F. and Leone, A. J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in Biology classrooms. In, S. M. Carver and D. Klahr (Eds) *Cognition and Instruction: Twenty-five Years of Progress*. Mahwah: Erlbaum.
- [17] Kotze, A. (2005). *Experimental Physics - Method and Theory*.
- [18] World Economic Forum – Global Competitiveness Report 2013 - 2014.

Division F – Applied Physics

A single DFB laser with multilevel directly modulated signal for high speed optical fibre communication system

**T V Chabata, E K Rotich Kipnoo, R R G Gamatham, A W R Leitch,
and T B Gibbon**

Physics Department, Nelson Mandela Metropolitan University, P. O. Box 77000, Port Elizabeth 6031, South Africa. Tel: +27 41-504-2141; Fax: +27 41-504-2573.
Email: Tichakunda.Chabata@nmmu.ac.za

Abstract. A multilevel modulation format that meets the high bandwidth requirements and increases the bit rate is experimentally investigated. A single distributed feedback (DFB) laser at 1550 nm is used to transmit a total of 20 Gb/s of data over a single mode optical fibre link. The data rate is doubled from binary 10 Gb/s by employing a multilevel technique. An economical, all electrical multilevel signal generation technique was designed. The multilevel format transmitter codes two bits in one symbol. This enables an increase in the bit rate of the system, though at the expense of receiver complexity. The two bits coded in a single symbol generate a four level signal that has to be decoded at the receiver. The designed complex receiver section utilizes the offline digital signal processing algorithms to evaluate the system performance through bit error rate (BER) measurements. The system performance was experimentally evaluated on back-to-back transmission.

1. Introduction

The current increase in the use of cloud based services, high definition (HD) TV streaming, on-line gaming and sharing of photos, videos and other files has greatly congested the already operational 10 Gb/s intensity modulation direct detection (IM/DD) optical networks. There is therefore a need to develop new technologies which can potentially provide the required increase in data capacity while maintaining the bandwidth and the form of existing links. A cost effective alternative is one based solely on optical amplitude modulation format allowing the use of existing receivers [1]. A technique to achieving high spectral efficiency in IM/DD optical links is to increase the number of amplitude levels in the signal. This modulation format is commonly referred to as M-level pulse amplitude modulation (M-PAM), where M refers to the number of levels in the signal. Each optical symbol will carry $\log_2(M)$ bits of information. Recently a 4-PAM modulation of a vertical cavity surface emitting laser (VCSEL) at 50 Gbps was reported [1].

In this paper, a multilevel signal transmitter and analyzer at the receiver is demonstrated. The advanced modulation technique employing 4 -pulse amplitude modulation (4-PAM) that doubles the transmission rate is designed and implemented using a directly modulated distributed feedback (DFB) laser. The narrow linewidth of the DFB laser is of importance in the design of the high speed optical link. The magnitude of temporal dispersion of the signal is directly proportional to the length of the fibre and the linewidth of the laser source. The generated multilevel signal is transmitted using a single DFB laser at 1550 nm at an effective data rate of 20 Gb/s. An offline digital signal processing (DSP), multilevel signal analyzer was developed to evaluate the quality of the received optical signal using the Q factor approach. The ultimate quality measure in an optical communication links is the bit error rate (BER) [2]. The quality factor and probability density function are the important parameters in analyzing and estimating the BER from an eye diagram plot.

2. Theory

Figure 1 shows the probability density function (pdf) of a received binary signal with symbol S_0 and symbol S_1 [3]. In the derivation, we assume that the pdf is Gaussian and is given by:-

$$P_n(x) = \frac{1}{\sigma_n \sqrt{2\pi}} \exp \left[-\frac{x^2}{2\sigma_n^2} \right] \quad [1]$$

The erroneous received bit is the overlapping region between pdf of symbol S_0 and symbol S_1 . The probability of error when S_0 is transmitted is the shaded area under the curve (blue) and is given by:-

$$P\left(\frac{e}{s_0}\right) = \frac{1}{\sqrt{2\pi}\sigma_0} \int_{Th_1}^{\infty} e^{-\left(\frac{Th_1 - U_0}{\sqrt{2}\sigma_0}\right)^2} = \frac{1}{2} \operatorname{erfc} \left[\frac{Th_1 - U_0}{\sqrt{2}\sigma_0} \right] \quad [2]$$

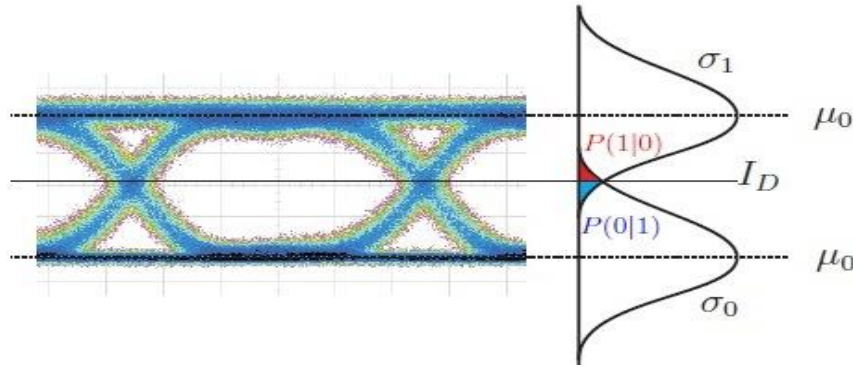


Figure 1: The Probability density functions of a received binary signal [3].

For similar threshold value, the probability of error when S_1 is transmitted is also the area under the curve (red) and is given by:-

$$P\left(\frac{e}{s_1}\right) = \frac{1}{\sqrt{2\pi}\sigma_1} \int_{-\infty}^{Th_1} e^{-\left(\frac{U_1 - Th_1}{\sqrt{2}\sigma_1}\right)^2} = \frac{1}{2} \operatorname{erfc} \left[\frac{U_1 - Th_1}{\sqrt{2}\sigma_1} \right] \quad [3]$$

where, $P(e/s_0)$ and $P(e/s_1)$ are contributing to total probability of error. The probability of error for binary (2 level signal, with unbalanced noise distribution, ($u_1 \neq u_0$)) is given by:

$$BER = P_e = P_{s_0} P\left(\frac{e}{s_0}\right) + P_{s_1} P\left(\frac{e}{s_1}\right) = \frac{1}{2} \operatorname{erfc} \left[\frac{Q_0}{\sqrt{2}} \right] + \frac{1}{2} \operatorname{erfc} \left[\frac{Q_1}{\sqrt{2}} \right] = \frac{1}{2} \operatorname{erfc} \left[\frac{u_1 - u_0}{\sqrt{2}(\sigma_1 + \sigma_0)} \right] \quad [4]$$

The Q- factor term is given by:-

$$Q = \frac{u_1 - u_0}{\sigma_1 + \sigma_0} \quad [4.1]$$

The same reasoning can be extended to any M-ary modulation format to estimating the BER.

For a 4 level modulation format, three optimum threshold decision values are applicable, each of the threshold values being between a pair of two adjacent symbols levels. The total error rate is driven from error rate measurements done on each of the threshold values [4].

$$BER = ER_1 + ER_2 + ER_3 = \frac{1}{2} \operatorname{erfc} \left[\frac{u_1 - u_0}{\sqrt{2}(\sigma_1 + \sigma_0)} \right] + \frac{1}{2} \operatorname{erfc} \left[\frac{u_2 - u_1}{\sqrt{2}(\sigma_1 + \sigma_2)} \right] + \frac{1}{2} \operatorname{erfc} \left[\frac{u_3 - u_2}{\sqrt{2}(\sigma_2 + \sigma_3)} \right] \quad [5]$$

3. Research design

The schematic diagram shown in Figure 2 is a complete multilevel optical transmission link comprising of an electrical transmitter, the medium and the optical receiver. The complimenting, P and N, arms of the programmable pattern generator (PPG), generates the pseudo random bit sequence (PRBS), (2^7-1) , which on combining directly modulates the (DFB) laser. The different electrical attenuations on the two outputs guarantee the existence of two data signals at different power levels. The N output attenuated at 13 dBm has the electrical data sequence delayed in time by an integer multiple of a bit period, decorrelating the two complemented bit sequences. After mixing the two decorrelated data sequences interfere constructively and destructively, producing a four level data sequence. The mixing doubles the transmission data rate from 10 Gb/s to 20 Gb/s. The multilevel transmitter therefore codes two bits in one symbol prior to modulation and transmission through the fibre.

The DFB laser signal at 1550 nm is directly modulated at a bit rates of 20 Gb/s by a Non-Return-to- Zero (NRZ) Pseudo-Random Binary Sequence (PRBS) signal and is transmitted through a back to back single mode fibre. The power meter (PM) measures the optical power at the back-end of the receiver after the optical signal power has been attenuated by a variable optical attenuator (VOA). The direct intensity modulated signal is directly detected (DD) by the photodiode (PD). The photodiode (PD) demodulates and converts the optical signal back to an electrical data signal. The front-end of the receiver comprises of an oscilloscope that captures and stores the received signal. The captured data is transferred to an offline digital signal processing (DSP) circuit. An Offline digital signal processing algorithm was developed to monitor and evaluate the performance of the designed multilevel optical link.

The developed DSP analyser utilises the loaded eye diagram information from the oscilloscope as the input data. Three optimum threshold values are accurately determined thereby demodulating the four signal symbols generated at the transmitter. Mean peak to peak voltage for zero and one level of each of the three eyes is calculated together with their respective standard deviations. The Q- factor is calculated from the mean peak to peak voltages and the standard deviation as given in equation [4.1]. The final and most important stage is the determination of the BER values using the already calculated Q- factor, utilizing equation [5].

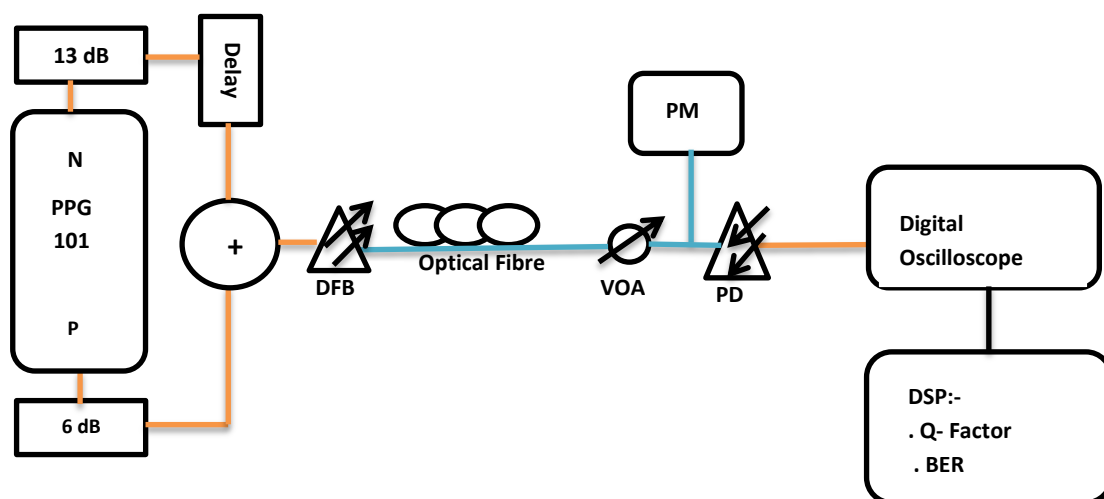


Figure 2: Experimental setup for a 4- Level pattern generator with differentially attenuated outputs and an integer bit period delay:

PPG: programmable pattern generator. DFB: distributed feedback laser. VOA: variable optical attenuator power PM: power meter. PD: photodiode and a DSP: digital signal processing circuit. BER: bit error rate, Quality factor technique.

4. Results and discussion

Two decorrelated P and N pseudo random bit sequences ($2^7 - 1$) each at 10 Gb/s are differentially attenuated before combining and doubling the transmission data rate to 20 Gb/s. Figure 3 shows the individual P and N signal before mixing and the combined signals after mixing. After mixing each of the four symbol comprises of two bits each (00; 01; 10 and 11). The graph clearly shows that the combined signal (red) is a result of the interference of the two individual signals (blue and black).

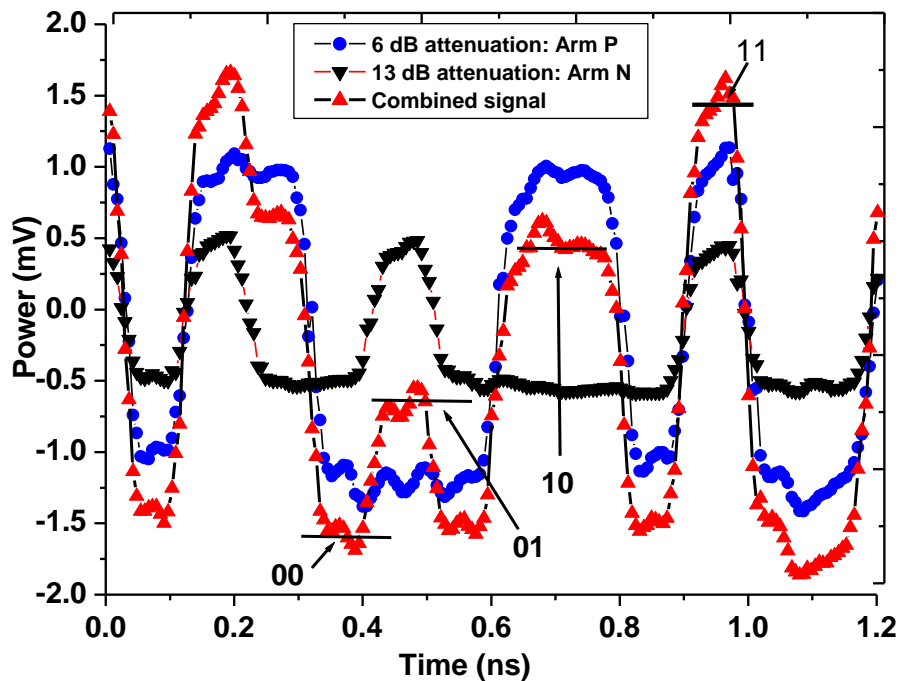


Figure 3: The individual and combined waveforms of the signal out of the multilevel transmitter.

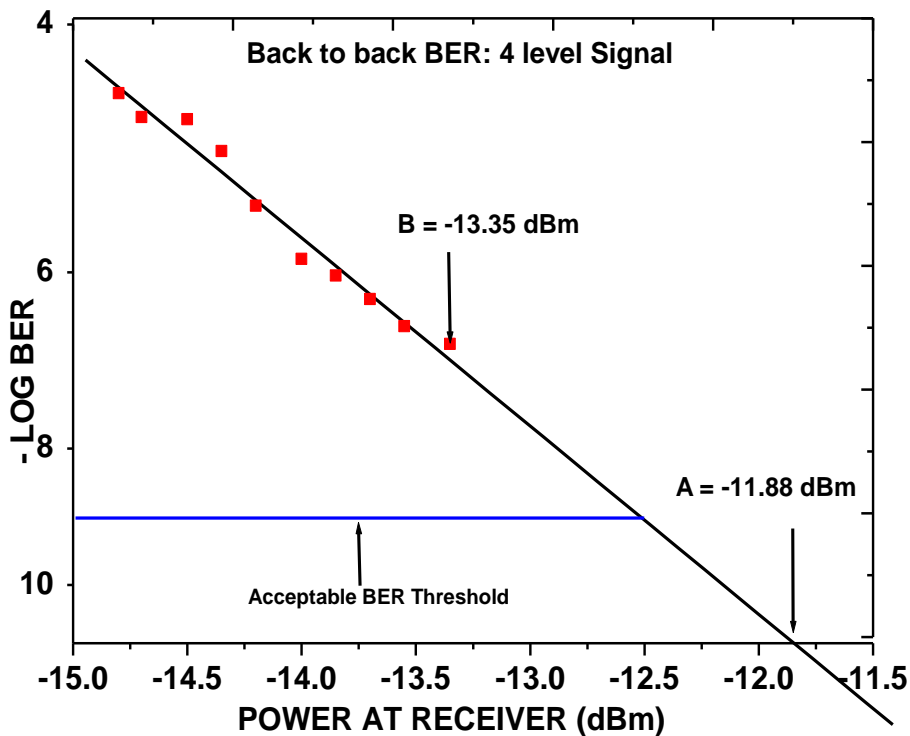


Figure 4: Back to back bit error rate (BER) for the four level signal

Figure 4 shows the BER curve for back to back transmission obtained using the statistical probability density function (pdf), given in equation (5). The equation uses the analogue information from the signal (Q factor) to accurately estimate the BER. A receiver sensitivity of -12.47 dBm at the acceptable BER of 10^{-9} is obtained, with a penalty of 7 dB compared to a two level signal reported in [6]. The penalty is brought about by modulating an additional bit within the same modulation depth prompting the need for a more sensitive receiver to demodulate the two bits per symbol.

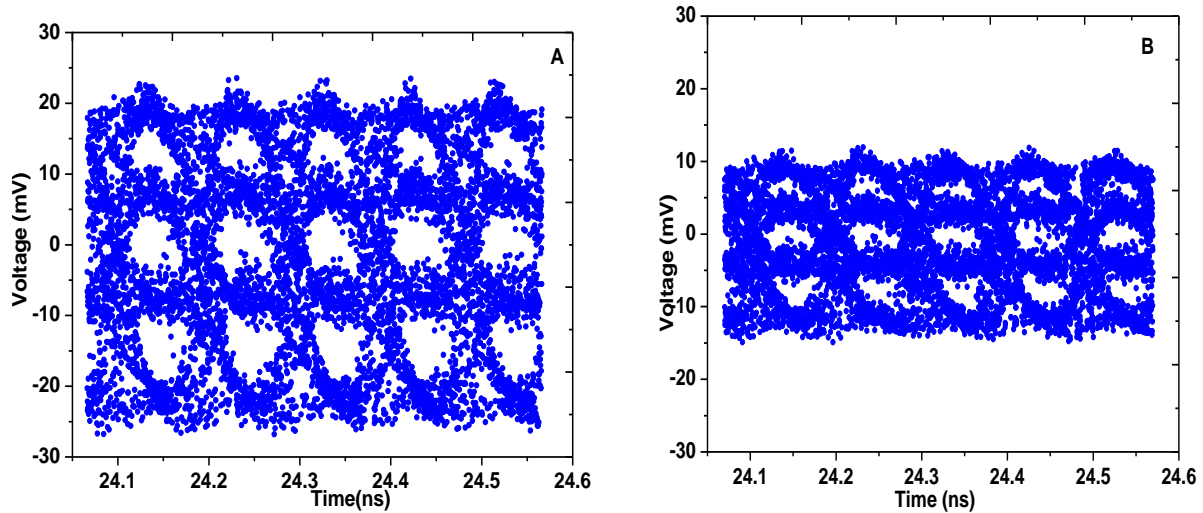


Figure 5 (A and B): Eye diagrams at -11.88 dBm and – 13.35 dBm respectively.

Figure 5 shows the open eye diagrams at different receiver powers. The diagram has three open eyes and four signal levels. The eye closure is attributed to the attenuation of the signal in the back to back single mode fibre transmission link. The DSP algorithm developed, using the pdf and Q factor is an attractive technology to evaluate and monitor the performance of a high order modulation scheme in the absence of the relevant hardware.

5. Conclusion

In this paper a four level signal analyser has been successfully implemented. The multilevel transmitter and the offline signal analyser designed and experimentally implemented are technologies that can help alleviate the ever growing demand for transmission speed in many optical interconnects and local area networks. The developed multilevel DSP algorithms signal analyser can be used in any M-ary modulation formats thereby promoting a cost effective way of evaluating the performance of any high optical communication link.

Acknowledgement

We are grateful for Research Funding from: Telkom, Dartcom, Ingoma, NRF, CSIR, THRIP, and Scholarship Funding from ALC and NMMU.

References

- [1]. J B Jensen, M I Olmeido and I T Monroy, (2013), "Modulation Formats for Beyond-100Gbps Ethernet Optical Links – A Review of Research", *DTU Fotonik, Department of Photonics Engineering, Technical University of Denmark*
- [2]. E Lach, W Idler 2011, "Modulation formats for 100G and beyond", Elsevier. *Opt. Fib. tech.* 17, pg 377-386
- [3]. W Freude, R Schmogrowl, B Nebendahl, M Winter, A. Josten, D Hillerkussl, S Koenig, J Meyer, M Dreschmann, M Huebner, C. Koos, J Becker, 2012. "Quality Metrics for Optical Signals: Eye Diagram, Q-factor, OSNR, EVM and BER", *Institute of Photonics and Quantum Germany, ICTON*,
- [4]. M F L Abdullah, R Talib .2012, "Multilevel signal analyzer tool for optical communication system". *Department of communication Engineering. University Tun Hussein Onn Malaysia. Institute of Advanced engineering and science.*
- [5]. S Krzysztof, W Pette, G Johan, H Åsa, K Johnny, K Magnus, A Peter; A Erik and L Anders, 2011, "30 Gbps 4-PAM transmission over 200m of MMF using an 850 nm VCSEL", *European Conference and Exposition on Optical Communications (ECOC), OSA.*
- [6]. T V Chabata, R R G. Gamatham , H Y S Kourouma, E K Rotich Kipnoo, A W R Leitch, and T B Gibbon (2013) , " Digital Signal Processing Algorithm for Signal Reconstruction and Performance Measurement for a Single Photo Detector Optical Communication System". *Proc. 16th annual Southern Africa Telecommunication Networks and Applications Conference (SATNAC) Spier, Stellenbosch, Western Cape, South Africa, pg 27-30, 1-4th Sept 2013.*
- [7]. T.B. Gibbon, K. Prince, T.T. Pham, A. Tatarczak, C. Neumeyr, E. Rönneberg, M. Ortsiefer, "VCSEL transmission at 10 Gb/s for 20 km single mode fibre WDM-PON without dispersion compensation or injection locking", *Elsevier. Opt. Fib. tech.* 17, pg 41-45 (2011).
- [8]. G P AGRAWAL .2007 Non Linear Fibre Optics Academic Press 4th Edition,
- [9]. HFAN -9.0.2 Application note (04/08), "Optical signal to noise ratio and the Q-Factor in Fiber Optic Communication Systems".
- [10]. R Rodes, J Estaran, B Li, M. Mueller, J B. Jensen, T Gründl, M. Ortsiefer, C Neumeyr, J Roskopf, K J Larsen, M C Amann, and I T Monroy (2012), "100 Gb/s single VCSEL data transmission link", *DTU Fotonik, Department of Photonics Engineering, Technical University of Denmark, DK2800 Kgs. Lyngby, Denmark* , OSA

Overview of the Mineral-PET run-of-mine Diamond bearing rock sorter

M Cook¹, M Tchonang¹, E Chinaka², M Bhamjee¹, F Bornman³, SH Connell¹

¹ University of Johannesburg, Johannesburg, South Africa

² Necsa, South Africa

³ Multotec, Johannesburg, South Africa

E-mail: shconnell@uj.ac.za

Abstract. Mineral-PET is a technology for the sorting of diamond bearing rock (kimberlite ore) based on a mineral analogue of the well-known medical Positron Emission Tomography (PET) imaging technique. The naturally occurring carbon in kimberlite needs to be activated via photonuclear transmutation before it can be imaged. For the R&D phase of the project, a technology demonstrator has been built. This is a planar PET array built around a conveyor belt using kimberlite phantoms. The phantoms consist of blocks of cement with the radioactive material (²²Na) uniformly distributed throughout it to simulate the homogenous background radiation from various non-diamond PET emitters. Diamonds are modeled in the phantom by the inclusion of a localized hot-spot of ²²Na. This system has been used to benchmark computational simulations and to explore the physics issues for the specification of a pilot scale plant at a mine. The review will provide new results and updates on the performance and outlook for Mineral-PET.

1. Introduction

The detection and imaging of diamonds within coarse crushed kimberlite rocks can be carried out in a run-of-mine scenario. This promises large savings and reduced diamond breakage. A schematic showing the Mineral-PET process is shown in figure 1. The device is inserted into the kimberlite processing chain after the coarse crush, which breaks rocks up into roughly 10 cm pieces. The stream is then sent to an irradiation hopper where an electron beam incident on a tungsten target creates gamma rays via bremsstrahlung. The beam is fanned out over a 1 m² area via perpendicularly oriented bending magnets. The gamma rays are incident on the kimberlite, which is several layers deep to maximise usage of the gamma beam. The gamma rays convert ¹²C in diamonds and within the surrounding material to the radioactive PET isotope ¹¹C, making use of the giant dipole resonance in the ¹²C(γ, n)¹¹C cross section. The rock is then stored for about 20 minutes in a hold hopper to allow the non-carbon background signal to die away. After the decay period, the kimberlite is directed onto a conveyor belt and the rocks then pass between two horizontally oriented planar PET detector arrays. Each node of a cluster computer is assigned a volume of space between the detectors in a revolving fashion. As coincident events are detected, corresponding trajectories are reconstructed, and images gradually built up in the volumes traversed by each trajectory. When the volume reaches the end of the detection region, the relevant node then analyses the 3D PET image, and decides

whether or not a diamond is present. This information is fed to an ejection system, which separates diamond bearing rock, allowing the remainder to be discarded.

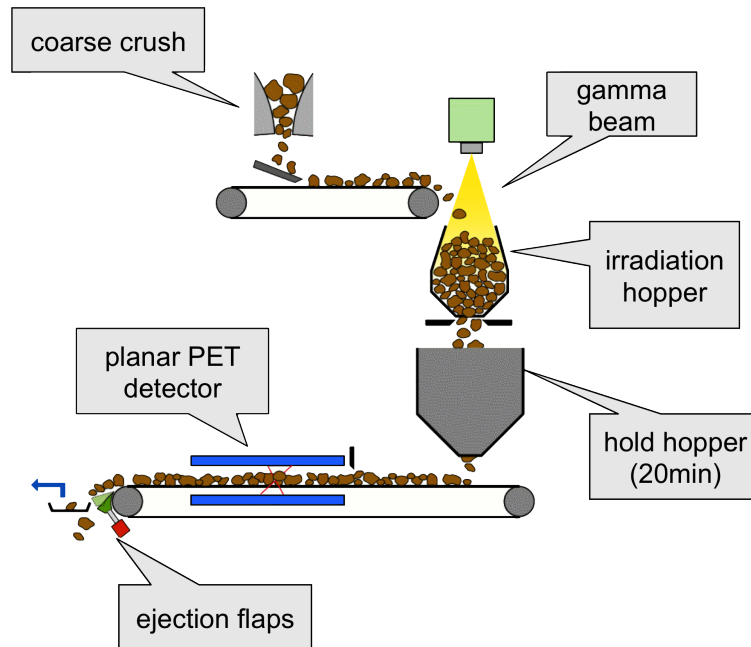


Figure 1. Run-of-mine Mineral-PET schematic [1].

Simulations and calculations have been made to optimise imaging algorithms, determine the expected activity levels and to optimise electron beam parameters and detector layout. Results show that a 5mA, 40MeV accelerator and a 2 m² bi-planar PET array could reliably detect diamonds smaller than 1 mm at a kimberlite throughput of 700 tonnes/hour, assuming 0.054% homogeneously distributed non-diamond carbon. Larger diamonds are more easily detected, with the detection probability growing faster than the diamond volume. An analysis of data taken at a 100 MeV electron accelerator shows that all non-PET activations are of a similar order of magnitude to the natural background activity. All the radiological implications reduce to below IAEA guidelines within two hours, and the material is radiologically equivalent to natural background within 2 days. A scaled down version of the Mineral-PET system has been constructed as a technology demonstrator, where the detector system has a reduced size and the activation system is represented by phantoms. It successfully identifies simulant diamonds within these phantoms and validates the simulations. The resolution is optimised by a custom genetic algorithm which adjusts various detector parameters in a hierarchical manner. Two complimentary techniques to determine radiological exposure are developed, which indicate that the exposure levels at Mineral-PET will be easily manageable. These aspects are all described in more detail in this reference [2].

2. Simulation of Activation and Minimum Detection Limits

The activation stage of the Mineral-PET has been simulated with GEANT4. One set of simulations studied the power efficiency of different electron accelerator energies, beam currents, converter materials and thicknesses. The power efficiency is defined as the number of useful gamma rays created per incoming electron, divided by the beam energy. Useful gamma rays are gamma rays likely to excite the giant dipole resonance of carbon ($^{12}\text{C}(\gamma, n)^{11}\text{C}$), taken as between 20MeV and 27MeV. The power efficiency plateaus after about 40MeV, so this is chosen

as the Mineral-PET beam energy. Gold and tungsten bremsstrahlung targets perform equally well, so a 3 mm tungsten target is chosen due to its superior strength and high melting point.

The mixed radiation field arising from both this accelerator-converter configuration and for the subsequent penetration into a total depth of kimberlite of 30cm was also modelled using GEANT4. This depth of kimberlite corresponds to about three radiation lengths, and can therefore be taken as the effective activation depth for the Mineral-PET process. The results of this calculation were used to determine the production of both the ^{11}C and ^{15}O PET isotopes. The dominant background for Mineral PET a few minutes after irradiation is the activity from ^{15}O as there are many oxides in the kimberlite ore. This component is no longer dominant after the residence period in the hold hopper. The ^{11}C activity arises from either the presence of a diamond or from the very dilute presence of homogenously distributed carbon (carbon dioxide and carbonates). The Mineral-PET detector system was also modelled by Monte Carlo to determine the overall detection efficiency. This calculation could also be used to optimise the planar detection geometry. A study was performed of overall Mineral-PET resolution for the reconstruction of hot spots of ^{11}C PET activity. This includes the detector position resolution, and the effective positron range and non-collinearity smearing. A novel hybrid diamond detection algorithm was developed which will be discussed elsewhere. Ultimately, a diamond is identified when the number of trajectories (the Line of Response (LoR) formed by the back-to-back decay photons) that intersect a region of interest (ROI) passes a given threshold. The threshold is dependent on the tolerated false positive rate, the size of the diamond to be detected, the resolution of the detection system, the level of ^{11}C activation and the amount of homogenously distributed non-diamond carbon in the kimberlite. Combining the results of an analytic formulation of the physics of the diamond detection with a Monte Carlo procedure to determine the parameters in the model, the diamond detection probability could be evaluated.

Figure 2 shows the simulated probability of detecting a diamond as a function of diamond length, for various different beam currents. Mineral-PET should have a 99.7% chance of finding a 1mm diamond with a beam current of 5mA when the non-diamond carbon concentration is below 0.2% CO_2 .

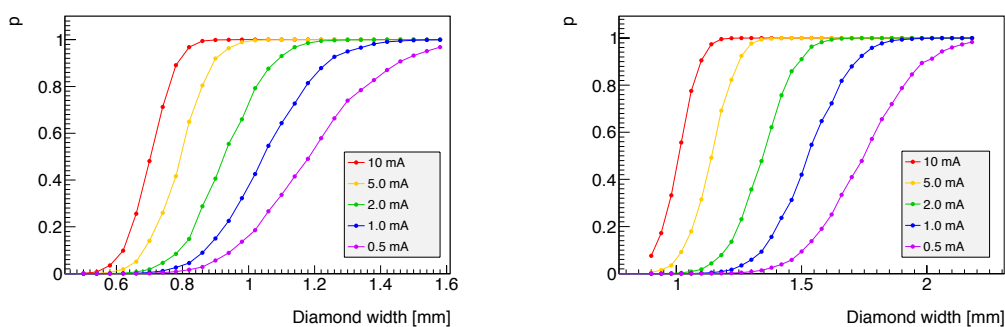


Figure 2. Probability to find diamonds in a 0.2% CO_2 (left) and a 2% CO_2 (right) background as a function of diamond size, for various beam currents.

3. Residual Activity Identification, Monte Carlo Validation

A sample of kimberlite irradiated at the Aarhus University microtron ($E_e = 100$ keV) is shown in figure 3. This is a two dimensional histogram, with elapsed time on one axis and energy on the other. The PET line at 511 keV is far larger than other sources of activity. The change of

slope of this line as the oxygen signal decays and the carbon signal starts to dominate is evident. Most other species quickly decay. The peak at just under 1500 keV is the potassium-40 peak.

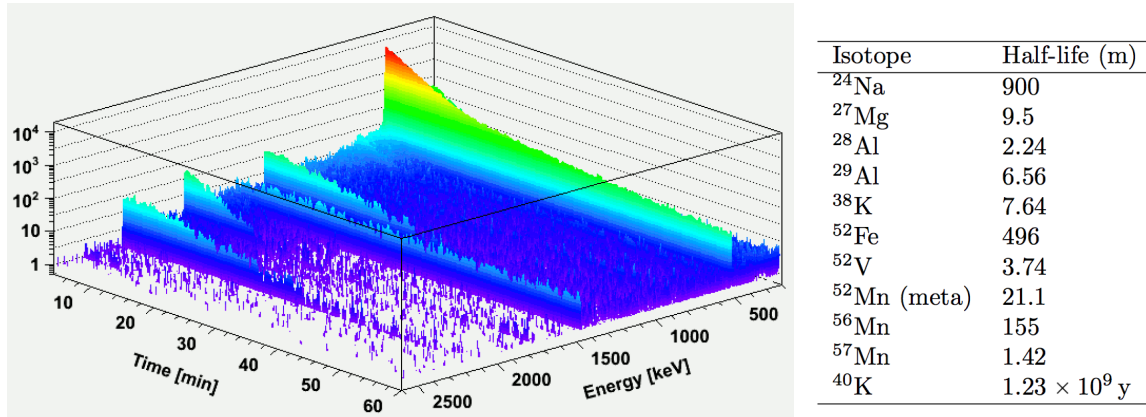


Figure 3. Left : Two dimensional energy - elapsed time spectrum of photon irradiated kimberlite showing induced PET activity and the activation of various isotopes. Right : The identification of isotopes responsible for the residual activity in using both peak energy and lifetime. [1].

The first goal was to comprehensively understand all the isotopes that are excited during irradiation. The second goal was to obtain well calibrated experimental data on the activation process for comparison to simulations including the detailed mixed radiation field at two different energy distributions. The first goal has been achieved, as the combined information of peak energy and lifetime enable a unique identification of the various isotopes which have been activated. Of these isotopes, the long-lived ones are ^{24}Na , ^{52}Fe , ^{56}Mn and ^{40}K . The potassium can be discounted, because it is present at roughly the same level in non-irradiated samples. The observed initial activities for 1 cc of kimberlite from the irradiation (about 60 times the dose envisaged for the Mineral-PET method) are: for ^{56}Mn : $1.41 \pm 0.04 \text{ Bq}$, for ^{24}Na : $2.81 \pm 0.02 \text{ Bq}$. The fact that the size of the naturally occurring potassium is comparable to that of the isotopes that have been activated is indicative of the fact that very little radiation has been induced. The ^{24}Na could be a concern, as it is easily metabolised. However, all the radiological implications reduce to below IAEA guidelines within two hours, and the material is radiologically equivalent to natural background within 2 days. The Mineral-PET method can therefore be considered radiologically safe.

With respect to the second goal, the GEANT4 simulation was used to model the mixed radiation field and from there the activation could be calculated. The simulation results were within a factor of 2 and 3 for the carbon and oxygen and 5 for the sodium with respect to the experimental results. The modelling of the residual activity is therefore partially successful, and the remaining disagreement is attributed to possible variations in the kimberlite composition, as well as uncertainties in the irradiated dose estimation.

4. Technology demonstrator

A scaled down version of the full Mineral-PET detector system built around a conveyor belt has been created as a laboratory demonstrator. This serves as a proof of concept, and allows experimentation with detector technology, detector alignment optimisation, software development etc. An accelerator has not yet been secured, so results using the laboratory demonstrator have been acquired using ^{22}Na sources, which have been incorporated into kimberlite, in such a way as to emulate the activity profile of photon irradiated kimberlite,

with and without incorporated diamonds. The technology demonstrator is able to locate a diamond simulant in the background of homogeneously distributed carbon simulant within a kimberlite phantom moving on a conveyor belt through the Mineral-PET detector array. The visualisation of the detection is evidenced by a digital presentation of the lines of response (LoR) recorded by the detectors in real time, and the painting of the location of the diamond on the kimberlite phantom by a data projector. The quantitative results are consistent with the simulations discussed above.

5. Radiological Exposure Modelling

One of the concerns for an industrial scale Mineral-PET plant is the radiation exposure of workers. The shielding requirements for the irradiation system have already been investigated in [3]. This included radiation shielding calculations, analysis, and optimisation with the aid of the MCNPX Monte Carlo code. The conclusion was that a 1.6 m thick shielding matrix of lead, iron, wax and boron carbide adequately shielded personnel from the irradiation system. If the accelerator, irradiation system and hold hopper were buried underground away from personnel, these requirements would be less stringent.

Once the kimberlite is out of the hold hopper, we would like to quantify the exposure received by Mineral-PET workers in common scenarios, for example nearby the conveyor belt or a pile of mine tailings. Two numerical techniques have been developed to calculate the exposure given a particular geometry of radioactive material. The first uses an attenuation model to arrive at an integral that can be solved numerically. This technique only applies to cuboid shaped bodies of material. The results of this are compared to a full Monte Carlo physics simulation that is more accurate, but far more computationally intensive. The second technique can, however, be applied to any geometry. The development of the radiological assessment tools have already been presented in [4].

Here we present additional results specifically for the case that will occur in a Mineral-PET plant. This is of a body of kimberlite 10 cm deep on a 1 m wide conveyor belt, after irradiation has occurred and the decay time has passed. Figure 4 shows the geometry, consisting of a 20 m long kimberlite stream, 1 m wide and 10 cm tall. The 50 cm radius water ball is shown one metre away from the kimberlite at a height of 30 cm. We will use the 0.2% carbon dioxide activity estimate, and assume that 29 minutes have passed since irradiation. At this time, the carbon and oxygen signals dominate, so we ignore the contribution from sodium. The combined specific activity of the carbon and oxygen is 33.9 Bq/cm^3 . Each decay creates two 511 keV photons, so the overall 511 keV gamma rate is 67.8 Hz/g .

Figure 4 on the right shows the activity absorbed by the target as a function of the position of the centre of the target. The conveyor belt extends from $x = -10 \text{ m}$ to $x = 10 \text{ m}$, and from $y = -0.5 \text{ m}$ to $x = 0.5 \text{ m}$. The edge of the target is 5 cm from the belt at its closest, and the centre of the target is 30 cm above the belt. This approximates a worker standing at various positions next to the conveyor belt. Both the activity in Bq/cm^3 and the gamma attenuation constant are multiplied by 0.7 to take the packing density into account.

The activity that hits the target when it is 5 cm from the middle of the belt is $2.40 \times 10^5 \text{ photons/s}$. Considering a photon energy of 511 keV and a gamma radiation quality factor of 1 [5], this corresponds to $70.8 \mu \text{ Sv/h}$. At a distance of 1 m, this has already fallen to $7.50 \times 10^4 \text{ photons/s}$ or $22.1 \mu \text{ Sv/h}$. Workers would have to spend six hours a day for 118 days at a distance of 5 cm, to reach the 50 mSv level set in the IAEA safety guide relating to radioactive material in mining [6]. Workers could be employed full time at a distance of 1 m.

Note that the kimberlite will be more active at the time of detection if the carbon dioxide concentration is 2%. In this case, worker access should be more limited, or shielding employed.

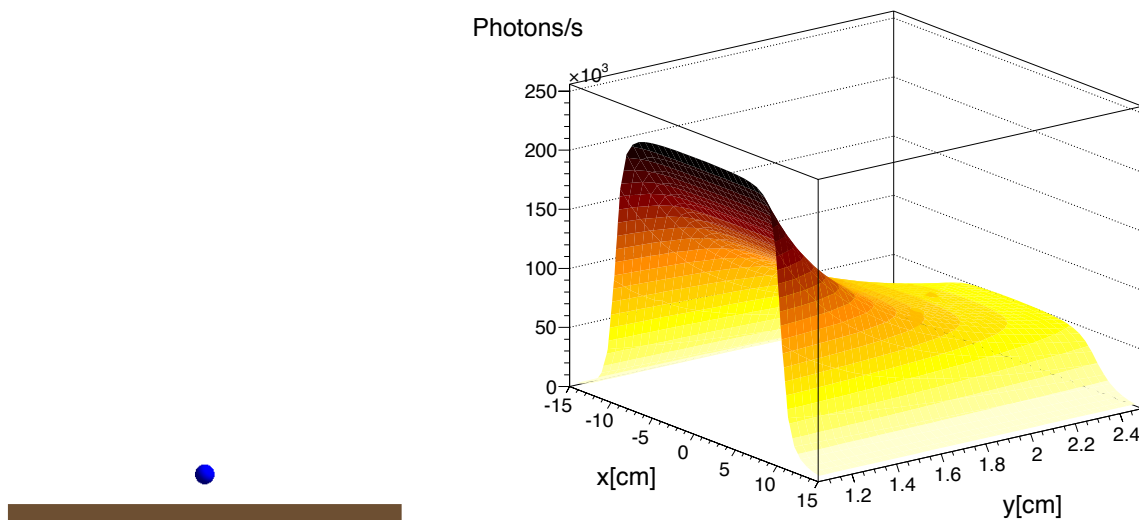


Figure 4. Left : Top and side views of the geometry in the kimberlite belt radiological exposure calculation. Right : Activity at different points (blue dot) along the side of a 20 m \times 1 m conveyor belt (brown dot) carrying kimberlite 10 cm deep.

6. Conclusion

An overview of the Mineral-PET run-of-mine diamond bearing rock sorter is presented. The paper described Monte Carlo simulation results for various aspects of the activation and detection stages. Minimum Detection Limits for the diamond size in a realistic plant were identified. The most recent results for the calculation of the residual activity of the irradiated kimberlite were presented. Finally, the previously developed radiological exposure modelling tools were applied to a Mineral-PET plant scenario.

References

- [1] Ballestrero S, Bornman F, Cafferty L, Caveney R, Connell S, Cook M, Dalton M, Gopal H, Ives N, Lee C A, Mampe W, Phoku M, Roodt A, Sibande W, Sellschop J P F, Topkin J and Unwucholaa D A 2010 *12th International Conference on Nuclear Reaction Mechanisms* ed Cerutti F and Ferrari A (Varenna, Italy) pp 589–602
- [2] Cook M 2014 *PhD Thesis : Remote Detection of Light Elements Using Positron Emission Tomography* (University of Johannesburg)
- [3] Chinaka E M, van Rooyen J, Zibi Z and Connell S H 2014 *To appear in the conference proceedings of the 2012 South African Institute of Physics Conference*
- [4] Cook M, Ballestrero S, Connell S and Tchonang M 2012 *To appear in the conference proceedings of the 2012 South African Institute of Physics Conference*
- [5] Leo W 1994 *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach* (Springer) ISBN 9783540572800 URL <http://books.google.co.za/books?id=hDEbAQAAIAAJ>
- [6] International Atomic Energy Agency 2002 *Series No. WS-G-1.2*

Interpretation of spectral electroluminescence images of photovoltaic modules using modelling

JL Crozier¹, EE van Dyk and FJ Vorster

Nelson Mandela Metropolitan University, Port Elizabeth, South Africa

E-mail: s207094248@nmmu.ac.za

Abstract. Electroluminescence (EL) is a useful solar cell and module characterisation technique as it is fast, non-destructive and sensitive to the effects of shunt and series resistance and recombination parameters. A solar cell in normal operation receives an optical input in the form of the incoming light and outputs an electrical current. Conversely, a LED receives an electrical input resulting in an optical output. EL occurs when a solar cell is forward biased receiving an electrical input and outputs an emission spectrum. The emission spectra from indirect band gap materials like silicon are in the infrared region of light while the direct band gap semiconductor materials used in LEDs emit light in the visible region. The intensity of the luminescence emitted is related to the applied voltage and the quantum efficiency of the cell material. The spectrum of the emitted luminescence can be related to the recombination properties of the cell such as surface recombination velocity and minority carrier diffusion length or lifetime. For EL imaging, a silicon CCD camera is commonly used because it has very good spatial resolution, however, this sensor is only sensitive to wavelength in the range of 300-1200 nm. There is an overlap in wavelengths from about 900 to 1100 nm allowing the EL to be detected. The spectrum of the detected EL is thus dependant on the sensitivity of the camera, the transmission of the filters and the emitted photon flux. In this study the emitted spectrum is modelled and the effects of recombination properties of the cell on the emitted spectrum are examined. The integration of short-pass filters into the experimental setup is theoretically modelled and discussed.

1. Introduction

1.1. Introduction

A solar cell and a light emitting diode (LED) can be represented as a device with electrical and optical terminals as shown in figure 1. A solar cell receives an optical input in the form of the incoming light and outputs an electrical current. Conversely a LED receives an electrical input resulting in an optical output. EL occurs in a solar cell when it is forward biased, receiving an electrical input and outputs an emission spectrum. The emission spectrum from silicon is in the infrared region of the electromagnetic spectrum[1].

A solar cell can be visualised as a spatially extended device where light is emitted at different points on the surface, each with their own local quantum efficiency and local junction voltage.

¹ To whom any correspondence should be addressed.

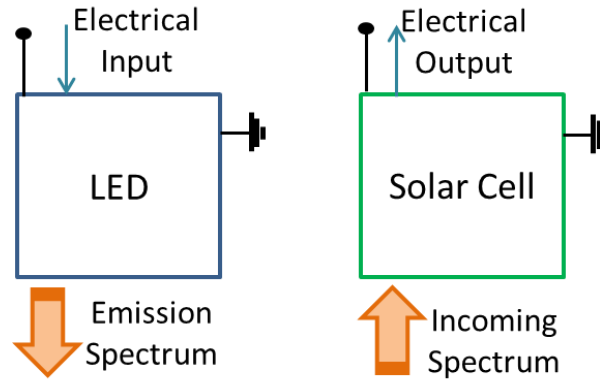


Figure 1: Solar cell represented as a device with optical and electrical terminals[2].

A detailed balance theory relating the carrier collections properties of a solar cell to the spectral EL emission is rigorously proven by U. Rau (2007)[3].

The intensity of the emitted EL at a point (x) on the cell is dependent on the reflectance ($r_f[E_v]$) of photons within the material, the local internal quantum efficiency ($Q_i[E_v]$) and the local junction voltages (U). These three factors relate to classical loss mechanisms namely, optical, recombination and resistive losses[2]. If the local junction voltage is assumed to be constant across the cell, the relationship between the optical and recombination properties and the spectrum of the emitted EL of the cell can be investigated.

1.2. Charge carrier distribution

When a forward bias is applied to a solar cell, minority carriers are injected at the p-n junction. Electrons are injected at the p-side of the p-n junction and holes at the n-side. Since the n-layer (emitter) is very thin, the amount of radiative recombination occurring is negligible[4]. Thus the distribution of electrons in the p-side of the p-n junction is discussed in this section.

The minority carrier concentration, n_e , at the junction ($z=0$) is given in equation (1) and is dependent on the intrinsic carrier concentration, n_i , the doping concentration of the base, N_A , applied junction voltage, U , and the thermal voltage, $k_B T/e$ [1]:

$$n_e(0) = \frac{n_i^2}{N_A} \exp\left(\frac{eU}{k_B T}\right). \quad (1)$$

The carrier concentration at the rear surface, $n_e(d)$, is related to the surface recombination velocity (S), the carrier diffusivity (D_e) and the charge carrier diffusion current density (j_e), as follows:

$$j_e(d) = -D_e \frac{dn_e}{dx}(d) = S n_e(d) \quad (2)$$

The carrier diffusion length is related to carrier diffusivity and carrier lifetime (τ_e):

$$L_e = \sqrt{D_e \tau_e} \quad (3)$$

These boundary conditions allow the minority carrier distribution, $n(z)$, to be determined as a function of applied voltage, diffusion length and rear surface recombination velocity. Electrons are injected at the junction ($z=0$) where the concentration is given by equation (1) and dependent on intrinsic carrier concentration, the doping concentration of the base, the applied junction voltage, U , and the thermal voltage, $k_B T/e$. An increase in the applied junction voltage will increase the carrier concentration which results in increase radiative recombination which leads to an increase in EL intensity. In this study the modelling of the EL spectrum is discussed with respect to recombination such as diffusion length and surface recombination velocity.

1.3. Photon Emission under forward bias

The rate of spontaneous photoemission by a semiconductor can be determined using a generalised form of Planck's law of radiation of non-black bodies[4].

The rate of spontaneous photoemission of photons of energy, E_γ , in a certain volume element is given by:

$$r_{sp}(E_\gamma) = \alpha(E_\gamma) \frac{E_\gamma^2}{\pi^2 \hbar^3 c^2} \frac{1}{\exp\left(\frac{E_\gamma - \Delta\eta}{k_B T}\right) - 1} \quad (4)$$

Where $\alpha(E_\gamma)$ is the energy-dependant absorption coefficient, \hbar is the reduced Planck's constant, c , is the speed of light, $\Delta\eta$ is the local separation of quasi-Fermi Energies, k_B is Boltzmann's constant and T is temperature in Kelvin.

The quasi-Fermi levels of n- or p- type Silicon can be related to the doping concentration (N_A), the intrinsic equilibrium charge carrier concentration (n_i) and the carrier distribution ($n(z)$).

Low- injection conditions are assumed with recombination occurring through band-band transitions [5]. The dependence of the recombination parameters on the carrier concentration is ignored in order to get a numerically solvable equation. Further assumptions are made that carrier diffusion does not occur laterally in the cell, only parallel to the depth. These assumptions allow the rate of spontaneous photoemission to be approximated using:

$$r_{sp}(E_\gamma, z) \approx \alpha(E_\gamma) \frac{E_\gamma^2}{\pi^2 \hbar^3 c^2} \exp\left(-\frac{E_\gamma}{k_B T}\right) \frac{N_A n(z)}{n_i^2} \quad (5)$$

Equation (5) predicts the local spectral distribution of spontaneously emitted photons within the sample.

1.4. Emitted Photon Detection

The rate at which of photons of an certain energy interval, $I(E_\gamma)$, are emitted from the surface of the cell sample can be obtained by integrating equation 4 across the thickness of the cell and is given by equation (6).

$$I(E_\gamma) = [1 - R_f(E_\gamma)] \int_0^d (r_{sp}(z, E_\gamma) \{ \exp[-\alpha(E_\gamma)z] + R_r(E_\gamma) \exp[-\alpha(E_\gamma)(2d - z)] \}) dz \quad (6)$$

The reflectance terms take into account a single reflection from the front surface (R_f) and at the rear surface (R_r). The inclusion of the absorption coefficient, $\alpha(E_\gamma)$, in the equation takes into account the reabsorption of photons in their optical path to the surface of the cell. The generation of electron-hole pairs due to reabsorption of these spontaneously emitted photons is ignored as non-radiative recombination is assumed to be the dominant form of recombination [1]. Equation (6) only considers photons emitted normal to the cell surface.

For EL imaging, a silicon CCD camera is commonly used because it has very good spatial resolution, however, it is only sensitive to wavelengths in the range of 300-1200 nm[6]. There is an overlap in wavelengths from about 900 to 1100 nm allowing the EL to be detected. InGaAs cameras can also be used for EL imaging and have a wider spectral detection range. However, they lack the spatial resolution of silicon and since they are more sensitive to longer wavelengths, they are more susceptible to light trapping within the sample which results in smearing of the image[1].

The spectrum of the detected EL is thus dependant on the sensitivity of the camera (Φ_{camera}), the transmission of filters (T_{filter}) used and the emitted photon flux[1] as shown:

$$I_{detect}(E_\gamma) = T_{filter}(E_\gamma) \Phi_{camera}(E_\gamma) I(E_\gamma) \quad (7)$$

2. Experimental Setup

2.1. Modeling

The results are modelled using a Mathematica routine. The parameters used are given in table 1. These parameters are used unless otherwise specified and are literature values or from Kirchartz et.al.[7].

Table 1. The parameters used for modelling the EL spectrum of Silicon.

Parameter	Symbol	Value used
Intrinsic carrier concentration	n_i	$9.69 \times 10^{15} \text{ m}^{-3}$ at 300 K
Doping concentration of the base	N_A	10^{22} m^{-3} at 300 K
Temperature	T	300 K
Local Junction Voltage	U	0.05 mV
Surface Recombination Velocity	S	10 m/s
Diffusion length	L_e	100 μm
Electron diffusivity	D_e	$27 \times 10^{-4} \text{ m.s}^{-1}$
Thickness of p-layer	d	250 μm
Speed of light	c	$2.998 \times 10^8 \text{ m.s}^{-1}$
Reduced Planck's constant	\hbar	$6.58 \times 10^{-4} \text{ eV.s}$
Boltzmann's constant	k_B	$8.617 \times 10^{-5} \text{ eV.K}^{-1}$

3. Results

3.1. Modeling of excess charge carrier distribution

The effects of recombination properties on the charge carrier distribution is modelled using equations 1 and 2 and the parameters of table 1. Figure 2 shows the excess carrier concentration of electrons in the p-side of a solar cell with a thickness of 250 μm . The distribution of minority carriers across the depth of the cell is determined by the diffusion length of the carriers and by increasing the diffusion length, the distribution of carriers across the cell increases. The surface recombination velocity, S , is the rate at which carriers recombine at the surface of the material and is set at 10 m.s^{-1} . The faster recombination occurs at the surface, the lower the carrier concentration at the surface ($z=250 \mu\text{m}$).

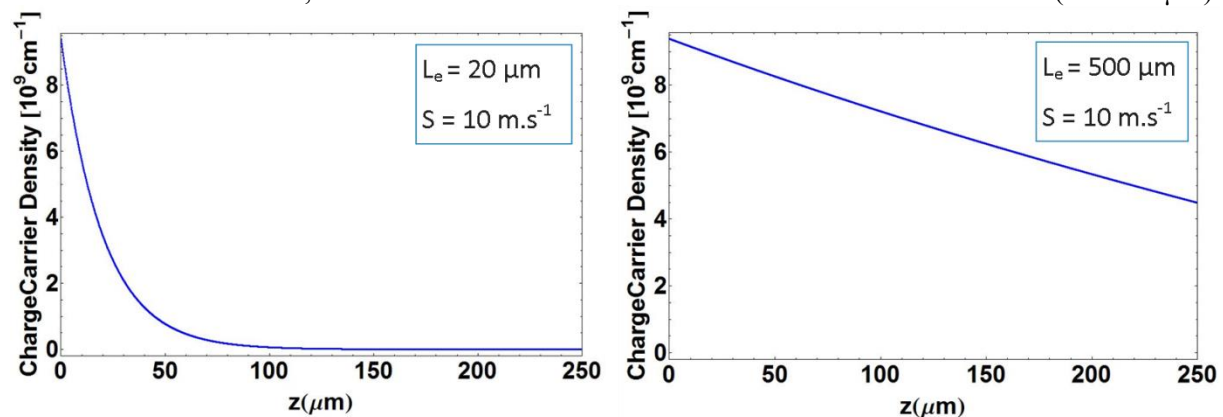


Figure 2: Excess charge carrier distribution across the depth of the cell modelled with diffusion lengths equal to 20 μm and 500 μm .

3.2. Modeling of detected EL

The EL intensity, which is the rate of photons emitted from a local surface element is simulated and plotted against wavelength in figure 3. Varying the diffusion length of the carriers has a more significant effect on the short wavelength region of the spectrum. High energy (short wavelength) photons are

more likely to be re-absorbed as silicon has a higher absorption coefficient in the low wavelength region. A short diffusion length means the carriers recombine deeper in the cell close to the p-n junction and the photons are more likely to be reabsorbed before they can be emitted at the surface. A longer diffusion length means carriers are able to recombine further away from the junction closer to the surface and thus the number of emitted photons of all energies is greater. This effect is especially noticeable in the short wavelength, high EL photon energy region.

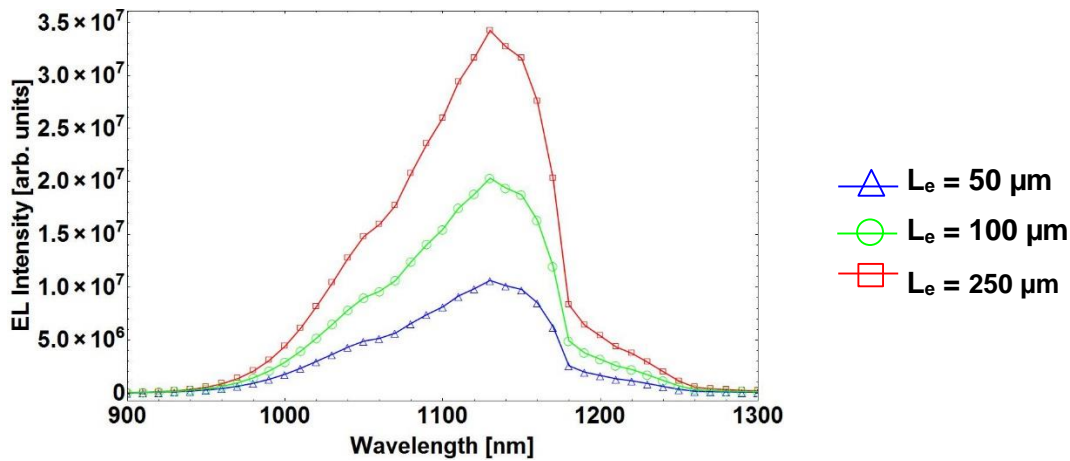


Figure. 3: The EL spectrum modeled at different diffusion lengths.

3.3. Detecting EL using a CCD camera

Figure 4 shows the overlap of the modeled EL spectrum and the CCD sensitivity[8], and short-pass filters with cut-offs at 900 nm and 1000 nm. The overlap between the CCD camera and the EL spectrum is only at short wavelengths and thus only a small percentage of the total EL can actually be detected. The incorporation of short-pass filters was suggested by Wurfel et al[1], in order to isolate only photons emitted from EL in the short wavelength region. These photons are generated near the surface as high energy photons generated deeper would be reabsorbed. The use of filters limits the spectrum of the detected EL and thus the image can be spectrally defined.

Figure 5 shows the CCD camera's quantum efficiency, the modelled EL and the detected EL. The EL and quantum efficiency are normalised. The detected EL is in the wavelength range of between 950-1150 nm which includes only a small portion of the total emitted EL in the region where the camera has low quantum efficiency.

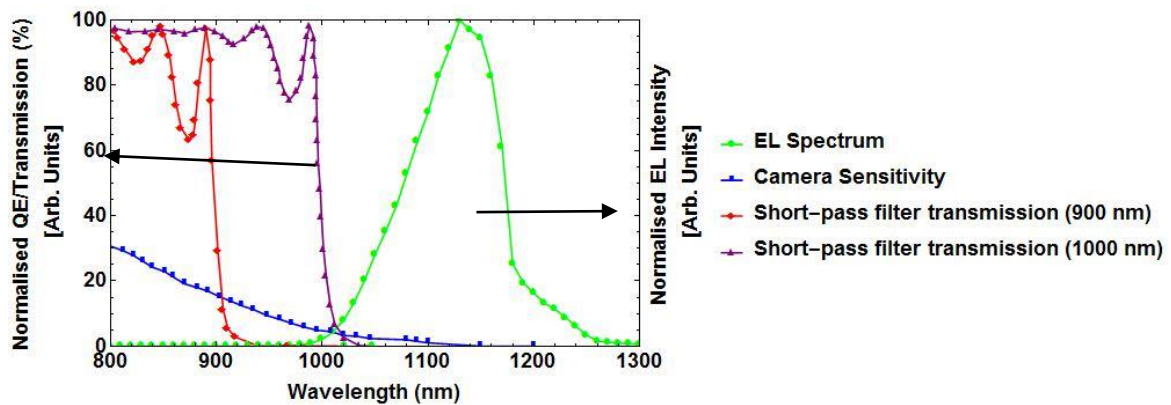


Figure 4: The EL spectrum overlaid with the CCD camera sensitivity and the transmission of the short-pass filters.

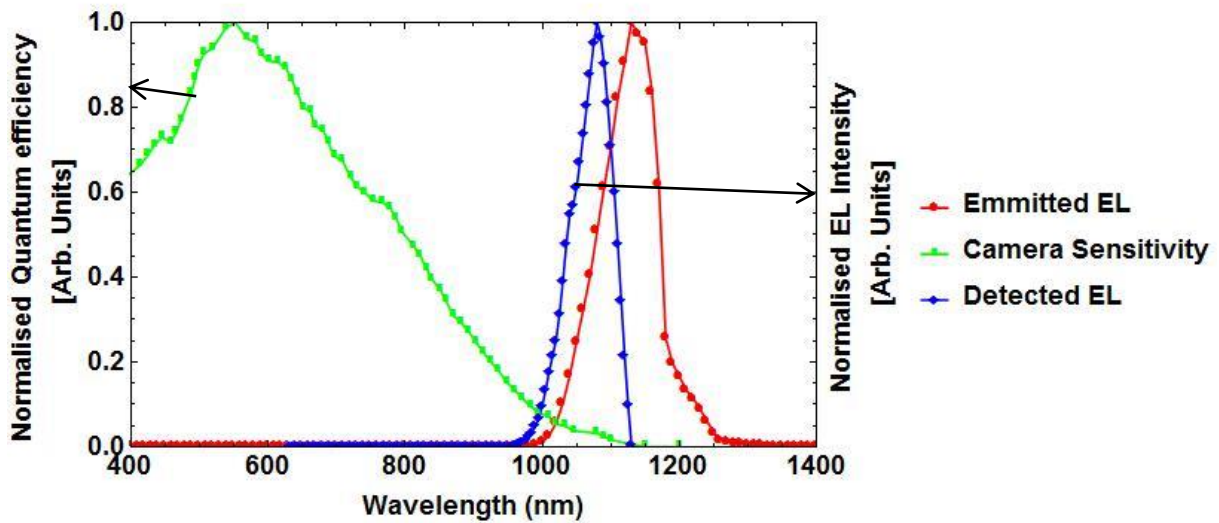


Figure 5: The EL spectrum of silicon modelled with parameters set in table 11 compared with the silicon CCD camera sensitivity. The normalised detected EL peak is also plotted.

4. Discussion and Conclusion

A solar cell can be visualised as a spatially extended device where light is emitted at different points on the surface, each with their own local quantum efficiency and local junction voltage.

A thorough understanding of the spectrum of emitted during EL is important. Modelling the effects of different recombination properties and filter integration allows for a better idea of the EL process. While the EL imaging with a silicon CCD camera is not able to measure the EL spectrum the use of filters allows some assumptions of the detected wavelength range to be made. The modeled EL spectrum illustrated that the minority carrier diffusion length has an effect on the EL intensity in the short wavelength region. As this region is within the detection sensitivity of the camera, the modeling of the effects of diffusion is important for interpreting the EL images. Future work includes integration of filters into experimental set-up and correlation between experimental and simulated results.

Lateral filter inhomogeneity in the short-pass filters can cause a significant variation in the transmission of the filter. This needs to be addressed in future experimental work, possibly by rotating the filter and averaging the images.

5. References

- [1] Würfel P, Trupke T, Puzzer T, Schäffer E, Warta W, Glunz SW 2007 *J Appl Phys* **101** 123110.
- [2] Kirchartz T, Helbig A, Reetz W, Reuter M, Werner JH, Rau U 2009 *Prog Photov* **17** 394-402.
- [3] Rau U 2007 *Phys Rev B: Condens Matter* **76**.
- [4] Giesecke JA, Kasemann M, Warta W 2009 *J Appl Phys* **106** 014907.
- [5] Rißland S, Breitenstein O 2013 *Sol Energy Mater Sol Cells* **111** 112-4.
- [6] Fuyuki T, Kitiyanan A 2008 *Appl Phys A* **96** 189-96.
- [7] Kirchartz T, Helbig A, Rau U 2008 *Sol Energy Mater Sol Cells* **92** 1621-7.
- [8] Sensovation. SensoCam Scientific Cameras. 2012 [cited 2014 10 July]; Available from: http://www.sensovation.com/bausteine.net/f/10339/web_SensoCamA4single.pdf?fd=2.

Reproducing observed solar radiation characteristics in tropical regions using stochastic theoretical models

Marie Chantal Cyulinyana and Hartmut Winkler

Department of Physics, University of Johannesburg, PO Box 524, 2006 Auckland Park, Johannesburg, South Africa

E-mail: kamirwa@gmail.com

Abstract. This paper discusses the framework for developing a theoretical model that characterizes solar radiation at ground level and evaluates the model using Monte Carlo simulations. It presents an overview of the methodology and calculations required for such a study, including the effect aerosols have on irradiance. It explores the results of previous studies that have provided broad band solar irradiance measurements at specific locations and that have developed empirical relationships between a series of solar parameters appropriate for African conditions. It seeks to reproduce these relationships using a theoretical modelling approach that incorporates radiative transfer processes, as well as Rayleigh and Mie scattering. It presents wavelength-dependent techniques that characterize the various components of solar radiation using theoretical models, and evaluates their suitability in this context. The application of Monte Carlo based atmospheric radiative transfer models is also briefly discussed.

1. Introduction

Solar radiation is energy distributed over a wide continuous spectrum ranging from ultraviolet to the infrared. The transmittance of the atmosphere to solar light is wavelength dependent, as is the efficiency of the solar device, which is a function of the technology employed.

The intensity of the Sun's rays reaching the outer surface of the Earth's atmosphere is approximately 1360 W/m^2 . The solar beam intensity at the Earth's surface is reduced by scattering and absorption due to molecules, aerosols, water vapour and naturally clouds. The beam of sunlight incident at the Earth's surface from the direction of the Sun is called direct solar radiation, while scattered sunlight incident from other directions, i.e. sky light, is referred as diffuse radiation [1]. Despite their significance, direct and diffuse solar radiation measurements are not abundantly available due to the cost and difficult maintenance of measuring equipment such as pyranometers.

In past decades, the characterization of solar radiation has been based on satellite derived data as well as empirical models based on, amongst other things, meteorological records. This includes a variety of studies of solar irradiance in African and other tropical localities [2-5]. These approaches rely on simplifying assumptions to deal with the mathematical complexities inherent in analytical determinations of the irradiance.

An alternative means of treating this problem is through stochastic computation, commonly referred to as a Monte Carlo simulation. This involves launching photons from the top of atmosphere as if they come from the Sun. Each photon is tracked through a sequence of interactions until it is absorbed in the atmosphere, observed at the surface, or scattered out of the atmosphere. The outcome of each interaction is governed by the application of randomly generated numbers to the physics of the process and the probability of achieving a particular outcome [6]. The true solar irradiance profile

is then sufficiently approximated through multiple computational repetition of the photon generation process.

The study presented here is a precursor to a wider PhD study that aims to determine the solar irradiance and solar energy potential of Rwanda. The methodology developed is however equally suited to other localities, when atmospheric parameters are suitably adjusted. In the case of Rwanda, the model to be eventually developed will be required to match not only the measured local annual average solar irradiation (4.3-5.2 kWh/m²/day, depending on the exact locality [2,5]), but also established cloud cover patterns and atmospheric turbidity (with seasonal trends).

In addition to the above, the Monte Carlo simulations will also be providing important details about the angular and spectral distribution of the scattered skylight. This is critical for the determination of the efficiency of solar power generating devices, and to determine their optimal configuration (orientation). This is particularly significant in view of the high equipment and maintenance costs that would be required to secure this information through experimental means.

2. The stochastic model employed

The equation of radiative transfer mathematically describes a light beam's energy losses to absorption, gains by emission, and redistribution by scattering. Radiative transfer modelling is the numerical prediction of quantities related to the radiative energy transfer in a medium such as the Earth's atmosphere. The approach here involves stochastic modelling of the scattering and absorption of sunlight in complex atmospheric conditions [6-8].

As solar light penetrates the atmosphere, the solar radiation is attenuated by the atmosphere through scattering, absorption and reflection. The attenuation is proportional to the density of the medium, the incoming beam, and the path length or optical path length, and is described by the Beer-Lambert equation.

$$I_{\lambda} = I_{0,\lambda} \exp(-\tau_{\lambda} m) \quad (1)$$

Where I_{λ} and $I_{0,\lambda}$ are the solar beam intensity per unit wavelength at ground level and at the top of the atmosphere respectively, τ_{λ} is the (wavelength dependent) optical depth and m is the airmass, i.e. the relative beam path through the atmosphere measured in units of the vertical atmospheric depth.

Scatterers in the atmosphere are usually divided into two main types - aerosols and molecules. The scattering mechanism largely depends on the scattering particle size. Rayleigh scattering occurs for particle sizes smaller than the wavelength (as is the case with air molecules), while Mie Scattering happens if the particle size is of the order of the wavelength (i.e. for a large fraction of aerosols).

For Rayleigh scattering, the relationship between the relative intensity of scattered light and the angle of deflection ξ is described by the phase function,

$$p_R(\xi) = \frac{3}{4} (1 + \cos^2 \xi), \quad (2)$$

which we note reaches a minimum for scattering at 90° to the original beam direction [9].

For Mie scattering, the phase function is far more complex, as it depends on aerosol shape, size and composition. A good way of approximating aerosol scattering is by means of the Henyey-Greenstein phase function

$$p_{aer}(\xi) = \frac{1 - g_{\lambda}^2}{(1 + g_{\lambda}^2 - 2g_{\lambda} \cos \xi)^{3/2}} \quad (3)$$

where g is a parameter referred to as the asymmetry factor [10].

3. Application of the stochastic computation model to specific scenarios

A random number generator is used to provide random numbers between 0 and 1 which are used to calculate the photon step size (i.e. the distance covered by the photon until it is absorbed or scattered), the nature of the interaction (absorption or scattering) and the scattering angle (deflection angle ξ and azimuth). In addition, through application of the spectral distribution of the solar beam above the atmosphere, one can even use random numbers to determine the photon wavelength. Where a

scattering event has taken place, the process is then repeated until the photon reaches the ground, is scattered back into space or becomes absorbed. Repeating this procedure to a very large sample of photons yields a good approximation of the solar irradiance observed on the ground. Through considering the zenith angle and azimuth from which the scattered photons strike the ground, one is able to model the angular light distribution of diffuse light, and even to estimate its spectral properties. Even cloud cover and weather patterns can be determined stochastically, though none of that is attempted here. The formalism for applying random numbers to the two phase functions described earlier has been previously established. Prah et al used a generated random number (r) to calculate the step size, Δs (see equation 4):

$$\Delta s = \frac{-\ln r}{u_t} \quad (4)$$

(where u_t is the total attenuation coefficient).

They also used generated random numbers to calculate the azimuthal angle, $\psi = 2\pi r$ and the scattering angle, ξ [11], which are the key elements to get the new direction of a moving photon:

$$\cos \xi = \frac{1}{2g} \left\{ 1 + g^2 - \left[\frac{1 - g^2}{1 - g + 2gr} \right]^2 \right\} \quad (5)$$

For the simulation that formed the basis to this paper, we made the simplifying assumption that the Sun is overhead. We computationally launched 21000 photons from the top of the atmosphere at three specific wavelengths and tracked the progress of each one through the atmosphere. We chose an asymmetry factor of $g = 0.2$, typical of many aerosols.

4. Results

The following results represent Rayleigh and aerosol scattering, brightness and the colour characteristic of the sky at different wavelengths as determined by our simulations.

4.1 Rayleigh and aerosol scattering

Figure 1 compares the deflection angle for the Rayleigh and Mie scattering processes. The preferred forward direction for scattering of aerosols as well as the ‘dumbbell’ shape of the Rayleigh phase function is clearly visible in the histogram.

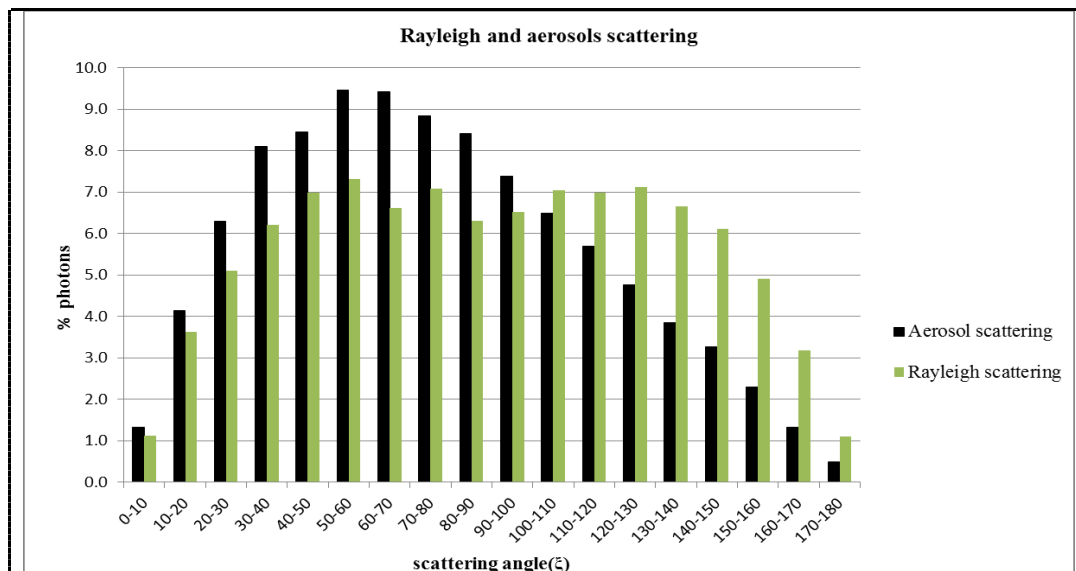


Figure 1. Comparison between Rayleigh and aerosol scattering. The graph displays the percentage of photons scattered during atmospheric transit as a function of ξ (in degrees)

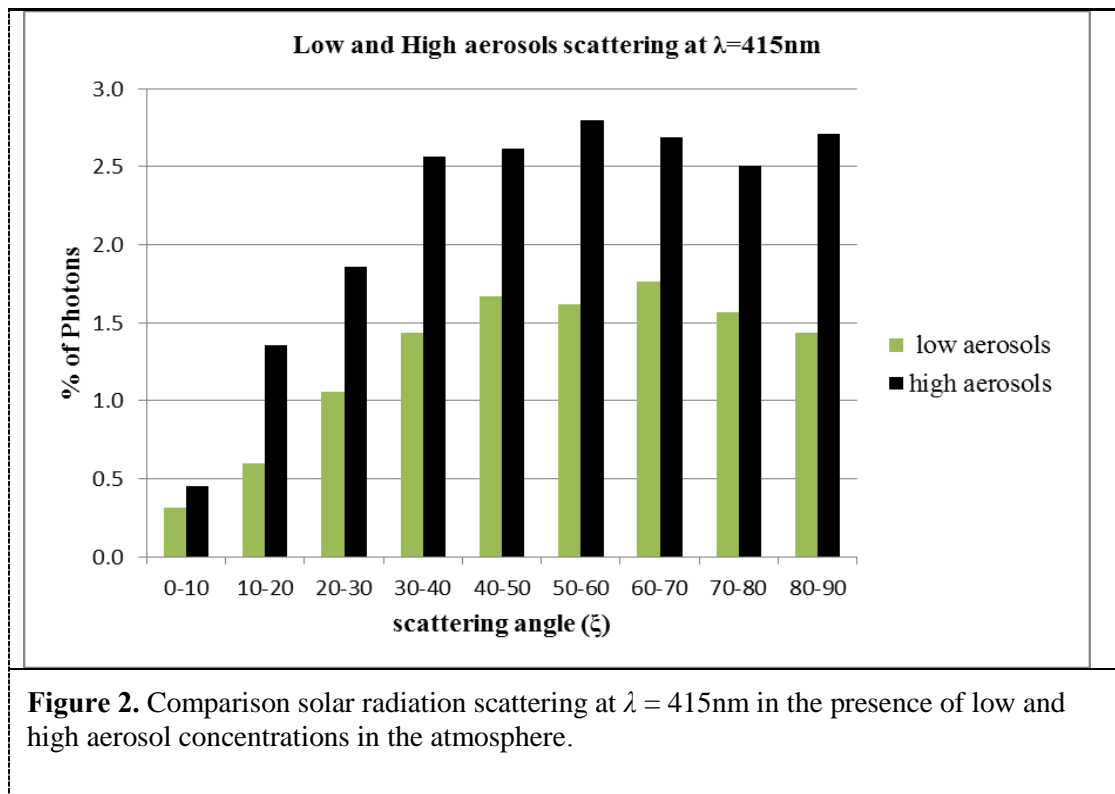
4.2. Scattered photons at different wavelength

The simulations in this paper were performed for the following wavelengths and optical depths:

Table 1. Wavelength range and optical depth

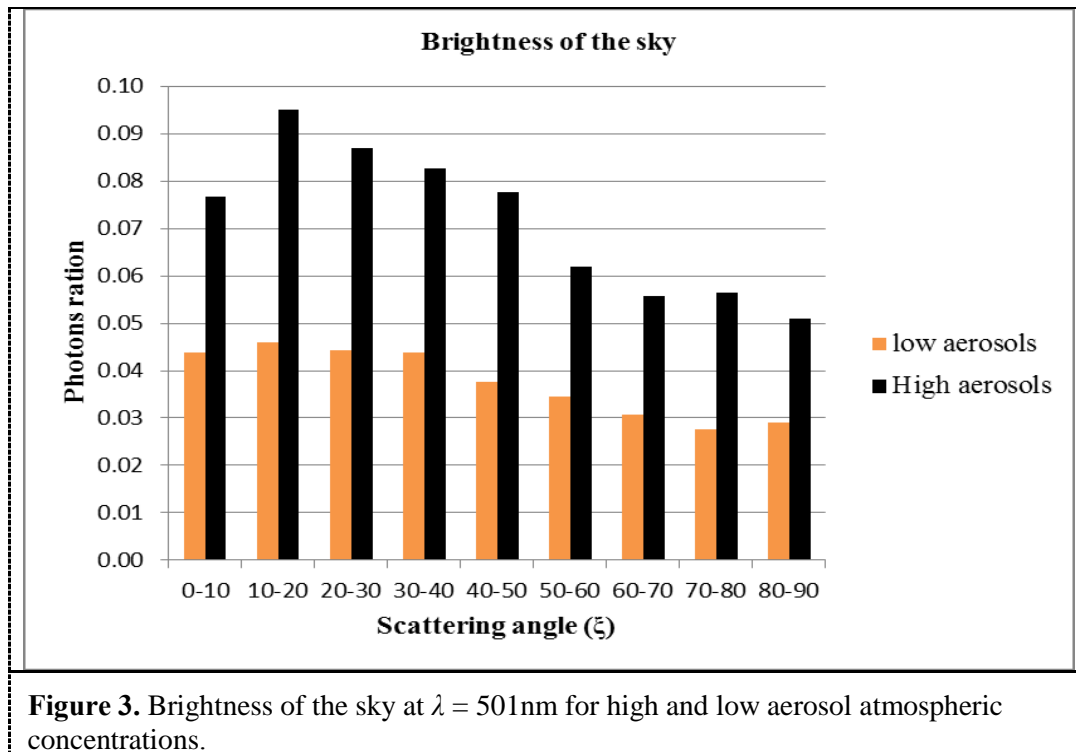
λ	τ (low)	τ (high)
415 nm	0.26	0.44
501 nm	0.14	0.26
615 nm	0.09	0.18

The first of these optical depths are typical of clear conditions at a rural site at altitude of ~ 1200 m above sea level, deduced from previous studies carried out over southern Africa [12]. The other optical depths represent an example of an atmosphere containing some aerosols generated by biomass burning. The wavelengths chosen correspond to standard sunphotometer bands.



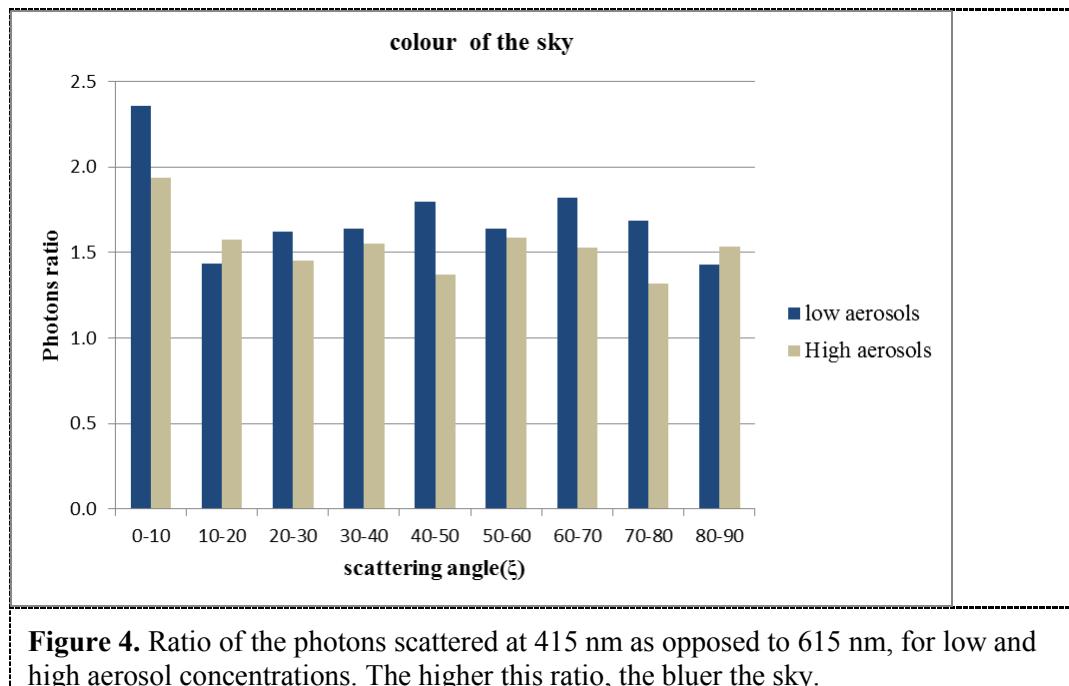
4.3 Brightness of the sky

Figure 3 displays the actual sky brightness, and differs from the distribution in Fig. 2 due to the fact that the solid angle corresponding to any particular deflection angle range is actually proportional to $\cos \xi$. As in the figure before that, the larger aerosol concentration leads to a greater sky brightness, as is expected in view of the known aureole around the Sun in hazy conditions. Here the photon ratio represents the fraction of photons scattered in a given range of scattering angle per total number of launched photons.



4.4 Colour of the sky

Finally, we test whether our simulation leads to colour differences between different scattering angle directions. In Fig. 4, photon ratio refers to the fraction of photons scattered at two different wavelengths (615nm and 415nm respectively) in the presence of low and high aerosol concentrations. The results of this test, shown in Fig. 4 below, are inconclusive. A much larger photon sample is required to detect any dependence on scattering angle or optical depth.



5. Further work

The results of the preliminary stochastic simulations that form the basis of this paper show the potential of applying such simulations to determine both direct and diffuse irradiance values, especially as these can be expanded to a whole spectrum of wavelengths. Accuracy can be improved through drastically increasing the photon number in the computations. The models are also easily adapted to different solar beam inclination angles, meaning that an irradiation time sequence can be developed. The possibility of refining the aerosol scattering characteristics to match local aerosol types make this a promising approach to the modelling of ground level solar irradiance, and hence determining the solar energy potential at a site.

References

- [1] Muneer T, Gueymard C and Kambezidis H 2004 *Solar Radiation and Daylight Models* 2nd Ed (Butterworth Heinemann: Elsevier)
- [2] Safari B and Gasore J 2009 *Asian Journal of scientific Research* **2(2)** 68-75
- [3] Quansah E, Amekudzi L K, Preko K, et al. 2014 *Journal of Solar Energy* ID 897970
- [4] Jakhiani A Q, Othman A K, Rigit A R H and Samo S R 2011 *World Applied Sciences Journal* **14** 83-90
- [5] Chineke T C 2008 *Renewable Energy* **33(4)** 827-831
- [6] Kattawar G W *Monte Carlo methods in radiative transfer* Center for 'Theoretical Physics Department of Physics, Texas University
- [7] Hay T D, Bodeker G E, Kreher K, et al. 2012 *J. Quantitative Spectroscopy & Radiative Transfer* **113** 721-738
- [8] Miesch C, Briottet X, Kerr Y H and Cabot F 1999 *Applied Optics* **38** 7419-7430
- [9] Bucholtz A 1995 *Applied Optics* **34** 2765-2773
- [10] Kocifaj M 2009 *Solar Energy* **83** 1914-1922
- [11] Prah S A, Keijzer M, Jacques S L and Welch A J 1989 *SPIE Institute Series* **IS 5** 102-110
- [12] Formenti P, Winkler H, Fourie P, et al. 2002 *Atmospheric Research* **62** 11-32

Characterization of plume and thrust for the corona ionization space propulsion system

P Ferrer

University of the Witwatersrand, 1 Jan Smuts Ave., Braamfontein 2000,
Johannesburg, South Africa

E-mail: philippe.ferrer@wits.ac.za

Abstract. The thrust production of a miniaturized space propulsion system using corona ionization of the propellant depends on the exhaust plume composition. A two fluid simulation was constructed to extract electrical properties of the plume, such as the electric field, potential and the ion and electron charge densities. The form of the potential is seen to depend on the propellant flow rate, which affects the non-uniform pressure profile appropriate for this system and hence ion mobility. The form of the pressure profile used introduces a space charge polarization close to the ionising needle, which dramatically increases thrust production. The thrust description, which shows the local thrust contributions from the plume, indicates that thrust can have even negative components in the polarization region, but is overall much larger than for the conventionally studied constant pressure background system.

1. Introduction

Since the launch of SANSA (South African space agency) in 2010, space related activities have received new impetus in southern Africa. Although the emphasis is on astronomy, science and engineering applications are also being pursued. One of the research directions concerns itself with electric propulsion systems (or “thrusters”) which can be used in outer space on deep space probes or satellites.

Lately, research on thruster miniaturisation has been gaining in importance, with the aim to reduce production and launch costs. At the same time, manufacture of smaller space components became desirable, so that the overall versatility of space vehicles can be increased.

In miniaturising electric propulsion systems, one aims for the high efficiency and long lifetime of larger systems. This can be difficult to achieve; the well-known gridded electrostatic ion engines, for instance, lose efficiency on reduction of the discharge chamber [1], while some progress with the popular Hall thrusters is made [2], especially on erosion problems.

A number of alternative miniaturized electric propulsion systems is being studied. Work is pursued on thrusters relying on the ejection of charged particles, such as pulsed plasma thrusters, the FEEP, colloid thrusters, Helicon thrusters, vacuum arc thrusters, micro-particle thrusters, hollow cathode thrusters and many of their derivatives (see [3] for a comprehensive discussion of miniaturised electric propulsion systems). Thrusters are studied relying on heating mechanisms of the propellant [4], such as laser ablation thrusters and arc-jet thrusters [5] or else on interaction with external fields, like the electro-dynamic tethers [6]. Many of these systems are unique candidates for specific mission requirements.

Here, we discuss an aspect of a thruster depending on corona ionisation. Since thrust is produced via electrical interactions between the thruster body and the plume, knowledge of the plume composition is vital.

2. Operation of the corona ionisation thruster

The thruster consists of two hollow needles separated by a short distance with a potential difference between them, and embedded in an insulator (figure 1). A propellant gas escapes into vacuum through the needles. At the needle tips, the gas experiences a strong electric field and corona ionization takes place. Electrons are produced at the negative needle tip, positive ions at the positive needle tip. The charged particles are mostly repelled from the needle they emerge from, with a small fraction forming a plasma bridge, ensuring continuing operation. Less than a percent of gas molecules are corona ionized in the gas, forcing the ions and electrons to move through a dominant neutral gas background.

Experimental data has been collected for such a system and is described in detail elsewhere [7].

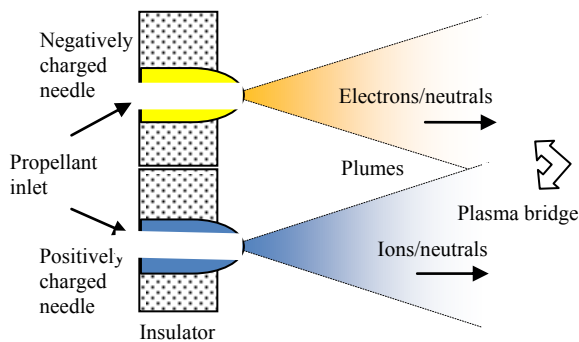


Figure 1. Oppositely charged needles create charged particle flows through a predominantly neutral propellant via corona ionization. A plasma bridge ensures continuing operation.

An analytic model for the radial electric field inside the positive plume was suggested in [7], which made some simplifying assumptions, such as ignoring electrons, using unrealistically large values for the plume spread angle (45 deg) and a very high secondary electron emission coefficient of 0.67.

The similarity in thrust and discharge behaviour of the two needle system of figure 1 and a simpler configuration involving a positively charged needle and a grounded plate (see figure 2) was noted in [7]. The simulation introduced in the next section is based on the simpler configuration of figure 2.

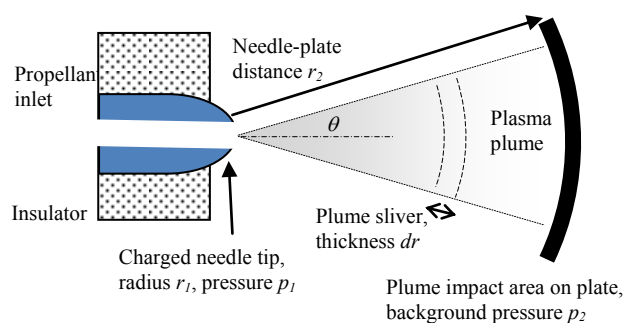


Figure 2. Discharge system of needle to plate configuration showing similar properties to the needle-needle thruster. The gas density of the plume decreases quadratically towards the plate.

3. Two fluid simulation

Non-intrusive direct measurements of plasma systems are difficult to perform on small scales, and analytic solutions of these can often not be found. A simulation can yield valuable insights in such a situation.

We use the well known 1-dimensional two fluid simulation [8] to obtain qualitative insight into the plume behaviour. The simulation considers two concentric spheres with a potential difference and a spatially varying pressure profile (equation (1)) between them. The inner sphere represents the

needle-tip and the outer sphere the discharge plate of figure 2. Since the gas expands as a cone into vacuum, an inversely quadratic relationship of the form

$$p(r) = a + \frac{b}{r^2}, \quad (1)$$

was assumed for the pressure decrease as a function of distance r from the needle-tip, as a first approximation [7,8]. Here, the parameters a and b were fixed to satisfy the boundary pressures at the needle tip and the background pressure at the discharge plate. At high potential difference, discharge occurs over the entire enclosed region. We use that fraction of the total current which goes through a cone representing the plume. This conical section is assumed to have uniform density along the angular directions, and besides yielding the measured current, must also satisfy the measured neutral gas flow through it. While more accurate models exist, this model reduces complexity and computational time, and allows direct comparison to the analytical model in [7]. The MATLAB program for this simulation was validated in [8], where all the simulation parameters are also listed.

4. Electric plume composition and propellant flow rate

The discharge occurs between the needle, which emits the propellant gas at a constant mass flow rate, and the discharge plate. In the simulation, their separation was taken as 40 mm with a potential difference of 1000V, and the background pressure was set as 0.6 Torr (these values were chosen because experimental data for them is available [7]). The needle exit pressure varied, simulating a changing flow rate. Experimentally, a flow rate of 1.4 mg/s was close to a needle pressure of 97 Torr, and 1,0 mg/s close to 70 Torr.

The potential, electric field and charge densities for the entire discharge region are shown in figure 3.

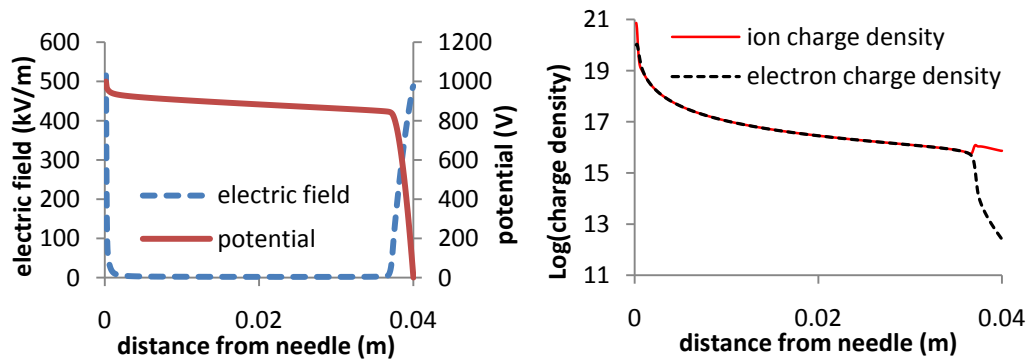


Figure 3. (Left) The potential and electric field throughout the discharge region is shown. (Right) The ion and electron charge distribution (taken for 70 Torr needle tip pressure, representative).

The electric field peaks close to the electrodes (at the anode and cathode sheaths), and is suppressed by the electron presence in between. The potential drops slowly at first, then rapidly within the cathode sheath. The charge densities cancel each other throughout most of the discharge region, except near the electrodes. At the plate, secondary electrons are ejected, accelerated through the gas and give rise to an electron avalanche until the ion charge density is neutralised. We are mainly interested in the region close to the needle, since the plate is an artefact of the simplified system of figure 2.

Figure 4 shows the electric field strength close to the needle for different needle-tip pressures (= mass flow rates). For a needle-tip pressure of 0.6 Torr, the background pressure is constant, and there

is no flow from the needle. The field there is suppressed, and similar systems have been studied as ionic wind phenomena. As the flow rate increases, the field increases significantly and a corresponding potential drop occurs close to the needle.

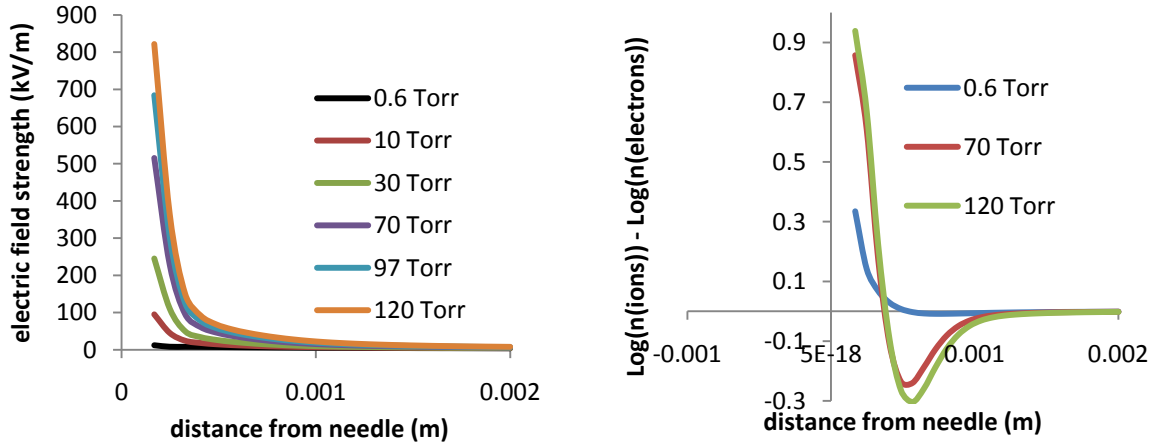


Figure 4. (Left) Electric field strength increases with mass flow rate. (Right) Increased mass flow rate sets up a polarised space charge region.

Figure 4 (right) shows that the decreasing pressure profile of equation 1 introduces a “polarisation” of the space charges close to the needle, which is responsible for the increase in electric field, and which increases with increasing mass flow rate.

5. Thrust characteristics of the plume

The likely origin of thrust for this system is the repulsion of ions in the plume. Other possibilities were investigated [8], but are at best secondary, less important effects. The ions move through the plume with a drift velocity given by a product of their mobility and the electric field, to a good approximation. Both these quantities change throughout the plume (see figure 3(left) for electric field variation and equation (1) for pressure, and hence, mobility variation), giving rise to effective local charge densities, discretised in the simulation by “plume slivers” of thickness $dr = 80$ microns (see figure 2). These charge densities, which remain constant during discharge once equilibrium has been reached, interact electrostatically with the positively charged needle, exerting a force on it. The force T (= thrust) was shown in [8] to be given by the expression

$$T = \frac{\pi}{2} (1 - \cos 2\theta) \int_{r_1}^{r_2} q(n_i - n_e) E r^2 dr \quad (2)$$

where θ is the plume half angle, q is positive electron charge, n_i and n_e are the local ion and electron charge densities respectively, and E is the local electric field magnitude. This information is available from the simulation, and allows us to graph the local thrust contribution from the plume (figure 5).

It is clear that the thrust magnitude responds to the polarization induced by the non-uniform pressure profile of equation (1). The thrust contribution to the total thrust from the region dominated by negative charge density is negative, but despite this, the total thrust increases significantly with propellant mass flow rate, since the positive space charge region more than compensates for this (see table 1).

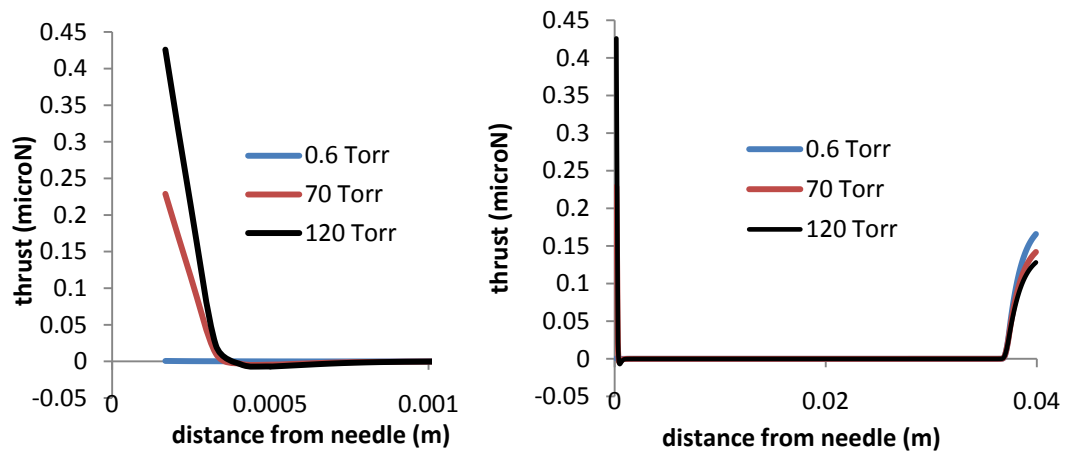


Figure 5. (Left) Thrust contribution from charge densities close to needle tip for three different needle-tip pressures. (Right) Thrust contribution from entire discharge region.

Table 1 shows the global results for the different mass flow rates given by the simulation for a plume angle of ~ 1.4 degrees, discharge voltage of 1000V and background pressure of 0.6 Torr.

Table 1. Needle tip pressure and simulation results for plume discharge current (= power) and total thrust calculated from equation (2).

Needle-tip pressure (Torr)	Plume discharge current (mA) and Power (W)	Total thrust (μN)
0.6	0.596	8.9×10^{-4}
10	0.581	2.16×10^{-2}
30	0.555	1.07×10^{-1}
70	0.510	3.39×10^{-1}
97	0.483	5.03×10^{-1}
120	0.460	6.35×10^{-1}

The discharge power decreases, while the thrust increases significantly with propellant flow rate. It must be noted that the plume angle is not independent of flow rate, which has not been taken into account here. This will affect the values of the plume discharge current. The variation in discharge current will however be much smaller than the drastic increase (3 orders of magnitude) in total thrust.

6. Conclusion

A two fluid simulation was used to extract electrical properties of the plume, such as the local electric field, potential and charge densities. A variation in needle exit pressure corresponded to different flow rates of propellant being corona ionised. A quadratically decreasing pressure profile, which was chosen to simulate the gas expansion of the plume into vacuum, exhibited increased space charge accumulations close to the needle-tip compared to a constant background pressure environment. These space charges must be a consequence of the higher pressure (= density) close to the needle-tip, which increases ion production rate and decreases mobility. Both positive and negative thrust contributions from the plume were simulated, with the positive contribution far dominating. Thrust was seen to increase drastically with increasing flow rate, with the discharge power varying little. This behaviour will have a large effect on the thruster efficiency and becomes an important design consideration.

References

- [1] Sengupta A 2009 Magnetic confinement in a ring-cusp ion thruster discharge plasma *Journal of Applied Physics* **105** 093303
- [2] Ahedo E and J. Gallardo J 2003 Scaling down hall thrusters *Proc. of the 28th International Electric Propulsion Conference Toulouse, France, IEPC-2011-* **104**
- [3] Wright W, Ferrer P 2014 Electric Micropropulsion Systems *Progress in Aerospace Sciences*, pp. 48-61, DOI information: 10.1016/j.paerosci.2014.10.003
- [4] Phipps C, Birkan M, Bohn W, Eckel HA, Horisawa H, Lippert T *et al.* 2010 Review: laser-ablation propulsion *J Propul Power*, **26 (4)** 609–37
- [5] Wollenhaupt B, Hammer A, Herdrich G, Fasoulas S and Roser H 2011 A very low power arcjet (VELARC) for small satellite missions *Proc. of the 32nd International Electric Propulsion Conference, Electric Rocket Propulsion Society (Wiesbaden, Germany) IEPC-2011-* **257**
- [6] Kawamoto S, Makida T, Sasaki F, Okawa Y and Nishida S 2006 Precise numerical simulations of electrodynamic tethers for an active debris removal system *Acta Astronautica* **59** 139 – 48
- [7] Ferrer P and Tchonang M 2011 Miniaturization of electrostatic ion engines by ionization and acceleration coupling *J. Phys. D. Appl. Phys.* **44** 335204
- [8] Ferrer P and Liedberg H 2014 Plume characterization for a miniaturized corona ionization propulsion system *Submitted to J. Phys. D Appl. Phys.*

Modelling Flow Phenomena in Time Dependent Store Release from Transonic Aircraft

D A MacLucas and I M A Gledhill

Defence, Peace, Safety and Security Unit,
CSIR, PO Box 395, Pretoria 0001 South Africa

E-mail: igledhil@csir.co.za

Abstract. Computational Fluid Dynamics is routinely used in clearance of stores for carriage and release from aircraft in the transonic range of flight. A well-known validation case is modelled in this study, for which aerodynamic loads have been compared with wind tunnel experimental data by other authors. In this study, having validated the numerical model, we apply more recent methodologies from flow dynamics to study the detailed flow field in the region of the store fins.

1. Introduction

In the Mach number range between 0.8 and 1.2, aerodynamic loads on aircraft and launch vehicles are very sensitive to the presence of shocks. This is of particular relevance in the carriage of stores on transonic aircraft such as the South African Air Force SAAB Gripen JAS-39C and D multi-role fighters, and the BAES Hawk 120 Lead-in Fighter Trainer. Notably, where store fins are closely positioned to the supporting pylon, choked flow may occur, leading to a low-pressure footprint on the upper aft surface of the store. The consequent significant nose-down pitch on release may result in contact between the store and the aircraft, risking the aircraft and pilot in the collision. Explosive Release Units (ERUs) are fitted in such cases. All stores must be cleared for carriage and release before flight test commences.

During the weapons integration and clearance phase, computational calculations and wind tunnel tests are conducted. While numerical models in the subsonic and supersonic ranges rely on a range of assumptions and can be executed relatively fast, transonic models are very sensitive to geometry and are computationally demanding.

Computational Fluid Dynamics (CFD) is complementary to wind tunnel test and flight test. CFD is routinely used for certification among NATO (North Atlantic Treaty Organization) countries [1], and is increasingly used for aerodynamic characterization, aeroelastic studies, risk reduction, and optimization. CFD is used in conjunction with faster methods such as panel methods (Green's function calculations), full potential models with closely coupled boundary layer solvers, and empirical methods, but is the only methodology providing capture of shocks and viscous phenomena through the complete transonic range. In carriage configurations, solution of the Navier-Stokes equations and a selected turbulence model is required. In time-accurate release cases, the inviscid Euler equations are usually a good enough approximation, and a six-degree-of-freedom (6DOF) trajectory solution with fully time-accurate flow is obtained. In contrast, Captive Trajectory Systems (CTS) in transonic wind tunnels must rely on a quasi-steady approximation, in which steady flow at sample points along the store trajectory is used to derive loads.

Validation of directly relevant cases in CFD is a rigorous requirement. The case of a double-ogive finned store dropped from a pylon beneath a delta wing has been studied experimentally [2] [3] and is a much-referenced case [4] in which aerodynamic loads on the store and translational and rotational displacements are used for validation. In this paper, we use this public domain case to study the detailed flow field and its time-dependence in the interference region. A full time-accurate inviscid computation of the case is performed in conjunction with a 6DOF solver with the aim of predicting the store trajectory and capturing the transient flow phenomena.

The wing is a clipped delta wing, based on the NACA 64A010 section, with a 45° leading edge sweep angle supported by a centreline sting. The store is a tangent ogive forebody and similar afterbody with a cylindrical centrebody, and truncated NACA0008 fins with a 60° leading edge sweep. The store has a full scale mass of 907kg. The afterbody is truncated to accommodate the sting support of the CTS. No turbulence tripping was used. Full scale geometry, in which the mean aerodynamic chord of the wing is 4.32 m, was scaled down to 5% for the wind tunnel test. The experimental Reynolds number was 2.4×10^6 per foot, which corresponds to 34 million based on the mean aerodynamic chord in the full scale model. The experimental tests of interest were conducted at the Arnold Engineering Development Center (AEDC) [3] at angle of attack $\alpha=0^\circ$ and Mach number $M = 0.95$. In the carriage position the separation distance of the store from the pylon was 0.070 inches. Data corrections may be found in the original report. The test case was conducted at a static temperature $T = 236.707$ K and static pressure $p \sim 36042$ Pa.

2. Methodology

Geometry was modelled at full scale. The simulation axes are illustrated below in Figure 1a. The flight axis and store-fixed axes are illustrated below in Figure 1b. The x axis is oriented in the direction of flow, the y axis from the centreline to the wing tip, and the z axis in the upwards direction in flight. The store and force moment coefficients are computed about the store body-fixed axis. A decision is required on whether to model the store sting or not. In this case, the sting is not modelled to match the flight condition as opposed to the wind tunnel condition.

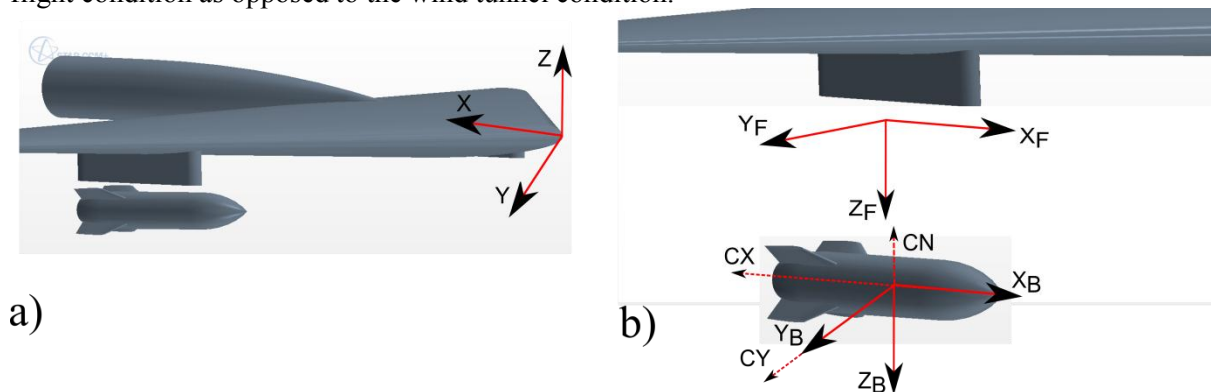


Figure 1: Axis systems. a) Simulation axes b) Flight axis set at the store centre of gravity in the carriage position below the pylon. Body-fixed axis fixed to the store centre of gravity. Positive sense for the normal force coefficient C_N , side force C_Y and axial force coefficient C_X illustrated.

A coupled density-based solver included in the CFD package Star-CCM+ v8.04 was used. Spatial discretisation is 2^{nd} order upwind in order to capture shock and discontinuities adequately, with solution by a coupled implicit solver. Inviscid flux discretization is the AUSM+ flux vector splitting scheme [5]. Since dispersive numerical effects degrade the solution at shocks in this scheme, the Venkatakrishnan gradient reconstruction limiter [6] is applied. A 6DOF model was included to solve for the store translation and rotation about all three axes during release. Far field boundary conditions

are set as follows relative to the simulation axes: 90 m from the wing upstream and downstream; 50 m in the cross-flow direction and 100m in the z direction.

A grid of approximately 9 million hexadedral-dominant cells was used to discretize the computational domain for the carriage model. In order to accommodate the 6DOF motion of the store, overset, or Chimera, grids were constructed. These allow motion of an independently-generated store grid through a background grid attached to the wing. Grids are illustrated in Figure 2 below. Refinement zones were also generated for both the parent and the release corridor. These assist with increasing the grid resolution in these areas of interest without significantly increasing the overall cell count. The grid size along the boundary of the overset region was matched to the uniform grid size spacing in the release corridor as required by the overset method [7,8]. A distance-weighted interpolation scheme was used to interpolate between the background and overset grids.

The CFL (Courant-Friedrichs-Lewy) number, which determines stability of the method and controls the speed of convergence, was set in carriage models at 50. The implicit integration scheme had 2nd order temporal accuracy where the time-step was set at 10 μ s. The number of inner iterations (pseudo-time steps) was set at 5. Convergence was monitored through both dependent variable residuals and the store force and moment coefficient histories. A settling time of 10 ms was included, prior to releasing the store, to ensure that the store force and moment coefficient values had sufficiently converged. Since the store ejectors were not modelled in this simulation, the 6DOF solver was initialised using the linear and angular velocities measured by the CTS at the end of the ejector stroke. Since the store has been constrained until the end of the carriage time, the aerodynamic forces on the store were introduced linearly over 5 ms once the carriage time had expired. The method is used to reduce any impact loads encountered by the store.

Due to the significant computation requirement of time-accurate simulations, the model was run over 6 days across 12 networked i7 quad-core machines (*i.e.* 48 cores). This resulted in 0.19 s of full-scale release time after the end of the nominal ejector stroke. This represents about a third of the remaining release time available from the experimental data. Further work will include a modelling running to 0.5 s, which is expected to take around 18 days. Much is to be said of managing the grid size since this has a significant effect on length of the simulation run. Providing the optimal grid resolution such that the necessary shock waves are captured whilst keeping the grid cell count low are critical elements in ensuring the simulation time is practical. In addition, selecting the optimal number of the quad core machines requires a thorough benchmarking exercise which is planned for future work.

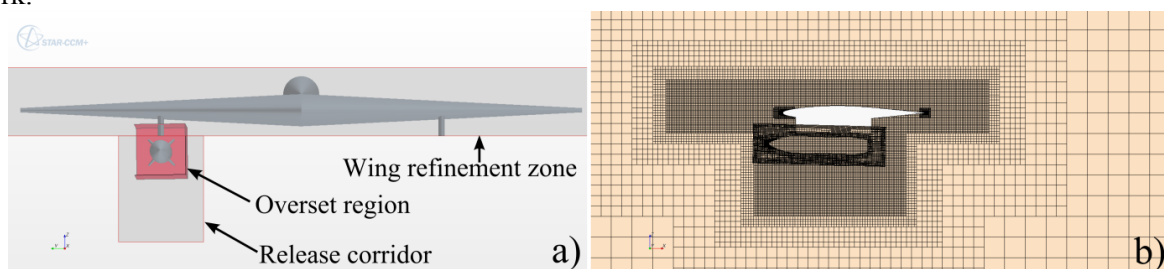


Figure 2 a) Background and parent grids. b) Mesh cross-section through the centre line of right wing's pylon. Note the parent and release corridor increased grid resolution in the refinement zones. Store is orientated to match attitude recorded by the CTS in the wind tunnel experiment.

3. Results

Figure 3 below illustrates the variation of the store force coefficients during carriage, including the force ramping time. This plot clearly indicates the requirements for the establishment of flow in the carriage configuration, since the forces take nearly 1000 time steps to reach approximately steady-state behaviour. By the end of this time, the side force coefficient C_Y is close to, but has not reached, a steady value, and therefore, in future models, this time will be at least doubled. Spikes are observed in

the normal force coefficient C_N , and the side force coefficient C_Y , at the start of the ramp time. This was not expected, since the ramping of the force was used to reduce impact loading on the store. However, these spikes dissipate quickly and do not appear to have a significant on the results that follow.

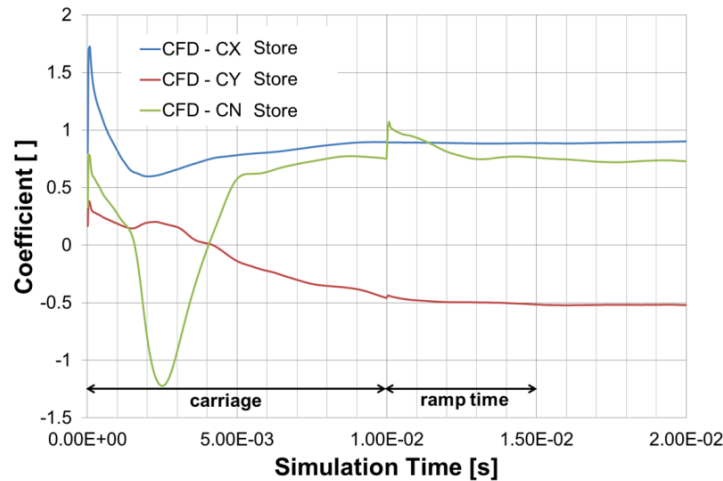


Figure 3 Store force coefficients variation during carriage and force ramping periods.

Figure 4 below shows the comparison between the trajectories recorded by the CTS and the computational result. The correlation between experiment and computation is excellent. It is noted that there are larger discrepancies for the z component of the trajectory. It is quite likely that this is a result of a small difference in the position of the store measured by the CTS just after the firing of the ejectors, and that used to initialize the store position in the simulation. It is also important to note that though it is possible to initialize the linear and angular velocities of the store after the firing of the ejectors, it is not possible to do so accurately in the surrounding highly transient flow field. Further work will require the addition of the ejectors to capture the unsteady flow field correctly.

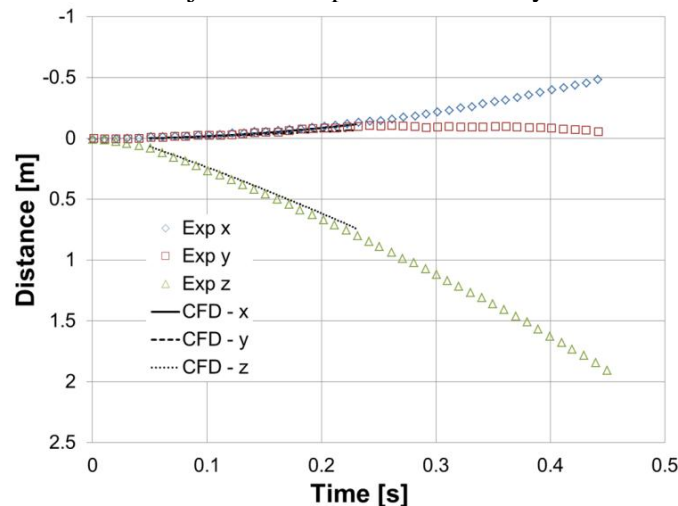


Figure 4 Computed store trajectory compared relative to pseudo-steady CTS wind tunnel results.

Figure 5 below illustrates the store force and moment coefficient histories plotted against the experimental results. Figure 5a illustrates a good correlation between the experimental and computational results despite the initial spike in the force coefficients due to the ramping (as described above). Interestingly, the axial force coefficient found in the simulations is in good agreement with the

experimental value despite the simulation being inviscid in nature, indicating that wave and form drag, rather than viscous drag, may be the primary effects. Since the sting is not included in the simulation, it is expected that the drag force would be higher than the wind tunnel configuration with the sting. This is most likely since the arrangement in the simulation results in a larger wake region, thereby increasing drag. Should the sting have been included in the simulation the axial force coefficient should be expected to be lower than the experimental results since viscous forces are not accounted for in the current simulation scheme.

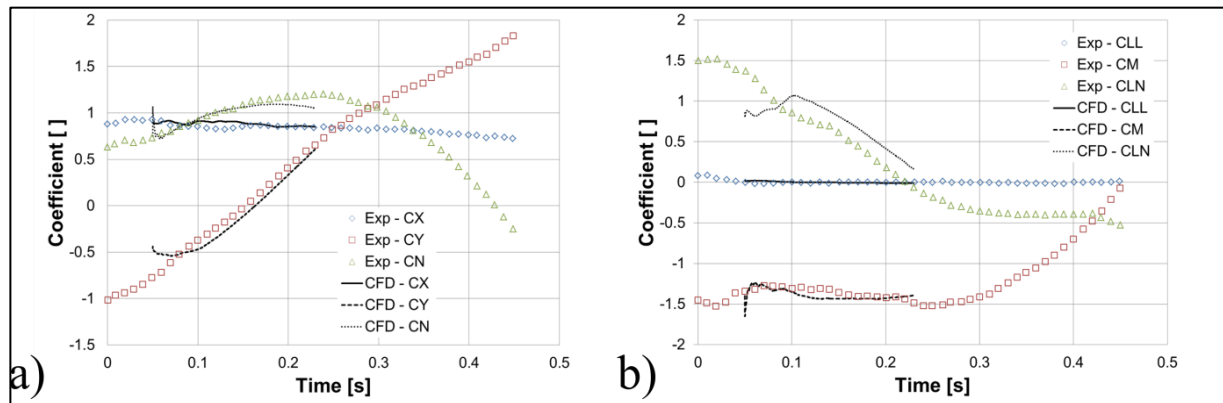


Figure 5 Store force and moment comparison with the wind tunnel results. a) Store force coefficient history, b) Store moment coefficient history.

The moment coefficients histories are illustrated above in Figure 5b. These also show relatively good agreement except for the yawing coefficient CLN. The computational results exhibit a nearly constant offset from the experimental results. This again may be a property of the initialization. Despite the flow outboard from the wing centre line of the delta wing, the rolling moment coefficient of the store is nearly zero, as seen in both sets of results. Agreement with the pitching moment CM is relatively good. This is a key indicator for flight clearance since a strong negative pitching moment, as seen in this case, could result in the rear of the store colliding with the parent as the store pitches downward. This curve illustrates the necessity for ejectors to displace the store from the parent during release.

It is important to note that in this work a comparison is made between pseudo-steady results from the wind tunnel CTS experiment, and a fully time-accurate unsteady computational approach. Therefore, full agreement between the two, since the experiment, with time steps of the order of 200 μ s, does not capture the unsteady flow field, as calculated computationally (and observed in test flight). Therefore, differences between the two sets of results should be expected.

Figure 6 below illustrates the complex flow field developed on the underside of the wing of the parent during release of the store. Two shocks, L_1 and L_2 , develop on the fore sections of the pylon. On the aft section of the pylon two additional shocks P are found on either side of the aft sections of the pylon. Interestingly, the shock P on the outboard section of the pylon is preceded by an expansion (see red contour adjacent to P). Note that there is a mutual interference between the store and pylon where numerous shocks are visible on the underside of the store. These shocks impinge on the fins of the store and they have significant effects on the yawing and pitching moment of the store during the early stages of release. Figure 6 also illustrates a wake region developing behind the pylon which interacts with the wing trailing edge shock W. The transient wake region tends to disrupt the shape and magnitude of the wing shock W. In addition, it is found that there is a transient interaction between the aft pylon shocks P and this wake region. Further work is required to quantify this transient behaviour on the loading the store experiences and establish whether it ultimately affects the store's trajectory.

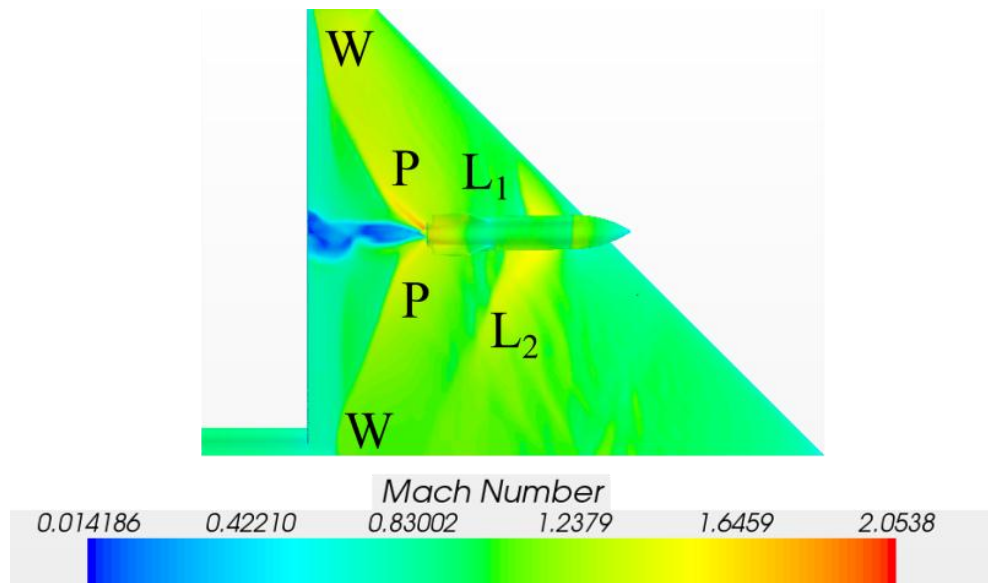


Figure 6 Contours of Mach number for the underside of the wing and the store during release stage.

4. Conclusions

Preceding simulations to the current work of the parent alone indicated that the flow beneath the wing is highly transient thereby necessitating the need for the current work of performing time-accurate time modelling of release. The time-accurate 6DOF release of a store from a parent was successfully conducted in Star-CCM+. A good level of agreement between computation and experiment was observed for both the trajectory and the store loads. However, there were deficiencies in the method which require further work. This includes modelling the ejectors and finding methods to minimize the total grid count while maximising the resolution of the shocks in the system which have significant effects on the store loads. The potential of adaptive gridding will be explored in future work since it would be ideal for this case. In addition, a benchmarking exercise is necessary to minimise the run time of the simulation on a cluster of machines. Further analysis of the flow field under the wing is also necessary in order to quantify the effect of these transient flow fields on the trajectory of the store.

References

- [1] Cenko A 1992 Store separation lessons learned during the last 30 years. In *27th Int. Conf. of the Aeronautical Sciences*
- [2] Fox J H 2000 Generic wing, pylon, and moving finned store *Technical report, Defence Technical Information Center ADP010735*
- [3] Heim R 1991 CFD wing/pylon/finned store mutual interference wind tunnel experiment *Technical report Arnold Engineering Development Center*
- [4] Lijewski L E and Suhs N E 1994 Time-accurate computational fluid dynamics approach to transonic store separation trajectory prediction *J. Aircraft* **31** 886–891
- [5] Liou M S 1996 A sequel to AUSM: AUSM+ *J. Comput. Phys.* **129** 364–382
- [6] Venkatakrishnan V 1995 Convergence to steady state solutions of the euler equations on unstructured grids with limiters *J. Comput. Phys.* **118** 120–130
- [7] Steger J L, Dougherty F C and Benek J A 1983 *A Chimera Grid Scheme, Advances of Grid Generation*, ed K N Ghia and U Ghia *American Society of Mechanical Engineers, ASME FED-5* (New York: ASME) 55-69
- [8] Benek J A, Steger J L, Dougherty F C 1983 A Flexible Grid Embedding Technique with Applications to the Euler Equations *American Institute of Aeronautics and Astronautics Paper 83-1944*

Polarization alignment system for quantum key distribution

M Mariola¹, A Mirza¹ and F Petruccione^{1,2}

¹University of KwaZulu-Natal, Westville Campus, Durban, South Africa

²National Institute for Theoretical Physics (KZN), Durban, South Africa.

E-mail: mmspazio7610@gmail.com

Abstract. To share a secret message between a sender and a receiver it is necessary to encrypt the message using a particular key and related algorithm. Generally a public key has an associated private key. The private key is computationally secure unless the eavesdropper has a quantum computer. In quantum cryptography the transmitter and receiver share a private key used to encrypt and decrypt the message. In one implementation of quantum cryptography the bits of the key are sent as series of polarized single photons. If an eavesdropper is present in the channel, the receiver receives a different bits of the key, because after the measurements by the eavesdropper, by the Heisenberg's Uncertainty Principle, the quantum state of the single photons may change. Since the value of the single bit depends on the polarization states, the polarization basis of the transmitter and receiver must be aligned. In this paper various solutions for automatic alignment control of the polarization basis are shown.

1. Introduction

The RSA protocol is a method to share a secret key between the transmitter, conventionally called Alice, and the receiver conventionally called Bob [1]. This numerical method is still computationally secure due to the calculation time in order to extract the private key contained in the public key. In 1997 Shor showed that quantum computers are able to factorize the public key in polynomial time [2] and hence the RSA protocol can be violated. The protocol above is known as an asymmetric key system. Quantum cryptography is a symmetric key system, where the private key is used to encrypt and decrypt the message. The power of quantum cryptography is the capacity to recognize the presence of the eavesdropper, conventionally called Eve, during the transmission of the key. The idea exploits the Heisenberg's Uncertainty Principle.

Quantum key distribution was proposed in 1984 and the first protocol is known as BB84 [3]. Alice sends to Bob a private key as a series of single polarized photons. The protocol uses two nonorthogonal basis, and the polarization state of the single photons represents the value of the bit of the private key. The basis are tilted 45 degrees. Alice, for each single photon, randomly chooses the polarization base and the value of the quantum bit. Bob randomly choose the measurement basis for the single photons. One basis is indicated with the symbol $+$ and represents the vertical and horizontal polarization. The photon sent with the vertical polarization (\uparrow) corresponds to the bit 1, and the photon sent with the horizontal polarization (\rightarrow) corresponds to 0. The second base, indicated with the symbol \times is tilted by 45 degrees with respect to the first one. The photon polarized with the direction \nwarrow corresponds to the bit 1, otherwise the bit sent with the direction of polarization \nearrow is 0. Once the key is transmitted,

Table 1. Alice sends to Bob the key using different basis. The bits are chosen when the basis of Alice and Bob are the same.

Alice	Base	+	+	×	+	×	×	+
	Quantum bit	→	↑	↖	↑	↗	↗	↑
	bit of the key sent	0	1	1	1	0	0	1
Bob	Base	+	×	×	×	×	×	+
	Quantum bit	→	↑	↖	↑	↗	↗	↑
	bit of the key received	0	1	1	N	0	1	1
Final key		0		1		0	1	1

Alice transmits via a public channel to Bob the basis used, and the final key is composed by the bits transmitted and received with the same basis. The process of the key exchange is represented in Table 1.

If the basis chosen are the same and the polarization base of the transmitter and the receiver are not aligned some, errors may still occur as indicated by the red bit **1** shown on the table 1. With the presence of a misalignment Alice and Bob may register the presence of Eve in the channel. In order to remove the errors the process called distillation is necessary. Bob should know the timing of the photon transmission and the basis must be aligned. The timing can be obtained by radio or optical synchronization [4]. The commercial systems for quantum cryptography use fiber optics as optical channel and can be used only when Alice and Bob are stationary. Quantum cryptography in free space permits to share a secret message between two stationary points or when Alice and Bob are two non stationary points.

2. Polarization alignment system using one laser beacon and one polarizer

In order to align the polarization basis of the transmitter and receiver, the system uses a vertically polarized laser beacon [5] and one polarizer. The wavelength of the laser beacon and the wavelength of the single photons are different. The polarization direction of the laser beacon is aligned with the polarization basis used by Bob for the quantum transmission. The polarizer is mounted on Alice's side, and the direction of polarization is aligned with the polarization basis of Alice for the quantum transmission. The system is shown in figure 1. Bob sends the polarised laser beacon to Alice, and Alice receives the signals affected by scintillation and wandering due to atmospheric turbulence. The turbulence effects on the polarization of the laser beacon can be neglected [6]. When a polarized laser beacon with the intensity I'_0 crosses a polariser, in output the intensity follows the Malus' law:

$$I_1 = I'_0 \cos^2(\theta), \quad (1)$$

where I'_0 is the intensity of the spot at the input of the polarizer and I_1 is the intensity of the spot at the output. θ is the angle between the direction of polarization of the laser and the direction of the polarization of the polarizer. When $\theta = 0$ It follows $I_1 = I'_0$. Since the intensity changes with the time, due the atmospheric turbulence, it is not possible to know in advance the value of I'_0 . As shown in figure 1, the value of I_1 can be obtained if the misalignment correspond to an angle θ_1 and exist an ambiguity of rotation between θ and θ_1 . In order to find the correct direction, the polarizer can be rotated until the maximum value of the intensity is measured. The Algorithm is shown in figure 2. The tracking system of Alice measures the intensity I_1 stored

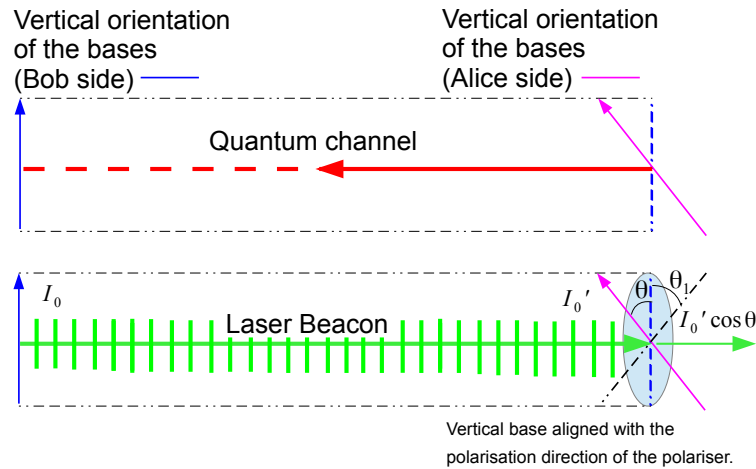


Figure 1. The blue dashed line and the green lines, correspond to the vertical direction of the laser beacon and the violet arrow correspond to the direction of the vertical polarization base of Alice. The vertical polarization direction of Alice is orientated in the same direction of the polariser of the alignment system. The black dashed line represents the ambiguity between the real angle of misalignment θ and the specular angle θ_1 . When the laser crosses the polarizer the intensity in output follows the equation (1) from which is possible calculate the angle of misalignment without information of the verse of rotation (θ or θ_1).

in the the variable $a1$. The system randomly chooses the direction of rotation, for example clockwise, and the power drive mechanically turns the basis (polarizer of the tracking system and the bases of the quantum channel). Alice measures the intensity of the laser beacon and store this value in the variable $a2$. If $a2$ is smaller than $a1$, Alice changes the verse of rotation to anticlockwise. If the condition is true the direction of rotation does not change from clockwise. This kind of system can not work properly when the atmospheric turbulence is strong. The basis can not be stable in one direction but an oscillation around the vertical position occurs. Increasing the sensitivity of the detectors the amplitude of the oscillations decrease and the system can be considered aligned also if a small oscillations are present. The experimental setup is shown in the figure 3. The side where the fixed polarizer is mounted represent the position of Bob. The fixed polarizer is used in order to create a polarized laser beacon. Initially the polarizers are not aligned and the shaft starts to rotate by the micro-controller command. For the experiment a stepper motor was used. The stepper motor is used in order to control the speed and the amplitude of rotation. At each step the micro-controller measures the intensity of the signal outcome from the rotating polarizer and decides the direction of rotation using the algorithm in figure 2. The system shown on the bottom can be modified as shown in figure 4, in order to calculate the angle θ . A beam splitter divides the signal in direction of the detector D1 and detector D2. The signal measured from the detector D1 can be used to calculate the initial intensity I_0 , and by the inverse of equation (1) it is possible calculate the angle θ . Since the ambiguity of the direction of rotation is still present, it is possible to follow the algorithm in figure 2. Instead of using the intensity, the test can be done on the angle θ . The angle θ can be calculated by the inverse of equation (1). The last configuration is the best way if the alignment system works in strong turbulent environment because the power I_0' received is known by the measurement in D1 after previous corrections.

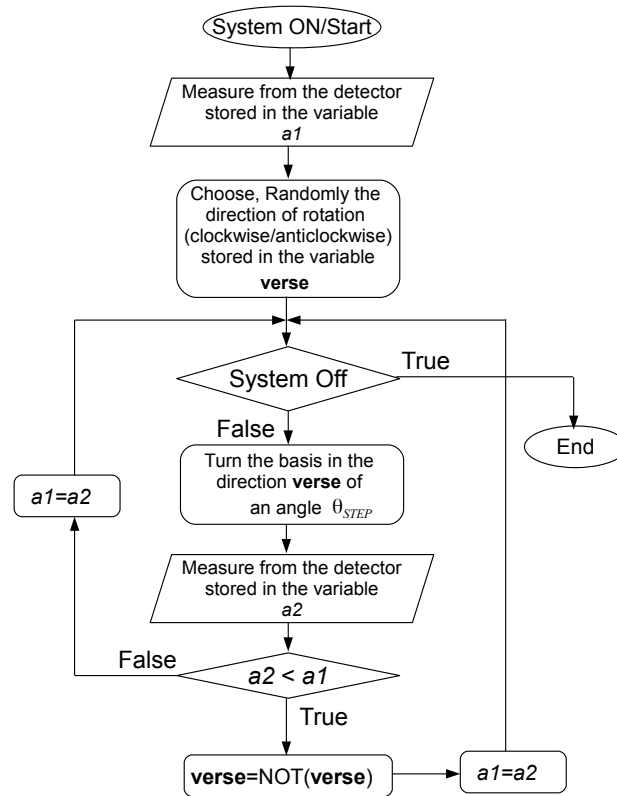


Figure 2. The flowchart shows the algorithm used to find the direction of the vertical polarization.

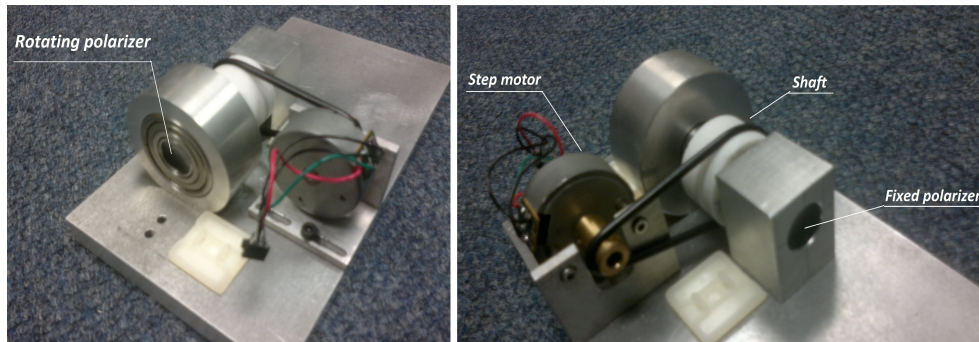


Figure 3. The experimental setup moves the rotating polarizer mounted on the shaft. The laser is incident on the fixed polarizer to create a reference signal. The signal is acquired from the detector fixed to the rotating polarizer. The micro-controller acquires from the detector the signal, and rotates the shaft until the polarizers are aligned.

3. Polarization alignment system using a polarizer beam splitter

The previous system is able to work when Alice and Bob are located at two fixed points. A recent patented system uses a polarising beam splitter as shown in figure 5 and is able to work in non-stationary conditions [7]. If the polarization of the beam is not aligned with the basis of Alice, the detector 1 receives a power $I_1(\theta) = I'_0 \cos^2 \theta$, where θ is the angle of misalignment. The detector 2 receives a power $I_2(\theta) = I'_0 \sin^2 \theta$. Theoretically we obtain the following curve in

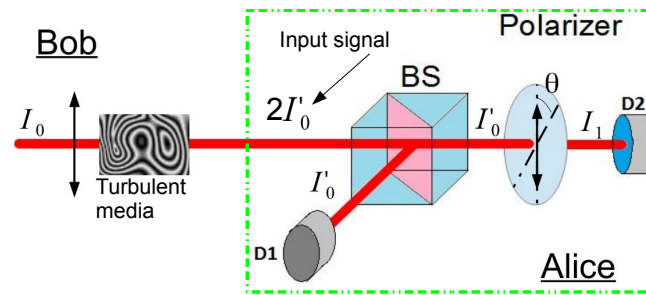


Figure 4. The beam is split at the beam splitter (BS), the value of I_0 can be determined from the value measured in detector D1. The angle can be calculated by the equation $\theta = \arccos \sqrt{\frac{I_1}{I_0}}$.

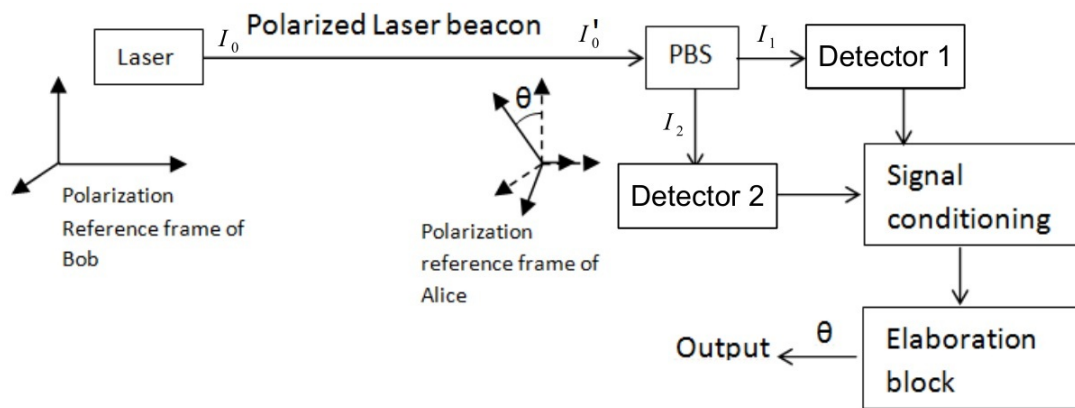


Figure 5. Patent pending application no. 2014/03405

function of the angle θ .

In order to align the system, the laser beacon is sent with a polarization of 45 degrees with respect to the vertical polarization. Alice receives the laser beacon and by the measurement of I_1 and I_2 is able to know the angle. The system is aligned when $I_1 = I_2$. It is possible to

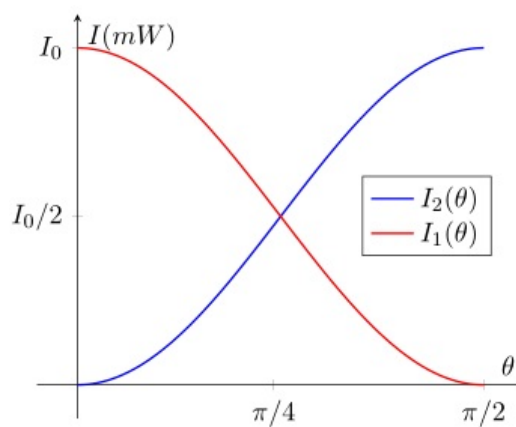


Figure 6. I_1 and I_2 from figure 5. The measurements of both channels make it possible to calculate the angle θ .

calculate the angle without ambiguity. Comparing the signals it is possible to know in which direction the systems should be rotated. This systems can work using a programmable logic unit or using the analog controls.

4. Conclusion

In this work a polarizer tracking system for quantum key distribution was proposed. The system proposed in the second section can be used when Alice and Bob are located at stationary points. This first system requires more effort with respect to the system proposed in the third section, because it continuously rotates the polarization basis around the exact alignment position. The system proposed in the third section works when Alice and Bob are located at stationary or non-stationary points. This is because this system calculates the “instantaneously” angle θ and the direction of rotation. The first system proposed was tested using a programmable logic unit (Atmega328P programmed by Arduino software). The third proposed system was tested with the micro-controller and with an analog electronic circuit. It was experimentally observed that the alignment system developed is relatively accurate for the purpose of free space QKD.

Acknowledgments

This work is based on research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation.

References

- [1] Bonavoglia P 2014 Il cifrario rsa URL <http://www.crittologia.eu/critto/rsa/rsa.html>
- [2] Shor P W 1997 *SIAM journal on computing* **26** 1484–1509
- [3] Bennett C H, Brassard G *et al.* 1984 *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* vol 175 (New York)
- [4] Mariola M, Abdul M and Francesco P 2011 vol ISBN:978-1-86888688-3 pp.403-408 (South African Institute of Physics)
- [5] Mariola M, Abdul M and Francesco P 2012 (South African Institute of Physics)
- [6] Fante R L 1979 *Proceedings of IEEE* vol 63 pp 1669–1692
- [7] Mariola M, Mirza A and Petruccione F 2014 System and method for determining angles between apparatuses, devices or systems

Simulation of the Egyptian 2nd Testing Research Reactor (ETRR-2) Experimental Benchmark in aid of Verification and Validation of the OSCAR-4 System

¹ M Mashau, ² B Erasmus, ² RH Prinsloo, ¹ SH Connell

¹ Department of Physics, University of Johannesburg, P.O Box 524, Auckland Park 2006, South Africa.

² South African Nuclear Energy Corporation (Necsa), P.O Box 582, Pretoria 0001, South Africa.

E-mail: mauricemashau1@gmail.com

Abstract.

This paper investigates the applicability of the current OSCAR-4 code system to simulate the ETRR-2 reactor. Various modelling approaches are applied to this benchmark, in order to quantify the capability of neutronic modelling in OSCAR-4. In particular, the modelling of the control rod calibration experiments poses challenges to the traditional deterministic calculational path and alternatives are investigated to address these shortcomings. Results indicate that an improved homogenization approach directly impacts on the accuracy of the full core diffusion solution.

1. Background and motivation

The reliable and safe operation of a nuclear reactor highly depends upon the ability to accurately predict the neutron flux distribution, which is required to determine quantities such as power distributions, control rod worths, shutdown margins and isotopic depletion rates – quantities required throughout the reactor operation [5]. As a means of predicting the above mentioned quantities and safety limits during reactor operation, computational reactor codes become invaluable. Safety being the key issue, it is of importance to perform safety analysis of research reactors through validating reactor codes against experimental data. Recently, the IAEA (International Atomic Energy Agency) published a set of experimental benchmarks, including neutronics and thermal-hydraulic benchmarks for, amongst others, the ETRR-2 reactor [3]. Previously, such an expansive, collated set of results had only been published for power reactors, which limited the benchmark tests for research reactors to code-to-code comparison.

For the purpose of this work, ETRR-2 was of particular interest, since some of its core components are similar in design as compared to the SAFARI-1 reactor (South African Fundamental Atomic Research Installation) operated by Necsa. Therefore, ETRR-2 was chosen as a test case to strengthen the validation basis for calculating research reactors with the OSCAR-4 (Overall

System for the Calculation of Reactors, generation 4) code system. OSCAR-4 is currently utilized to simulate the SAFARI-1 reactor [6] at Necsa. OSCAR-4 is a nodal diffusion based code which is used to perform day-to-day reactor calculations in support of the SAFARI-1 reactor. As a means of improving the capabilities of OSCAR-4, there is a need to solve more benchmark problems in aid of verification and validation of the code system for research reactor applications. Verification means the precision of the code is tested against other standard codes for reactor analysis, whilst validation is comparing the accuracy of the code against experimental data.

OSCAR-4 is a deterministic code system, and its standard calculational path introduces a number of typical approximations in the modelling of the physical phenomena present during the operation of a reactor. As a result of these approximations within the code system, certain modeling scenarios, particularly in regions near strongly absorbing materials (such as control rods) pose challenges to these standard methods and careful adaptations should be considered. Deterministic code systems are typically used for reactor operation support calculations as their computer simulation time is significantly lower than Monte Carlo based full core solvers. Hence, the main focus of this study is on the modelling approach of the ETRR-2 control rods, and then in particular on how to establish or propose a better model to simulate control rod worth experiments. From the experimental benchmarks published for ETRR-2 [3], Core SU-29 was chosen as the base core configuration for simulating the proposed experiments.

A full core OSCAR-4 model was built and tested for its precision using a more accurate Monte-Carlo based code (SERPENT [4]), as a reference model. Furthermore, using a newly developed link between OSCAR-4 and SERPENT, the existing calculational path can be modified or improved by generating cross sections from SERPENT. As a way of resolving the challenge, associated with modelling the ETRR-2 control rods, a modified OSCAR-4 model was built using cross sections generated from SERPENT for selected core components. These models were compared with experimentally measured data and the results will be discussed. The SERPENT code is thus used both as full-core reference solver, and cross section generation tool.

2. Theoretical background and OSCAR-4 overview

In order to determine the neutron flux distribution in a reactor core, the neutron transport equation must be solved, which describes the movement, distribution and interaction of neutrons with matter. In typical reactor simulation tools, calculations are performed based on steady-state conditions, assuming that the system evolves slowly with time in such a way that the time-independent solution of the neutron transport equation can be used to predict the required quantities during reactor operation.

Below is the neutron transport equation [1] and each symbol has its standard meaning in reactor analysis.

$$\begin{aligned} \Omega \cdot \nabla \Phi(\mathbf{r}, E, \Omega) + \Sigma_t(\mathbf{r}, E, \Omega) \Phi(\mathbf{r}, E, \Omega) = & \int dE' \int d\Omega' \Sigma_s(\mathbf{r}, E' \rightarrow E, \Omega' \rightarrow \Omega) \Phi(\mathbf{r}, E', \Omega') \\ & + \chi(E) \int dE' \int d\Omega' \nu \Sigma_f(\mathbf{r}, E', \Omega') \Phi(\mathbf{r}, E', \Omega') \quad (1) \end{aligned}$$

Equation 1 is derived based on mechanisms which are responsible for neutron production and neutron loss. The first term describes the net leakages, the second term describes the removal of neutrons due to absorption or scattering. The third term describes the gain of neutrons due to in-scattering from all other energy E' and angles Ω' and the last term is the source term.

There exist two classes of transport solution methods to calculate the neutron flux distribution for reactor analysis. These methods are classified as Stochastic and Deterministic.

The scale of the problem is often too large to solve Equation 1 directly with all six independent variables ($r=(x,y,z)$, $\Omega = (\theta, \phi), E$). Deterministic based methods solve Equation 1 by discretizing each independent variable to form a set of algebraic equations that can be solved numerically [1]. The most commonly-employed approach approximates the solution of Equation 1 by solving the multi-group diffusion equation.

$$-\nabla D_g^n \nabla \Phi_g^n(\mathbf{r}) + \Sigma_{t,g}^n \Phi_g^n(\mathbf{r}) = \sum_{g'=1}^G \Sigma_{g' \rightarrow g}^n \Phi_{g'}^n(\mathbf{r}) + \chi_g \sum_{g'=1}^G \nu \Sigma_{f,g'}^n \Phi_{g'}^n(\mathbf{r}) \quad (2)$$

Here Equation 2 is written in discretized form for n nodes, assuming constant cross sections per node (core components can be treated as nodes).

The diffusion approximation assumes that the angular flux distribution is at most linearly anisotropic and scattering is isotropic[1] (not valid in certain regions). For a given core component or mesh, 2D detailed neutron transport calculations (solving equation 1) in fine energy group structure are performed, to obtain the heterogeneous flux shape. The flux shape is then used as a weighting factor, to produce accurate flux-volume weighted cross sections and equivalence parameters (this technique is collectively known as spatial homogenization and energy group condensation)[2]. These equivalence parameters must be defined in such a way that component-average reaction rates and surface-averaged net leakages are preserved when the multi-group diffusion equation is solved[5].

The success or failure of this approach is often found in the quality of the 2D transport solution, as well as the relevance of the boundary conditions used to perform such calculations. Applying Equation 2 to perform full core simulations in 3D does not necessarily weaken the solution, since cross sections and associated equivalence parameters act as correction factors in regions where diffusion approximation falls apart.

In this work we apply the OSCAR-4, which is the current version of the OSCAR code system. To generate flux-volume weighted cross sections and equivalence parameters, a collision probability transport code, HEADE, is used to solve a 2D fine-group problem for a given core component. HEADE is based on a low-order response matrix formalism, applying isotropic partial current boundary conditions on internal boundaries. As such it is quite applicable to the modelling of fuel elements (since fission dominate components have near-isotropic angular flux distributions). However, due to its approximations, HEADE is not necessarily suited to components such as control rods and/or reflectors. As an alternative to HEADE, SERPENT is used as a 2D transport code, to improve the quality of the transport solution needed to generate cross sections and required parameters. These cross sections are then tabulated against state parameters (such as Fuel Temperature, Burn-up, Moderator density, etc) for later use. This is a once-off process and the results for all core components are linked into a single run-time cross-section library. The linked cross sections and associated parameters are then passed to the global diffusion solver, Multi-Group Reactor Analysis Code (MGRAC). MGRAC uses the Multi-Group Analytic Nodal Method (MANM)[8], to solve the multi-group diffusion equations for a 3D full core model.

3. Experiments and model description

ETRR-2 is a 22 Mega-Watt pool-type research reactor. It is fuelled with low enriched uranium (19.7 %), cooled, moderated with light water and reflected by beryllium. In this paper we focus on the ETRR-2 benchmark experiments, in particular the control rod worth or calibration experiments. To perform these experiments, a reactor core is adjusted to be critical at low power to avoid feedback effects. The reactor is made super-critical by extracting the rod to be calibrated by a certain distance and the period method is used to measure the reactivity. The core configuration is again adjusted to be critical by inserting one of the rods not being calibrated into the core, to compensate for the change in reactivity. These steps are repeated until the calibrated rod is fully extracted from the core. In this regard, a typical outcome of such an experiment, and associated calculation, is the differential rod worth curve for each calibrated rod. Additionally, the so-called control rod s-curve, used to characterise the absorbing capability of the rod as a function of extraction position, may be produced.

The aim of this work then is to quantify the improvement provided on this benchmark by addressing two particular shortcomings identified in the standard deterministic calculational path as applied in OSCAR-4. These are:

- The shortcomings in HEADE to calculate the reference transport solution for non-fuel components
- The impact of utilizing, for non-fuel components, explicit component environments as boundary conditions for cross-section generation, as opposed to the standard approach of typical generic, mini-core environments

To facilitate this investigation, three models were built to simulate rod worth experiments. Their description is summarized in the following table.

Table 1. Models

Model	Description
Reference (SERPENT)	Full core 3D heterogeneous Monte-Carlo calculations with SERPENT.
Standard OSCAR-4	HEADE-based cross sections were generated in 6 energy groups (e.g Fuel, reflectors, control rods and water box). For fuel components, cross sections are generated from an infinite environment with reflective boundary conditions, and for non-fuel components, cross sections are generated from a mini-core environment with fuel acting as a driver zone.
Modified OSCAR-4	For the selected non-fuel components, cross sections were generated in 6 energy groups from SERPENT. environment. A full-core 2D SERPENT model was utilized to generate cross-sections, with the 2D model representing a case where all control rods are inserted. Fuel element cross-sections were generated using HEADE, utilizing the standard model.

A note is needed here. The construction of the modified OSCAR-4 model, described in Table 1 above, has particular significance with regard to practical reactor core calculations. The usage of infinite (or reflective) boundary conditions for the cross-section generation models of fuel elements is quite typical, and in actual fact preferable, as compared to utilizing their explicit environments (given their respective core locations). In other words, one generic environment for all fuel is often a more practical option than unique cross-sections per fuel element, particularly since fuel elements are move around and shuffled throughout the core during their lifetimes. Non-fuel elements however, do not generally move around the core structure from cycle to cycle and as

such more detailed environmental conditions may be utilized during the cross-section generation process for such components. For this reason both the Standard and Modified OSCAR-4 models utilize infinite environment for fuel elements, but differ with respect to both the code (HEADE vs SERPENT) and environmental accuracy (approximate 2D vs full-core 2D), respectively.

4. Results and discussion

Table 2. Results

	k_{eff} rods in	k_{eff} rods out	Critical k_{eff}	Max power error	Rod 5 total worth (\$)	Rod 6 total worth (\$)
Reference (SERPENT)	0.86311	1.02542	1.00294 ± 38.8 pcm	–	2.89	1.03
Standard OSCAR-4	0.86489	1.03160	1.01254 ± 128 pcm	4.94%	3.28	0.95
Modified OSCAR-4	0.85560	1.02028	0.99850 ± 50 pcm	2.73%	3.06	1.05
Experimental value	N/A	N/A	1.000	N/A	2.17	0.86

Table 2 provides an overview of the primary performance of each of the three models considered. In particular, we compare the two OSCAR-4 models (Standard and Modified) with the reference SERPENT results. We compare the models to one-another with respect to k_{eff} for the all rods in and out cases, as well as for maximum assembly power errors (here for the rods in case). Furthermore, we compare all three models to experimental values for criticality prediction and control rod worth experiments (for rods 5 and 6). Note that a large number of critical cases were analysed and as such results are given in terms of average predicted k_{eff} and standard deviation.

In particular we notice that both OSCAR-4 models predict the k_{eff} of the all rods in and out cases to around 500 pcm of the reference SERPENT result, but that the maximum power error is noticeably improved for the Modified OSCAR-4 model. This indicates that the Modified model captures the local effects within the reactor core to a greater degree of accuracy. This is expected, since the HEADE code may not correctly capture the sharp gradients which occur at the interface between, for example, fuel and control elements, during cross-section generation, whereas the Monte Carlo based SERPENT code should resolve these regions accurately.

With respect to the experimental comparisons, we notice that the estimation of criticality ($k_{eff} = 1$) by the two OSCAR-4 models once again show a marked improvement for the case of the Modified OSCAR-4 model. The critical estimation improves from $k_{eff}=1.01254$ (error of 1254 pcm) to 0.99850 (error of 150 pcm). The standard deviation for the modified model is also significantly lower. The comparisons with the two control rod calibration experiments show a more complicated behaviour. The predicted total control rod worth of the Modified OSCAR-4 model is closer to the SERPENT result for both the rod 5 and rod 6 calibration experiments, as compared to the Standard OSCAR-4 model. However, all models overestimate the total control rod worth of both rods 5 and 6 as compared to their respective experimental values of 2.17\$ and 0.86\$. Although this could indicate some bias in all the models, it is also possible that some normalisation, such as β_{eff} or prompt neutron lifetime, used to convert doubling times into reactivity values, may not have been reported correctly by the benchmark supplier. Further investigation into this issue revealed that other institutions who have calculated this experiment noted similar concerns, while calculating similar control rod worth values as given here [7].

Nevertheless, the accuracy of the Modified OSCAR-4 model as compared to the full-core SERPENT solution again shows a noticeable improvement. In particular for rod 5, we notice that the Modified model improves the total rod worth estimate by more than 20 cents. To illustrate this in more detail,

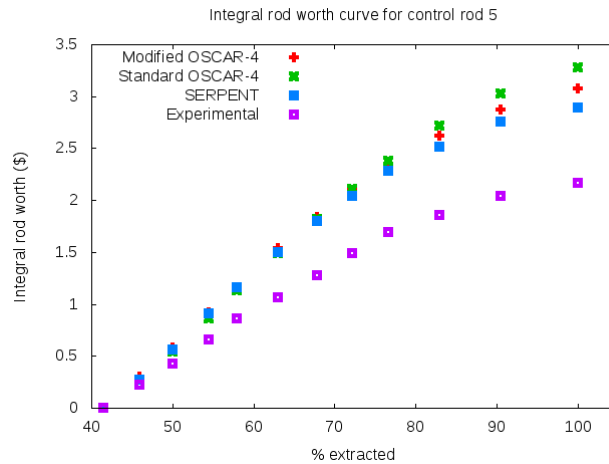


Figure 1. Integral rod worth curve for control rod 5 calibration

Figure 1 shows the control rod 5 calibration curves for the reference SERPENT model, the Standard OSCAR-4 model and the Modified OSCAR-4 model as well as the experimental results. Here it is once again clear that the Modified model improves upon the Standard model, but also that some general offset exists between all the models and the experimental curve. As noted before, further investigation into the experimental procedure is needed to resolve this.

5. Conclusions

From the results presented above, it can be seen that the Modified OSCAR-4 model shows improvement in most of the calculated parameters, with the exception of the k_{eff} value for the case with all rods in the core. Furthermore, the Modified model shows a more stable prediction of k_{eff} values for critical reactor configurations. This shows that the improved transport solution in the non-fuel elements (specifically the control rods), and hence improved cross sections and equivalence parameters, impact greatly on the accuracy of the full core diffusion model, as expected. In the case of ETRR-2, these improvements are required to predict the reactor state to an acceptable degree of accuracy.

An important result to note, is the fact that the use of fuel cross sections and equivalence parameters, generated from an infinite lattice environment, yields acceptable results. This is important from an operational point of view, as position dependent parameters need not be generated for fuel elements. During reactor operation, the core configuration changes from cycle to cycle, meaning fuel elements are not bound to a specific position as the non-fuel elements are. If position dependent parameters were required for fuel elements as well, it would be required to generate these parameters for each new core configuration, making the current approach not feasible.

Finally, the good agreement between the Modified OSCAR-4 model and SERPENT reference results, makes the ETRR-2 benchmark (core SU-29) a good case to be added to the verification set of the OSCAR-4 system. This shows that the current OSCAR-4 system, with the link to

SERPENT, is capable of simulating ETRR-2 with comparable precision, for the chosen parameters, to a full-core Monte Carlo transport simulation. As far as validation is concerned, further investigation into the experimental results is required to determine the accuracy of the models. This can be done by simulating different core configurations and their associated experiments with both the Modified model as well as SERPENT, and again comparing the results with the supplied experimental results.

- [1] Duderstadt J.J, Hamilton L.J 1976 *Nuclear Reactor Analysis*, John Wiley & Sons New York, N.Y
- [2] Hébert A 2009 *Applied Reactor Physics*. Presses Internationales Polytechnique
- [3] IAEA 2013 *Proceedings Series on (CRP 1496) "RR Benchmarking Database: facility description and experiments"* ETRR-2 Nuclear Reactor Experimental Data.
- [4] Leppänen J 2012 *Psg2/Serpent-A Continuous-Energy Monte Carlo Reactor Physics Burn-up Calculation Code* (VTT Technical Research centre of Finland)
- [5] Smith K.S 1986 *Progress in Nuclear Energy*, 17 303
- [6] Stander G, Prinsloo R.H, Müller E.Z, Tomašević D.I 2008 *International Conference on Reactor Physics, Nuclear Power: A sustainable Resource* (Interlaken, Switzerland)
- [7] Villarino E.A, Lecot C, Enany A 1999 *In proceedings of the Meeting of the International Group on Research Reactors (IGORR'99)* (Bariloche, Argentina)
- [8] Vogel D.L, Weiss Z.J 1992 *International Topical Meeting on Advances in Reactor Physics* (Charleston, South Carolina, USA)

Analysis of a thermal conductivity measurement technique formulated as an inverse heat conduction problem

Vishal Ramnath

Department of Mechanical and Industrial Engineering, University of South Africa (UNISA), Private Bag X6, Florida 1710, Gauteng Province, South Africa

E-mail: ramnav@unisa.ac.za

Abstract. Thermal analysis and solution of heat problems most often utilizes known thermal conductivity material data which is typically experimentally determined from heat flux measurements through the application of Fourier's law. The challenge posed by this approach is the need for known thermal conductivity reference materials which may be inhomogeneous and have large associated uncertainties in industrial physics applications. In this paper we investigate the feasibility of developing a thermal conductivity measurement system that utilizes known radiometric input sources and temperature output measurements which may have smaller relative uncertainties by formulating the system as an inverse heat conduction problem utilizing recently reported research results from the fields of geophysics and mathematical optimization.

1. Introduction

Thermal physics analysis and solution of heat problems often utilizes known thermal conductivity material data which is typically experimentally determined from heat flux measurements through the application of Fourier's law $q'' = -k\nabla T$, where $q''/[\text{W m}^{-2}]$ is the heat flux, $k/[\text{W m}^{-1} \text{K}^{-1}]$ is the thermal conductivity, and $T/[\text{K}]$ is the absolute temperature following the nomenclature in [1]. The use of Fourier's law presents a straight forward mechanism to define and infer the thermal conductivity through the ratio of the heat flux measured using standard techniques as discussed in [2], and the temperature gradient for a specified direction of heat flow where the unknown thermal conductivity may be expressed in terms of known reference quantities such as thermal conductivity, wall thickness and wall temperature amongst other experimental measurements for homogeneous isotropic materials as illustrated in figure 1.

Unfortunately materials such as aluminium/steel thermal properties may be inhomogeneous and/or non-isotropic either from physical effects in the smelting process or manufacturing effects from the fabrication of structures requiring knowledge of thermal conductivity such as reflectors/concentrators in for example solar plants [6] where in addition the thermal conductivity may vary with temperature [7, 8] as illustrated in figure 2 for a selection of materials.

As a result it is desirable in applied industrial research contexts to infer thermal conductivity information directly by an inverse problem formulation using direct temperature measurements which are more practical and experimentally convenient, either when reference thermal

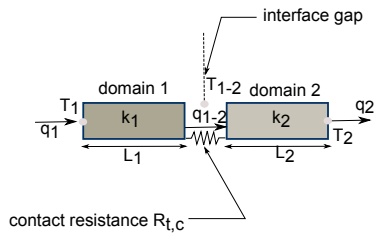


Figure 1. Composite wall system modelled as a 1D thermal circuit. In this model the unknown thermal conductivity k_2 is expressed as $k_2 = \frac{L_2}{A} \left(\frac{T_1 - T_2}{q} - \frac{L_1}{k_1 A} - R_{t,c} \right)^{-1}$ and the uncertainty may be estimated by standard techniques such as the GUM and GS1 as discussed in [3, 4, 5]

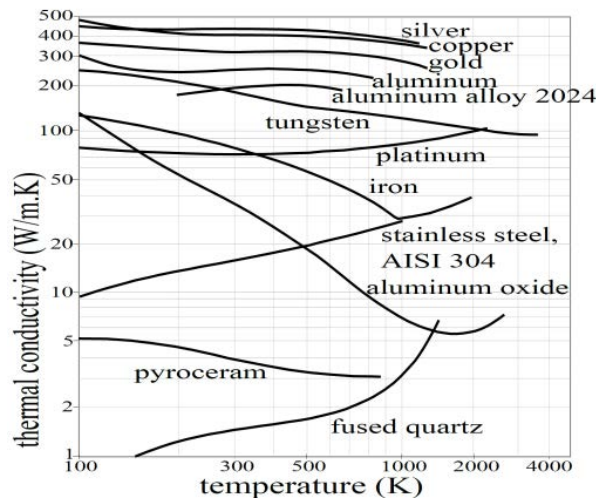


Figure 2. Illustration of thermal conductivity temperature dependence for selection of industrial/engineering materials (Graphic source: [1, page 47])

conductivity materials are unavailable for prior testing, impractical due to prevailing operating conditions, or when in situ process or condition monitoring measurements are necessary.

In general accurate heat sources such as calibrated lasers or electrical resistance heating elements, and temperature measurement devices such as thermocouples or resistance thermometers, are simple and straightforward to procure and utilize. As a result in this paper we opt to analyze a system in which the thermal conductivity may be inferred through an Inverse Heat Conduction Problem (IHCP) modelling approach utilizing just radiometric/electrical heating sources and temperature measurements which are easily accessible and which avoids the need for specific reference thermal conductivity material components, and specialist heat flux instruments and blackbody equipment [9].

2. Literature Review

Within the statistical literature simulation problems may typically be classified as either direct/forward or indirect/inverse where in general terms the former are cases in which PDE parameters are known and one utilizes this information to solve for the PDE solution, whilst in the latter one utilizes the PDE solution to infer the underlying PDE parameters.

In the context of thermal physics a direct/forward problem would correspond to using known thermophysical properties such as thermal conductivity and/or specific heat capacity as inputs with suitable boundary conditions for a boundary $\Gamma = \partial\Omega$ to solve the heat diffusion equation $\nabla \cdot (k \nabla T) + \dot{q} = \rho c_p \frac{\partial T}{\partial t}$ for some problem domain Ω whilst an indirect/inverse problem would in a certain sense “work backwards” to use the PDE temperature field solution to infer the corresponding thermophysical parameters as discussed in [10]. Traditionally inverse problem

studies have been common in the geophysical sciences to determine rock densities, gravitational field strengths, and in oil and gas exploration as discussed in [11, 12, 13] with newer insights into Monte Carlo based inverse parameter uncertainties reported in [14], and which has generally been formulated and solved using the Monte Carlo and Levenberg-Marquardt techniques as outlined in [15, 14].

For IHCP studies methods of solution have mainly to date traditionally consisted of the Levenberg-Marquardt and conjugate gradient techniques for mixtures of problems in laboratory thermophysical testing and space vehicle atmospheric reentry design and research studies after the development of techniques such as Tikhonov regularization which were able to modify and reformulate the original ill-posed/unstable inverse problems as approximate well-posed problems.

More recently investigations in inverse theory problems across different fields has involved applications of newly developed techniques such as the variational iteration method or (VIM) [16], the method of fundamental solutions or MFS [17, 18, 19, 20, 21, 22, 23, 24, 25], the lattice Boltzmann method or LBM [26], heuristic approaches drawing from a mixture of techniques [27, 28], and meshfree approaches [29]. Finite difference methods for IHCP studies are further discussed in [30, 31] who explore how the Tikhonov regularization method may be used to regularize the ill-conditioned linear system of equations for a non-steady two dimensional heat conduction problem, and a iteration approach to determine the regularization parameter was explored in [32]. A measurement methodology for infrared thermography of inverse models that used a maximum entropy principle was reported in [33] for a die forging application in order to deduce the unknown boundary condition of a heat flux using a finite difference discretization.

3. Mathematical Development

Based on the literature review we opt for simplicity since the underlying problem is nonlinear to use a Levenberg-Marquardt optimization in our formulation to avoid a Tikhonov regularization. Utilizing the heat diffusion equation $\nabla \cdot (k \nabla T) + \dot{q} = \rho c_p \frac{\partial T}{\partial t}$ which reduces to a generalized two dimensional Poisson equation $\frac{\partial}{\partial x}(k \frac{\partial T}{\partial x}) + \frac{\partial}{\partial y}(k \frac{\partial T}{\partial y}) = -\dot{q}$ in the special case for steady-state conditions for a planar domain for the particular geometry of the experimental system illustrated in figure 2 where the domain Ω is a circular region of diameter D where the boundary Γ is held at a constant temperature T_f , where $T/[K]$ is the temperature, $k/[W m^{-1} K^{-1}]$ the thermal conductivity of the sample, and \dot{q} is an energy sink/source term which is positive if energy is generated within the medium and negative if energy is being consumed.

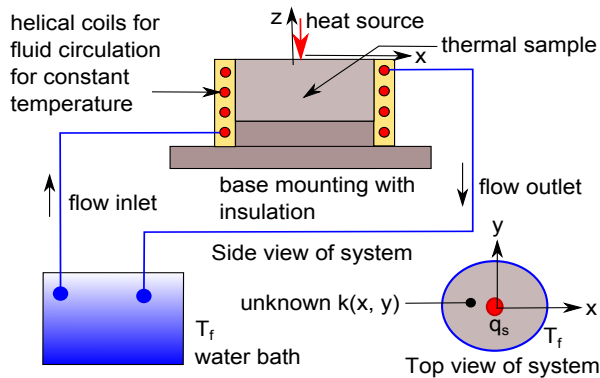


Figure 3. Experimental inverse heat conduction problem (IHCP) system to test for thermal conductivity

Adopting a radial basis function discretization following the discussions in [34, 35, 29] we approximate the thermal conductivity as $k = \sum_{i=1}^{N_k} \alpha_i \varphi_k(\|\mathbf{x} - \mathbf{x}_i\|)$, and the temperature field as $T = \sum_{j=1}^{N_T} \beta_j \varphi_T(\|\mathbf{x} - \mathbf{x}_j\|)$. For the temperature field let N_T be the total number of points composed of N_I interior points and N_B boundary points such that $N_T = N_I + N_B$. In this

approach it is not necessarily the case that $N_k = N_T$ however this will simplify subsequent calculations. The terms α_i ($i = 1, \dots, N$) and β_j ($j = 1, \dots, N$) are coefficients that are used to build up the $k(\mathbf{x})$ and $T(\mathbf{x})$ fields for $\mathbf{x} \in \Omega$ in terms of the RBF's φ_k and φ_T respectively. Utilizing Gaussian RBF's for simplicity to illustrate the approach adopted of form $\varphi_k = e^{-\varepsilon_k^2 r^2}$ and $\varphi_T = e^{-\varepsilon_T^2 r^2}$ where $r = \sqrt{(x - x_i)^2 + (y - y_i)^2}$ and where ε_k and ε_T are suitable shape parameters for the thermal conductivity and temperature fields respectively, and substituting into the generalized two dimensional Poisson equation we then have using the short hand $r_i = r$ for points in the interior that $0 = \frac{\partial k}{\partial x} \frac{\partial T}{\partial x} + k \frac{\partial^2 T}{\partial x^2} + \frac{\partial k}{\partial y} \frac{\partial T}{\partial y} + k \frac{\partial^2 T}{\partial y^2} + \dot{q}$ where $k = \sum_{i=1}^{N_k} [\alpha_i e^{-\varepsilon_k^2 r_i^2}]$, $\frac{\partial k}{\partial x} = \sum_{i=1}^{N_k} [-2\varepsilon_k^2 (x - x_i) \alpha_i e^{-\varepsilon_k^2 r_i^2}]$, $\frac{\partial^2 k}{\partial x^2} = \sum_{i=1}^{N_k} [-2\varepsilon_k^2 \alpha_i \{1 - 2\varepsilon_k^2 (x - x_i)^2\} e^{-\varepsilon_k^2 r_i^2}]$, and similar expressions for the corresponding temperature field.

Let T_k be the experimental measured temperature and θ_k the estimated temperature for some choice of assumed parameter \mathbf{P} , where in our particular problem the parameter is used to specify the thermal conductivity. Constructing an objective function $S(\mathbf{P}) = \sum_{k=1}^N [T_k - \theta_k(\mathbf{P})]^2$ it follows that we solve the IHCP by determining the choice of parameter such that $S(\mathbf{P})$ is minimized. The objective function may be written in matrix form by specifying the measured and estimated temperatures as a column vectors $\mathbf{T}^\top = [T_1, \dots, T_N]$ and $\boldsymbol{\theta}^\top = [\theta_1, \dots, \theta_N]$ and then constructing the objective function as $S(\mathbf{P}) = [\mathbf{T} - \boldsymbol{\theta}(\mathbf{P})]^\top [\mathbf{T} - \boldsymbol{\theta}(\mathbf{P})]$. A solution will occur when $\nabla S(\mathbf{P}) = 0$ which after algebraic manipulations reduces to the following equations

$$-2\mathbf{J}^\top(\mathbf{P})[\mathbf{T} - \boldsymbol{\theta}(\mathbf{P})] = \mathbf{0}$$

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \theta_1}{\partial P_1} & \frac{\partial \theta_1}{\partial P_2} & \frac{\partial \theta_1}{\partial P_3} & \cdots & \frac{\partial \theta_1}{\partial P_M} \\ \frac{\partial \theta_2}{\partial P_1} & \frac{\partial \theta_2}{\partial P_2} & \frac{\partial \theta_2}{\partial P_3} & \cdots & \frac{\partial \theta_2}{\partial P_M} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \theta_N}{\partial P_1} & \frac{\partial \theta_N}{\partial P_2} & \frac{\partial \theta_N}{\partial P_3} & \cdots & \frac{\partial \theta_N}{\partial P_M} \end{bmatrix}$$

where as previously mentioned N is the total number of measured/estimated temperatures, M is specified as the total number of parameters such that $\mathbf{P}^\top = [P_1, \dots, P_M]$, and $\mathbf{J}(\mathbf{P}) = [\frac{\partial \boldsymbol{\theta}^\top(\mathbf{P})}{\partial \mathbf{P}}]^\top$ is the corresponding Jacobian matrix of the objective function system. For linear problems the unknown parameter to be determined may be calculated as $\mathbf{P} = (\mathbf{J}^\top \mathbf{J})^{-1} \mathbf{J}^\top \mathbf{T}$ however for a nonlinear inverse problem it may be observed that the Jacobian matrix, also known as the sensitivity matrix, has a functional dependence on the parameter \mathbf{P} and as a result must be solved iteratively using a local linearization. In order for a suitable iteration procedure to converge it is necessary that $|\mathbf{J}^\top \mathbf{J}| \neq 0$ otherwise the IHCP is ill-conditioned. Whilst Tikhonov regularization is possible to solve a resultant ill-conditioned linear system $\boldsymbol{\Lambda} \boldsymbol{\theta} = \mathbf{B}$ of form $\boldsymbol{\theta}_\alpha = [\boldsymbol{\Lambda}^\top \boldsymbol{\Lambda} + \alpha(\mathbf{R}^{(s)})^\top]^{-1} \boldsymbol{\Lambda}^\top \mathbf{B}$ where $\mathbf{R}^{(s)}$ for $s = 0, 1, 2$ is a Tikhonov regularization matrix following the overview presented in [31] where α is the regularization parameter.

Different approaches are possible for estimating α and whilst there is no definitive method the L -curve method appears common within the mathematical literature but is difficult to implement. To avoid these difficulties we opt to use the Levenberg-Marquardt method which is also amenable for potentially ill-conditioned systems and which has the advantage of being well established with existing implementations [36] that are relatively straightforward to implement and which incorporate newer mathematical optimization techniques such as the homotopy approach [37].

Based on our RBF discretization set $N_k = N_T$ to simplify the algebra when constructing the corresponding Jacobian matrix and note that our approach then specifies the parameters as $\mathbf{P}^\top = [\alpha_1, \dots, \alpha_M]$ and the estimated temperature field as $\boldsymbol{\theta}(\mathbf{P})$ as the solution of the generalized Poisson equation as previously indicated. The algorithm to implement the IHCP

formulated thermal conductivity measurement technique may now be conveniently summarized as follows:

- (i) specify the spatial coordinates (x_i, y_i) for the points in the interior Ω and for the points (x_b, y_b) on the boundary Γ
- (ii) assign the associated measured interior temperatures as T_i along with associated heat source/sink terms \dot{q}_i taking note of the known boundary temperatures $T_f(x_b, y_b)$
- (iii) construct the Jacobian matrix \mathbf{J} either symbolically/numerically using the assumed values from the RBF parameterization
- (iv) solve the nonlinear inverse problem using the Levenberg-Marquardt method for the given measured temperature data and assumed parameter values and iterate until convergence to the required accuracy level is achieved
- (v) the associated parameter uncertainties $u(\alpha_i)$ ($i = 1, \dots, M$) in the RBF formulation of the thermal conductivity field may be constructed from Monte Carlo simulations and statistical post-processing of the results using simulation inputs obtained by sampling from the measured temperature values and their associated reported uncertainties
- (vi) the parameter uncertainties may then be used as inputs in a GUM/GS1 uncertainty quantification calculation for the thermal conductivity uncertainty at any particular spatial coordinate within the domain

4. Discussion

In this paper we have investigated how to use an IHCP formulation as a means to determine thermal conductivity values for materials from experimental temperature measurements and numerical simulations. Potential benefits of the method that have been investigated are that there is no need for specialist laboratory reference thermal conductivity material standards, the method can be applied in industrial/plant environments with standard equipment/instruments with minimal physical complexity, and that the numerical algorithm to implement the method may be utilized to accurately relate and estimate the thermal conductivity uncertainties in terms of the input experimental data.

References

- [1] Incropera F P and DeWitt D P 1996 *Fundamentals of Heat and Mass Transfer* 4th ed (John Wiley & Sons)
- [2] Tsai B K, Gibson C E, Murthy A V, Early E A, DeWitt D P and Saunders R D 2004 Heat flux sensor calibration Tech. rep. National Institute of Standards and Technology nIST Special Publication 250-65 (37 pages)
- [3] JCGM/WG1 2008 Evaluation of measurement data – guide to the expression of uncertainty in measurement (JCGM 100:2008) Tech. rep. BIPM/IEC/IFCC/ILAC/ISO/IUPAC/IUPAP/OIML
- [4] JCGM/WG1 2008 Evaluation of measurement data – supplement 1 to the “guide to the expression of uncertainty in measurement” – propagation of distributions using a Monte Carlo method (JCGM 101:2008) Tech. rep. BIPM/IEC/IFCC/ILAC/ISO/IUPAC/IUPAP/OIML
- [5] JCGM/WG1 2011 Evaluation of measurement data – supplement 2 to the “guide to the expression of uncertainty in measurement” – extension to any number of output quantities (JCGM 102:2011) Tech. rep. BIPM/IEC/IFCC/ILAC/ISO/IUPAC/IUPAP/OIML
- [6] Hasuike H, Yoshizawa Y, Suzuki A and Tamaura Y 2006 *Solar Energy* **80** 1255–1262
- [7] Rodriguez-Sanchez M R, Soria-Verdugo A, Almendros-Ibanez J A, Acosta-Iborra A and Santana D 2014 *Applied Thermal Engineering* **63** 428–438
- [8] Powell R W, Ho C Y and Liley P E 1966 Thermal conductivity of selected materials Tech. rep. National Institute of Standards and Technology (previously NBS) <http://www.nist.gov/data/nsrds/NSRDS-NBS-8.pdf>
- [9] Miklavac A, Pusnik I, Batagelj V and Drnovsek J 2013 *Measurement Science and Technology* **24** 1–8
- [10] Ozisik M N and Orlande H R B 2000 *Inverse Heat Transfer – Fundamentals and Applications* (Taylor and Francis)
- [11] Mosegaard K and Tarantola A 1995 *Journal of Geophysical Research* **100** 12431–12447

- [12] Mosegaard K and Sambridge M 2002 *Inverse Problems* **18** R29–R54
- [13] Mosegaard K 2011 *Geophysics* **76** W51–W61
- [14] Fernandez-Martinez J L, Fernandez-Muniz Z, Pallero J L G and Pedruelo-Gonzalez L M 2013 *Journal of Applied Geophysics* **98** 62–72
- [15] Pujol J 2007 *Geophysics* **72** W1–W16
- [16] Geng F and Lin Y 2009 *Computers and Mathematics with Applications* **58** 2098–2102
- [17] Hon Y C and Wei T 2004 *Engineering Analysis with Boundary Elements* **28** 489–495
- [18] Hon Y C and Wei T 2005 *CMES Comput Model Eng Sci* **7** 119–132
- [19] Jin B and Marin L 2007 *International Journal for Numerical Methods in Engineering* **69** 1570–1589
- [20] Johansson B T, Lesnic D and Reeve T 2011 *International Journal of Computer Mathematics* **88** 1697–1713
- [21] Johansson B T, Lesnic D and Reeve T 2012 *International Communications in Heat and Mass Transfer* **39** 887–895
- [22] Karageorghis A, Lesnic D and Marin L 2011 *Inverse Problems in Science and Engineering* **19** 309–336
- [23] Marin L, Karageorghis A and Lesnic D 2011 *Engineering Analysis with Boundary Elements* **35** 342–354
- [24] Mierzwiczak M and Kolodziej J A 2011 *Inverse Problems in Science and Engineering* **19** 777–792
- [25] Reeve T and Johansson B T 2013 *Engineering Analysis with Boundary Elements* **37** 569–578
- [26] Kameli H and Kowsary F 2012 *International Communications in Heat and Mass Transfer* **39** 1410–1415
- [27] Liao W, Dehghan M and Mohebbi A 2009 *Journal of Computational and Applied Mathematics* **232** 351–360
- [28] Frackowiak A, Botkin N D, Cialkowski M and Hoffman K H 2010 *International Journal of Heat and Mass Transfer* **53** 2123–2127
- [29] Rashedi K, Adibi H, Amani J and Parand K 2014 *Engineering Analysis with Boundary Elements* **40** 1–21
- [30] Pourgholi R, Rostamian M and Emamjome M 2010 *Inverse Problems in Science and Engineering* **18** 1151–1164
- [31] Pourgholi R, Esfahani A and Abtahi M 2013 *Journal of Applied Mathematics Computation* **41** 61–79
- [32] Qian Z and Feng X 2013 *Inverse Problems in Science and Engineering* **21** 467–484
- [33] Lair P, Dumoulin J and Millan P 1998 *Numerical Heat Transfer, Part A: Applications: An International Journal of Computation and Methodology* **33** 267–277
- [34] Fasshauer G E 2007 *Meshfree Approximation Methods with MATLAB* (World Scientific Publishing)
- [35] Fasshauer G E and Zhang J G 2007 *Numerical Algorithms* **45** 345–368
- [36] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 2007 *Numerical Recipes – The Art of Scientific Computing* 3rd ed (Cambridge University Press)
- [37] Burden R L and Faires J D 2001 *Numerical Analysis* 7th ed (Brooks/Cole)

A quantum circuit modeling toolkit for high performance computing

Makhamisa Senekane¹, Bheki Zulu¹ and Francesco Petruccione^{1,2}

¹ Centre for Quantum Technology, School of Chemistry and Physics, University of KwaZulu-Natal, P/Bag X54001 Durban, South Africa

² National Institute for Theoretical Physics (NITheP_KZN), P/Bag X54001 Durban, South Africa

E-mail: 211560527@stu.ukzn.ac.za, 210556489@stu.ukzn.ac.za, petruccione@ukzn.ac.za

Abstract. Theoretically, quantum computers are known to solve a certain class of problems more efficiently than their classical counterparts. This is due to parallelism which is inherent in quantum algorithms. However, a full-scale quantum computer has not been realised as yet. Therefore, in order to validate and debug quantum circuits, a classical computer is used. Since most of these circuits are simulated using personal computers (PCs), quantum circuits with a limited number of quantum bits (qubits) can only be simulated, due to computational limitations of PCs. In this work, we report the simulation of quantum circuits for a high performance platform using message passing interface for the Python (mpi4py) package.

1. Introduction

Quantum computation uses quantum mechanical principles such as entanglement, superposition and tunneling to process information. It dates back to the ideas of Manin [1], Feynman [2] and Bernioff [3] in the 1980s. It was mainly motivated by the limitations imposed on the conventional classical computation by the fundamental laws of physics. Additionally, the first universal quantum machine was proposed by Deutsch [4] in 1985. However, a major breakthrough came in 1994, with the work of Shor [5], who showed that a quantum computer can be used to break some of the conventional cryptographic algorithms. Another contribution was by Grover [6] two years later, who invented a quantum search algorithm which offers a quadratic speed-up over the best known classical search algorithm. Since then, there has been a concerted effort to develop more quantum algorithms for different applications [7, 8, 9].

In a classical computer paradigm, computer scalability can be governed by Moore's law [10]. However, as the transistors get smaller, the quantum effects would be too pronounced, and Moore's law would cease to apply. Therefore, since Moore's law does not apply anymore in the quantum regime, scalability of quantum computers cannot be determined by Moore's law. On the contrary, scalability in the quantum regime is dependent on the ability to isolate the quantum system (hence mitigate decoherence) and the advancement of fault-tolerant quantum computation mechanisms.

Although quantum computation offers many interesting computational capabilities, the major drawback is that a scalable quantum computer is yet to be realized. So, in order to debug and validate quantum algorithms, a classical computer is used. Furthermore, the Gottesman-Knill theorem [11], makes it possible for the classical simulation of some quantum circuits, since it

states that a certain class of quantum circuits, known as stabilizer circuits, can be simulated efficiently on a classical computer. Many classical simulators of quantum circuits have been simulated [12]. A comprehensive list of these simulators can be found in Ref. [13]. It can be observed that most of the simulators developed thus far are intended for a stand-alone personal computer (PC). Some attempts have been made to develop simulators that make use of high performance computing (HPC) [14, 15, 16]. We report in this work a quantum circuit modeling toolkit which is intended for use on a high-performance computing platform. This simulator is developed using Python programming language, which ease of use and rapid development. Additionally, Python is freely available, with an open-source licensing. Finally, since Python is a vectorized language, it is suitable for high performance scientific computing. A python module that is used to develop this toolkit is message passing interface for Python (mpi4py).

The remainder of this manuscript is arranged as follows. The next section provides a background information on quantum computation and high performance computing. Section 3 discusses the implementation details of the simulator. It is followed by Section 4, which discusses the simulation results obtained using this simulator. The last section concludes this manuscript.

2. Background Information

2.1. Quantum Computation

As opposed to a classical computer, which has a binary digit (bit) as its basic unit of information, a quantum computer has a quantum bit (qubit) as its unit of information. Additionally, a bit, which can only exist in state 0 or 1, a qubit can exist in the superposition of states $|0\rangle$ and $|1\rangle$, where states $|0\rangle$ and $|1\rangle$ are known as computational basis states. These computational basis state can be represented as column vectors, where

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (1)$$

and

$$|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2)$$

Mathematically, a qubit can be represented as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad (3)$$

where α and β , which are known as probability amplitudes, satisfy the condition

$$|\alpha|^2 + |\beta|^2 = 1. \quad (4)$$

Analogous to their classical counterparts, the building blocks of quantum circuits are quantum gates. However, unlike some of the classical gates, which are not reversible, all quantum gates are reversible. This owes to the fact that basically, a quantum gate is any unitary operation, and unitaries are invertible, hence reversible. Some of the common quantum gates are

a) Hadamard gate, H

$$H = 1/\sqrt{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

b) Phase shift gate, R

$$R = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}.$$

2.2. High Performance Computing

Although it is generally believed that classical computers cannot efficiently simulate universal quantum computers, it is nevertheless possible to classically simulate some of the quantum circuits, especially if such circuits are small (using few qubits). Because a stand-alone PC has limited memory and computing speed, and computational resources for simulating quantum circuits on a classical computer grow exponentially with an increase in input qubits, a PC puts huge limitations on the quantum circuits that can be efficiently simulated. These limitations can be overcome by classical simulation of quantum circuits on a high performance computing platform. HPC deploys the aggregation of computing power among different nodes so as to yield high overall computational power. It makes use of supercomputers and parallel processing schemes to solve compute-intensive problems [17].

High performance computing can be implemented using either of the two computing architectures, namely shared memory or distributed memory architectures. In the former architecture, all the processors have access to the entire memory of the machine. In contrast, in the latter architecture, each machine has a processing element (PE), which contains its processor and its memory, where each processor could only access the memory in its PE.

2.3. Message Passing Interface

MPI is an industry-wide standard protocol for passing messages between parallel processors [18] though generally, it can also run on shared memory machines. It is a *de facto* standard for parallel programming on distributed memory architecture systems [19]. Its goals are high performance, scalability and portability. Furthermore, message interface passing supports both point-to-point and collective communication.

It is worth noting that MPI is in itself neither an implementation nor a language, but rather a specification or an application programmer interface (API). MPI has a number of components such as user-defined datatypes, communication ports, communication operations and communication contexts [20, 21]. Typical communication operations include broadcast, reduce, scatter and gather [21].

2.4. Message Passing Interface for Python

One of the implementations of MPI standard for Python programming language is mpi4py [22]. It provides an object oriented approach to message passing which provides MPI bindings for Python. In the work presented in this manuscript, mpi4py package was used because it offers a number of benefits. These benefits include: clean and efficient MPI interface for Python, extensible and compatible implementation and easy of use.

3. Implementation

The toolkit was implemented using a single physical host, with a dual-processor PC at the clock frequency of 2.7 Gigahertz per processor, with a random access memory (RAM) of 4 Gigabytes. However, three virtual machines (VMs) were created and clustered together on this physical host machine, with each VM having a RAM of 1 Gigabytes. Different quantum circuits were then simulated, using various virtual nodes. Some of the circuits simulated include Einstein-Podolsky-Rosen (EPR) pair generation circuit, Greenberger-Horne-Zeilinger (GHZ) state preparation circuit and a three-qubit quantum Fourier Transform circuit [7, 12]. The QFT circuit was further expanded to simulate up to nine qubits, so as to make a more realistic comparison with the PC-based results obtained in [12]. The simulations were implemented using Python programming language, using mpi4py for MPI implementation.

Quantum gate operations can include either multiplication of state vectors with unitary matrices or tensor product for composite systems. As an example, consider the EPR pair generation circuit in figure 1. Also, consider that the two inputs are $|a\rangle = 0$ and $|b\rangle = 0$ respectively. Then

this can be mathematically represented as

$$\begin{aligned} \text{epr_generation}(|00\rangle) &\rightarrow 1/\sqrt{2} * \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \left[\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] \\ &= 1/\sqrt{2} * \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}. \end{aligned}$$

This result can be naturally decomposed into smaller tensor matrices, which can then be distributed over different processors for HPC [14, 23, 24]. A similar approach was taken for other quantum circuits. The simulation results obtained are provided in the next section.

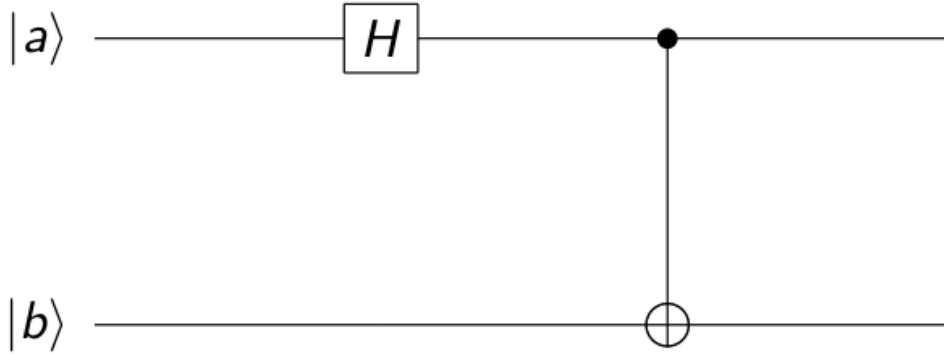


Figure 1. A circuit diagram of EPR pair generation circuit. It consists of two inputs $|a\rangle$ and $|b\rangle$, together with two quantum gates (Hadamard and control NOT (CNOT) gates).

4. Evaluation

Efficiency tests were run for this high performance computing quantum circuit simulation toolkit. Three quantum circuits were used, and the execution times of each circuit with respect to the number of processes were recorded. The timer functionality *Wtime()* was used to record the execution times. Figures 2, 3 and 4 show the results obtained for EPR pair generation, GHZ state generation and three-qubit QFT circuits respectively. The results obtained underline the utility of the simulator for quantum circuit simulation, since an increase in the number of processes used results in a significant decrease in execution time. Furthermore, the results demonstrate the feasibility of a portable, parallel simulation toolkit for quantum computing, written in Python. A measure that is employed for the utility of the simulator reported in this manuscript is the speed-up as compared to the stand-alone PC. This is given as:

$$\text{speedup} = \frac{T_{\text{serial}}}{T_{\text{parallel}}}, \quad (5)$$

where T_{serial} is the execution time for a stand-alone PC given in [12], while T_{parallel} is the execution time obtained in figures 2, 3 and 4. From the results, the speed-up of the EPR pair generation circuit can be up to 3.4, that of GHZ circuit can be up to 3.2 while that of three-qubit QFT can be up to 2.5. These speed-ups provide a good support for the utility of this quantum circuit modeling toolkit.

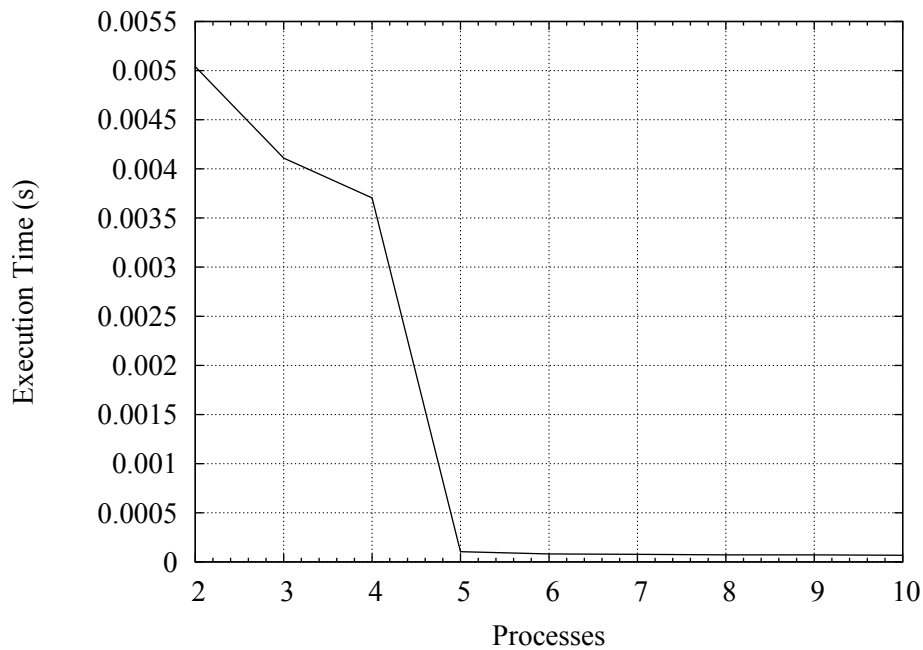


Figure 2. Execution time for EPR pair generation circuit as a function of number of processes.

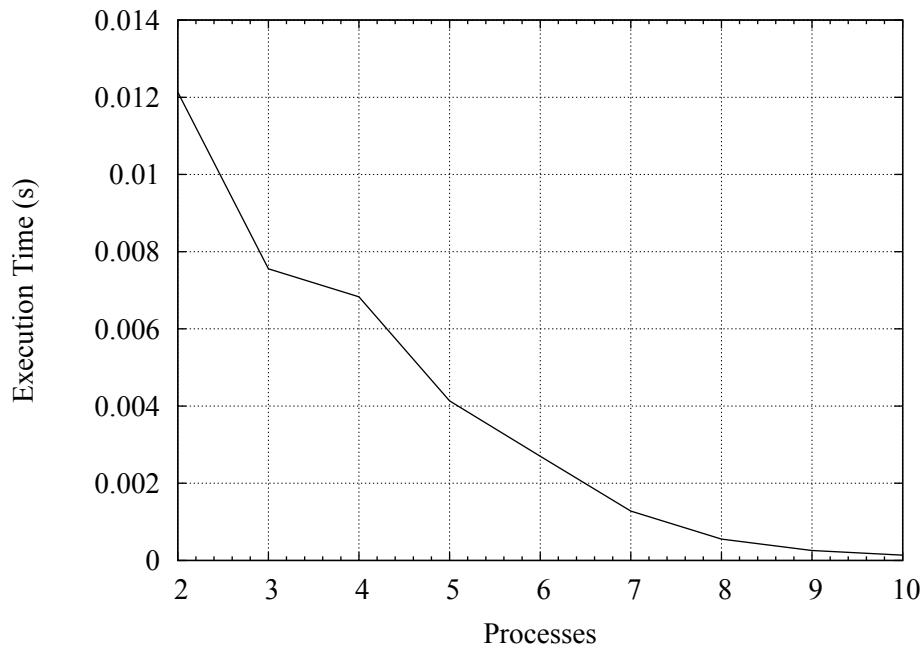


Figure 3. Execution time for GHZ state generation circuit as a function of number of processes.

Finally, coupled with the speed-up mentioned above, since this simulator uses Python programming language, it can be easily extended to simulate other quantum circuits, with a very rapid development time. Additionally, since Python is one of the three leading programming languages in high performance scientific computing (the other two being C/C++ and Fortran), coupled with the fact that it is the only one of the three that is vectorized, it makes it a good candidate for quantum computing simulation toolkit. This circuit can then be used to simulate different type of quantum circuits mentioned in this manuscript.

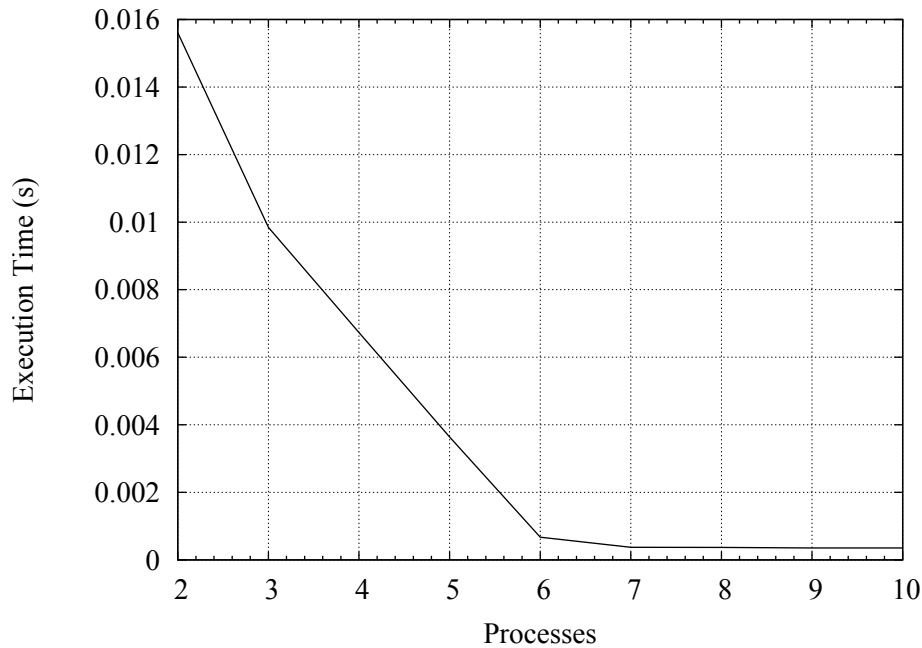


Figure 4. Execution time for three-qubit QFT circuit as a function of number of processes.

It is worth noting that this current simulator in its current form is limited in terms of functionality, since very few quantum circuits can be efficiently simulated. Additionally, the focus is only on the increase in the number of processes in order to effect a speed-up. The work is still on-going to increase the number of quantum circuits that can be simulated using this toolkit, and the running of tests on a physical (as opposed to virtual) HPC platform with multiple number of nodes and communication delays between the nodes.

5. Conclusion

In this manuscript, we have reported a quantum circuit modeling toolkit for HPC. This toolkit was evaluated using standard stabilizer circuits. The simulation results confirm the utility of the toolkit. However, this simulator still has some limitations. The quantum circuit that can be simulated by this toolkit are very few. This is because the work presented in this manuscript is still in progress. The next focus will be on the optimization of the simulator code, in order to make it extensible and scalable. Additionally, it (future work) will focus on extending the simulated circuits to any generalized quantum digit circuit, and implementing such circuits on a physical HPC platform.

Acknowledgments

This work is based on research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation.

References

- [1] Knill E, Laflamme R, Barnum H, Dalvit D, Dziarmaga J, Gubernatis J, Gurvits L, Ortiz G, Viola L and Zurek W 2002 *Los Alamos Science* **27**
- [2] Feynman R P 1982 *International journal of theoretical physics* **21** 467–488
- [3] Benioff P 1980 *Journal of Statistical Physics* **22** 563–591
- [4] Deutsch D 1985 *Proceedings of the Royal Society of London A. Mathematical and Physical Sciences* **400** 97–117

- [5] Shor P W 1994 *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on* (IEEE) pp 124–134
- [6] Grover L K 1996 *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing* (ACM) pp 212–219
- [7] Nielsen M A and Chuang I L 2010 *Quantum computation and quantum information* (Cambridge University Press)
- [8] Childs A M and Van Dam W 2010 *Reviews of Modern Physics* **82** 1
- [9] Bacon D and VAn DAm W 2010 *Communications of the ACM* **53** 84–93
- [10] Moore G E *et al.* 1998 *Proceedings of the IEEE* **86** 82–85
- [11] Gottesman D 1998 *arXiv preprint quant-ph/9807006*
- [12] Senekane M, Mohapi L and Petruccione F 2013 *AFRICON, 2013* (IEEE) pp 1–4
- [13] “Quantiki” “2010 (Accessed: March 16, 2014)” “List of QC simulators” URL ‘ ‘http://www.quantiki.org/wiki/List_of_QC_simulators”
- [14] Obenland K M and Despain A M 1998 *arXiv preprint quantph9804039*
- [15] Niwa J, Matsumoto K and Imai H 2002 *Unconventional Models of Computation* (Springer) pp 230–251
- [16] De Raedt K, Michielsen K, De Raedt H, Trieu B, Arnold G, Richter M, Lippert T, Watanabe H and Ito N 2007 *Computer Physics Communications* **176** 121–136
- [17] Hager G and Wellein G 2010 *Introduction to high performance computing for scientists and engineers* (CRC Press)
- [18] Gropp W, Lusk E and Thakur R 1999 *Using MPI-2: Advanced features of the message-passing interface* (MIT press)
- [19] Jin H, Jespersen D, Mehrotra P, Biswas R, Huang L and Chapman B 2011 *Parallel Computing* **37** 562–575
- [20] Pacheco P S 1997 *Parallel programming with MPI* (Morgan Kaufmann)
- [21] Pacheco P 2011 *An introduction to parallel programming* (Elsevier)
- [22] Dalcín L, Paz R, Storti M and D’Elía J 2008 *Journal of Parallel and Distributed Computing* **68** 655–662
- [23] Patz G 2003 *A parallel environment for simulating quantum computation* Ph.D. thesis Massachusetts Institute of Technology
- [24] Steeb W H 1997 *Matrix calculus and Kronecker product with applications and C++ programs* (World Scientific)

The Alan Cousins Telescope - A Robotic Multi-Purpose Telescope for Sutherland

P van Heerden^{1,2}, P Martinez¹ and C Engelbrecht²

¹ South African Astronomical Observatory, Observatory Rd, Observatory, Cape Town 7925 South Africa

² Department of Physics, University of Johannesburg, Kingsway, Auckland Park, Johannesburg 2092, South Africa

E-mail: pierre@sao.ac.za

Abstract. Much work has been done on the Alan Cousins Telescope over the last few years in terms of testing and commissioning the telescope system and to prepare it for regular operation. Many problems were encountered and overcome in the process. The result is a system that is essentially ready for use by remote observers and that will soon be ready for regular, automatic operation.

1. Introduction

The Alan Cousins Telescope is a 0.75 – m multi-purpose, robotic and remote access telescope at the Sutherland site of the South African Astronomical Observatory. It has a modular photoelectric photometer, currently with a Johnson-Cousins UBVRI filter set. Unlike the other telescopes in Sutherland of a similar size, the ACT was designed and built to be controlled by a computer, thus eliminating the need for an observer to be present.

Efforts to bring the telescope system to a point where it can be used regularly either under automatic control of the computer or under manual control of a remote observer started in 2010 with the development of the first version of the ACT control software suite. In 2011 the control software suite was installed on the control computer in the dome along with a new Programmable Logic Converter (PLC) to remove many of the lower-level operational burdens from the computer - in particular communicating with and managing the various devices that form part of the telescope system (e.g. filter and aperture wheels, telescope focus, dome rotation, dome shutter).

During the testing and commissioning process, many hardware and software related issues were encountered - all of which (except for the most recent issues) have been solved, resulting in a system that is essentially ready for operation with a remote user, albeit with limited efficiency. Although a detailed discussion of all of these problems and their solutions fall beyond the scope of this document, a brief overview of the more recent hurdles and their solutions are given here.

2. Flopping Acquisition Mirror

As with most telescope systems, the ACT has a folding mirror (acquisition mirror) with a circular cut out in one half of the mirror in order to be able to direct the light within the telescope's

field of view (or part thereof) to either the science instrument or to a camera used to acquire the target (depending on the position of the mirror).

Some large offsets were occasionally encountered while moving the telescope to a star and centering it on the field of view. After some investigation it was noted that moving the acquisition mirror in and out of the beam caused the star field to shift by tens of arcseconds - much too large to be able to use the telescope efficiently and collect meaningful photometry.

This “flopping” effect was caused by the motorised screw mechanism used to move the acquisition mirror back and forth and the two rails on which the mirror moved not being sufficiently rigid. This issue was solved by tightening the screws holding the rails to the casing, introducing a third rail to reduce the torsion on the existing two rails and connecting the rails together at the free end.

3. Spurious Photons on PMT

While performing the first commissioning tests of the science instrument on the ACT (a photomultiplier tube), some sharp spikes were occasionally recorded in the photon counts detected by the system. The peaks were so strong, in fact, that the photon count rate reported by the PMT were very nearly at or greater than the maximum allowed by the control software, which causes the control software to close the science instrument shutter and move the acquisition mirror into the beam in an effort to protect the PMT from overillumination and eventual temporary or permanent damage.

At the time, the spikes in the count rate reported by the PMT was thought to be noise originating from the cables connecting the various devices that form part of the telescope, instrument and dome to the PLC that controls them. PMTs are typically quite susceptible to electronic noise from devices in the area. The spikes seemed to occur most frequently while moving the acquisition mirror. It was noted that unplugging the cable for the acquisition mirror from the PLC made the spikes disappear, however the system would not be usable if the acquisition mirror could not be articulated.

Many attempts were made by the SAAO electronic technicians to remove (or at least reduce) the noise, to the point where the cable was replaced completely. These attempts were unsuccessful, but further investigation revealed that the spikes were in fact not electronic noise, but real light emanating from the acquisition mirror. Opto-isolator switches are used on the rails of the acquisition camera to sense when the acquisition mirror is at the limit of the rails.

The opto-isolator switches on the acquisition mirror rails were replaced with Hall-effect sensors, which eliminated the noise problem.

4. Simplified Scheduler Programme

The scheduler programme is a component of the ACT control software suite which determines what observations in the queue to observe whenever a block of observations are completed. For the first version of the scheduler programme, it was decided to allow the user to schedule observing blocks on-the-fly. This caused many problems, as it meant that the scheduler programme needed to effectively multiplex between observing blocks in the queue and those scheduled by the user. It also required that the other components of the software suite be informed of whether or not an intelligent observer is present when the system is effectively performing an observation automatically - this is because different safety and sanity checks are performed by all software in the suite depending on whether or not an intelligent observer is present, which is determined based on whether a command is received from the scheduler programme or whether commands are given directly to the software components by the user.

This issue was resolved by significantly stripping down the scheduler programme, to the point where it was simply a programme that reads the details of observing blocks from a database and issues the necessary commands to complete the observations. In other words, the user

can no longer issue any commands to the scheduler programme, except to cancel the block of observations currently being completed or to modify the observing queue in the database, which is done with an external programme that is not connected to the control software suite.

5. Pointing and Tracking Errors

Pointing errors are caused by misalignments and manufacturing flaws in the mount. For telescopes with an equatorial mount (such as the ACT) it is quite common for the polar axis to be offset from the Earth's rotation axis. Another fairly common source of pointing errors is a non-perpendicularity between the right ascension and declination axes. Such errors make it difficult to find targets and so most telescope systems have a numerical model in the software that corrects for them. Depending on the regularity and size of the errors, between a handful and several dozen terms may be required in a pointing model that corrects for these pointing errors and anything between tens to thousands of data points may be required in order to construct a reliable model.

In the ACT's case, the errors are quite large (up to ca. $10'$), but fairly regular and stable - see figure 1. This means that a fairly simple pointing model should be able to correct for the pointing errors. However, given the size of the errors, misidentification of stars is not uncommon, which greatly impedes progress in this regard.

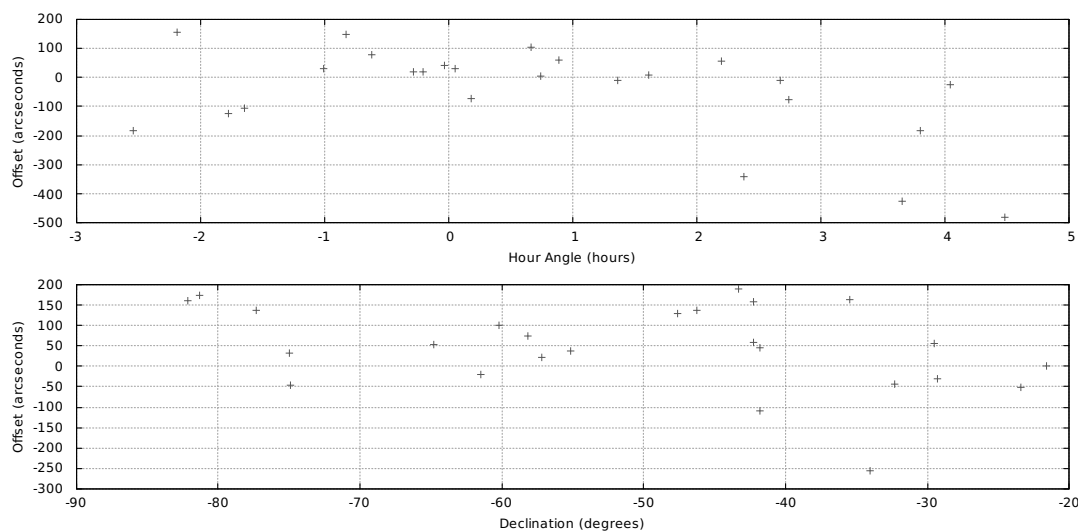


Figure 1. Hour angle and declination pointing errors. Note the parabolic profile in the hour angle offsets, which is indicative of a misalignment in the polar axis. The declination errors are more complex, but a major contributing factor is the non-circularity of the declination axis running surface.

In addition to the pointing model, the ACT control software has an advanced pattern matching system that can match the star pattern in the camera's field of view with that of a recorded template field. Offline tests of the pattern matching algorithm has shown that it can effectively correct for initial pointing offsets of up to $4 - 5'$. It should not be difficult to establish a pointing model that can reduce the pointing errors to this level, however a moderate number of reliable pointing offset measurements (on the order of 50) is required in order to establish a sufficiently accurate pointing model.

Due to the large pointing errors, the telescope is also not able to track a star and keep it in the aperture of the science instrument for more than a few minutes at a time. This problem will also

be solved by establishing an effective pointing model. Additionally, there are some small-scale tracking errors (minor shifts in hour angle while the telescope is tracking a star) which cannot be effectively removed with a good pointing model. However, these errors are small enough (a handful of arcseconds at most) to still allow the telescope system to collect photometry. The small-scale tracking errors are also highly regular, with frequencies that can be easily identified from a few hours of tracking offset data.

6. Conclusion

In summary, the ACT has enjoyed much attention since 2010, with several thousands of hours being spent testing the system, troubleshooting any identified problems and finding appropriate solutions. Although many problems have been encountered, the system is now reaching a level of stability and reliability that will enable it to operate automatically and without human supervision in the near future.

The large pointing errors is the only issue at the moment that is preventing the telescope from being operated automatically on a regular basis. However, it is at the moment possible for a human observer to use the telescope system (albeit remotely) and collect photometry with it. The observer will need to scout for the target star and the telescope will need to be re-centred on the star every few minutes, but it is strictly speaking possible.

Acknowledgments

The authors would like to thank the National Research Foundation and the University of Johannesburg for their generous support during the course of this project.

Using a VCSEL to accurately measure the chromatic dispersion in single mode fibre by the phase shift technique

S Wassin, E K Rotich Kipnoo, R R G Gamatham, A W R Leitch and T B Gibbon

Physics Department, Nelson Mandela Metropolitan University, P. O. Box 77000, Summerstrand South, Port Elizabeth-6031, South Africa.

Email: Shukree.Wassin@live.nmmu.ac.za

Abstract – The demand for high-speed data transmission and higher bandwidth is increasing rapidly due to the growing consumer need for advanced telecommunication technology. All fibre optic cables have an inherent transmission limiting factor, known as chromatic dispersion. In this paper, a method for characterizing the chromatic dispersion in single mode fibre is described. Our approach is based on the phase shift technique, where the phase difference between two sinusoidal modulated signals is measured. A vertical cavity surface emitting laser (VCSEL) source was implemented to characterize the chromatic dispersion along different lengths of G.652, G.655 (+) and G.655 (-) single mode fibre, around the 1550 nm wavelength region. A dispersion coefficient D , between 16.5 ps/nm.km to 19.1 ps/nm.km for the G.652 single mode fibre, 2.6 ps/nm.km to 4.2 ps/nm.km for the G.655 (+) single mode fibre and -2.8 ps/nm.km to -3.2 ps/nm.km for the G.655 (-) single mode fibre was measured. The experimental results are in close agreement to those obtained in literature.

Keywords: Chromatic dispersion, phase shift technique, VCSEL.

1. Introduction

From as early as 500 B.C, when the ancient Greeks developed a telegraph system made up of beacon fires, smoke signals and mirrors, up until the early 1970's when Corning Glass Works developed their first fibre optic cable for long haul transmission, mankind has relied heavily on communication across distance. Recent findings revealed that the gross domestic product (GDP) of a country is correlated to the development of the telecommunication infrastructure of a country [1]. The development of optical communications network systems has facilitated the growth in many areas such as, social, economic, business, education and politics [2]. Transmission speeds and fibre bandwidth capacity is limited by pulse broadening, due to the dispersive properties of an optical fibre. Chromatic dispersion is one aspect of the fibre causing an optical pulse to distort as it travels along the fibre [3]. In this paper, we experimentally demonstrated a method for measuring the chromatic dispersion in single mode fibre using a Vertical Cavity Surface Emitting Laser (VCSEL). Our technique is based on the phase shift measurement scheme. The phase shift technique has been used for the precise measurements of the optical path length and for characterizing the chromatic dispersion in multimode and single mode fibres [4].

2. Theory

The majority of all chromatic dispersion measurements in single mode fibres are obtained by determining the group delay τ_g , with respect to a specific wavelength λ . Suppose an optical signal propagating along a fibre of length l , with a group velocity expressed [2] as

$$v_g = \frac{d\omega}{d\beta} \quad (1)$$

arrives at a receiver after a certain time. The time here is referred to as the group delay τ_g , defined [2] as

$$\begin{aligned} \tau_g &= \frac{l}{v_g} \\ &= l \frac{d\beta}{d\omega} \end{aligned} \quad (2)$$

An optical signal having a spectral width of $\Delta\omega$ experiences pulse broadening since the different spectral components, with respect to distinct wavelengths propagates at different group velocities, as illustrated in figure 1. Chromatic dispersion, intramodal dispersion or group velocity dispersion (GVD) refers to the fact that the group velocity is dependent on the wavelength. The expression quantifying the spreading of the signal can be mathematically described as [2]

$$\begin{aligned} \Delta\tau_g &= \frac{d\tau_g}{d\omega} \Delta\omega \\ &= \frac{d}{d\omega} \left(l \frac{d\beta}{d\omega} \right) \Delta\omega \\ &= l \frac{d^2\beta}{d\omega^2} \Delta\omega \end{aligned} \quad (3)$$

The term $\beta_2 = \frac{d^2\beta}{d\omega^2}$ is known as the group velocity dispersion (GVD) parameter and represents the amount of broadening the pulse experiences as it propagates along the fibre [2, 5]. Since the spectral width $\Delta\omega$ is related to the linewidth of the optical source, $\Delta\lambda$, (3) can be rewritten [2] as

$$\Delta\tau_g = \frac{d}{d\lambda} \left(l \frac{d\beta}{d\omega} \right) \Delta\lambda \quad (4)$$

where

$$D(\lambda) = \frac{d}{d\lambda} \left(l \frac{d\beta}{d\omega} \right) \quad (5)$$

By considering (4) and (5), it follows that

$$\begin{aligned} \Delta\tau_g &= D(\lambda) \cdot l \cdot \Delta\lambda \\ \therefore D(\lambda) &= \frac{\Delta\tau_g}{l \cdot \Delta\lambda} \end{aligned} \quad (6)$$

where $D(\lambda)$ is known as the *chromatic dispersion coefficient* and expresses the broadening of the optical signal as a function of wavelength and is given in ps/nm.km [6]. The overall chromatic dispersion along a fibre of length l is mathematically expressed $D(\lambda) = D_m + D_w$, where D_m and D_w are defined as the material and waveguide dispersion, respectively. Material dispersion is attributed to the wavelength dependence on the refractive index of the optical fibre whereas waveguide dispersion in optical fibre arises as a result of the geometric shape, structural design of the optical fibre [6]. By correctly adapting the material and waveguide dispersion properties, the chromatic dispersion across the fibre length can be reduced, since material dispersion is negative in the transmission region below the zero-dispersion wavelength and positive in the wavelength region above 1.3 μm [5]. Waveguide dispersion is negative in the 1.1 μm to 1.7 μm spectrum, as illustrated in figure 2 [5]. A brief summary of the measurement techniques used for characterizing the chromatic dispersion in single mode fibre is given in table 1.

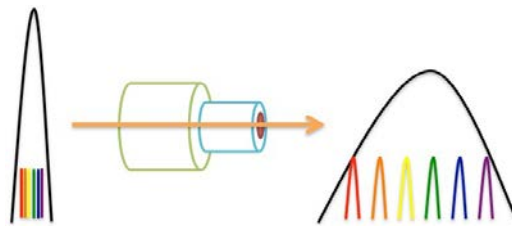


Figure 1. Pulse broadening caused by the variation in speed in the time domain.

For phase shift measurements, the change in phase $\phi(\lambda)$ of a sinusoidal modulated signal travelling along a fibre of length l is measured due to changes in the group velocity with respect to the wavelength.

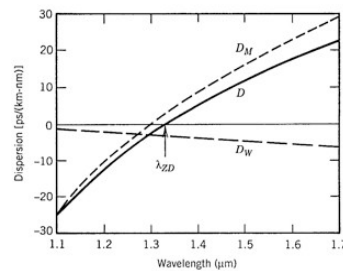


Figure 2. Material and waveguide dispersion slopes for typical single mode fibre [5].

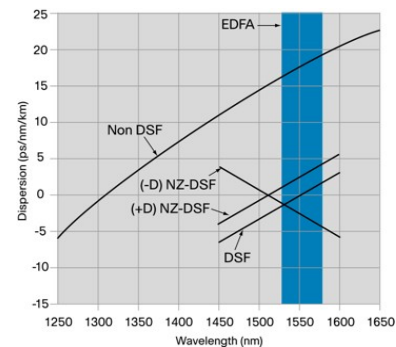


Figure 3. Chromatic dispersion slopes of the various single mode fibres [7].

Table 1. Properties of the chromatic dispersion measurement technique [8].

Measurement Technique	Measurement Resolution	Fibre length
Pulse delay/ Time-of-flight Technique	50 ps to 100 ps	Exceeding 0.5 km
Phase shift Technique	10 ps to 20 ps	Exceeding 0.5 km
Interferometric Technique	0.1 ps	Less than 5 m

The mathematical description of the phase difference is given as [9]

$$\varphi(\lambda) = 2\pi f_m \tau_g(\lambda) \quad (7)$$

where f_m is the modulation frequency. The measurement accuracy of the phase shift technique improves when considering higher modulation frequencies as well as longer fibre lengths [4].

3. Experimental Setup

Figure 4 shows the experimental configuration. A Vertical Cavity Surface Emitting Laser (VCSEL) was used for the characterization of chromatic dispersion along the single mode fibre. The directly modulated VCSEL is designed to operate in the 1545 nm to 1550 nm wavelength region. Two different types of single mode fibre were considered, the first type being G.652 standard single mode fibre.

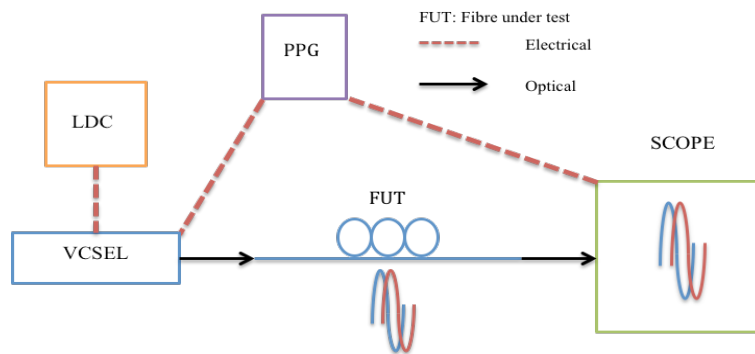


Figure 4. Schematic of the phase shift experimental setup.

The second category of single mode fibre measured was the G.655 Non-zero dispersion shifted fibre (NZDSF) positive and negative. A summary of the lengths of fibre used in the phase shift experiment is tabulated in table 2. A Laser Diode Controller (LDC) externally powered the VCSEL and by varying the bias current of LDC the wavelength of operation of the VCSEL was shifted.

Table 2. Description of the fibre lengths (km) used.

G.652 Standard Single Mode Fibre	G.655 NZDSF (+)	G.655 NZDSF (-)
6.1	26.6	25.5
11.5	51.4	
12.2	76.7	
18.3		
23.0		

A Programmable Pattern Generator (PPG) was required to sinusoidally modulate the VCSEL at a modulation frequency f_m of 8.5 GHz. The phase difference between the reference and test signals of different wavelengths was measured with a wide band oscilloscope. This particular oscilloscope has an 8.5 GHz optical channel bandwidth and a -21 dBm optical sensitivity in the 1550 nm region.

4. Results and Discussions

Figures 5 and 6 illustrate the direction of the shift occurring along the 26.6 km G.655 NZDSF (+) and 25.5 km G.655 NZDSF (-) respectively. The reference wavelength selected in both

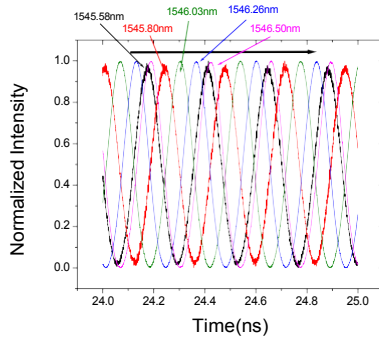


Figure 5. Phase shift along 26.6 km G.655 NZDSF (+).

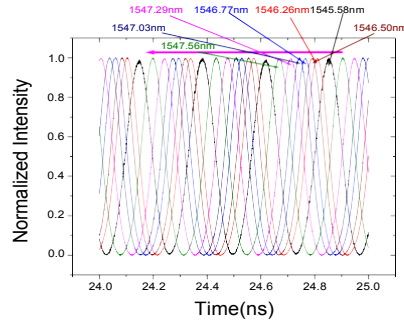


Figure 6. Phase shift along 25.5 km G.655 NZDSF (-).

instances was 1545.58 nm. From figure 5, the shift is occurring to the right of the reference signal as the wavelength is increased, whereas in figure 6 the shift occurs to the left of the reference signal with increasing wavelength. These occurrences are due the design of the optical fibre during manufacturing, where the refractive index of the fibre material is tailored in such a way as to define the sign of the chromatic dispersion along the fibre. If we consider the shift in figure 6 to be moving to the right of the reference wavelength, 1545.58 nm, as seen in figure 5. From figure 6, the sequence of wavelengths after the reference wavelength is given as 1547.56 nm, 1547.29 nm, 1547.03 nm, 1546.77 nm, 1546.50 nm and 1546.26 nm respectively. From this it can be inferred that, the longer wavelength signals travels slower along the G.655 NZDSF (-) fibre than the signals of shorter wavelength. This validates the property of the G.655 NZDSF (-), where it allows longer wavelengths to propagate slower and shorter wavelengths faster.

Figures 7 and 8 show chromatic dispersion curves measured along the G.652 standard single mode fibre and G.655 NZDSF (+) for the respective fibre lengths tabulated in table 2. An indirect approach was formulated for determining the phase difference between two signals at different wavelength, since the oscilloscope referred to in figure 4 is not capable of directly measuring and calculating phase changes. The phase difference, with reference to a specific wavelength, between sinusoidally modulated signals was calculated according to

From (8), T refers to the period of the clock signal and t_{sep} refers to the separation time between the signals, given by the oscilloscope. After calculating the phase difference it is inserted into (6), in accordance with (7). The chromatic dispersion curves for the respective lengths of G.652 and G.655 single mode fibre experimentally derived, are in good agreement with the theoretical fit as displayed in figures 7 and 8. A minor deviation from the theoretical fit was observed for the chromatic dispersion coefficient experimentally measured at the initial wavelength, in the 6.1 km G.652 SMF and in the 26.6 km and 76.7 km G.655 NZDSF (+). This observation can be attributed to the small wavelength spacing between the reference and test signal. As the wavelength spacing increased, the distance between the experimental curve and the theoretical curve reduced. Generally, the accuracy improved as the distance between the reference and test signal increased.

$$\varphi(\lambda) = \frac{2\pi \cdot t_{sep}}{T} \quad (8)$$

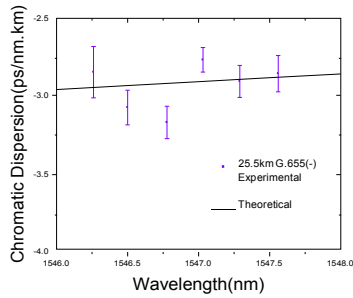


Figure 9. Chromatic dispersion spectra experimentally derived for the G.655 NZDSF (-).

calculated chromatic dispersion curve and falls within the -1.4 ps/nm.km to -4.8 ps/nm.km range stipulated by OFS Optics. The phase shift occurring in the 25.5 km G.655 NZDSF (-) illustrated in figure 6 authenticates the negative sign for the experimental data given in figure 9.

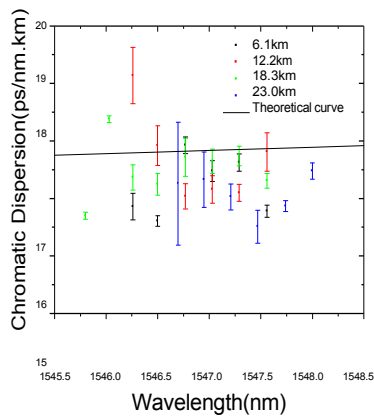


Figure 7. Dispersion curves of chromatic dispersion measurement made in G.652 standard single mode fibre.

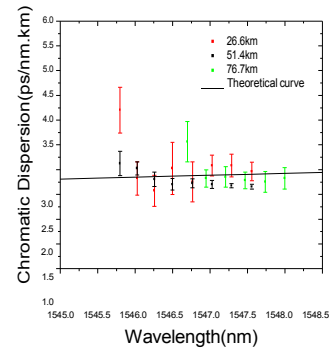


Figure 8. Dispersion curves of chromatic dispersion measurement made in G.655 NZDSF (+).

5. Conclusion

We have successfully described an accurate system for characterizing the chromatic dispersion along any single mode fibre. The high modulation frequency selected was a major contributor to the accuracy of our phase shift technique. The results indicated that longer fibre lengths also improved the accuracy of the measurement technique. A range of fibre lengths was used for demonstrating chromatic dispersion measurements by the phase shift technique, from 6.1 km to 76.7 km. Therefore the measurement system describe in this work can be used for characterizing the chromatic dispersion along the last-mile fibre network systems for FTTx application. We showed that the direction of propagation along a G.655 NZDSF (-) is opposite to the conventional forward movement of the signal in the G.655 NZDSF (+) or the G.652 standard single mode fibre. By utilizing the properties of the NZDSF (-), it is possible to achieve a total chromatic dispersion of 0 ps/nm or a value very close to 0 ps/nm across an optical fibre network system.

6. References

- [1] Patel S and Edwards M 2012 Broadband Telecommunication Stimulates Economic Growth-Optical Fibre is a Key Enabler Amcham Update Available from: www.amchamindia.com/pdf/newsletter Issue 4 [Accessed 19 May 2014]
- [2] Kazovsky L G et al 2001 *Broadband Optical Access Networks* (New Jersey:John Wiley & Sons, Inc) pp 47-51
- [3] Watanabe S et al 1993 Compensation of Chromatic Dispersion in a Single-Mode Fiber by Optical Phase Conjugation *IEEE Phot. Tech. Lett.* **5** (1) 92-95
- [4] Tanaka S and Kitayama Y 1984 Measurement Accuracy of Chromatic Dispersion by the Modulation Phase Technique *J. Lightwave Tech.* **LT-2** (6) 1040-044
- [5] Agrawal G P 2002 *Fiber-Optic Communication Systems* (New York:John Wiley & Sons, Inc) pp 38-42
- [6] Keiser G 2000 *Optical Fiber Communication* (New York: McGraw-Hill Companies, Inc) p104
- [7] Cisco White Paper 2008 Fibre Types in Gigabit Optical Communication (USA: Cisco Systems) pp 13-14
- [8] Cohen L G 1985 Comparison of Single-Mode Fiber Dispersion Measurement Techniques *J. Lightwave Techn.* **LT-3** (5) 958-66
- [9] Costa B, Mazzoni D, Puleo M and Vezzoni E 1982 Phase Shift Technique for the Measurement of Chromatic Dispersion in Optical Fibres Using LED's *IEEE J. Quant Elect.* **QE-18** (10) 1509-1515

Calculating solar irradiance to determine optimal angles of solar cells for De Aar

Graham Webber and Hartmut Winkler

Department of Physics, University of Johannesburg, PO Box 524, 2006 Auckland Park, Johannesburg, South Africa

E-mail: 201337952@student.uj.ac.za

Abstract. Solar energy harvesting is a growing industry in South Africa. De Aar is a favoured location for solar power stations, as it has high potential yield and is close to some of South Africa's largest power lines. This paper uses standard methods of solar irradiance estimation to calculate the potential yield with respect to wavelength for De Aar. It is necessary to take wavelength into account as light is not extinguished uniformly with respect to wavelength. De Aar was chosen for this paper because there are many years of total surface irradiance data available and some data for irradiance in specific wavelength bands. Comparison of these values with actual data collected in De Aar was done to determine the accuracy of these models for the conditions in De Aar. These estimations were done for a typical midsummer's day and a typical winter's day, employing four different photovoltaic device spectral response functions. We use these calculations to determine the optimum panel alignments.

1. Introduction

Solar irradiance research is necessary to provide information and knowledge regarding the optimum and expected energy yields to developers of photovoltaic solar power systems, whether these are on a domestic, medium or full solar power station scale. The specific location and meteorological conditions play a major role in the calculation of a photovoltaic system's yield. The ability to evaluate the impact of these factors is vital in determining the economic viability of a proposed installation and the optimal design, and prediction of the system performance. The best way of determining the amount of solar irradiance at a site is through long-term data monitoring, but because of practical and cost considerations these data are measured only at a few stations, and mostly only as integrated flux rather than as wavelength specific intensity.

In this paper we focus on the direct solar beam irradiance, which is under clear, moderately turbid skies by far the largest contributor to the total solar irradiance collected at ground level. Photovoltaic devices in addition also respond to the so-called diffuse light component, i.e. scattered sunlight reaching the ground from directions other than the solar position. This component is also a function of azimuth and zenith angle, and is particularly difficult to predict accurately. A full study of the solar irradiance at a site requires the full consideration of diffuse irradiance as well, and is in progress. Furthermore, while this paper restricts itself to photovoltaic devices, the end results of this study can also be applied to concentrated solar power (which only utilises the direct beam).

For this study we chose to apply irradiance models to the estimation of direct solar radiation at De Aar in the Northern Cape, a town near an important electric power line node that has, in view of its ample sunshine, been chosen as the site of several major solar power station developments. De Aar has

the added advantage that it hosts a well-developed weather station which includes a node of the Baseline Surface Radiation Network (BSRN) [1]. Furthermore, there has in the past been an extensive study of the solar irradiance as a function of wavelength in De Aar [2]. Theoretical data from these calculations can therefore be compared to the recorded data for each wavelength-band.

2. Theory

The direct beam spectral irradiance E_b is the solar light (in units of power per unit area per unit wavelength) that passes through the atmosphere and reaches the panel without being scattered away or absorbed by gases or aerosols. Direct beam irradiance is thus that part of the extra-terrestrial spectral irradiance E_0 incident on a ground level panel normal to the beam. We adopted for E_0 the synthetic/composite extra-terrestrial spectrum of Gueymard [3]. The factors that determine the magnitude of E_b are the atmospheric transmittances for the different extinction processes including Rayleigh and Mie (from aerosol) scattering, ozone, uniformly mixed gases and water vapour absorption. The direct beam irradiance for a specific wavelength is calculated as [4]:

$$E_b = E_0 T_R T_a T_o T_g T_w \quad (1)$$

where T_R , T_a , T_o , T_g and T_w are the transmittance factors for Rayleigh scattering, aerosol, ozone, mixed gasses and water vapour respectively, and are all functions of the wavelength λ . The transmittance factors are described by Bouguer's Law [4]:

$$T = \exp(-m \tau) \quad (2)$$

where τ is the optical thickness, and m is the optical airmass.

We describe the solar position in terms of the zenith angle ζ and the azimuth γ . Unless the Sun is close to the horizon, the airmass is then well approximated by the expression $m = \sec \zeta$.

The Rayleigh optical thickness has been calculated from the equation below, which was developed from a least-squares curve fit to the theoretical expression and deviates from this by 0.01% or less throughout the spectrum [4]:

$$\tau_{R\lambda} = (P/1013.25 \text{ mbar})(117.2594 \lambda^4 - 1.3215 \lambda^2 + 3.2073 \times 10^{-4} - 7.6842 \times 10^{-5} \lambda^{-2})^{-1} \quad (3)$$

To determine the ozone, mixed gas and water vapour optical thickness, we adopted the expressions and constants of the SMARTS2 model of Gueymard [4]. Water vapour has minimal effect on the transmission outside the near infrared. There however it is by far the most important absorber. The total precipitable water in the atmosphere may vary rapidly with time and thus so does the water vapour optical thickness.

Detailed complete spectral determinations of the aerosol optical thickness are difficult and hence only rarely available. Hence the aerosol transmissivity can often only be estimated with the aid of climatological information. Alternatively, estimates of turbidity are sometimes based on meteorological visibility data. The lack of detailed knowledge of the aerosol properties justifies the use of a simplified methodology, the modified Angström approach, which considers two different spectral regions, below and above $\lambda_0 = 0.5 \mu\text{m}$. The aerosol transmittance is obtained from [4]:

$$\tau_{a\lambda} = (\beta_i \lambda)^{-\alpha_i} \quad (4)$$

where $\alpha_i = \alpha_1$ if $\lambda < \lambda_0$ and α_2 otherwise, and $\beta_i = 2^{\alpha_2 - \alpha_1} \beta$ if $\lambda < \lambda_0$ and $\beta_i = \beta$ otherwise. β is the Angström coefficient and expresses the turbidity. α is the wavelength exponent it is related to the optical characteristic of the aerosols. The values of the wavelength exponents α_1 and α_2 were obtained by fitting the spectral optical coefficients of different reference aerosol models [4].

3. Procedure

We determined the solar zenith and azimuth angles through the PSA algorithm, which has average deviation in the solar vector of 0.147 minutes of arc [5].

We then examined data obtained from multifilter rotating shadowband radiometer measurements in De Aar for 4 August 2000 and 25 December 2000, previously presented by Winkler et al. [2]. These data contain measurement at 60 second intervals and for each of ~ 10 nm wide equivalent wavelength bands centred at 416 nm, 501 nm, 613 nm, 670 nm, 867 nm and 939 nm. We consequently also calculated the model fits at one minute intervals for each band.

The multifilter rotating shadowband radiometer tracks the Sun so the direct beam irradiance is calculated as if the panel in the model is normal to the solar vector. Thus the surface in the model was set to be normal to the solar vector.

The total precipitable water, which also depends on temperature and atmospheric pressure, only has an effect on one of the wavelength bands we are interested in: the one centred at 939 nm. Average temperatures for the days in question were determined from archived weather station data [6]. Air pressure was estimated from the altitude at the site. The parameters determining T_a and T_w were adjusted to find the best fit for the calculated data against each of the measured wavelength bands. The α_1 and α_2 values are under most conditions such that the resulting aerosol transmissivity is at its minimum at the shorter wavelengths.

The total daily ground level direct beam irradiance incident on a tilted panel was calculated for the two days examined by summing the amounts for each wavelength step and each time step. If the panel tilt angle is η and the panel azimuth is γ_p , the panel inclination relative to the solar beam θ is then given by [7]

$$\cos \theta = \sin \zeta \cos(\gamma - \gamma_p) \sin \eta + \cos \zeta \cos \eta \quad (5)$$

The response curve for some typical solar cells was applied to the function of the incident light and the optimal stationary panel orientation was then calculated [8]. The azimuth angle of the panel was taken as zero, as this model is symmetrical around due north. The optimal stationary panel tilt was calculated by gradually varying the zenith angle of the panel and redoing the yield calculation until the maximum daily yield is reached.

4. Results

The model predictions closely match the available empirical data. We display two examples of direct beam irradiance vs. time plots, for 4 August 2000, in figure 1 and figure 2. The discrepancies at approximately 8h30 are also visible in the other bands, and we believe these to be caused by an inaccuracy in the provided calibration of the instrumental angular response function for a particular solar zenith angle.

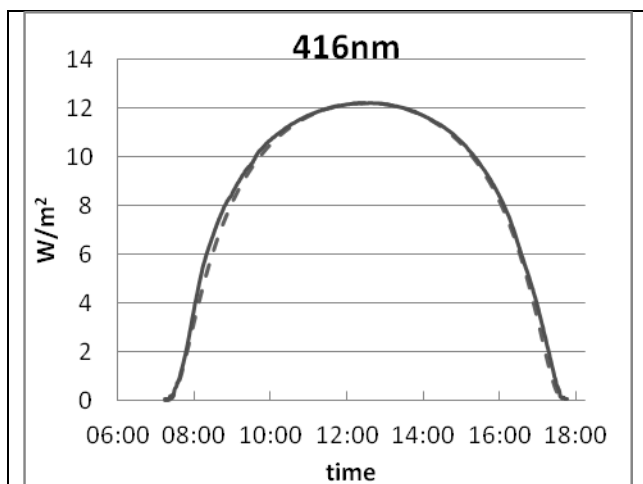


Figure 1: Empirical (solid line) and calculated data (dashed line) for 416 nm for 4 August 2000

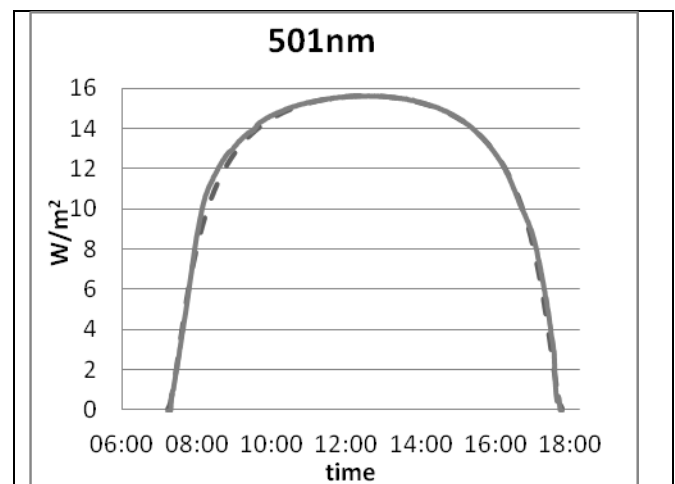


Figure 2: Empirical (solid line) and calculated data (dashed line) 501 nm for 4 August 2000

The optical thicknesses calculated were assumed not to vary in the course of the day. For ozone we adopted a seasonal average, which is justified in view of the ozone transparency being equal or close to 1 in all wave bands used. For this study we also ignore the mixed gas optical depth, as CO₂ and O₂ only absorb light at specific wavelengths, and these wavelengths do not correspond to the bands used in this study. The optical thickness for water vapour is zero for all the measured bands except at 939 nm where the values of 0.143 and 0.142 were determined for August and December respectively. The values for the Angström coefficient that yielded the best fits to the direct irradiance vs. time curves for 4 August 2000 and 25 December 2000 were 0.0122 and 0.0223 respectively.

The total daily direct beam irradiance was calculated to be 13064 Wh.m⁻² for 25 December and 8798 Wh.m⁻² for 4 August. This would correspond to the yield of a 'perfect panel' (with 100% quantum efficiency) able to track and align itself to the Sun throughout the day. Because of cost considerations, solar panels are usually fixed rather than Sun tracking. They are then set up at a specific tilt angle and orientation that is considered optimal for maximum power generation, either annually or seasonally. This optimum configuration depends on site latitude, altitude, prevailing meteorological and atmospheric conditions as well as the panel's efficiency (which is a function of wavelength).

The optimal tilt angle for a fixed, 'perfect', north-south aligned solar panel was calculated as 52.6° north for 4 August and 4.4° south for 25 December, yielding 6930 Wh.m⁻² and 9037 Wh.m⁻² respectively. The ratio between the yield, at their optimum tilts, for 25 December and that of 4 August is 1.30. The apparent contradiction of deriving an optimal south-facing tilt angle in midsummer for a latitude south of the Tropic of Capricorn is explained by the fact that the early morning and late afternoon solar azimuths are south-east and south-west respectively at that time of the year.

We repeated the calculations using relative spectral response functions typical for three photovoltaic technologies using CdS, CuInS and c-Si. The results of this are summarised in Table 1.

Table 1. Optimal tilt angles and relative energy yield ratios calculated for different solar panel technologies.

	perfect panel	CdS	CuInS	c-Si
optimal tilt: 25 Dec 2000	4.4° S	1.7° S	3.8° S	5.4° S
optimal tilt: 4 Aug 2000	52.6° N	51.4° N	52.3° N	53.0° N
optimal tilt yield ratio: 25 Dec vs. 4 Aug	1.30	1.45	1.30	1.25

5. Discussion

The work presented in this paper is the precursor of an ongoing wider study that seeks to determine optimal solar panel orientations and projected energy yields for photovoltaic technologies under South African conditions, through the construction of site-specific solar spectral irradiance models and the convolution thereof with instrumental spectral response characteristics.

The results show that daily spectral irradiance curves calculated for specific wavelengths for selected days through the adopted model correlate to a good approximation to empirical data. Although the wavelength bands adopted in this study only represent a small part of the solar spectrum, they are well distributed across that part of the spectrum where photovoltaic cells are most sensitive. The greatest uncertainty is in those parts of the spectral irradiance curve affected by water vapour, such as near 939 nm, especially because of the sometimes rapid humidity fluctuations. The spectral regions most influenced by water vapour are however deeper in the infrared, where most panels show little or no sensitivity.

The case study presented here is based on calculations for two days that represent different times of year and thus different atmospheric conditions. The model's data fit both times of year well. This gives

us confidence to continue developing this model and expanding the test sites and probing other time periods. Some error is expected as the model is using average aerosol characteristics of the atmosphere while in reality high turbidity episodes occur periodically [2].

The Angström coefficient monthly averages determined by Power and Willmott [9] for Grootfontein Agricultural College, 140 km from De Aar, are 0.021 for December and 0.013 for August. These are in good agreement with the Angström coefficients that gave us the best fits for the days investigated. If these are typical for the site and season, as we believe they are, the Angström coefficient at De Aar is generally very low by international standards, confirming De Aar as a good location for solar power stations.

The difference between the optimal tilt yield ratio of each photovoltaic technology and that of the ‘perfect’ solar panel highlights the accuracy that may be gained when taking wavelength into account for solar power estimations. For a typical c-Si solar cell there is a lower fluctuation in yield with time of year than a ‘perfect’ solar panel (which represents a wavelength independent solar cell).

For 25 December there is a difference of $\sim 1^\circ$ (and $\sim 0.5^\circ$ on 4 August) between the optimal fixed tilt angles of a ‘perfect’ solar panel and a typical c-Si solar cell. This may only be a small value, but it does highlight how detailed calculations that fully consider wavelength dependence can lead to different outcomes, both in terms of the optimum tilt angle of a panel as well as the total energy harvested.

References

- [1] Ohmura A, Dutton E G, Forgan B et al 1998 *Bulletin American Meteorol. Soc.* **79** 2116
- [2] Winkler H, Formenti P, Esterhuyse D J, Swap R J, Helas G, Annegran H J and Andreae M O 2008 *Atmospheric Environment* **42** 5569
- [3] Gueymard C A 2004 *Solar Energy* **76** 423
- [4] Gueymard C A 2001 *Solar Energy* **71** 325
- [5] Blanco-Muriel M, Alarcon-Padilla D C, Lopez-Moratalla T and Lara-Coira M 2001 *Solar Energy* **70** 431
- [6] Tutiempo Network Retrieved April 23, 2014, from Tutiempo: <http://www.tutiempo.net>
- [7] Chang T P 2009 *Solar Energy* **83** 1274
- [8] Xiaoyong H, Sanyang H, Wei H and Xiaogang L 2013 *Chemical Society Review* **42** 173
- [9] Power H C and Willmott C J 2001 *International Journal of Climatology* **21** 579

Division G – Theoretical and Computational Physics

Evolution of quark masses and flavour mixings in the 2UED

Ammar Abdalgabar¹ and Alan S Cornell²

National Institute for Theoretical Physics; School of Physics and Mandelstam Institute for Theoretical Physics, University of the Witwatersrand, Wits 2050, South Africa

E-mail: ¹ammr.abdalgabar@students.wits.ac.za, ²alan.cornell@wits.ac.za

Abstract. The evolution equations of the Yukawa couplings and quark mixings are performed for the one-loop renormalisation group equations in six-dimensional models compactified in different possible ways to yield standard four space-time dimensions. Different possibilities for the matter fields are discussed, that is where they are in the bulk or localised to the brane. These two possibilities give rise to quite similar behaviours when studying the evolution of the Yukawa couplings and mass ratios. We find that for both scenarios, valid up to the unification scale, significant corrections are observed.

1. Introduction

A theory of fermion masses and the associated mixing angles is unexplained in the Standard Model (SM) providing an interesting puzzle and a likely window to physics beyond the SM. In the SM one of the main issues is to understand the origin of quark and lepton masses, or the apparent hierarchy of family masses and quark mixing angles. Perhaps if we understood this we would also know the origins of CP violation. A clear feature of the fermion mass spectrum is [1, 2]

$$m_u \ll m_c \ll m_t, \quad m_d \ll m_s \ll m_b, \quad m_e \ll m_\mu \ll m_\tau. \quad (1)$$

Apart from the discovery of the Higgs boson at the Large Hadron Collider (LHC), another important goal of the LHC is to explore the new physics that may be present at the TeV scale. Among these models those with extra spatial dimensions offer many possibilities for model building and TeV scale physics scenarios which can be constrained or explored. As such, there have been many efforts to understand the fermion mass hierarchies and their mixings by utilizing the Renormalization Group Equations (RGEs) especially for Universal Extra Dimension (UED) models and their possible extensions (see [3, 4, 5] and references therein).

In these UED models each SM field is accompanied by a tower of massive states, the Kaluza-Klein (KK) particles. An extension of this scenario is to consider a type of model with two extra dimensions. This extension is non-trivial and brings further insight to extra-dimensional scenarios. It is theoretically motivated by specific requirements, such as, they provide a dark matter candidate, suppress the proton decay rate, as well as anomaly cancellations from the number of fermion generations being a multiple of three [6]. Different models with two extra dimensions have been proposed, such as T^2/Z_2 [5], the chiral square T^2/Z_4 [7], $T^2/(Z_2 \times Z'_2)$ [8], S^2/Z_2 [9], the flat real projective plane RP^2 [10], the real projective plane starting from the

sphere [11]. For simplicity, in this paper we assume that the two extra space-like dimensions have the same size, that is $R_5 = R_6 = R$. However, this simpler case provides the opportunity to compute in detail the RGEs and study the evolution of mass ratios, the renormalisation invariance R_{13} and R_{23} , and $\sin \beta$.

The four-dimensional chiral zero modes of the SM fermion are obtained by imposing a discrete Z_2 symmetry, this eliminates one 4-dimensional (4D) degree of freedom and allows us to have a 4D chiral fermion [7]. However, this can also be obtained directly from the properties of the orbifold, as in [10]. Higher massive modes are then vector-like fermions. Each of the gauge fields have six components and decompose into towers of 4D spin-1 fields and two towers of real scalars belonging to the adjoint representation [12].

The one-loop correction to the gauge couplings are given by

$$16\pi^2 \frac{dg_i}{dt} = b_i^{SM} g_i^3 + \pi \left(S(t)^2 - 1 \right) b_i^{6D} g_i^3, \quad (2)$$

where $i = 1, 2, 3$, $t = \ln(\frac{\mu}{M_Z})$, $S(t) = e^t M_Z R$ is the summation of all KK excited states, μ being the energy, that is, for the evolution between $M_Z < \mu < \Lambda$ (Λ is the cut-off scale, where we have set M_Z as the renormalisation point). More details about the calculation of the $S^2(t)$ factor can be found in [12, 13]. The numerical coefficients appearing in equation (2) are given by:

$$b_i^{SM} = \left[\frac{41}{10}, -\frac{19}{6}, -7 \right], \quad (3)$$

and

$$b_i^{6D} = \left[\frac{1}{10}, -\frac{13}{2}, -10 \right] + \left[\frac{8}{3}, \frac{8}{3}, \frac{8}{3} \right] \eta, \quad (4)$$

η being the number of generations of fermions propagating in the bulk. Therefore, in the two cases we shall consider: that of all fields propagating in the bulk, $\eta = 3$; and for all matter fields localized to the brane $\eta = 0$.

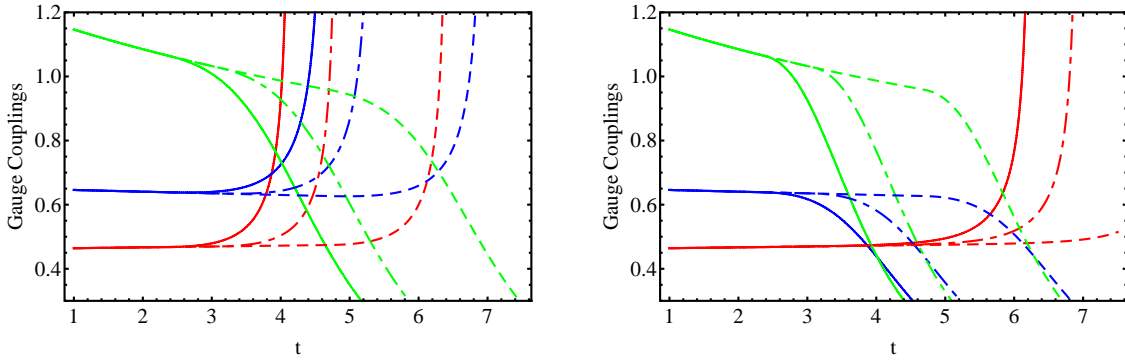


Figure 1. (Colour online) Gauge couplings g_1 (red), g_2 (blue), g_3 (green) with: in the left panel, all matter fields in the bulk; and the right panel for all matter fields on the brane; for three different values of the compactification scales ($R^{-1} = 1$ TeV (solid line), 2 TeV (dot-dashed line), and 10 TeV (dashed line)) as a function of the scale parameter t .

The evolution of the Yukawa couplings were derived in [12, 13], where the one-loop RGEs in the 2UED we study are:

$$16\pi^2 \frac{dY_i}{dt} = \pi \left(S(t)^2 - 1 \right) Y_i [T_i - G_i + T], \quad (5)$$

Table 1. The terms present in the various Yukawa evolution equations, see equation (5).

Scenarios	G_u	G_d	G_e	$T_u = -T_d$	T_e
Bulk	$\frac{5}{6}g_1^2 + \frac{3}{2}g_2^2 + \frac{32}{3}g_3^2$	$\frac{1}{30}g_1^2 + \frac{3}{2}g_2^2 + \frac{32}{3}g_3^2$	$\frac{27}{30}g_1^2 + \frac{3}{2}g_2^2$	$3(Y_d^\dagger Y_d - Y_u^\dagger Y_u)$	$3Y_e^\dagger Y_e$
Brane	$4(\frac{17}{20}g_1^2 + \frac{9}{4}g_2^2 + 8g_3^2)$	$4(\frac{1}{4}g_1^2 + \frac{9}{4}g_2^2 + 8g_3^2)$	$4(\frac{9}{4}g_1^2 + \frac{9}{4}g_2^2)$	$6(Y_d^\dagger Y_d - Y_u^\dagger Y_u)$	$6Y_e^\dagger Y_e$

where $i = u, d, e$, $T = 2(3\text{Tr}(Y_d^\dagger Y_d) + 3\text{Tr}(Y_u^\dagger Y_u) + \text{Tr}(Y_e^\dagger Y_e))$ and the values of G_i and T_i are given in table 1. That is, when the energy scale $\mu > \frac{1}{R}$ or when the energy scale parameter $t > \ln(\frac{1}{M_Z R})$, we shall use equation.(5), however, when the energy scale $M_Z < \mu < \frac{1}{R}$, the Yukawa evolution equations are dictated by the usual SM ones, see [12, 13, 14].

Yukawa coupling matrices can be diagonalised by using two unitary matrices U and V , where

$$UY_u^\dagger Y_u U^\dagger = \text{diag}(f_u^2, f_c^2, f_t^2); \quad VY_d^\dagger Y_d V^\dagger = \text{diag}(h_d^2, h_s^2, h_b^2).$$

The CKM matrix appears as a result (upon this diagonalisation of quark mass matrices) of $V_{CKM} = UV^\dagger$. The variation of the CKM matrix and its evolution equation for all matter fields in the bulk is [15, 16]:

$$16\pi^2 \frac{dV_{i\alpha}}{dt} = -6(\pi(S^2 - 1) + 1) \left[\sum_{\beta, j \neq i} \frac{f_i^2 + f_j^2}{f_i^2 - f_j^2} h_\beta^2 V_{i\beta} V_{j\beta}^* V_{j\alpha} + \sum_{j, \beta \neq \alpha} \frac{h_\alpha^2 + h_\beta^2}{h_\alpha^2 - h_\beta^2} f_j^2 V_{j\beta}^* V_{j\alpha} V_{i\beta} \right]. \quad (6)$$

For all matter fields on the brane, the CKM evolution is the same as equation (6) but multiplied by 2.

The mixing matrix V_{CKM} satisfies the unitarity condition, providing the following constraint

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \quad (7)$$

that is, we have a triangle in the complex plane and the three inner angles α , β and γ are given by

$$\sin \beta = \frac{J}{|V_{td}||V_{tb}^*||V_{cd}||V_{cb}^*|}, \quad \sin \gamma = \frac{J}{|V_{ud}||V_{ub}^*||V_{cd}||V_{cb}^*|}, \quad (8)$$

with $\alpha = \pi - \beta - \gamma$. The shape of the unitarity triangle can be used as a tool to explore new symmetries or other interesting properties that give a deeper insight into the physical content of new physics models.

On the other hand, in the quark sector both the mass ratios are related to mixing angles as

$$\theta_{13} \sim \frac{m_d}{m_b}, \quad \theta_{23} \sim \frac{m_s}{m_b}. \quad (9)$$

In [2, 17] a set of renormalisation invariants is constructed

$$R_{13} = \sin(2\theta_{13}) \sinh \left[\ln \frac{m_b}{m_d} \right] \sim \text{constant}, \quad R_{23} = \sin(2\theta_{23}) \sinh \left[\ln \frac{m_b}{m_s} \right] \sim \text{constant}. \quad (10)$$

2. Results

For our numerical calculations we set the compactification radii to be $R^{-1} = 1$ TeV, 2 TeV and 10 TeV. Only some selected plots will be shown and we will comment on the other similar cases not explicitly presented. We quantitatively analyse these quantities in 2UED models, though we observed similar behaviours for all values of R^{-1} . The initial values we shall adopt at the M_Z scale can be found in [18].

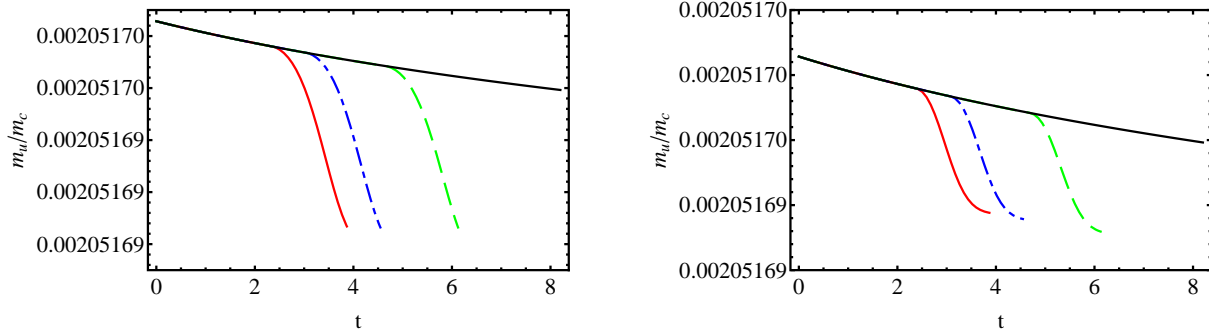


Figure 2. (Colour online) Evolution of the mass ratio $\frac{m_u}{m_c}$ with: in the left panel all matter fields in the bulk; and the right panel for all matter fields on the brane. Three different values of the compactification radius have been used $R^{-1} = 1$ TeV (solid line), 2 TeV (dot-dashed line), and 10 TeV (dashed line), all as a function of the scale parameter t .

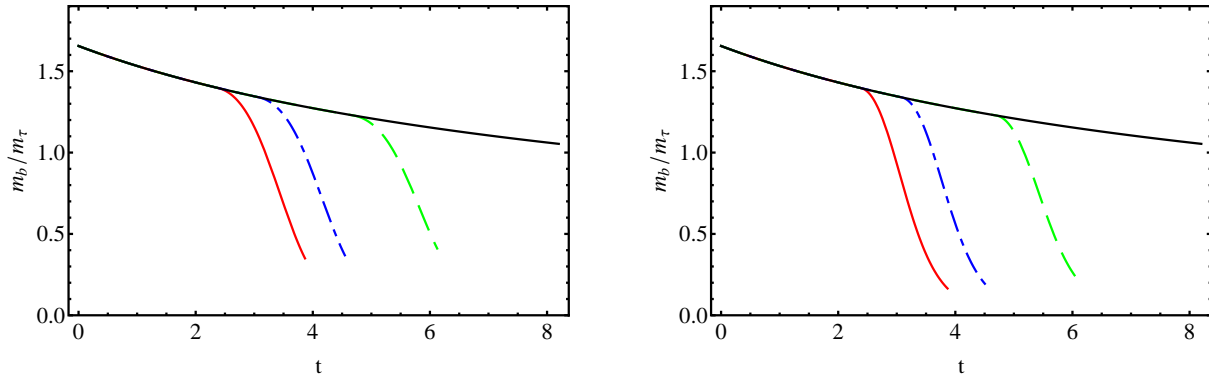


Figure 3. (Colour online) Evolution of the mass ratio $\frac{m_b}{m_\tau}$, with the same notations as Fig.2

3. Discussions and Conclusions

As illustrated in figures 1 2, 3, the gauge couplings and mass ratios evolve in the usual logarithmic fashion when the energies are below 1 TeV, 2 TeV and 10 TeV respectively. However, once the first KK threshold is reached the contributions from the KK states become increasingly significant and the effective 4D SM couplings begin to deviate from their normal trajectories. One finds that the running behaviours of the mass ratios are governed by the combination of the third family Yukawa couplings and the CKM matrix elements. This implies that the mass ratios of the first two light generations have a slowed evolution well before the unification scale. Beyond that point, their evolution diverges due to the faster running of the gauge couplings,

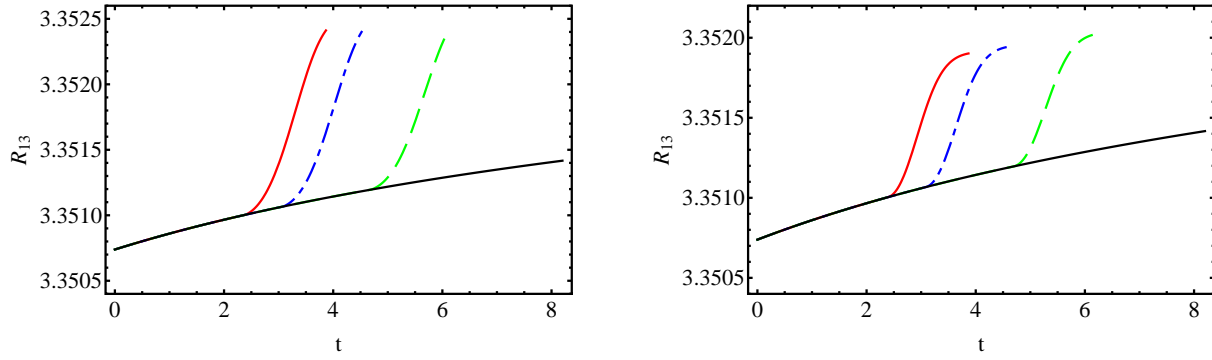


Figure 4. (Colour online) Evolution of the R_{13} , with the same notations as Fig.2

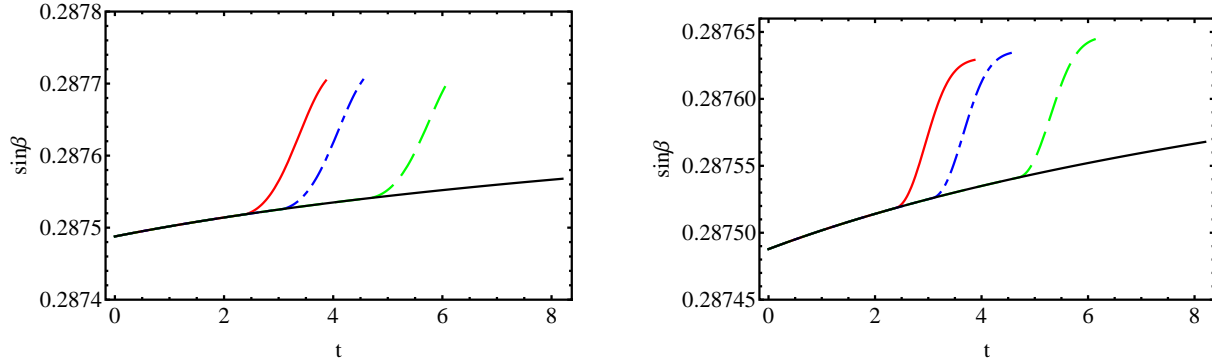


Figure 5. (Colour online) Evolution of $\sin \beta$, with the same notations as Fig.2

where any new physics would then come into play, and we find the scaling dependence of $\frac{m_d}{m_s}$ and $\frac{m_e}{m_\mu}$ is very slow.

On the other hand, Grand Unification Theories (such as $SU(5)$ and $SO(10)$) imply the well-known quark-lepton symmetric relation for fermion masses $m_d = m_e$. Due to power law running of the Yukawa couplings, the renormalisation effects on these relations can be large for $\frac{m_b}{m_\tau}$, for both scenarios, see figure 3. We have shown by numerical analysis of the one-loop calculation that the mass ratio $\frac{m_b}{m_\tau}$, as one crosses the KK threshold at $\mu = R^{-1}$ for both scenarios, results in a rapid approach to a singularity before the unification scale is reached, which agrees with what is observed in the SM, however, the mass ratios decrease at a much faster rate. Note that we observed similar behaviour for $\frac{m_d}{m_e}$ and $\frac{m_s}{m_\mu}$.

Let us now focus on the evolution of the set of renormalisation invariants R_{13} and R_{23} that describe the correlation between the mixing angles and mass ratios to a good approximation. With a variation of the order of λ^4 and λ^5 under energy scaling respectively, as shown in figure 4, the energy scale dependence is weak because the increase of the mixing angles are compensated by the deviation of the mass ratios. Therefore the effect is not large.

In figure 5 we present the evolution of the inner angle from the electroweak scale to the unification scale by using the one-loop RGE for the 2UED model, and demonstrate that the angle has a small variation against radiative corrections. To be more precise, the relative deviation for $\sin \beta$ is only up to 0.05% in the whole range studied. Similar analysis can also be found for the angles α and γ . This result makes sense, since both the triangle's sides and area become larger and larger when the energy scale increases, the unitarity triangle (UT) is only rescaled and its

shape does not change much during the RG evolution. The fact that inner angles are rather stable against radiative corrections indicates that it is not possible to construct an asymptotic model with some simple, special form of the CKM matrix from this simple scenario. The stability against radiative corrections suggests that the shape of the UT is almost unchanged from RGE effects. In this regards, the UT is not a sensitive test of this model in current and upcoming experiments.

In conclusion, the mass ratios in the 2UED model, with different possibilities for the matter fields, were discussed, where they are either bulk propagating or localised to the brane. We found that the 2UED model has substantial effects on the scaling of fermion masses for both cases, including both quark and lepton sectors. We quantitatively analysed these quantities for $R^{-1} = 1$ TeV, 2 TeV and 10 TeV, observing similar behaviours for all values of the compactification radius. We have shown that the scale dependence is not logarithmic, it shows a power law behaviour. We also found that for both scenarios the theory is valid up to the unification scale, leading to significant RG corrections.

Acknowledgments

We would like to thank our collaborators Aldo Deandrea and Ahmad Tarhini for the helpful discussions. This work is supported by the National Research Foundation (South Africa).

References

- [1] D. Falcone 2002 Fermion masses and mixings in gauge theories *Int. J. Mod. Phys. A* **17** 3981 (Preprint hep-ph/0105124)
- [2] L. X. Liu 2010 Renormalization invariants and quark flavor mixings *Int. J. Mod. Phys. A* **25** 4975 (Preprint hep-ph/0910.1326)
- [3] A. S. Cornell and L. X. Liu 2011 Evolution of the CKM matrix in the universal extra dimension model *Phys. Rev. D* **83** 033005 (Preprint hep-ph/1010.5522)
- [4] A. S. Cornell, A. Deandrea, L. X. Liu and A. Tarhini 2012 Scaling of the CKM matrix in the 5D MSSM *Phys. Rev. D* **85** 056001 (Preprint hep-ph/1110.1942)
- [5] T. Appelquist, H. C. Cheng and B. A. Dobrescu 2001 Bounds on universal extra dimensions *Phys. Rev. D* **64** 035002 (Preprint hep-ph/0012100)
- [6] B. A. Dobrescu and E. Poppitz 2001 Number of fermion generations derived from anomaly cancellation *Phys. Rev. Lett.* **87** 031801 (Preprint hep-ph/0102010)
- [7] B. A. Dobrescu and E. Ponton 2004 Chiral compactification on a square *JHEP* **0403** 071 (Preprint hep-th/0401032)
- [8] R. N. Mohapatra and A. Perez Lorenzana 2003 Neutrino mass, proton decay and dark matter in TeV scale universal extra dimension models *Phys. Rev. D* **67** 075015 (Preprint hep-ph/0212254)
- [9] N. Maru, T. Nomura, J. Sato and M. Yamanaka 2010 The universal extra dimensional model with $S^{*2}/Z(2)$ extra-space *Nucl. Phys. B* **830** 414 (Preprint hep-ph/0904.1909)
- [10] G. Cacciapaglia, A. Deandrea and J. Llodra Perez 2010 A dark matter candidate from Lorentz invariance in 6D *JHEP* **1003** 083 (Preprint hep-ph/0907.4993)
- [11] H. Dohi and K. y. Oda 2010 Universal extra dimensions on real projective plane *Phys. Lett. B* **692** 114 (Preprint hep-ph/1004.3722)
- [12] A. Abdalgabar, A. S. Cornell, A. Deandrea and A. Tarhini 2013 Evolution of Yukawa couplings and quark flavour mixings in 2UED models *Phys. Rev. D* **88** 056006 (Preprint hep-ph/1306.4852)
- [13] A. Abdalgabar, A. S. Cornell, A. Deandrea and A. Tarhini 2014 Higgs quartic coupling and neutrino sector evolution in 2UED models *Eur. Phys. J. C* **74** no. 5 2893 (Preprint hep-ph/1307.6401)
- [14] A. S. Cornell, A. Deandrea, L. X. Liu and A. Tarhini 2013 Renormalisation running of masses and mixings in UED models *Mod. Phys. Lett. A* **28** no. 11 1330007 (Preprint hep-ph/1209.6239)
- [15] A. Abdalgabar and A. S. Cornell 2013 Evolution of Yukawa couplings and quark flavour mixings in the 5D MSSM *J. Phys.: Conf. Ser.* **455** 012050 (Preprint hep-ph/1305.3729)
- [16] K. S. Babu 1987 Renormalization group analysis Of the Kobayashi-Maskawa matrix *Z. Phys. C* **35** 69
- [17] L. X. Liu and A. S. Cornell 2010 Scaling of Yukawa couplings and quark flavor mixings in the UED model *PoS KRUGER* **2010** 045 (Preprint hep-ph/1103.1527)
- [18] Z. z. Xing, H. Zhang and S. Zhou 2008 Updated values of running quark and lepton masses *Phys. Rev. D* **77** 113016 (Preprint hep-ph/0712.1419)

Structural and electronic properties of iron doped technetium sulphide

Mahmud Abdulsalam and Daniel P Joubert

The National Institute for Theoretical Physics, School of Physics and
Mandelstam Institute for Theoretical Physics, University of the Witwatersrand,
Johannesburg, Wits 2050, South Africa, tel.: +27 (0)11 717 6804, fax: +27
(0)11 717 6830.

E-mail: mahmudaaa@yahoo.co.uk

Abstract. In this paper Density Functional Theory with the inclusion of van der Waals corrections is used to study the effect of Iron doping on the structural and electronic properties of TcS_2 (technetium disulphide) in the $P\bar{1}$ structure. We show that Fe doped TcS_2 is stable against decomposition to its constituent solid components for the configurations examined and that it is possible to modify the bandgap of these indirect bandgap materials by doping. Substitutional Fe doping of TcS_2 decreases the bandgap from the un-doped configuration. The degree of decrease depends on the percentage doping and the doping configuration. We find that at 25 substitutional doping percentage the system is a magnetic semiconductor.

1. Introduction

Transition metal dichalcogenides (TMDCs) are a class of inorganic materials with the formula MX_2 where M is a transition metal (Tc, Mo, W, Nb, Re, Ti and Ta) and X is a member of the chalcogen family (S, Se, Te) [1–3]. TMDCs tend to crystallise in a layered structures that arises from the stacking of planes composed of transition metal atoms sandwiched between two chalcogenide atoms in the form X-M-X as shown in figure 1 for TcS_2 in the $P\bar{1}$ structure. This sandwich layer is internally strongly bonded while the interaction between layers is mainly through van der Waal's forces [4]. The weak inter-layer forces makes it possible to fabricate ultra thin layer by exfoliation or cleavage. TMDCs have a wide range of useful properties, ranging from optical, catalytic, electronic, mechanical, thermal and chemical and this makes them materials of interest for researchers [5–7].

The electronic properties of TMCs range from metallic to semiconducting [8]. To achieve the most practical application of TMDCs in nano-devices, solar energy application and other electronic applications, their electronic and optical properties can be modulated by: the application of external fields, forces, alloying [9, 10], modified layer stacking or doping [11–17]. Recent investigations have shown that application of strain can change the band gap of a TMDC from indirect to direct, while transition from semi-conductor to metal has also been observed in some TMDCs [18, 19]. The band gap of these materials can be tuned by the application of external electric fields [20, 21].

Another method by which the electronic property of a TMDCs can be tuned is doping [16, 17]. It is the purpose of this paper to investigate the structural and electronic properties of pure and Fe doped TcS_2 in the $P\bar{1}$ structure [22], a structure TcS_2 has been synthesized in. To the best

of our knowledge, this is the first time such an investigation is reported. In order to achieve a possible modification of the electronic properties of TcS_2 , we use TcS_2 with the substitution of Fe for Tc. Based on Hick's [23] rule for solubility, the ionic radius of the host and intercalating metal should be close and the doped metal should come from a group adjacent to the host metal in the periodic table. Based on this, we decided investigate substitutional doping of TcS_2 by replacing selected Tc atoms by Fe atom, an element from the adjacent group in the periodic table, having one electron more than Tc in its d-shell. We calculated the DFT electronic band structure of $\text{Tc}_{1-x}\text{SFe}_x$ and $\text{Tc}_{1-x}\text{Fe}_x\text{S}$ ($x = 0.25$ or 0.5), corresponding to 25% and 50% Fe doping respectively by considering modifications of the primitive unit cell only.

The electronic properties of TcS_2 have been studied by some researchers using theoretical and experimental approaches [24–27]. Wilson et. al [27] performed optical measurements which yielded a band gap of 1.3 eV. Fang. et.al [25] using the localized-spherical-wave (LSW) method and Weck et.al. [26] using GGA-DFT reported on the electronic properties of TcS_2 structures. Oviedo-Roa et. al. [24] used a periodic density functional approach to investigate catalytic properties of 4d-transition-metal sulfides. The reported experimental band gap is very close to the optimal band gap of 1.4 eV for solar energy applications. It is the focus of this paper to investigate the possibility of engineering the electronic properties of the TcS_2 for optimal solar energy application through doping. However, due to page restrictions, only modifications of the primitive unit cell of the structure will be considered. An in depth study of an extensive range of configurations will be reported elsewhere.

2. Computational Method

We used Density Functional Theory (DFT) [28, 29] with the inclusion of van der Waals forces as proposed by Grime (DFT-D2) [30]. Van der Waals corrections are necessary for the accurate description of the structural properties of layered materials. The Projector Augmented Wave (PAW) [31] as implemented in *Vienna ab initio Simulation Package* (VASP) [32, 33] is used to mimic electron-ion interaction and the Perdew, Burke and Ernzerhof (PBE) [34–36] parametrization of the generalised gradient approximation (GGA) [37–39] is used for the exchange correlation potential. The van der Waals correction in DFT-D2 adds a correction to the PBE approximation [30].

A $5 \times 5 \times 4$ Γ -centred Monkhorst-Pack mesh was used for sampling the Brillouin zones and the energy cut-off used in the calculations is 520 eV. These parameters were found to be adequate to energy convergence of less than 0.1 meV.

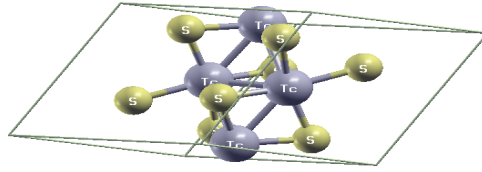
3. Results and discussions

3.1. Structural Properties

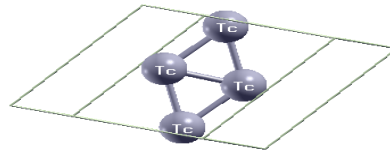
In order to determine the equilibrium properties of each of the studied structures, we calculated the cohesive energy per atom (E_{coh}) as a function of volume per atom V (eV) and fitted a third order Birch-Murnaghan equation of state (EOS) [40] to the calculated values. The volume which corresponds to the lowest cohesive energy is taken as the equilibrium volume of the structure and all subsequent calculations of the structure are carried out at this volume.

In the $P\bar{1}$ structure, the chemical unit per unit cell is Tc_4S_8 (figure 1a). In the $P\bar{1}$ equilibrium structure the Tc layer is sandwiched between two layers of S atoms. The four Tc atoms do not lie on a plane, but are slightly buckled. There are two Tc atomic positions that have different environments, labelled Tc1, Tc4 and Tc2, Tc3, respectively, in figure 1b.

For the 25% substitutional doping case there are 2 symmetrically distinct possibilities for replacing the Tc atom by Fe while for 50% doping there are 3 symmetrically distinct possibilities as shown in table 1.



(a) TcS_2



(b) Positions of Tc atoms.

Figure 1: Primitive unit cell for Technetium Sulphide in the $P\bar{1}$ structure and the positions of the four Technetium atoms.

Table 1: Configurations for substitutional doping of Tc by Fe of TcS_2 in the $P\bar{1}$ structure. Numbers in the top row refer to the labelling of atomic sites in figure 1b

	1	2	3	4
un-doped	Tc	Tc	Tc	Tc
d1.0.25	Fe	Tc	Tc	Tc
d2.0.25	Tc	Fe	Tc	Tc
d1.0.50	Fe	Fe	Tc	Tc
d2.0.50	Fe	Tc	Tc	Fe
d3.0.50	Tc	Fe	Fe	Tc

In table 2 the equilibrium parameters for the Technetium Sulphide and Iron doped structures are listed.

Table 2: Technetium Dichalcogenides equilibrium lattice parameters, bulk modulus, Cohesive energy E_{coh} per atom, volume, Formation energy (E_f) per atom and magnetic moment per unit cell in Bohr magnetons.

Material	a(Å)	c/a(Å)	B_0 (GPa)	B'_0	E_{coh} (eV)	V_0 (Å ³)	E_f (eV)	Magnetic moment
TcS ₂ (un-doped)	6.37 (6.37[a])	1.06 (1.04 [a])	28.80	12.78	-5.22	217.08 (213.39[a])	-0.62	0.00
50 percent doping								
d1_0.50	6.32	1.20	29.09	9.98	-4.75	204.72	-0.51	0.00
d2_0.50	6.28	1.20	30.29	8.12	-4.72	205.92	-0.48	0.00
d3_0.50	6.28	1.20	32.34	11.15	-4.69	204.24	-0.45	0.00
25 percent doping								
d1_0.25	6.40	1.31	29.63	10.12	-4.94	211.80	-0.52	1.00
d2_0.25	6.41	1.31	30.32	12.70	-4.95	211.92	-0.53	1.00

a = Experimental values, reference [22]

From the table 2 it is clear that the obtained TcS₂ values for the lattice constants and volume of the unit cell that our results agree well with the available literature. Since the cohesive energy and formation energy (the quantities that are used to determine the stability of the studied structures) follow the same trend, we will therefore resort to the use of formation energy which is the energy liberated when the system is decomposed into its constituent solid components, as measure of stability of the structures.

The formation energy is given by

$$E_f(S_m Tc_n Fe_o^{solid}) = E_{coh}(S_m Tc_n Fe_o^{solid}) - \frac{mE_{coh}(S^{solid}) + nE_{coh}(Tc^{solid}) + oE_{coh}(Fe^{solid})}{m + n + o} \quad (1)$$

where m, n and o are integers representing the number of sulphur, technetium and iron involved in the formation the system in question.

A positive value of formation energy at a given temperature and pressure signifies instability in the structure and tendency of the structure to decompose to its constituent solid components. On the other hand, the more negative the value of the formation energy, the more stable the structure is. Our results suggest that the doped structures are stable with respect to decomposition to their component solids and the energetically most favourable doped structures are in the order of d1_0.50, d2_0.25, d1_0.25, d1_0.50, d2_0.50 and d3_0.50.

We note that there is induced magnetisation for 25% doping. Roughly half to the total magnetisation is from the Fe atom and one Tc atom in a ferro magnetic arrangement, in both cases. In the case of d1_0.25, there is induced magnetization on Tc_2 due to the presence of Fe_1. For d2_0.25, the induced magnetization is on Tc_3 due Fe_2. Magnetic moments associated with other atoms are negligible. For the 50% doping, there is no net magnetization and no magnetic moments associated with any atom in the structures. The magnetic semiconductors are materials of interest in the field of spintronics.

3.2. Electronic Properties

TcS₂ and its Fe doped derivatives studied in this report are indirect band gap materials. In table 3 the band gaps for the different configurations are listed. We can see that from the table that the dominant contributing orbital character remains Sulphur p-type, S(p), in all cases, apart from the d2_0.50 configuration, despite the relatively high doping levels considered. The Fe orbitals have little contribution at the band edges. However, as a result of hybridisation with the Fe s-orbitals in all cases apart from the d1_0.50 configuration where the dominant hybridisation is with Fe(d), the band gap of the doped system is narrowed in all cases.

Table 3: Band gaps DFT-D2 and MBJ (in brackets) for un-doped and Fe doped TcS₂ systems with their dominant contributing orbitals at the band edges

Configuration	Bandgap (eV)	others	Dominant orbitals at the band edge	
			VBM	CBM
un-doped	0.94 (1.20)	0.9[b]	S(p)	S(p)
d1.0.25	0.22 (0.60)		S(p)	S(p)
d2.0.25	0.31 (0.54)		S(p)	S(p)
d1.0.50	0.36 (0.48)		S(p)	S(p)
d2.0.50	0.20 (0.41)		S(p)	S(s)
d3.0.50	0.18 (0.42)		S(p)	S(p)

b = reference [26]

The formation energies of the different doping configurations (see table 2) are relatively close and it is likely that all configurations will be found in a real system. Local bandgaps of doped systems will probably differ from region to region. From table 3 we note that substitutional doping of TcS₂ with Fe can significantly change the bandgap. In all cases studied here the bandgap is decreased from that of the un-doped configuration.

4. Conclusions

In this preliminary report we numerically investigated the feasibility of substitutionally doping TcS₂ with Fe. We only explored a sub-set of possible doping configurations for a 25% and 50% substitutional doping. Other configurations are possible if larger unit cells are considered and this is part of ongoing work.

We found that by including van der Waals corrections to the PBE approximation, accurate structural properties for TcS₂ in the $P\bar{1}$ structure, a layered structure, can be determined from numerical simulations. This encourages us to believe that accurate predictions of the structural properties for the Fe doped TcS₂ configurations can be determined using the same approach. The sub-set of doped configurations we considered are all predicted to be stable against decomposition to their component solids. We are testing for stability against structural deformations.

The indirect DFT-D2 bandgaps are sensitive to doping with the bandgaps of all the doped configurations smaller than the bandgaps of the un-doped systems.

DFT bandgaps are not accurate predictors of the fundamental or optical bandgaps, though the MBJ values tend to be reasonably accurate. Challenging numerical techniques such as GW [41] and BSE [42] approaches must be used for accurate predictions of band gaps and electronic structure. The general trend found here, that doping decreases the band gap, is expected to prevail when more accurate GW and BSE techniques are used.

We found that the 25% substitutionally Fe doped TcS₂ structures are magnetic semiconductors. These systems merit further investigation as materials of interest to spintronics.

In summary, we have shown numerically that Fe doped TcS₂ can be stable and that it is possible to engineer the band gap by doping.

5. Acknowledgement

The support of the DST-NRF Centre of Excellence in National institute for theoretical physics (NITheP) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the CoE. We also wish to acknowledge Centre for high performance computing, South Africa, for providing us with computing facilities.

References

- [1] H. R. Gutiérrez, N. Perea-López, A. L. Elías, A. Berkdemir, B. Wang, R. Lv, F. López-Urías, V. H. Crespi, H. Terrones, and M. Terrones *Nano letters*, vol. 13, no. 8, pp. 3447–3454, 2012.
- [2] A. Splendiani, L. Sun, Y. Zhang, T. Li, J. Kim, C.-Y. Chim, G. Galli, and F. Wang *Nano letters*, vol. 10, no. 4, pp. 1271–1275, 2010.
- [3] A. Kumar and P. Ahluwalia *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 85, no. 6, pp. 1–7, 2012.
- [4] J. Verble, T. Wietling, and P. Reed *Solid State Communications*, vol. 11, no. 8, pp. 941–944, 1972.
- [5] J. Wilson and A. Yoffe *Advances in Physics*, vol. 18, no. 73, pp. 193–335, 1969.
- [6] A. Yoffe *Advances in Physics*, vol. 42, no. 2, pp. 173–262, 1993.
- [7] A. Yoffe *Annual Review of Materials Science*, vol. 3, no. 1, pp. 147–170, 1973.
- [8] Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, and M. S. Strano *Nature nanotechnology*, vol. 7, no. 11, pp. 699–712, 2012.
- [9] D. O. Dumcenco, H. Kobayashi, Z. Liu, Y.-S. Huang, and K. Suenaga *Nature communications*, vol. 4, p. 1351, 2013.
- [10] J. Kang, S. Tongay, J. Li, and J. Wu *Journal of Applied Physics*, vol. 113, no. 14, p. 143703, 2013.
- [11] Y.-W. Son, M. L. Cohen, and S. G. Louie *Nano letters*, vol. 7, no. 11, pp. 3518–3522, 2007.
- [12] L. Kou, C. Tang, W. Guo, and C. Chen *ACS nano*, vol. 5, no. 2, pp. 1012–1017, 2011.
- [13] Y. Mao and J. Zhong *Nanotechnology*, vol. 19, no. 20, p. 205708, 2008.
- [14] V. W. Brar, R. Decker, H.-M. Solowan, Y. Wang, L. Maserati, K. T. Chan, H. Lee, Ç. O. Girit, A. Zettl, S. G. Louie, *et al.* *Nature Physics*, vol. 7, no. 1, pp. 43–47, 2010.
- [15] D. Ma, Z. Lu, W. Ju, and Y. Tang *Journal of Physics: Condensed Matter*, vol. 24, no. 14, p. 145501, 2012.
- [16] J. Honolka, A. Khajetoorians, V. Sessi, T. Wehling, S. Stepanow, J.-L. Mi, B. B. Iversen, T. Schlenk, J. Wiebe, N. Brookes, *et al.* *Physical review letters*, vol. 108, no. 25, pp. 256811–256816, 2012.
- [17] J.-M. Zhang, W. Zhu, Y. Zhang, D. Xiao, and Y. Yao *Physical review letters*, vol. 109, no. 26, p. 266405, 2012.
- [18] S. Bhattacharyya and A. K. Singh *Physical Review B*, vol. 86, no. 7, p. 075454, 2012.
- [19] E. Scalise, M. Houssa, G. Pourtois, V. Afanasev, and A. Stesmans *Nano Research*, vol. 5, no. 1, pp. 43–48, 2012.
- [20] A. Ramasubramaniam, D. Naveh, and E. Towe *Physical Review B*, vol. 84, no. 20, p. 205325, 2011.
- [21] K. Dolui, C. D. Pemmaraju, and S. Sanvito *ACS nano*, vol. 6, no. 6, pp. 4823–4834, 2012.
- [22] H.-J. Lamfers, A. Meetsma, G. Wiegers, and J. De Boer *Journal of alloys and compounds*, vol. 241, no. 1, pp. 34–39, 1996.
- [23] W. Hicks *Journal of The Electrochemical Society*, vol. 111, no. 9, pp. 1058–1065, 1964.
- [24] R. Oviedo-Roa, J.-M. Martínez-Magadán, and F. Illas *The Journal of Physical Chemistry B*, vol. 110, no. 15, pp. 7951–7966, 2006.
- [25] C. Fang, G. Wiegers, C. Haas, and R. De Groot *Journal of Physics: Condensed Matter*, vol. 9, no. 21, p. 4411, 1997.
- [26] P. F. Weck, E. Kim, and K. R. Czerwinski *Dalton Transactions*, vol. 42, no. 43, pp. 15288–15295, 2013.
- [27] J. Wilson, F. DiSalvo, and S. Mahajan *Phys*, vol. 24, p. 117, 1975.
- [28] P. Hohenberg and W. Kohn *Physical review*, vol. 136, no. 3B, pp. B864–B871, 1964.
- [29] W. Kohn and L. J. Sham *Physical Review*, vol. 140, no. 4A, p. A1133, 1965.
- [30] S. Grimme *Journal of computational chemistry*, vol. 25, no. 12, pp. 1463–1473, 2004.
- [31] G. Kresse and D. Joubert *Physical Review B*, vol. 59, no. 3, p. 1758, 1999.
- [32] G. Kresse and J. Hafner *Physical Review B*, vol. 49, no. 20, p. 14251, 1994.
- [33] G. Kresse and J. Hafner *Physical Review B*, vol. 47, no. 1, p. 558, 1993.
- [34] J. P. Perdew, K. Burke, and M. Ernzerhof *Phys. Rev. Lett.*, vol. 77, pp. 3865–3868, Oct 1996.
- [35] J. P. Perdew, K. Burke, and M. Ernzerhof *Phys. Rev. Lett.*, vol. 78, pp. 1396–1396, Feb 1997.
- [36] M. Ernzerhof and G. E. Scuseria *The Journal of chemical physics*, vol. 110, no. 11, pp. 5029–5036, 1999.
- [37] A. D. Becke *Phys. Rev. A*, vol. 38, pp. 3098–3100, Sep 1988.
- [38] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais *Phys. Rev. B*, vol. 46, pp. 6671–6687, Sep 1992.
- [39] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais *Phys. Rev. B*, vol. 48, pp. 4978–4978, Aug 1993.
- [40] F. Birch *Phys. Rev.*, vol. 71, pp. 809–824, Jun 1947.
- [41] M. Shishkin and G. Kresse *Phys. Rev. B*, vol. 75, p. 235102, 2007.
- [42] A. Ramasubramaniam *Phys. Rev. B*, vol. 86, p. 115409, Sep 2012.

An Open Quantum Systems Approach to Avian Magnetoreception

Betony Adams, Ilya Sinayskiy, Francesco Petruccione

Quantum Research Group, School of Chemistry and Physics, University of KwaZulu-Natal
Durban, KwaZulu-Natal, 4001, South Africa and National Institute for Theoretical Physics
(NITheP), KwaZulu-Natal, 4001, South Africa

E-mail: betony@gmail.com

Abstract. The emerging field of Quantum Biology centres on the possibility that biological systems might employ nontrivial quantum effects in their day to day behaviour. This surprising result has given rise to the investigation of such quantum effects in areas as diverse as photosynthesis and magnetoreception. In the case of avian magnetoreception, experiment supports the role of a radical pair mechanism in how birds sense the magnetic field. Following from radical pair theory and using the theory of open quantum systems we have completed the analytical derivation of the master equation in the Born-Markov approximation for the simple case of two electrons, each interacting with an environment of N nuclear spins as well as the external magnetic field, then placed in a boson bath and allowed to dissipate. We have then solved the master equation and analysed the dynamics of the radical pair.

1. Introduction

It is generally well accepted that birds employ the earth's magnetic field in their remarkable feats of migration. The exact mechanism of this use is less well understood. Of the two main hypotheses, one suggests that ferromagnetic crystals in the birds' beaks align themselves in the earth's magnetic field and thus allow the birds to orientate themselves [1]. However, experiment demonstrates that these crystals sense magnetic intensities that are orders of magnitude bigger than the weak geomagnetic field [2]. Current consensus is that birds employ another more sensitive mechanism, a chemical compass which utilises quantum effects to 'see' the magnetic field. The detail of this compass, in the form of a radical-pair mechanism, was first proposed by Schulten *et al.* in 1978 [3]. The mechanism relies on the photo-activated creation of a radical pair in a singlet spin state, which then undergoes singlet-triplet mixing to result in different chemical signatures. This hypothesis is supported by evidence that efficient avian magnetoreception is light-dependent, requires an undamaged visual system [4] and is sensitive to weak magnetic fields [8]. One of the most convincing pieces of evidence in support of the radical pair mechanism is the experimental demonstration that oscillating radiofrequency fields cause birds to be disoriented. It was first noted in 2004 that birds are disoriented in oscillating fields of MHz range, most remarkably at a frequency of 1.315 MHz [5, 6]. This, being the Larmor frequency of a free electron in a geomagnetic field, led to speculation that one of the radicals of the pair has no hyperfine interactions. Ritz *et al.* went on to specify the other specific frequencies at which birds are disoriented, although to a much lesser degree, concluding that such disorientation would be explicable by the fact that for a radical with no hyperfine interactions the Larmor

resonance would have greater effect than for a frequency corresponding to only one of the various possible energy-level splittings of the hyperfine interaction [7]. Recent experiments conducted on migratory birds over seven years at Oldenburg in Germany have also demonstrated that birds' ability to navigate effectively is compromised by anthropogenic electromagnetic radiation over a frequency range from 50 kHz to 5MHz [9].

The work documented in this paper revisits radical pair theory through an open quantum systems approach. Firstly we derive a master equation in the Born-Markov approximation which describes the dynamics of the radical pair as it dissipates in a boson bath approximating the environment. We then solve the master equation and analyse the dynamics of the radical pair. This approach is useful in its theoretical verification of experimental parameters used in other models of the avian compass, such as the lifetime of the radical pair and the rates at which singlet-triplet mixing occurs [10][11]. By varying the number of nuclear spins interacting with the radical pair we can also investigate the effects of different nuclear environments and thus conclude something about the structure of cryptochrome, the biological molecule in which it is thought that the radical pair reaction takes place [14].

2. Modelling the Radical Pair Mechanism

2.1. The System

The heart of the magnetic compass, the radical pair, results from the photo-excitation of an electron to form a spatially separated but spin-correlated electron pair. The radical pair begins in a singlet state $\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$, but the hyperfine interaction of each electron with its nuclear environment, as well as with the magnetic field, induces singlet-triplet mixing. For the purposes of this research we will neglect the dipole-dipole and exchange interactions between the two electrons due to sufficient spatial separation [11][12]. We also neglect the interaction of nuclear spins as the gyromagnetic ratio is small compared to the electronic case [12].

The Hilbert space of the system is given by $\mathcal{H}_s = \mathcal{H}_e^{(1)} \otimes \mathcal{H}_e^{(2)} \otimes \mathcal{H}_n^{(1)} \otimes \mathcal{H}_n^{(2)}$ where \mathcal{H}_e refers to the Hilbert space of either electron and is two dimensional while \mathcal{H}_n refers to the Hilbert space of the nuclear spin environment and is of dimension $(2j + 1) \times (2j + 1)$. In the following we already take into account that, for example, $S_z^{(1)} = S_z^{(1)} \otimes I_e^{(2)} \otimes I_n^{(1)} \otimes I_n^{(2)}$ where I_e and I_n are the identity matrices for electron and nuclei. Thus an expression such as $S_+^{(1)} I_-^{(1)}$ uses ordinary matrix multiplication. The Hamiltonian for this system models the two electrons, labelled (1) and (2), each interacting separately, and can be written as

$$H_s = \gamma_e(\vec{B}\mathbf{S}^{(1)} + \vec{B}\mathbf{S}^{(2)}) + \lambda \sum_{k=1}^2 \sum_{n,l=1}^3 A_{nl}^{(k)} \mathbf{S}_n^{(k)} \mathbf{I}_l^{(k)} \quad (1)$$

where \vec{B} is the magnetic field vector and

$$\mathbf{S} = (S_x, S_y, S_z) \quad \mathbf{I} = (I_x, I_y, I_z)$$

are the vectors of spin operators for electron and nuclear spin respectively and $\gamma_e \vec{B}\mathbf{S}$ is the Zeeman interaction with electron gyromagnetic ratio $\gamma_e = -g\mu_B$, where $g = 2$ and μ_B is the Bohr magneton. The sum runs over k to include both electrons. The first electron is anisotropically coupled to its spin environment while the second is isotropically coupled, with respective hyperfine coupling tensors

$$A^{(1)} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad A^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and hyperfine coupling constant given by λ .

Following from work done in [13] we choose as a basis for the electron spin states $\{|0\rangle, |1\rangle\}$, with $|0\rangle$ the ground state and $|1\rangle$ the excited state. From this we can express the raising and lowering operators as

$$S_+ = |0\rangle\langle 1| \quad S_- = |1\rangle\langle 0|$$

S_z in this basis is

$$S_z = \frac{1}{2}(|0\rangle\langle 0| - |1\rangle\langle 1|)$$

For the nuclear spins, j is the total angular momentum of the N spin-half nuclei, which gives $2j + 1$ possible states, from $m = -j, \dots, j$ where m is the magnetic quantum number. For N spins the basis can be taken as $\{|j, m\rangle, m = -j, \dots, j\}$. In this basis the raising and lowering operators are

$$I_+ = \sum_{m=-j}^{j-1} A_{jm}^+ |j, m+1\rangle\langle j, m| \quad I_- = \sum_{m=-j+1}^j A_{jm}^- |j, m-1\rangle\langle j, m|$$

where $A_{jm}^\pm = \sqrt{j(j+1) - m(m \pm 1)}$. I_z in this basis is

$$I_z = \sum_{m=-j}^j m |j, m\rangle\langle j, m|$$

We now look again at the system Hamiltonian for a single electron, and, using $S_x I_x + S_y I_y = \frac{1}{2}(S_+ I_- + S_- I_+)$, rewrite it as

$$H_S^{(1)} = \gamma_e B_0 S_z^{(1)} + \lambda(S_+^{(1)} I_-^{(1)} + S_-^{(1)} I_+^{(1)}) + \frac{\lambda}{2}(S_z^{(1)} I_z^{(1)})$$

where we have taken the magnetic field as pointing along the positive z axis. Using the relevant expressions above we arrive at a Hamiltonian that is not diagonal in the basis

$$\{|0, j, j\rangle, |1, j, -j\rangle, |1, j, m\rangle, |0, j, m-1\rangle\}$$

By finding eigenvalues and eigenvectors we can write the new diagonal Hamiltonian in the basis of its eigenvectors

$$\{|0, j, j\rangle, |1, j, -j\rangle, |\lambda_{jm}^- \rangle, |\lambda_{jm}^+ \rangle\}$$

In this basis the system Hamiltonian looks like

$$\begin{aligned} H_S^{(1)} = & \frac{1}{2}[(\gamma_e B_0 + \lambda j) |0, j, j\rangle\langle 0, j, j| - (\gamma_e B_0 - \lambda j) |1, j, -j\rangle\langle 1, j, -j|] \\ & + \sum_{m=-j+1}^j [v_1(j, m) |\lambda_{jm}^- \rangle\langle \lambda_{jm}^-| + v_2(j, m) |\lambda_{jm}^+ \rangle\langle \lambda_{jm}^+|] \end{aligned} \quad (2)$$

The diagonal system Hamiltonian for the second electron is found in the same way.

As the evolution of the system is happening at physiological temperatures it is natural to assume the system is embedded in a dissipative bosonic environment [16]. This can be written as

$$H_I = \sum_n [g_n a_n + \bar{g}_n a_n^\dagger] \otimes [\alpha(S_x^{(1)} + S_z^{(1)}) + \alpha(S_x^{(2)} + S_z^{(2)})] \quad (3)$$

where a_n^\dagger and a_n are creation and annihilation operators. The action of these operators will be described in the discussion of the derivation of the master equation below. Typically in the theory of open quantum systems one would use S_x or S_z coupling only, leading to either dissipation or decoherence. As there is currently no evidence which is the dominant interaction we include both here.

2.2. The Transition Operators

Following the approach suggested in [15] the jump operators which describe the transitions between the possible levels the system can occupy are found by taking the commutators

$$[H_S, V_k] = -\omega_k V_k \quad [H_S, V_k^\dagger] = \omega_k V_k^\dagger$$

where $V = S_x + S_z$ follows from the interaction Hamiltonian. Transition operators V_k and V_k^\dagger for the first electron take the form

$$V_1 = |\lambda_{j,j}^- \rangle \langle 0, j, j|, V_2 = |1, j, -j \rangle \langle \lambda_{j,-j+1}^+ | \dots V_k$$

where k here labels the number of transition operators and $|0, j, j \rangle, |1, j, -j \rangle, |\lambda_{jm}^- \rangle, |\lambda_{jm}^+ \rangle$ is the diagonal basis of the system Hamiltonian for the first electron, as described above. Corresponding transition frequencies $\omega_1, \omega_2 \dots \omega_k$ are expressed in terms of the magnetic field and the hyperfine coupling constant where negative frequencies denote energy leaving the system and vice versa. Transition operators and frequencies for the second electron are calculated in a similar manner with the isotropy of the hyperfine interaction in this case resulting in slightly different transition parameters.

The decoherence in the system is found by taking the commutator

$$[H_S, V_0] = 0$$

where the transition frequency of zero reflects the fact that there is no energy flow in or out of the system.

2.3. The Master Equation

In the interaction picture we can rewrite the interaction Hamiltonian using the transition operators as

$$\begin{aligned} H_I(t) = \sum_n [g_n a_n e^{-i\omega_n t} + \bar{g}_n a_n^\dagger e^{i\omega_n t}] \otimes [(\sum_{j=1}^{N_T} \alpha(V_j^{(1)} e^{-i\omega_j^T t} + V_j^{\dagger(1)} e^{i\omega_j^T t}) + \alpha V_0^{(1)}) \\ + (\sum_{j=1}^{N_T} \alpha(V_j^{(2)} e^{-i\omega_j^T t} + V_j^{\dagger(2)} e^{i\omega_j^T t}) + \alpha V_0^{(2)})] \end{aligned} \quad (4)$$

Here the sum runs to N_T which is the number of transition operators. As the number of nuclei increase the number of transition operators also increase, for example with only a single nucleus in the system there are only five transition operators whereas for all ten nuclei there will be fifty.

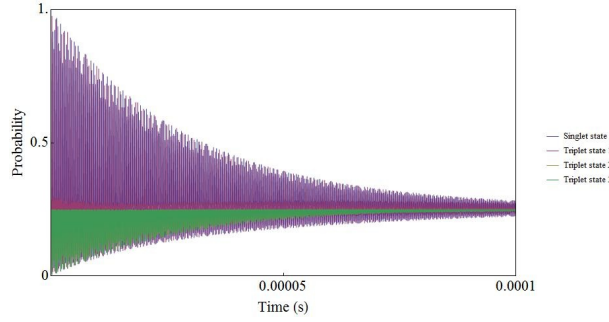


Figure 1. Radical pair dynamics for $j = \frac{1}{2}$

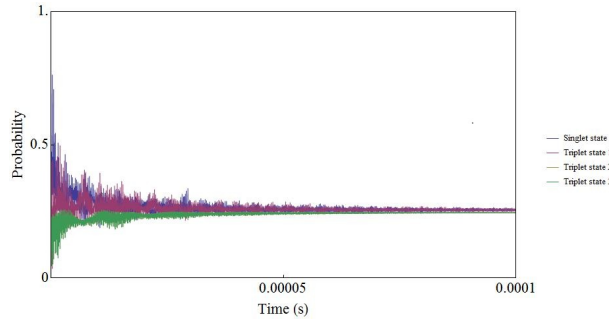


Figure 2. Radical pair dynamics for $j = 1$

In the derivation of the master equation the bath correlation functions determine the rates at which transitions happen, with the only contributions coming from

$$\langle a_n a_m^\dagger \rangle = \text{Tr}_B[a_n a_m^\dagger \rho_B] = \delta_{nm}(n(\omega_n) + 1) \quad \langle a_n^\dagger a_m \rangle = \text{Tr}_B[a_n^\dagger a_m \rho_B] = \delta_{nm}n(\omega_n)$$

where $n(\omega_n)$ is the Planck distribution at a specific frequency. After applying the Born-Markov approximations and completing the derivation the master equation can be expressed as

$$\begin{aligned} \frac{d}{dt} \rho_S^I = & \sum_{k=1}^2 \gamma_D [V_0^{(k)} \rho_S^I V_0^{(k)} - \frac{1}{2} \{V_0^{(k)} V_0^{(k)}, \rho_S^I\}] + \sum_{k=1}^2 \sum_{j=1}^{N_T} \gamma_j [V_j^{(k)} \rho_S^I V_j^{(k)} - \frac{1}{2} \{V_j^{(k)} V_j^{(k)}, \rho_S^I\}] \\ & + n(\omega_j^T) [V_j^{(k)} \rho_S^I V_j^{\dagger(k)} - \frac{1}{2} \{V_j^{\dagger(k)} V_j^{(k)}, \rho_S^I\}] + n(\omega_j^T) [V_j^{\dagger(k)} \rho_S^I V_j^{(k)} - \frac{1}{2} \{V_j^{(k)} V_j^{\dagger(k)}, \rho_S^I\}] \end{aligned} \quad (5)$$

where γ_j is the rate of spontaneous dissipation, γ_D the rate of decoherence and

$$n(\omega_j^T) = \frac{1}{\exp\left(\frac{\hbar \omega_j^T}{k_B T}\right) - 1}$$

gives the number of thermal photons (bosons) in a mode of frequency ω_j^T at a given temperature T and k_B is the Boltzmann constant. Here the sum also runs over k in order to include both electrons. The first term describes the decoherence while the last three describe the dissipation. The first of these dissipation terms accounts for spontaneous emission and the last two terms account for stimulated emission and absorption processes respectively, these are due to thermal fluctuations at a given temperature.

3. Solving the Master Equation

3.1. Parameters

In order to investigate the dynamics of the radical pair we took the geomagnetic field to be $47\mu\text{T}$ [10]. In the relevant literature hyperfine coupling constants for organic molecules range in value. We selected $\lambda = 30\text{MHz}$ [11] for the purposes of this paper. To accurately apply the model to biological systems we took the temperature to be 300K . To calculate the appropriate frequency-dependent rates of dissipation we used

$$\gamma_j = \frac{\omega_j^3}{3\epsilon_0\pi\hbar c^3}|d|^2 \quad (6)$$

and by approximating our transitions as those of the equivalent Rydberg atom we were able to calculate the transition dipole moment to be of the order of $10^{10}ea_0$. With these parameters we were then able to simulate the evolution of the radical pair.

3.2. Preliminary Results and Conclusions

The results for two possible variations of the radical pair mechanism can be seen above. Figure 1 demonstrates singlet-triplet mixing where each electron in the pair interacts with a single nuclear spin only, while Figure 2 shows singlet-triplet mixing for the electrons each interacting with two nuclear spins, taking into account the different possible alignments of these spins. In both instances the lifetime of the radical pair is of the order of microseconds, a surprisingly long-lived coherence which nonetheless verifies values arrived at through various methods elsewhere in the literature where coherence times have even been suggested to be of the order of milliseconds [17]. The difference in our approach as compared, for example, to Gauger *et al.*, [10] being that instead of estimating rates from experiment we have derived our model from first principles. The agreement of our derived lifetime with those estimated in the literature suggests the feasibility of our open quantum approach.

Being as there are only two of the ten proposed cases reported here it is too early for any absolute conclusions. One of the intentions of the research is to compare, for example, the effects of half integer as opposed to integer spin for the nuclear environment. As it stands it would appear that dissipation happens slower in the case of half integer spin, even though the lifetimes are comparable. This would have to be verified for all ten cases. The dissipation in the $j = \frac{1}{2}$ case also appears smoother whereas in the $j = 1$ case there is an initial more rapid decline in the singlet state. These results might suggest that the presence of additional nuclear spins or their configuration in relation to the radical pair either enhance or destroy the coherence.

As mentioned already in the introduction to this paper Mouritsen *et al.* have recently published a report gleaned from seven years of double-blind experiments that demonstrated the disorientation of migratory birds under the influence of anthropogenic electromagnetic radiation over a frequency range from 50 kHz to 5MHz [9]. This is important new evidence that radiation at frequencies deemed safe can functionally disrupt complex biological organisms. The use of quantum theory as applied to avian magnetoreception in the radical pair mechanism offers a possible model to explain this effect and, as such, has a valuable contribution to make.

References

- [1] Kirschvink J and Gould J 2000 *BioSystems*, **13** 181–201.
- [2] Mouritsen H and Ritz T 2005 *Current Opinion in Neurobiology*, **15** 406–414.
- [3] Schulten K, Swenberg CE, and Weller A 1978 *Z. Phys. Chem.*, **NF111** 1–5.
- [4] Wiltchko W *et al.* 2002 *Nature*, **419** 467–470.
- [5] Thalau P, Ritz T, Stapput K, Wiltchko R, Wiltchko W, 2004, *Naturwissenschaften* **92**, 86–90

- [6] Ritz T, Thalau P, Phillips JB, Wiltchko R and Wiltchko W, 2004, *Nature* **429**, 177–179
- [7] Ritz T, Wiltchko R, Hore PJ, Rodgers CT, Stapput K, Thalau P, Timmel CR and Wiltchko W 2009, *Biophysical Journal Volume* textbf96 3451–3457
- [8] Ritz T, Adem S and Schulten K 2000 *Biophys. J.*, **78** 2 707–718.
- [9] Mouritsen H *et al* 2014 *Nature*, **509** 353–356.
- [10] Gauger EM *et al* 2011 *Physical Review Letters*, **106** 040503.
- [11] Rodgers CT and Hore PJ 2009 *PNAS*, **106** 2 353–360.
- [12] Tiersch M and Briegel HJ 2012 *arXiv preprint:1204.4179v2* [*physics.bio-ph*].
- [13] Tarrant J 2011 NITheP Internship Project supervised by Goldstein K, Petruccione F and Sinayskiy I.
- [14] Solov'yov IA, Domratcheva T, Shahi ARM and Schulten K 2012 *J. Am. Chem. Soc.*, **134** 18046.
- [15] Sergi A, Sinayskiy I and Petruccione F 2009 *Physical Review A*, **80** 1 012108.
- [16] Breuer HP and Petruccione F 2002 *The Theory of Open Quantum Systems* Oxford University Press, Oxford.
- [17] Walters ZB 2013 *arXiv preprint:1208.2558v3* [*physics.bio-ph*].

Computer assisted ‘proof’ of the global existence of periodic orbits in the Rössler system

André E Botha¹ and Wynand Dednam^{1,2}

¹Department of Physics, University of South Africa, Science Campus, Private Bag X6, Florida Park 1710, South Africa

²Departamento de Física Aplicada Universidad de Alicante, San Vicente del Raspeig, E-03690 Alicante, Spain

E-mail: botha@unisa.ac.za

Abstract. The numerical optimized shooting method for finding periodic orbits in nonlinear dynamical systems was employed to determine the existence of periodic orbits in the well-known Rössler system. By optimizing the period T and the three system parameters, a , b and c , simultaneously, it was found that, for any initial condition $(x_0, y_0, z_0) \in \mathbb{R}^3$, there exists at least one set of optimized parameters corresponding to a periodic orbit passing through (x_0, y_0, z_0) . After a discussion of this result it was concluded that its analytical proof may present an interesting new mathematical challenge.

1. Introduction

The three-dimensional Rössler system (RS) was originally conceived as a simple prototype for studying chaos [1]. With only one nonlinear term it can be thought of as a simplification of the well-known Lorenz system [2] and as a minimal model for continuous-time chaos [3].

Over the past 38 years the RS has been studied extensively, both in its own right [4–15] and as a model, to illustrate various nonlinear phenomena and different types of chaos [16–19]. Surprisingly, despite the literally thousands of scientific articles that have been devoted to the RS, it still poses several open questions. For example, although the conjecture by Leonov [20], about the Lyapunov dimension of the RS (and two other types of Rössler systems), has recently been verified numerically [21], the exact Lyapunov dimension for the RS is not known analytically. Leonov [22] has recently also provided an analytical estimate of the attractor dimension in the RS, however, similar estimates for slightly more general Rössler systems remain a difficult open problem. There is also an ongoing discussion about whether or not the RS can indeed produce a strange attractor, in the same sense as that of the Lorenz system [17, 23].

More pertinent to the present study are the recent articles on the existence or non-existence and/or classification of the periodic orbits of the RS, in terms of its three control parameters a , b and c . Starkov and Starkov [24] have shown that, when the parameters satisfy the condition $c^2 - 4ab < 0$, no periodic solutions can exist. They have also proved that, when $ac < 0$ ($ac > 0$) and $ab > 0$, the periodic orbits exist entirely in the negative (positive) half space, i.e. $z < 0$ ($z > 0$).

There have been several attempts to demonstrate the existence of orbits more generally; such as those by Genesio and Ghilardi [25], in which the existence of the quasi-periodic orbits in

third-order systems like those of Rössler, Lorenz and Chua *et al.* [26] are proved. More recently, instead of the focus of such studies being on the unstable orbits, the experimentally observable chaotic and stable periodic orbits of simple chaotic flows, including the RS, have been classified very generally in terms of their control parameters [16]. In the parameter space, families of stable periodic solutions have been found to self-organize about the so-called isolated periodicity hubs of the RS. (See Ref. [16] and the references therein.)

Another active area of related research is on the integrability of the RS [10–12, 14, 27, 28]. In Ref. [12] it was shown that Darboux integrability of the RS for various parameter values leads to surfaces in state space containing periodic orbits. The bifurcations and routes to chaos in the RS have also been investigated [9, 29] and there is at least one exhaustive study on the complete parametric evolution of the system [6]. (Also see Ref. [5] and the references therein.)

The present work developed as a result of using the RS as a test case for our recently-developed numerical technique (called the optimized shooting method) for finding periodic solutions in nonlinear dynamical systems [30]. During the course of testing our method we discovered that it could optimize the system parameters to find at least one periodic orbit for *any* initial condition. At first this finding was very surprising to us, and we assumed that it was most likely due to a coding error. We therefore tested the codes more thoroughly and also experimented with a variety of different integration schemes and platforms, in order to verify the results independently. After confirming that our numerical results were indeed accurate, we turned to the literature to establish if a similar result had been reported elsewhere for the RS. However, in all of the related literature we found, there appeared to be no explicit proof (or mention) of the result we had obtained. Hence we consider it worthy of a separate brief report.

We are certainly not the first to use numerical computations to ‘prove’, i.e. motivate, the existence of certain analytical properties. Pilarczyk [18], for example, provides a computer assisted ‘proof’ of the existence of a periodic orbit in the RS. Wilczak and Zgliczyński [19] have also made use of a numerical method to ‘prove’ the existence of two period-doubling bifurcations in Rössler’s system, as well as the existence of a branch of period two points connecting them.

The remaining material in this article is organized as follows. In Sec. 2 we provide a brief description of the optimized shooting method. In Sec. 3 we state the main result in the form of a conjecture and describe the procedure that was followed in order to motivate it. Section 4 ends with a short discussion and conclusion.

2. Optimized shooting method

Today there are a variety of numerical methods available for finding the periodic orbits of nonlinear dynamical systems [31]. For the present purpose we make use of the optimized shooting method [30], which originally enabled us to discover the reported new property of the periodic orbits in the RS [1]. Briefly, this method makes use of Levenberg-Marquardt optimization to find the periodic orbits by minimizing a residual (or error function) [32, 33].

To apply the method to the RS, we re-write the system equations as

$$\dot{x} = T(-y - z), \quad \dot{y} = T(x + ay), \quad \dot{z} = T(b + z(x - c)), \quad (1)$$

where T is the (as yet) unknown period of the desired solution and the overdot indicates differentiation with respect to the dimensionless time $\tau = t/T$. Since one period corresponds to integration over τ from zero to one, we define the residual as

$$\mathbf{R} = (\mathbf{x}(1) - \mathbf{x}(0), \mathbf{x}(1 + \Delta\tau) - \mathbf{x}(\Delta\tau), \dots, \mathbf{x}(1 + p\Delta\tau) - \mathbf{x}(p\Delta\tau)), \quad (2)$$

where $\Delta\tau$ is a fixed integration step size, $p = 0, 1, 2, \dots$ and $\mathbf{x} \equiv (x, y, z)$. The residual is a function of the initial conditions, period, and three system parameters; since it depends on these through the solution to (1). Furthermore, since $\mathbf{x}(\tau) = \mathbf{x}(\tau + 1)$ for periodic solutions, it

can immediately be seen that such solutions correspond to a vanishing residual, i.e. the periodic solutions are found by optimizing \mathbf{R} to be as close as possible to zero.

3. Results

We state the main result of this work in the form of a conjecture.

Conjecture: For any initial condition $(x_0, y_0, z_0) \in \mathbb{R}^3$, there exists real parameters, a , b and c , for (1), such that its solution is periodic.

For initial conditions on plane $z = 0$, inspection of (1) shows that the conjecture is trivially satisfied: for $a = b = c = 0$ the solutions are circles, with periods equal to 2π . The analytical proof of the integrability of the system, for this case, is given in Refs. [15] and [11].

Off the $z = 0$ plane the validity of the conjecture is established numerically by performing optimization of the four parameters, T , a , b and c , for sets of randomly selected initial conditions. To facilitate the selection of these initial conditions it is convenient to rewrite (1) in terms of scaled (primed) coordinates, defined by $x = \alpha x'$, $y = \alpha y'$, and $z = \alpha z'$. Here $\alpha \geq 1$ is the scale factor. In terms of the primed coordinates (1) is given by

$$\dot{x}' = T(-y' - z'), \quad \dot{y}' = T(x' + ay'), \quad \dot{z}' = T(b/\alpha + \alpha z'(x' - c/\alpha)). \quad (3)$$

For a given α , the procedure for selecting the initial conditions then consisted of generating 100 distinct random points, within or on the unit sphere. For each initial condition the optimized shooting method was used to find the four parameters that produce a periodic orbit. Starting from $\alpha = 1$, the value of the scale factor was increased, in steps of 1, up to the maximum value of 100. This procedure produced a total of 100×100 random initial conditions. All 10^4 initial conditions were successfully optimized, i.e. in each case the magnitude of \mathbf{R} was successfully optimized to be below the set numerical tolerance ($|\mathbf{R}| < 10^{-12}$ in this calculation).

Table 1 lists the optimized parameters for 13 arbitrarily chosen periodic orbits. For ease

Table 1. Thirteen of the 10^4 periodic orbits found for the Rössler system via the optimized shooting method. The column headings are as follows: n is the orbit number (used to refer to the orbit in the main text), α is the scale parameter introduced in (3), x'_0 , y'_0 and z'_0 are the primed coordinates of the randomly selected initial conditions, T is the period, and the last three columns give the optimized system parameters.

n	α	x'_0	y'_0	z'_0	T	a	b	c
1	1	-0.322	0.283	-0.827	6.285	-0.01832876699	-37.861974659	45.448602039
2	1	0.083	0.498	0.756	6.282	0.01659280123	34.3119795334	45.496814976
3	1	0.271	-0.033	-0.492	6.284	-0.01074867143	-22.256338664	45.495734090
4	1	-0.843	-0.154	-0.338	6.283	-0.00756927443	-15.661740755	45.483293583
5	1	0.920	-0.101	-0.062	6.285	-0.00133598606	-2.7619826425	45.463987246
6	10	0.274	-0.843	-0.019	6.280	-0.00346652925	-9.6387456311	53.299901841
7	10	-0.431	0.429	-0.132	6.287	-0.02714795455	-75.519091787	52.954713982
8	10	-0.762	0.094	-0.023	6.283	-0.00504377136	-13.929997631	52.959008822
9	10	0.721	0.212	0.532	6.303	0.09058831072	242.356061832	52.926271854
10	10	0.212	-0.008	-0.946	6.412	-0.18855060455	-483.78516338	53.074125658
11	100	0.370	0.525	-0.489	9.221	-0.72151348243	-2913.5197570	96.641076651
12	100	0.465	-0.177	0.111	6.506	0.27456932042	1.25094211568	35.351207642
13	100	0.045	-0.881	0.206	1.670	0.28422944418	0.06378881294	25.148204118

of presentation, the randomly chosen initial conditions were first rounded off to three decimal places, before applying the optimized shooting method. Thus the initial conditions, as listed in columns 3-5 of table 1, are exact. On the other hand the periods listed in column 6 are given approximately to three decimal places. Since they are not required to reconstruct the orbits, their full 11 digit accuracy is not required here.

Two other features of table 1 are also worth mentioning. First, we notice that the signs of the parameters a and b are always the same as that of z'_0 . This is a property of all the found orbits and it confirms the analytical result proved by Starkov and Starkov [24]: when $ac < 0$ ($ac > 0$) and $ab > 0$, the periodic orbits exist entirely in the negative (positive) half space. Second, in Ref. [24] it was also shown that for $c^2 - 4ab < 0$, no periodic orbits can exist. In table 1, and indeed for all the found orbits, $c^2 - 4ab$ is positive. Thus our results are consistent with the known classifications for periodic orbits in the RS.

In total more than 100 spot checks were made on randomly selected orbits taken from the set of 10^4 . Such checks were made by plotting and visually inspecting the selected orbits. Figure 1 shows the plots for four orbits that are listed in table 1, i.e. orbits numbers 1, 9, 12 and 13.

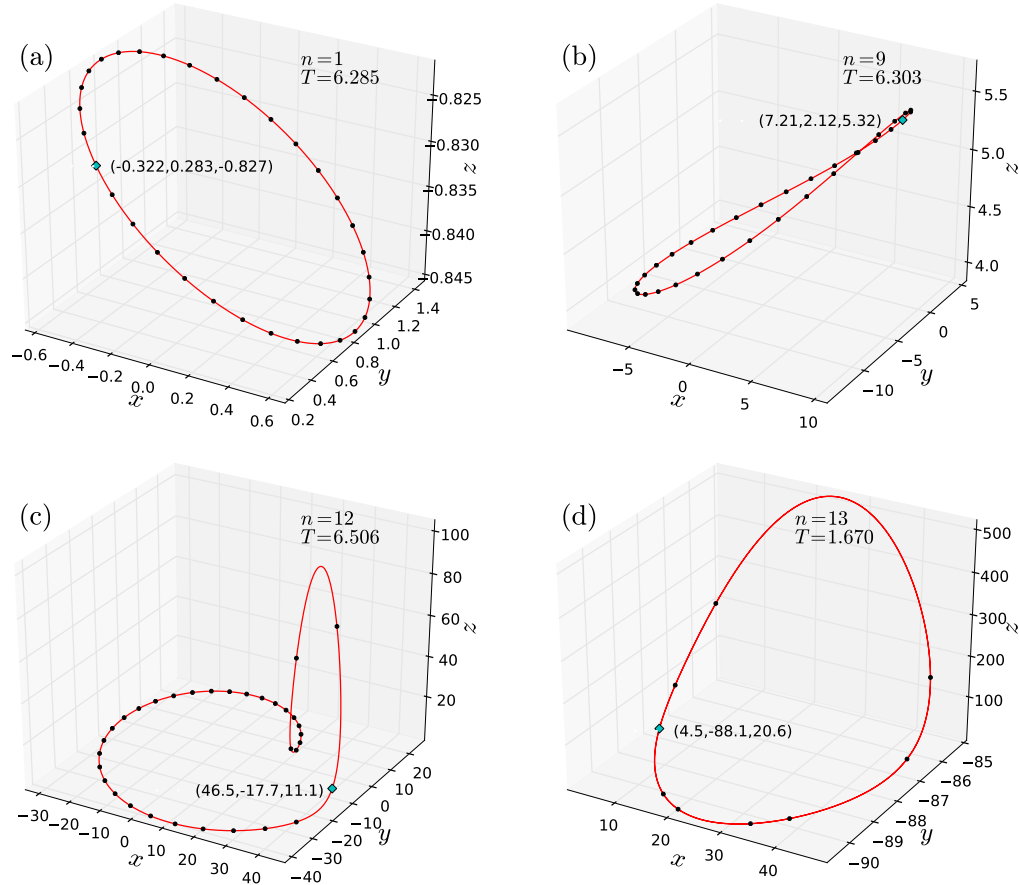


Figure 1. Phase portraits of four arbitrarily selected periodic orbits, chosen from table 1. The initial condition for each orbit is indicated by a blue square marker. Black filled circular markers indicate points along the orbit, every 200 integration time steps. Since a fixed time step of $\Delta\tau = 1/1024$ was used, the circular markers provide an indication of the velocity $(\dot{x}(\tau), \dot{y}(\tau), \dot{z}(\tau))$ along each orbit.

In figure 1(a) it can be seen that orbit one is confined to lie within the negative half space

$\mathbb{R}^3 - \{z \geq 0\}$, while orbits 9, 12 and 13, shown in figures (b)-(d), occur in the positive half space $\mathbb{R}^3 - \{z \leq 0\}$. This observation is consistent with the form of (1). If periodic orbits existed that crossed from one half space into the other, they would have to intersect the plane $z = 0$ more than once. Moreover, at $z = 0$ the sign of \dot{z} would have to be positive (negative) in going from the negative (positive) half space to the positive (negative) half space. However, the form of (1) clearly excludes this possibility, since for $z = 0$ the sign of \dot{z} is determined by the parameter b . Therefore, for a given set of parameters, the sign of \dot{z} at $z = 0$ cannot change. Hence, the non-trivial periodic orbits are confined to lie entirely within one half space or the other.

Finally, as an alternative way (i.e. without using the residual) of estimating the accuracy to which the optimized shooting method is capable of determining the periodic solutions, several initial conditions with $z = 0$ were also examined. In all such cases the magnitudes of the largest optimized system parameters were found to be less than 10^{-11} , with the expected period of 2π also given accurately to eleven decimal places.

4. Discussion and conclusion

We have established a computer assisted ‘proof’ of the global existence of periodic solutions in the Rössler system. By global we mean that, for any initial condition $(x_0, y_0, z_0) \in \mathbb{R}^3$, our numerical method was able to optimize simultaneously the period T and system parameters, a , b and c , in order to find at least one periodic solution passing through the initial condition. We have stated this result as a conjecture, as opposed to a proposition, because the latter term is generally reserved for mathematical results that can be proved rigorously. The term ‘proof’, which we have written consistently in single quotes, should thus be understood to mean a numerical justification/motivation, rather than a mathematical proof. This is also the sense in which the same term is used elsewhere in the literature on computer assisted proofs. (See, for example, Refs. [18] and [19].)

It is interesting to note that, in some cases, more than one solution passing through a given initial condition could be found, i.e. two or more distinct sets of the optimized parameters could be found for a given initial condition. For example, in the case of orbit number 12 (see figure 1(c) and table 1), there also exists a different period orbit passing through the same initial point, with precisely twice the period, i.e. $T = 13.0116560813$, at slightly different system parameters. This multiplicity, of period doubled orbits all passing through the same point, offers an interesting new possibility. Whereas conventionally the universal period doubling route to chaos (the so-called Feigenbaum scenario [34, 35]) is usually achieved by varying one system parameter at a time, we see here that it may also be possible (by following a very specific (three dimensional) path in the parameter space) to obtain a special sequence of period doubling orbits *which all pass through the same point in the phase space*. It remains to be seen whether such a special sequence would also lead to chaos and follow Feigenbaum’s universal scaling laws.

In conclusion, we have formulated a conjecture about the global existence of periodic solutions in Rössler’s system. Our conjecture is numerically supported (i.e. ‘proved’) and it postulates the existence of periodic orbits passing through *any* point in the phase space, for a suitable choice of the system parameters. While it is in principle possible to establish this result analytically, to date there does not appear to be any mention in the literature of either the result itself or its proof. Unfortunately we were not able to construct such an analytical proof ourselves. However, in view of the compelling computational evidence which we have provided, we hope that some theoretically inclined readers may find it interesting to devise a rigorous mathematical proof of our conjecture.

Acknowledgments

The financial assistance of the National Research Foundation of South Africa towards this research is hereby acknowledged.

References

- [1] Rössler O E 1976 An equation for continuous chaos *Phys. Lett. A* **57** 397–8
- [2] Lorenz E N 1963 Deterministic nonperiodic flow *J. Atmos. Sci.* **20** 130–41
- [3] Gaspard P 2005 Rössler systems *Encyc. Nonlin. Sci.* ed Scott A (New York: Routledge) pp 808–11
- [4] Algaba A *et al.* 2007 Resonances of periodic orbits in Rössler system in presence of a triple-zero bifurcation *Int. J. Bifurcation Chaos* **17** 1997–2008
- [5] Barrio R, Blesa F, Dena A and Serrano S 2011 Qualitative and numerical analysis of the Rössler model: bifurcations of equilibria *Computers and Mathematics with Applications* **62** 4140–50
- [6] Barrio R, Blesa F and Serrano S 2009 Qualitative analysis of the Rössler equations: bifurcations of limit cycles and chaotic attractors *Physica D* **238** 1087–1100
- [7] Castro V *et al.* 2007 Characterization of the Rössler system in parameter space *Int. J. Bifurcation Chaos* **17** 965–73
- [8] Galias Z 2006 Counting low-period cycles for flows *Int. J. Bifurcation Chaos* **16** 2873–86
- [9] Gardini L 1985 Hopf bifurcations and period-doubling transitions in Rössler model *Il Nuovo Cimento B* **89** 139–60
- [10] Lima M F S and Llibre J 2011 Global dynamics of the Rössler system with conserved quantities *J. Phys. A: Math. Theor.* **44** 365201 (15pp)
- [11] Lliubre J and Valls C 2007 Formal and analytic integrability of the Rössler system *Int. J. Bifurcation Chaos* **17** 3289–93
- [12] Lliubre J and Zhang X 2002 Darboux integrability for the Rössler system *Int. J. Bifurcation Chaos* **12** 421–28
- [13] Teryokhin M T and Panfilova T L 1999 Periodic solutions of the Rössler system *Russian Mathematics* **43** 66–9
- [14] Tudoran R M and Girban A 2012 On the completely integrable case of the Rössler system *J. Math. Phys.* **53** 052701 (10pp)
- [15] Zhang X 2004 Exponential factors and Darboux integrability for the Rössler system *Int. J. Bifurcation Chaos* **14** 4275–83
- [16] Gallas J A C 2010 The structure of infinite periodic and chaotic hub cascades in phase diagrams of simple autonomous flows *Int. J. Bifurcation Chaos* **20** 197–211
- [17] Lozi R 2013 *Topology and Dynamics of Chaos* ed Letellier C and Gilmore R (Singapore: World Scientific)
- [18] Pilarczyk P 1999 Computer assisted method for proving existence of periodic orbits *Topological Methods in Nonlinear Analysis* **13** 365–77
- [19] Wilczak D and Zgliczyński P 2009 Period doubling in the Rössler system – a computer assisted proof *Found. Comput. Math.* **9** 611–49
- [20] Leonov G A 2012 Lyapunov functions in the attractors dimension theory *J. Appl. Math. Mech.* **76** 129–41
- [21] Kuznetsov N V, Mokaev T N and Vasilyev P A 2014 Numerical justification of Leonov conjecture on Lyapunov dimension of Rössler attractor *Commun. Nonlinear Sci. Numer. Simulat.* **19** 1027–34
- [22] Leonov G A 2014 Rössler systems: estimates for the dimension of attractors and homoclinic orbits *Doklady Mathematics* **89** 1–3
- [23] Zgliczyński P 1997 Computer assisted proof of chaos in the Rössler equations and in the Hénon map *Nonlinearity* **10** 243–52

- [24] Starkov K E and Starkov K K 2007 Localization of periodic orbits of the Rössler system under variation of its parameters *Chaos, Solitons and Fractals* **33** 1445–49
- [25] Genesio R and Ghilardi C 2005 On the onset of quasi-periodic solutions in third-order nonlinear dynamical systems *Int. J. Bifurcation Chaos* **15** 3165–80
- [26] Chua L, Komuro M and Matsumoto T 1986 The double scroll family *IEEE Trans. Circuits Syst.* **33** 1073–118
- [27] Chandrasekar V, Senthilvelan M and Lakshmanan M 2009 On the complete integrability and linearization of nonlinear ordinary differential equations. III. Coupled first-order equations *Proc. R. Soc. A* **465** 585–608
- [28] Hitchin N 1997 *Integrable systems: An Introduction* (Oxford: Notes from the Mathematical Institute)
- [29] Nikolov S and Petrov V 2004 New results about route to chaos in Rössler system *Int. J. Bifurcation Chaos* **14** 293–308
- [30] Dednam W and Botha A E 2014 Optimized shooting method for finding periodic orbits of nonlinear dynamical systems *Engineering with Computers* DOI: 10.1007/s00366-014-0386-6
- [31] Abad A, Barrio R and Dena A 2011 Computing periodic orbits with arbitrary precision *Phys. Rev. E* **84** 016701 (6pp)
- [32] Levenberg K 1944 A method for the solution of certain problems in least squares *Quart. Appl. Math.* **2** 164–68
- [33] Marquardt D W 1963 An algorithm for least-squares estimation of nonlinear parameters *SIAM J. Appl. Math.* **11** 431–41
- [34] Feigenbaum M J 1978 Quantitative universality for a class of nonlinear transformations *Journal of Statistical Physics* **19** 25–52
- [35] Feigenbaum M J 1979 The universal metric properties of nonlinear transformations *Journal of Statistical Physics* **21** 669–706

Homogeneous Open Quantum Walks on a Line

Roland Cristopher Caballar^{1,2}, Ilya Sinayskiy¹ and Francesco Petruccione¹

¹National Institute for Theoretical Physics (NITheP-KZN) and School of Chemistry and Physics, University of KwaZulu-Natal, Durban, KwaZulu-Natal, 4001, South Africa

²National Institute of Physics, College of Science, University of the Philippines, Diliman, 1101 Quezon City

E-mail: Caballar@ukzn.ac.za, sinayskiy@ukzn.ac.za, Petruccione@ukzn.ac.za

Abstract. We analyze the dynamics of homogeneous open quantum walks on a line for a system with two internal degrees of freedom by analytically computing the probability distribution of the system. These distributions are plotted numerically, allowing for the interpretation of the dynamics of the system at any timestep. We also compute for the system's steady point at each point of the line on which the open quantum walk takes place.

1. Introduction

Recently, the open quantum walk was introduced in [1, 2], wherein the system undergoes a random walk, and its internal state changes with each step. The dynamics is driven by the coupling between the system and its associated environment. This open quantum walk may occur over discrete or continuous time steps in a finite or infinite graph, with the continuous-time limit investigated in [3, 4]. Previous work on open quantum walks resulted in the derivation of a central limit theorem for its asymptotic probability distribution, which is shown to be a normal distribution [5, 6]. Aside from showing rich dynamical behavior, possible applications for open quantum walks have been discussed in [1]. Those possible applications are dissipative quantum state preparation of single-and multiple-qubit gates, implementation of quantum logic gates for single and multiple qubits, and efficient quantum transport of excitations.

A particular type of open quantum walk was studied in [7]. The open quantum walk under consideration was a homogeneous open quantum walk on a line, with one jump operator corresponding to one direction of motion for the system. The jump operators were assumed to be simultaneously diagonalizable. The resulting distribution was shown to be, for intermediate timesteps, a binomial distribution, which converges for large timesteps to a Gaussian, distribution. The approach then allows us to determine the dynamics of the system undergoing the open quantum walk at any instant of time.

In this paper, we extend the work done in [7] to consider systems with 2 internal degrees of freedom undergoing homogeneous open quantum walks on the line with at least one of the jump operators diagonalizable. We analyze the dynamics of the system undergoing the open quantum walk, and we also derive the steady state of the system at each point of the line on which this open quantum walk takes place.

2. The Homogeneous Open Quantum Walk

A schematic diagram of the homogeneous open quantum walk is shown in figure (1). We designate the Hilbert space corresponding to the position of the system undergoing the open quantum walk as $\mathcal{H}_s = \mathbb{Z}$, which we designate as position space, and the Hilbert space corresponding to the internal degrees of freedom of the system at each node as \mathcal{H}_c , which we designate as node space. The total Hilbert space corresponding to the open quantum walk is then $\mathcal{H} = \mathcal{H}_c \otimes \mathcal{H}_s$. If the open quantum walk occurs over discrete timesteps, then at timestep n , the density matrix describing the system undergoing this open quantum walk is given as

$$\rho^{(n)} = \sum_{x=-M}^M \rho_x \otimes |x\rangle \langle x|, \quad (1)$$

where ρ_x is the density matrix that describes the system's internal degrees of freedom at node x , $|x\rangle$ is the ket in position space \mathcal{H}_s , and $2M + 1$ is the total number of nodes on which the system can move at timestep n . We note that the density matrices ρ_x satisfy the condition $\sum_{x=-M}^M \text{Tr}(\rho_x) = 1$.

In \mathcal{H}_c , we define two bounded operators B and C that satisfy the condition

$$B^\dagger B + C^\dagger C = I, \quad (2)$$

with this condition ensuring that probability is conserved at all timesteps n . These operators B and C correspond to the change in the system's internal degrees of freedom as it makes a transition from node x to a neighboring node $x \pm 1$. For these jump operators, we define a linear mapping \mathcal{L} on \mathcal{H}_c as

$$\mathcal{L}(\rho) = B\rho B^\dagger + C\rho C^\dagger \quad (3)$$

Lifting this mapping from \mathcal{H}_c to $\mathcal{H} = \mathcal{H}_s \otimes \mathcal{H}_c$, and iteratively applying the resulting map, we obtain an equation of motion for the system undergoing the open quantum walk at timestep n and node x . In particular, the jump operators in the system's Hilbert space \mathcal{H} will now have the form $B \otimes |x+1\rangle \langle x|$ and $C \otimes |x-1\rangle \langle x|$, while the mapping now has the form

$$\mathcal{M}(\rho^{(n-1)}) = \sum_{x=-n}^n \rho_x^{(n)} \otimes |x\rangle \langle x|, \quad (4)$$

where

$$\rho_x^{(n)} = B\rho_{x-1}^{(n-1)}B^\dagger + C\rho_{x+1}^{(n-1)}C^\dagger, \quad (5)$$

which is the time evolution equation for the density matrix at node x at timestep n . Also,

$$P_x^{[n]} = \text{Tr}(\rho_x^{[n]}) \quad (6)$$

is the probability that node x is occupied at timestep n .

If B is diagonalizable via a unitary transformation U , (2) can be transformed into the following form:

$$\tilde{B}^\dagger \tilde{B} + \tilde{C}^\dagger \tilde{C} = I, \quad (7)$$

where $\tilde{B} = U^\dagger B U$ and $\tilde{C} = U^\dagger C U$. In particular, \tilde{B} has the form

$$\tilde{B} = b_1 |1\rangle \langle 1| + b_2 |2\rangle \langle 2|, \quad (8)$$

where b_j are the eigenvalues of B . We note that the states $|1\rangle$ and $|2\rangle$ are internal states of the system, which are 2×1 unit vectors in the node Hilbert space \mathcal{H}_c . On the other hand, \tilde{C} will have the following form:

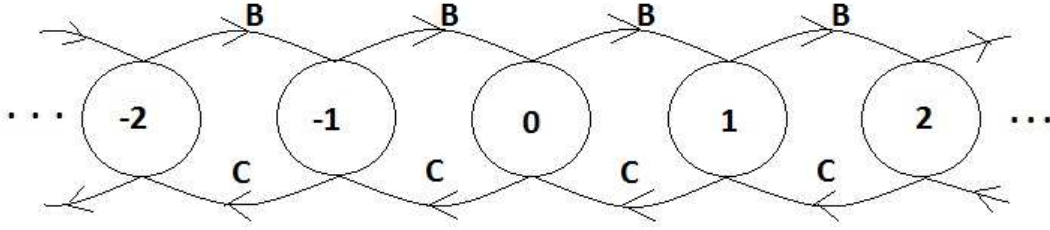


Figure 1. A homogeneous open quantum walk on a line.

$$\tilde{C} = \sqrt{1 - |b_1|^2}(\alpha |1\rangle \langle 1| - \beta^* |2\rangle \langle 1|) + \sqrt{1 - |b_2|^2}(\beta |1\rangle \langle 2| + \alpha^* |2\rangle \langle 2|), \quad (9)$$

where α, β are complex numbers that obey the condition $|\alpha|^2 + |\beta|^2 = 1$. In general, the jump operators for a homogeneous open quantum walk do not commute, except for the case where both B and C are simultaneously diagonalizable, which has been covered in Ref. [7]. It is still possible, as we demonstrate for a particular case below, the probability distributions can be computed exactly for arbitrary timesteps n .

3. Computation of probability distributions

In this section we present two methods for computing the probability distributions for systems undergoing open quantum walks. The first method, which uses the central limit theorem, allows us to determine the asymptotic probability distribution over time for the system. The second method, which makes use of Fourier transforms, gives us a brute force method for computing the system's probability distribution.

3.1. Central Limit Theorem for Jump Operators

Ref. [5] gives us a central limit theorem that tells us the asymptotic behavior of the probability distribution for a system undergoing an open quantum walk. In particular, for a homogeneous open quantum walk, the theorem can be stated as follows:

Theorem (Attal et al, Ref. [5], Theorem 5.2). *Consider the stationary open quantum random walk on Z associated to the jump operators B and C . We assume that the completely positive map $\mathcal{L}(\rho) = B\rho B^\dagger + C\rho C^\dagger$ admits a unique invariant state ρ_∞ . Let (ρ_n, X_n) be the quantum trajectory process to this open quantum walk. Then $\frac{X_n - nm}{\sqrt{n}}$ converges in law to the Gaussian distribution $\mathcal{N}(0, C)$ in \mathbb{R} , with mean $m = \text{Tr}(B\rho_\infty B^\dagger) - \text{Tr}(C\rho_\infty C^\dagger)$ and covariance $\sigma^2 = \text{Tr}(B\rho_\infty B^\dagger + C\rho_\infty C^\dagger) - m^2 + 2\text{Tr}(B\rho_\infty B^\dagger L - C\rho_\infty C^\dagger L) - 2m\text{Tr}(\rho_\infty L)$ where L is the solution to the equation $L - \mathcal{L}^\dagger(L) = B^\dagger B - C^\dagger C - I$.*

In the theorem, (ρ_n, X_n) is the Markov chain with values on $\mathcal{E}(\mathcal{H}_c) \times \mathbb{Z}$, where $\mathcal{E}(\mathcal{H}_c)$ is the space of all density matrices on \mathcal{H}_c , associated with the quantum trajectories of the mapping \mathcal{M} .

For intermediate timesteps, on the other hand, we can determine the dynamical behavior of the system undergoing the open quantum walk by analytically computing for it. Also, using the mapping given by (3), we can compute for the form of the steady state for the open quantum walk at each node of the line by solving the system of equations specified by the following equation:

$$\rho_\infty = \mathcal{L}(\rho_\infty) = B\rho_\infty B^\dagger + C\rho_\infty C^\dagger. \quad (10)$$

3.2. Fourier Transform Method for Computing Probability Distributions

To compute the probability distributions for the open quantum walk generated by the jump operators B and C , we use the method of Fourier transforms and dual processes first described in Ref. [6]. First, we compute for the following dual process:

$$Y_n(k) = (e^{ik}L_{B^\dagger}R_B + e^{-ik}L_{C^\dagger}R_C)^n(I), \quad (11)$$

where for a given operator A acting on U , $L_A U := AU$ and $R_A U := UA$. In (11), I is the $N \times N$ identity matrix. This dual is defined in momentum space K , where $k = \{-\pi, \pi\}$. Once we have determined the dual $Y_n(k)$, we then compute for the trace of the product $\rho_0 Y_n(k)$, where ρ_0 is the density matrix for the initial state of the system, with explicit form

$$\rho_0 = P_1|1\rangle\langle 1| + P_2|2\rangle\langle 2| + q_{12}|1\rangle\langle 2| + q_{12}^*|2\rangle\langle 1|. \quad (12)$$

The trace of ρ_0 is then $P_1 + P_2 = 1$. Finally, we compute for the probability distribution of the system in position space at point x and timestep n by taking the Fourier transform of $\text{Tr}(\rho_0 Y_n(k))$ as follows:

$$P_x^{(n)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} dk e^{-ikx} \text{Tr}(\rho_0 Y_n(k)). \quad (13)$$

3.3. Computation of the probability distribution for a particular form of jump operators

To demonstrate how we can compute for the probability distribution of a system undergoing an open quantum walk analytically, let us consider transformed jump operators of the form

$$\tilde{B} = b_1(|1\rangle\langle 1| - |2\rangle\langle 2|), \quad \tilde{C} = \sqrt{1 - |b_1|^2}(|1\rangle\langle 2| + |2\rangle\langle 1|). \quad (14)$$

With this form of B and C , it can easily be seen that B and C both satisfy (2), and B and C will not commute. Let us now compute for the probability distribution $P_x^{(n)}$. Evaluating (11), then substituting the resulting dual form and the initial state given by (12) in (13), we obtain the following probability densities at timestep n and position x :

$$P_x^{(n)} = \binom{n}{(n-x)/2} (|b_1|^2)^{(n-x)/2} (1 - |b_1|^2)^{(n+x)/2}. \quad (15)$$

This is a binomial distribution, which converges to a normal distribution as $n \rightarrow \infty$. Hence, for this very special case, we have demonstrated that it is possible to compute for the analytic form of the probability distribution for a system undergoing a homogeneous open quantum walk at any instant of time.

4. Numerical Analysis of the Probability Distributions of a Homogeneous Open Quantum Walk

In general, even if B and C have been transformed into the forms given by Eqs. (8) and (9), the resulting analytic form of the probability distribution is not as simple as that given in (15). In fact, the resulting expression only reduces to well known simple probability distributions at any instant of time n only for certain special cases. Thus, to properly analyze the probability distribution, we must plot it at a given instant of time n and see how the distribution evolves numerically, in order to obtain a physical interpretation for the system's dynamical behavior.

The first important feature to note is that if neither b_1 nor b_2 are equal to either zero or 1, the distribution converges numerically to a normal distribution with a well-defined peak for large timesteps. This is not surprising in light of Attal et al's central limit theorem for open quantum

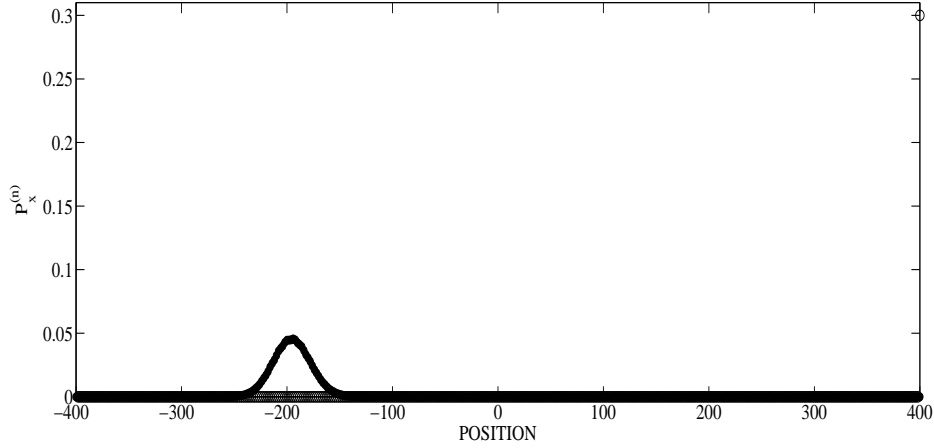


Figure 2. Probability distributions for a homogeneous open quantum walk at timestep $n = 400$, with B and C given by Eqs. (8) and (9), respectively. The solid circles denote a distribution corresponding to \tilde{B} with eigenvalues $b_1 = 0.45$ and $b_2 = 0.55$, while the unshaded circles denote a solitonic distribution corresponding to \tilde{B} with eigenvalues $b_1 = 1$ and $b_2 = 0$. For both distributions, the initial state $\rho_0^{(0)}$'s matrix elements are $\rho_{11} = 0.3$, $\rho_{22} = 0.7$, $\rho_{12} = 0.33 + 0.11i$ and $\rho_{21} = \rho_{12}^*$.

walks. However, by plotting the analytic form of the distribution for any given timestep, we can see how the distribution will evolve towards a normal distribution.

Second, we note that if $b_1 = 1$ and $b_2 = 0$, or vice versa, in (8), then $\alpha = 1$ and $\beta = 0$, or vice versa, in (9). We then obtain a solitonic distribution, which is an infinitely narrow peaked distribution with constant height moving with constant speed to the left or to the right. We illustrate this in Fig. (2), where we plot a normal distribution and a solitonic distribution, the latter represented by the single point on the right hand side.

5. Steady States for a Homogeneous Open Quantum Walk in \mathcal{H}_c space

We now turn to the question of determining the steady states ρ_∞ for the linear mapping \mathcal{L} in \mathcal{H}_c space of this homogeneous open quantum walk, and in so doing determine whether the central limit theorem for open quantum walks holds for this case. To do so, we must solve (10), and to simplify our task, we transform (10) by diagonalizing B , such that B and C will be transformed into the forms \tilde{B} and \tilde{C} given by Eqs. (8) and (9) for the diagonal blocks of \tilde{C} , respectively. Then (10) will have the form

$$\tilde{\rho}_\infty = \tilde{B}\tilde{\rho}_\infty\tilde{B}^\dagger + \tilde{C}\tilde{\rho}_\infty\tilde{C}^\dagger, \quad \tilde{\rho}_\infty = U\rho_\infty U^\dagger. \quad (16)$$

We assume that $\tilde{\rho}_\infty$ has the general form

$$\tilde{\rho}_\infty = \rho_{11} |1\rangle \langle 1| + \rho_{12} |1\rangle \langle 2| + \rho_{21} |2\rangle \langle 1| + \rho_{22} |2\rangle \langle 2|. \quad (17)$$

Substituting Eqs. (8), (9) and (17) into (16) and simplifying, we obtain a system of 4 equations in 4 unknowns for the matrix elements ρ_{jk} of $\tilde{\rho}_\infty$:

$$\begin{aligned}
 0 &= (1 - |b_1|^2)(|\alpha|^2 - 1)\tilde{\rho}_{11} + (1 - |b_2|^2)|\beta|^2\tilde{\rho}_{22} + \alpha\beta^*\sqrt{(1 - |b_1|^2)(1 - |b_2|^2)}\tilde{\rho}_{12} \\
 &\quad + \alpha^*\beta\sqrt{(1 - |b_1|^2)(1 - |b_2|^2)}\tilde{\rho}_{21} \\
 0 &= |\beta|^2(1 - |b_1|^2)\tilde{\rho}_{11} + (1 - |b_2|^2)(|\alpha|^2 - 1)\tilde{\rho}_{22} - \alpha\beta^*\sqrt{(1 - |b_1|^2)(1 - |b_2|^2)}\tilde{\rho}_{12} \\
 &\quad - \alpha^*\beta\sqrt{(1 - |b_1|^2)(1 - |b_2|^2)}\tilde{\rho}_{21} \\
 0 &= -\alpha\beta(1 - |b_1|^2)\tilde{\rho}_{11} + \alpha\beta(1 - |b_2|^2)\tilde{\rho}_{22} + \left[b_1b_2^* - 1 + \alpha^2\sqrt{(1 - |b_1|^2)(1 - |b_2|^2)}\right]\tilde{\rho}_{12} \\
 &\quad - \beta^2\sqrt{(1 - |b_1|^2)(1 - |b_2|^2)}\tilde{\rho}_{21} \\
 0 &= -\alpha^*\beta^*(1 - |b_1|^2)\tilde{\rho}_{11} + \alpha^*\beta^*(1 - |b_2|^2)\tilde{\rho}_{22} - (\beta^*)^2\sqrt{(1 - |b_1|^2)(1 - |b_2|^2)}\tilde{\rho}_{12} \\
 &\quad + \left[b_1^*b_2 - 1 + (\alpha^*)^2\sqrt{(1 - |b_1|^2)(1 - |b_2|^2)}\right]\tilde{\rho}_{21}
 \end{aligned} \tag{18}$$

However, the second of these four equations is actually linearly dependent on the first, so that at this point, we have 3 equations in 4 unknowns. In order to return the system to 4 equations in 4 unknowns, we impose the following additional constraint, following Ref. ([1]):

$$\text{Tr}(\tilde{\rho}_\infty) = \rho_{11} + \rho_{22} = 1 \tag{19}$$

Thus, solving the resulting system of linear equations gives us a unique solution for $\tilde{\rho}_\infty$:

$$\tilde{\rho}_\infty = \frac{(1 - |b_1|^2)(1 - |b_2|^2)}{2 - |b_1|^2 - |b_2|^2} \left(\frac{1}{1 - |b_1|^2} |1\rangle \langle 1| + \frac{1}{1 - |b_2|^2} |2\rangle \langle 2| \right). \tag{20}$$

6. Conclusion

We have shown in this work that, for a particular homogeneous open quantum walk for a system with 2 internal degrees of freedom, with one of the jump operators diagonalizable, it is possible to compute analytically for the probability distribution of the open quantum walk for a given timestep n . At the same time, we have also shown numerically that such a probability distribution will converge for large timesteps to a normal distribution with 1 peak or a solitonic distribution if the non-diagonalizable jump operator for this open quantum walk is transformable to a particular form given by (9). We also determined a unique steady state of the mapping \mathcal{L} at each node of the line for this homogeneous open quantum walk.

Acknowledgements

This work is based upon research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation.

References

- [1] S Attal, I Sinayskiy and F Petruccione (2012), *Phys. Lett. A* **376**, 1545
- [2] S Attal, F Petruccione, C Sabot and I Sinayskiy (2012), *J. Stat. Phys.* **47**, 832
- [3] I Sinayskiy and F Petruccione (2012), *Quant. Inf. Proc.* **11**, 1301
- [4] S Attal, J Deschamps and C Pellegrini (2013), arXiv:1312.3731
- [5] S Attal, N Guillotin-Plantard and C Sabot (2012), arXiv:1206.1472
- [6] N Konno and H-J Yoo (2013), *J. Stat. Phys.* **150**, 299
- [7] I Sinayskiy and F Petruccione (2012), *Phys. Scr.* **T151**, 014077

Properties of the quark-gluon plasma observed at RHIC and LHC

W. A. Horowitz

Department of Physics, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa

E-mail: wa.horowitz@uct.ac.za

Abstract. Puzzles and discoveries abound in the results from the Relativistic Heavy Ion Collider (RHIC) and from the relativistic heavy ion collisions at the Large Hadron Collider (LHC) including what seems to be the creation of the world's most perfect fluid and the stunning disappearance of large momentum particles into a dense, opaque quark-gluon plasma (QGP). Surprisingly, the methods of string theory appear to provide a better description of the QGP for observables associated with lower momentum particles while the completely opposite approach with an application of perturbative quantum chromodynamics (pQCD) works best for particles at the highest momenta. We discuss work attempting to bridge the divide between these two opposing descriptions of the properties of the QGP, the state of the universe a microsecond after the Big Bang.

1. Introduction

The great difficulty currently facing heavy ion physics is the apparent contradiction between the interpretation of low- p_T and high- p_T observables. A consensus has formed in which the distribution of low momentum particles is the result of rapid thermalization followed by nearly ideal hydrodynamic evolution. The best explanation of the early onset of thermalization [1] and nearly ideal fluid flow [2] is the existence of a strongly-coupled fluid best described by the methods of the AdS/CFT correspondence. On the other hand, naive application of the AdS/CFT correspondence to high- p_T probes yield results in contradiction with data. At the same time leading order pQCD predictions predicated on a weakly-coupled plasma weakly coupled to a high momentum probe [3] systematically describe the high- p_T data within a factor of 2 [4]; higher order correction seem likely to lead to an even better description of data [5]. Leading order [6] and sophisticated next-to-leading order [7] calculations based on the same weak-coupling perturbative picture of the plasma, though, yield a thermalization time and a viscosity to entropy ratio an order of magnitude larger than suggested by data. A hybrid strong-weak approach might reconcile these two pictures.

2. Energy Loss in AdS/CFT

2.1. Heavy Quarks

The now well-known leading order analytic energy loss formula for heavy quarks strongly-coupled to a strongly-coupled $\mathcal{N} = 4$ SYM plasma is

$$\frac{dp}{dt} = -\mu p, \text{ where } \mu = \frac{\pi\sqrt{\lambda}T^2}{2M_q}. \quad (1)$$

$\lambda = g^2 N_c$ is the 't Hooft coupling for the theory, T is the temperature of the plasma, and M_q is the mass of the heavy quark [8, 9]. The form of this energy loss is very different from that found assuming a probe weakly coupled to a weakly-coupled plasma: incoherent Bethe-Heitler bremsstrahlung energy loss [10] goes as

$$\left. \frac{dp}{dt} \right|_{BH} \sim -\frac{T^3}{M_q^2} p, \quad (2)$$

but the radiative energy loss in the deep-LPM regime is

$$\left. \frac{dp}{dt} \right|_{LPM} \sim -L T^3 \log(p/M_q), \quad (3)$$

where L is the length of the medium through which the heavy quark has passed [11].

Now $\mathcal{N} = 4$ SYM in the $N_c \rightarrow \infty$ and λ large and fixed limit is not QCD. However one hopes that the results from AdS/CFT can provide some useful insight into QCD processes. One of the complications of the dissimilarity of the two theories is that there is not a unique, reasonable mapping of the parameters in QCD to those in AdS/CFT. Using a set of these reasonable mappings, one finds a good description of the suppression of heavy quark decay fragments seen by RHIC experiments; see figure 1 (a). (A more detailed description of the model and parameters used to compute the figure can be found in [12].)

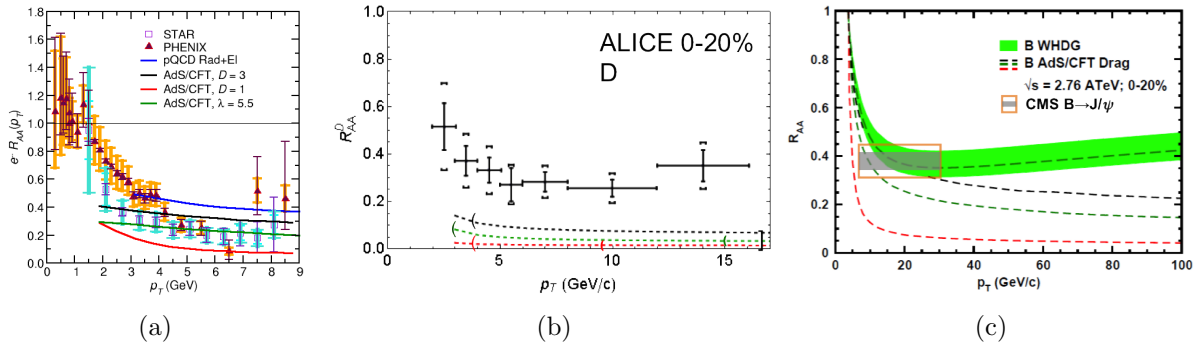


Figure 1: (a) AdS/CFT (and pQCD) predictions [12] for non-photonic electron decay products of heavy c and b quarks at RHIC [13, 14] and (b) D meson and (c) B meson suppression predictions from AdS/CFT [15] at LHC [16, 17]

Using this exact same set of mappings one may make predictions for D and B meson suppression at LHC, shown in figure 1 (b) and (c). As seen in (c), given the current uncertainties in the theoretical predictions and experimental measurements, the B meson predictions are consistent with data. The D mesons, however, as shown in (b), are falsifiably oversuppressed compared to ALICE data.

Previous calculations [12] included estimates of a “speed limit” for the applicability of the heavy quark drag calculations. Several independent lines of reasoning [18, 19] imply that the formalism does not apply to heavy quarks propagating faster than

$$\gamma \lesssim \gamma_{critical} = \left(1 + \frac{2M_q}{\sqrt{\lambda}T}\right)^2 \sim \frac{4M_q^2}{\lambda T^2}. \quad (4)$$

This speed limit is parametrically smaller than an estimate for the momenta at which the fluctuations in momentum loss become important [18],

$$\gamma_{fluc} \sim \frac{M_q^2}{4T^2}; \quad (5)$$

however, it turns out that numerically this latter speed limit is reached first. Efforts are underway to quantify the importance of momentum fluctuations in the suppression of high momentum heavy quarks.

2.2. Light Flavors

A critical test of any energy loss formalism is a simultaneous description of both heavy and light flavor suppression. The original calculations of light flavor energy loss require difficult numerics as the endpoints of the string are allowed to dynamically fall in the 5th dimension [20]. More recent work using an alternative setup yields a simple analytic solution [21]. However, it is not yet entirely clear what is the most appropriate setup in the AdS space to model light flavor energy loss. In the following, we will attempt to infer the physical consequences of the original light flavor energy loss calculation.

One of the first observations of the original setup was a generic Bragg peak in the energy loss such that the maximum stopping distance scaled as

$$\Delta x_{max} \sim \left(\frac{E}{\sqrt{\lambda}T}\right)^{1/3} \frac{1}{T}. \quad (6)$$

One can create a very naive energy loss model based on this maximum stopping distance [22]: assume that any light flavor created with $L < x_{max}$ gets out of the plasma unaltered while flavors created with $L > x_{max}$ are completely absorbed. There are large uncertainties in this model; in addition to the usual unknown mapping from QCD to AdS/CFT, one also does not know which single value of T to plug into Eq. (6). A maximal uncertainty band can be created by taking two extreme values for T : 1) the temperature at the point of creation at the moment of thermalization or 2) the transition temperature between the deconfined and confined phases of QCD matter. The predictions resulting from these two extremes are shown as a band in figure 2 () and (); the data at RHIC and LHC fall within the very large theoretical uncertainties.

Since the most naive calculation is not obviously falsified by the data, it is worth pursuing a more precise theoretical model. Preliminary investigations [25] show that results depend very sensitively on the “jet” prescription chosen, to the extent that one can even make the Bragg peak in the energy loss appear and disappear. Additional, large sensitivity comes from the precise initial string profile propagated from early times in the collision; see figure 2 (). It turns out that very little of the $2 \times \infty$ dimensions of the space of initial conditions has been explored. We will return to these issues later in this proceedings.

3. pQCD Energy Loss

One may also choose to determine the consequences of alternative picture of a weakly-coupled plasma weakly coupled to a high momentum probe. Jacksonian intuition suggests that at the

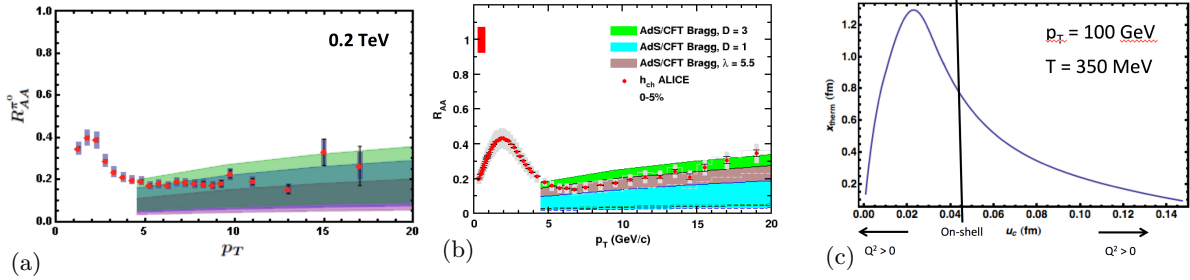


Figure 2: Predictions from a simple Bragg peak model for AdS/CFT light flavor energy loss [22] compared () to PHENIX data at RHIC [23] and () ALICE data at LHC [24]. () The stopping distance in AdS/CFT for light flavors depends strongly on the initial conditions.

GeV scale radiative energy loss dominates over collisional. However detailed calculations [26] show that the elastic energy loss is of the same order of magnitude as inelastic at the energy regimes applicable at RHIC and LHC; see figure 3. The PHENIX collaboration performed

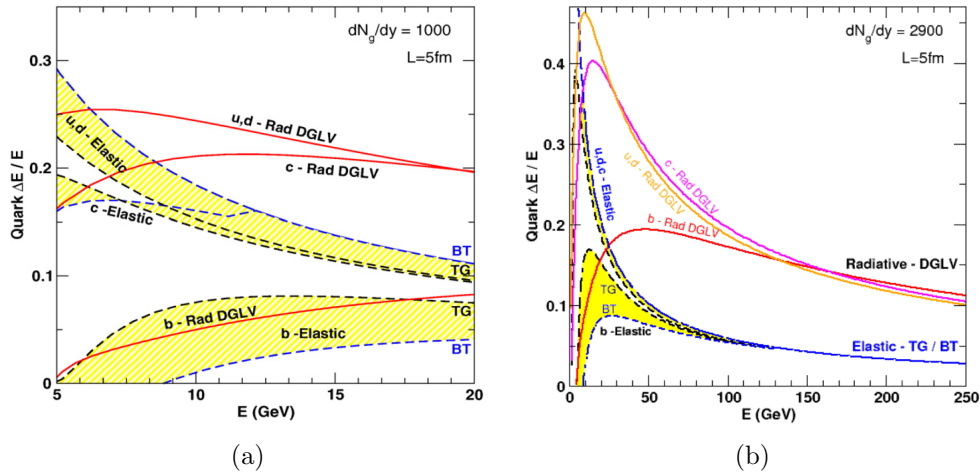


Figure 3: Comparison of magnitude of radiative and collisional energy loss in pQCD for (a) RHIC and (b) LHC [26]. Due to the LPM effect, elastic energy loss is the same order of magnitude as inelastic for all particle energies.

[27] a rigorous statistical analysis of the WHDG energy loss model [3] that incorporates both radiative and collisional energy loss in a reasonable geometric background that extracted the one free parameter in the calculation: the proportionality constant between the participant density and the number density of the color deconfined medium produced in heavy ion collisions. The value found by PHENIX corresponds to a central gluon rapidity density of $dN_g/dy = 1400^{+200}_{-375}$. Keeping the proportionality constant fixed, varying the medium density at different centralities and center of mass energies only by the measured multiplicities, the model robustly describes qualitatively a wealth of high- p_T observables; see figure 4, in which the theoretical uncertainty band is due only to the $1-\sigma$ range of values from the PHENIX proportionality constant extraction.

There are a very large number of sources of theoretical uncertainty not shown in the results above. Some of these sources of uncertainty include higher order contributions in: coupling α_s ; collinearity, or k_T/xE , where k_T is the radiated gluon's perpendicular momentum, and

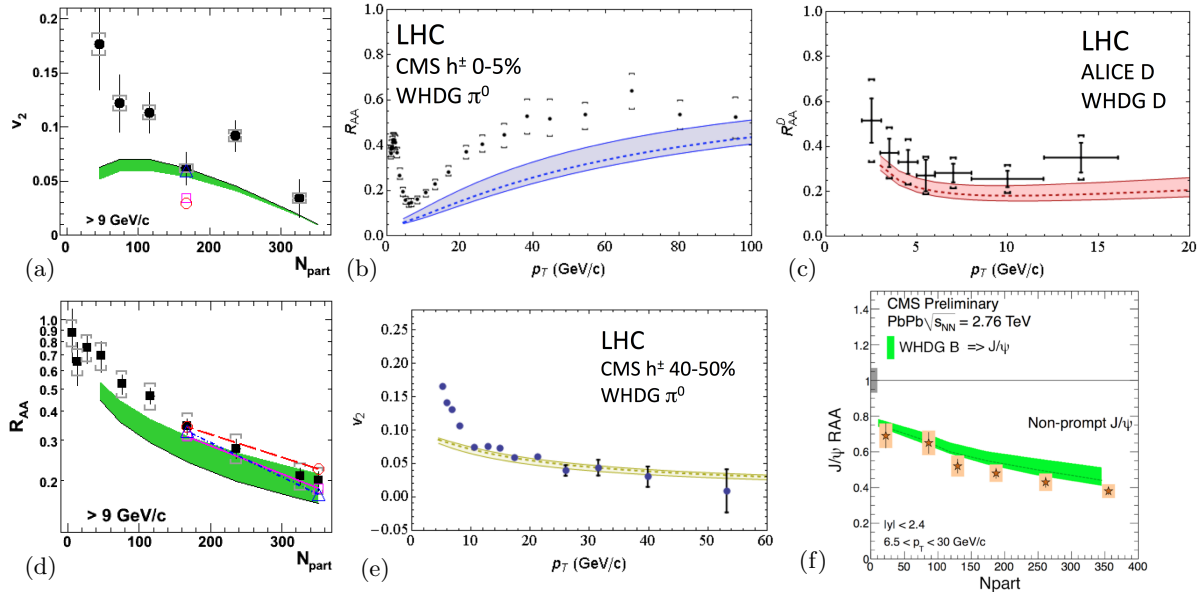


Figure 4: Constrained zero parameter WHDG predictions compared to data for (a) $v_2^{\pi^0}(N_{part})$ at RHIC [23, 3], (b) 0-5% centrality $R_{AA}(p_T)$ for light flavors at LHC [28, 29], (c) $R_{AA}^D(p_T)$ at 0-20% centrality at LHC [16, 22], (d) $R_{AA}^{\pi^0}(N_{part})$ at RHIC [23, 3], (e) $v_2(p_T)$ at LHC for light flavors at 40-50% centrality [30, 22], and (f) $R_{AA}^{B \rightarrow J/\psi}(N_{part})$ at LHC [31, 15].

xE is the fraction of the leading parton's initial energy carried away by the emitted gluon; softness, x ; quark mass to energy of the leading heavy quark, M_q/E ; and opacity, the ratio of the mean free path to the pathlength, λ_{mfp}/L . An early attempt [32] to estimate the sensitivity of the calculation to higher order contributions in α_s varied the value of α_s from 0.2 to 0.4 and found a strong dependence of the suppression on the value of α_s chosen, not surprisingly as $dp/dt_{coll} \sim \alpha_s^2$ and $dp/dt_{rad} \sim \alpha_s^3$, although the amount of dependence absorbed by reevaluating the proportionality constant as α_s was varied was never explored. More recent work with a running coupling ansatz found in fact a better agreement with pion suppression as a function of p_T than the fixed coupling calculation [5]. Others showed that [33] pQCD calculations rather significantly violate the assumption of collinearity at RHIC and LHC energies and are very sensitive to the treatment of wide angle radiation. Predictions appear stable, i.e. not sensitive, once the proportionality constant is fixed for a given prescription for the treatment of the wide angle radiation [29]. However the inferred properties of the medium depend on the proportionality constant, which may vary by a factor of 3 due to the uncertainty in the treatment of wide angle radiation [33]; quantitative information regarding the medium therefore requires a detailed understanding of the higher order corrections in collinearity.

4. Discussion and Conclusions

How can we move forward to resolve the seeming contradiction between the weakly- and strongly-coupled pictures? How can we narrow down the theoretical uncertainties due to 1) the jet definition prescription and 2) the initial conditions in AdS space setup? We hope to address both of these issues simultaneously with a hybrid weak-strong energy loss model. Perhaps the early energy loss evolution is dominated by weak-coupling physics but later evolution, as the medium cools, is dominated by strong-coupling physics. We are in the process of creating a model that interfaces these two regimes by matching the finite time energy-momentum tensor of a high momentum colored object created in a heavy ion collision as calculated in pQCD to

that calculated in AdS/CFT.

Acknowledgments

The author wishes to thank the South African National Research Foundation and SA-CERN for support.

5. References

- [1] Chesler P M and Yaffe L G 2011 *Phys.Rev.Lett.* **106** 021601 (*Preprint* [1011.3562](#))
- [2] Kovtun P, Son D and Starinets A 2005 *Phys.Rev.Lett.* **94** 111601 (*Preprint* [hep-th/0405231](#))
- [3] Wicks S, Horowitz W, Djordjevic M and Gyulassy M 2007 *Nucl.Phys.* **A784** 426–442 (*Preprint* [nucl-th/0512076](#))
- [4] Horowitz W 2013 *Nucl.Phys.A904-905* **2013** 186c–193c (*Preprint* [1210.8330](#))
- [5] Buzzatti A and Gyulassy M 2013 *Nucl.Phys.A904-905* **2013** 779c–782c (*Preprint* [1210.6417](#))
- [6] Danielewicz P and Gyulassy M 1985 *Phys.Rev.* **D31** 53–62
- [7] Chen J W, Deng J, Dong H and Wang Q 2013 *Phys.Rev.* **C87** 024910 (*Preprint* [1107.0522](#))
- [8] Gubser S S 2006 *Phys.Rev.* **D74** 126005 (*Preprint* [hep-th/0605182](#))
- [9] Herzog C, Karch A, Kovtun P, Kozcaz C and Yaffe L 2006 *JHEP* **0607** 013 (*Preprint* [hep-th/0605158](#))
- [10] Bethe H and Heitler W 1934 *Proc.Roy.Soc.Lond.* **A146** 83–112
- [11] Djordjevic M and Gyulassy M 2004 *Nucl.Phys.* **A733** 265–298 (*Preprint* [nucl-th/0310076](#))
- [12] Horowitz W and Gyulassy M 2008 *Phys.Lett.* **B666** 320–323 (*Preprint* [0706.2336](#))
- [13] Dion A (PHENIX collaboration) 2009 *Nucl.Phys.* **A830** 765C–768C (*Preprint* [0907.4749](#))
- [14] Bielcik J (STAR Collaboration) 2006 *Nucl.Phys.* **A774** 697–700 (*Preprint* [nucl-ex/0511005](#))
- [15] Horowitz W 2012 *AIP Conf.Proc.* **1441** 889–891 (*Preprint* [1108.5876](#))
- [16] Abelev B *et al.* (ALICE Collaboration) 2012 *JHEP* **1209** 112 (*Preprint* [1203.2160](#))
- [17] Chatrchyan S *et al.* (CMS Collaboration) 2012 *JHEP* **1205** 063 (*Preprint* [1201.5069](#))
- [18] Gubser S S 2008 *Nucl.Phys.* **B790** 175–199 (*Preprint* [hep-th/0612143](#))
- [19] Casalderey-Solana J and Teaney D 2007 *JHEP* **0704** 039 (*Preprint* [hep-th/0701123](#))
- [20] Chesler P M, Jensen K, Karch A and Yaffe L G 2009 *Phys.Rev.* **D79** 125015 (*Preprint* [0810.1985](#))
- [21] Ficin A and Gubser S S 2014 *Phys.Rev.* **D89** 026002 (*Preprint* [1306.6648](#))
- [22] Horowitz W and Gyulassy M 2011 *J.Phys.* **G38** 124114 (*Preprint* [1107.2136](#))
- [23] Adare A *et al.* (PHENIX Collaboration) 2010 *Phys.Rev.Lett.* **105** 142301 (*Preprint* [1006.3740](#))
- [24] Aamodt K *et al.* (ALICE Collaboration) 2011 *Phys.Lett.* **B696** 30–39 (*Preprint* [1012.1004](#))
- [25] Morad R and Horowitz W A *in preparation*
- [26] Horowitz W A 2010 (*Preprint* [1011.4316](#))
- [27] Adare A *et al.* (PHENIX Collaboration) 2008 *Phys.Rev.* **C77** 064907 (*Preprint* [0801.1665](#))
- [28] Chatrchyan S *et al.* (CMS Collaboration) 2012 *Eur.Phys.J.* **C72** 1945 (*Preprint* [1202.2554](#))
- [29] Horowitz W and Gyulassy M 2011 *Nucl.Phys.* **A872** 265–285 (*Preprint* [1104.4958](#))
- [30] Chatrchyan S *et al.* (CMS Collaboration) 2012 *Phys.Rev.Lett.* **109** 022301 (*Preprint* [1204.1850](#))
- [31] Mironov C (CMS Collaboration) 2013 *Nucl.Phys.* **A904-905** 194c–201c
- [32] Wicks S 2008
- [33] Horowitz W and Cole B 2010 *Phys.Rev.* **C81** 024909 (*Preprint* [0910.1823](#))

Finite-size key security of Phoenix-Barnett-Chefles 2000 quantum-key-distribution protocol

Mhlambululi Mafu¹, Kevin Garapo¹ and Francesco Petruccione^{1,2}

¹ Centre for Quantum Technology, School of Chemistry and Physics, University of KwaZulu-Natal, P/Bag X54001 Durban 4000, South Africa

² National Institute for Theoretical Physics (NITheP), KwaZulu-Natal, South Africa

E-mail: mafum@ukzn.ac.za

Abstract. A post-selection technique was introduced by Christandl, König and Renner [*Phys Rev. Lett.* 102, 020504 (2009)] in order to simplify the security of quantum key distribution schemes, yet it has not been worked out in detail for any specific and realistic protocol. Here, we demonstrate how it can be applied to study the security of the Phoenix-Barnett-Chefles 2000 trine state (PBC00) protocol, a symmetric version of the Bennett 1992 (B92) protocol.

1. Introduction

Quantum key distribution (QKD) enables secure real-time distribution of a cryptographic key between two parties, Alice and Bob, who are connected by a quantum channel and an authenticated classical channel in the presence of a competent malicious party, an eavesdropper called Eve [1]. This technique is based on two laws of quantum mechanics, namely the uncertainty principle and the no-cloning theorem [1]. The first, and most established, operable QKD protocol, BB84, was proposed by Bennett and Brassard [2] in 1984, basing on Wiesner's idea of conjugate coding [3]. In 1991, Ekert [4] extended the idea by introducing quantum entanglement and the violation of Bell's theorem [1]. Ever since, several protocols have been invented by both theorists and experimentalists. These include: Bennett 1992 (B92) [5], six state [6], the Scarani-Acín-Ribordy-Gisin 2004 (SARG04) [7], the differential-phase shift (DPS) [8], the coherent-one-way (COW) [9] and the Phoenix, Barnett and Chefles 2000 (PBC00) protocol [10]. The PBC00 protocol is based on the B92 protocol and the main difference between the two is that the latter uses two states whilst the former uses three states. Consequently, the PBC00 protocol is more symmetrical (in the sense that either Alice or Bob, but not both, can declare unused observables) and shows lower qubit losses than the B92 protocol.

Since the first unconditional security proof by Mayers [11], various techniques for proving the security of QKD protocols have been developed [12]. An unconditional security proof is a proof that considers an unbounded adversary. The construction of security proofs generally depend on the steps of the protocol and also on their practical implementation. For example, the unconditional security proof for the BB84 based protocols have long since been introduced [13]. This is because they share a common property of being symmetrical. However, the security proofs for the class of distributed-phase-reference (which consist of DPS and COW) protocols

still remain unknown [12], mainly because their construction and encoding deviates from the usual symmetry that exists in BB84-type protocols [8]. The study of security proofs of QKD protocols in the asymptotic key length has been studied in several papers [14, 15, 16, 17, 18, 19].

The unconditional security proof of the PBC00 protocol independent of the channel's qubit loss rate loss has been given in Ref [20]. The security proof was done by transforming the protocol to a secure QKD protocol based on the entanglement distillation protocol initiated by state rotations and a local filtering operation, followed by error correction. In this paper we compute the secret key rates against coherent attacks for the PBC00 protocol by using the post-selection technique when given a finite amount of resources.

In order to prove security against coherent attacks, one can use either the de Finetti theorem [21] or the post-selection technique [22]. The former technique was developed by Renner where one uses the fact that a permutation invariant state is close to an independent and identically distributed (i.i.d) state [23]. The latter technique gives a way to bound the diamond distance between two maps from a quantity evaluated for a very specific quantum state, called a “de Finetti state”. The two techniques are independent, therefore one can use either of them to prove security of a protocol provided the protocol is permutation invariant [22]. However, it has been shown that the post-selection technique gives better bounds than the de Finetti theorem [16], so one can always proceed by using the post-selection technique. Moreover, the symmetry, hence permutation invariance that exists in the PBC00 protocol enables us to directly apply the post-selection technique.

2. PBC00 QKD protocol

The PBC00 protocol proceeds as follows:

Preparation and Measurement. Alice prepares randomly with equal probability each qubit in one of the mutually non-orthogonal states $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle\}$ and encodes her bit on the states as in the original B92 protocol [24]. These states are defined as $|\psi_1\rangle \equiv \frac{1}{2}|0_x\rangle + \frac{\sqrt{3}}{2}|1_x\rangle$, $|\psi_2\rangle \equiv \frac{1}{2}|0_x\rangle - \frac{\sqrt{3}}{2}|1_x\rangle$ and $|\psi_3\rangle \equiv |0_x\rangle$, where $\{|0_x\rangle, |1_x\rangle\}$ is the X basis of a qubit state. The Z basis is defined by $\{|j_z\rangle \equiv [|0_x\rangle + (-1)^j|1_x\rangle]/\sqrt{2}\}$ where $(j = 0, 1)$. She creates a large trit string r and a large bit string b of the same length N . For each r_i , the i th trit value of the trit string r , she chooses the set $\{|\psi_1\rangle, |\psi_2\rangle\}$ (if $r_i=0$), $\{|\psi_2\rangle, |\psi_3\rangle\}$ (if $r_i=1$), and $\{|\psi_3\rangle, |\psi_1\rangle\}$ (if $r_i=2$). If the i th bit value b_i is 0, she prepares the first state of the chosen pair, however if the bit is 1, she prepares the second state. Alice sends all the prepared states to Bob.

On the receiving side, Bob performs measurements on each qubit he receives. These measurements are described by the POVM

$$\left\{\frac{2}{3}|\bar{\psi}_1\rangle\langle\bar{\psi}_1|, \frac{2}{3}|\bar{\psi}_2\rangle\langle\bar{\psi}_2|, \frac{2}{3}|\bar{\psi}_3\rangle\langle\bar{\psi}_3|\right\}. \quad (1)$$

These states can be defined as $|\bar{\psi}_1\rangle = \frac{\sqrt{3}}{2}|0_x\rangle - \frac{1}{2}|1_x\rangle$, $|\bar{\psi}_2\rangle = \frac{\sqrt{3}}{2}|0_x\rangle + \frac{1}{2}|1_x\rangle$ and $|\bar{\psi}_3\rangle = |1_x\rangle$ and are orthogonal to $|\psi_1\rangle, |\psi_2\rangle$ and $|\psi_3\rangle$ respectively.

Sifting. In this sifting step, both parties agree which signals to discard. Bob announces when all his measurements are done, and Alice in turn announces the trit string r . Bob regards the i^{th} measurement outcome $|\bar{\psi}_1\rangle$ (if $r_i = 0$), $|\bar{\psi}_2\rangle$ (if $r_i = 1$), and $|\bar{\psi}_3\rangle$ (if $r_i = 2$) as the bit value 0. Bob also considers $|\bar{\psi}_2\rangle$ (if $r_i = 0$), $|\bar{\psi}_3\rangle$ (if $r_i = 1$), and $|\bar{\psi}_1\rangle$ (if $r_i = 2$) as the bit value 1. All other events are considered as inconclusive and they discard the data. Bob announces whether his measurement outcome is inconclusive or not. Alice and Bob keep the data when Bob's outcome is conclusive, discarding the rest.

Parameter estimation. The role of parameter estimation is to decide whether the input given to the protocol can be used to distill a secret key. On the conclusive measurements, Alice randomly chooses $m < N$ samples as test bits in order to estimate the Quantum Bit Error Rate (QBER) on the code bits, and announces her selection to Bob. If the error rate is small, they use the post-processing methods to extract the key otherwise if the error rate is too high Alice and Bob abort the protocol. If the statistics λ_m are obtained by measuring m samples of ρ_{AB} according to a POVM with d possible outcomes and if $\lambda_\infty(\rho_{AB})$ denotes the perfect statistics in the limit of infinite measurements, then for any state ρ_{AB}

$$\Gamma_\xi := \{\rho_{AB} : \|\lambda_m - \lambda_\infty(\rho_{AB})\| \leq \xi(\varepsilon_{PE, m_i})\}. \quad (2)$$

Γ_ξ is a set of states from which a key is extracted with a non-negligible probability. By the law of large numbers [23]

$$\xi(\varepsilon_{PE, m_i}) := \sqrt{\frac{\ln(1/\varepsilon_{PE}) + 2 \ln(m+1)}{2m}}. \quad (3)$$

where ε_{PE, m_i} is the measured QBER in the direction $|\psi_1\rangle, |\psi_2\rangle$ and $|\psi_3\rangle$.

Error correction. In this step Alice and Bob apply a one-way error correction protocol by using authenticated classical communication channel to correct their strings. As result they will exchange L_{EC} bits on the channel and is given as

$$L_{EC} = f_{EC} n h(Q) + \log_2 \left(\frac{2}{\varepsilon_{EC}} \right), \quad (4)$$

where n is the length of the raw key, f_{EC} is a constant larger than 1 which represents a deviation of the real protocol from the asymptotic one, $h(Q) = -Q \log_2 Q - (1-Q) \log_2 (1-Q)$ is the binary Shannon entropy and Q is the QBER and ε_{EC} is the error in error correction step.

Privacy amplification. In this step, the quantity of the correct information which the eavesdropper may have obtained about Alice's and Bob's reference raw key is minimized by the use of a two-universal hash function resulting in a string called a key [23].

3. Security against collective attacks

The PBC00 protocol belongs to the class of prepare and measure schemes. It employs polarization-encoded qubits to transmit information. This simply involves a sequential exchange of n signals. Based on that, Alice and Bob's systems can be described by an n -partite density operator which is permutation invariant. This allows us to analyze the protocol based on the fact that it consists of states which are independent and identical copies of each other. More formally, the states $|\psi_1\rangle, |\psi_2\rangle$ and $|\psi_3\rangle$ all occur with the same probability.

The aim of the eavesdropper is to extract as much classical information as possible from the strings that are held by Alice and Bob. In the PBC00 protocol, the probability that Eve maximizes her probability of correctly distinguishing between the exchanged states is expressed as $P_D = \sum_i \mathcal{X}_i P_{(i,i)}$, where P_D is the discrimination probability, \mathcal{X}_i represents the a priori probability of the state $|i\rangle$ and $P_{(i,i)}$ denotes the probability that the state $|i\rangle$ is sent and that the Eve's measurement reveals the result i . For symmetric states, the maximum discrimination probability is equal to $P_D^{\max} = 2/3$. Based on this probability, without losing generality the number of states that can be unambiguously discriminated satisfies the inequality $N \leq \frac{2}{3} \binom{n+d-1}{n}$, where d is the photon Hilbert space i.e., $d=2$ for the PBC00 protocol. This gives us the necessary condition for unambiguous discrimination between N states each spanning a d -dimensional space when given n -copies of the state. The dimension of the symmetric subspace, $g_{n,d}$ can be expressed

as $g_{n,d} = \frac{2}{3} \binom{n+d-1}{n} \leq \frac{2}{3} (n+1)^{d-1}$. Since the PBC00 protocol uses symmetric encoding as in the BB84 protocol, it is then possible to parametrize the most general attacks. This leads to the unconditional security. The expression for the secret key rate $r_{N,\text{col}}$ against collective attacks is written as

$$r_N = r_{N,\text{col}} - \frac{2(d^4 - 1) \log_2(N+1)}{N}. \quad (5)$$

The bound for collective attacks for the PBC00 protocol is expressed as

$$\varepsilon \leq \frac{2}{3} \cdot 2^{-c\delta^2 N + (d^4 - 1) \log_2(N+1)} \quad (6)$$

and is calculated via the total security parameter ε defined later in (13), where c for $(c \geq 0)$ is the cost of reducing the key size by fraction δ [22].

4. Finite key analysis

Scarani and Renner have provided a general security formula to calculate the lower bound on the secure key rate of the BB84 protocol [14]. This has been followed by the derivation of many bounds for different protocols based on the BB84 encoding. In the same spirit, we use this intuition of finite size key analysis to derive the upper bound for the secure key rate against collective attacks for the PBC00 protocol.

Let ρ_{KE} be the quantum state that describes the classical key K of length ℓ distilled at the end of the run of the QKD protocol. Then for any $\varepsilon \geq 0$, a final key K is said to be ε -secure with respect to an adversary Eve if the key ρ_{KE} satisfies

$$\min_{\rho_E} \|\rho_{KE} - \tau_K \otimes \rho_E\|_1 \leq \varepsilon, \quad (7)$$

where $\|\cdot\|_1$ is the trace distance, $\rho_{KE} = \sum_{k \in \mathcal{K}} P_k(k) |k\rangle\langle k| \otimes \rho_E^k$ where P_k is the probability distribution of the key K and $\{|k\rangle\}_{k \in \mathcal{K}}$ is an orthonormal basis of some Hilbert space \mathcal{H}_k , $\tau_K = \sum_{k \in \mathcal{K}} P_k \frac{1}{|\mathcal{K}|} |k\rangle\langle k|$ is a fully mixed state on \mathcal{H}_k , ρ_E is the state held by an eavesdropper [23]. The parameter ε represents the maximum failure probability of the key extraction procedure. Therefore, the classical key K is indistinguishable from a random and uniform key with probability $1 - \varepsilon$.

In order to arrive at our security bounds for the PBC00 protocol, we follow closely the definitions found in Ref. [23] and the formalism in Ref. [14]. If \mathcal{H} denotes a finite-dimensional Hilbert space and $\mathcal{P}(\mathcal{H})$ is the set of positive semidefinite operators on \mathcal{H} then the set $\mathcal{S}(\mathcal{H}) := \{\rho \in \mathcal{P}(\mathcal{H}) : \text{tr} \rho = 1\}$ represents normalized states and the set $\mathcal{S}_{\leq}(\mathcal{H}) := \{\rho \in \mathcal{P}(\mathcal{H}) : \text{tr} \rho \leq 1\}$ represents the set of sub-normalized states on the Hilbert space. Let $\rho_{AB} \in \mathcal{S}_{\leq}(\mathcal{H}_{AB})$ and $\sigma_B \in \mathcal{S}(\mathcal{H}_B)$, then the min-entropy of A conditioned on B of the state ρ_{AB} relative to σ_B is defined as

$$H_{\min}(A|B)_{\rho|\sigma} := \max_{\sigma} \sup \{\lambda \in \mathbb{R} : \rho_{AB} \leq 2^{-\lambda} \mathbb{1}_A \otimes \sigma_B\}, \quad (8)$$

where the maximum is taken over the states $\sigma_B \in \mathcal{S}(\mathcal{H}_B)$. Furthermore we define,

$$H_{\min}(A|B)_{\rho} := \max_{\sigma_B \in \mathcal{S}(\mathcal{H}_B)} H_{\min}(A|B)_{\rho|\sigma}. \quad (9)$$

For some $\varepsilon \geq 0$, it has been found that the achievable length of secret key can be expressed as [25]

$$\ell \leq H_{\min}^{\bar{\varepsilon}}(X^n|E^n) - L_{\text{EC}} - 2 \log_2(1/\varepsilon), \quad (10)$$

where $\bar{\varepsilon} = (\varepsilon/8)^2$ is the smoothing parameter. In above expression, $H_{\min}^{\bar{\varepsilon}}(X^n|E^n)$ gives the amount of a key that can be extracted from a string X when given E , the uncertainty of the

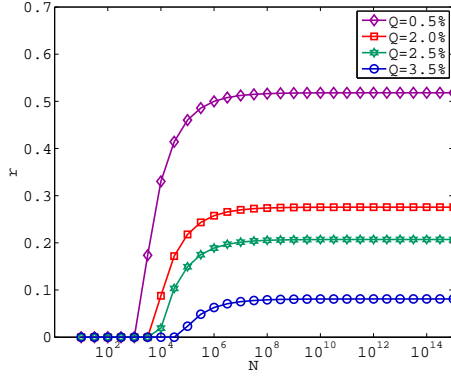


Figure 1. (Color online) Upper bound on the secret key fraction, r , for the finite PBC00 protocol as a function of the exchanged quantum signals N for bit error rates $Q = 0.5\%$, 2% , 2.5% , 3.5% , $\varepsilon = 10^{-5}$ and $\varepsilon_{EC} = 10^{-10}$.

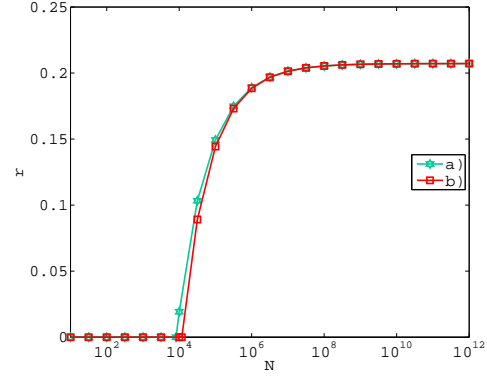


Figure 2. (Color online) Upper bound on the secret key fraction, r , for the finite PBC00 protocol as a function of the exchanged quantum signals N for bit error rate $Q = 2.5\%$, $\varepsilon = 10^{-5}$, $\varepsilon_{EC} = 10^{-10}$ for (a) collective attacks (b) post-selection technique.

adversary about X measured by using the smooth min-entropies. The smooth min-entropy of the state shared between Alice and the eavesdropper is lower bounded as

$$H_{\min}^{\bar{\varepsilon}}(X^n|E^n) \geq n(\min_{\sigma_{XE} \in \Gamma} H(X|E) - \Delta), \quad (11)$$

where $\Delta = (2 \log_2 d + 3) \sqrt{[\log_2(2/\bar{\varepsilon})]/n}$. For a finite number of signals, the achievable secure key rate was derived to be [14]

$$r = \frac{n}{N} \left[\min_{\sigma_{XE} \in \Gamma} H(X|E) + \Delta(n) - L_{EC} \right] + \frac{2}{N} \log_2(2\varepsilon_{PA}). \quad (12)$$

The total security parameter, ε of a QKD scheme depends on the sum of probabilities of failures of the classical post-processing protocols which can be written as

$$\varepsilon = \bar{\varepsilon} + \varepsilon_{PA} + \varepsilon_{EC} + \varepsilon_{PE}, \quad (13)$$

where $\bar{\varepsilon}$ denotes the error in the smooth min-entropy. In the PBC00 protocol, the secret key rate is obtained by measuring both the bit (e_1) and phase error (e_2) rates. The secret key rate for the bit error rate and the phase error rate is expressed as $p_c[1 - h(e_1) - h(e_2)]$, where p_c is the probability of the conclusive outcome. It has been found that the protocol is secure for a bit error rate of up to 9.81%. For an asymptotic formula $e_1 = e_2 = Q$, the lower bound for the protocol can be expressed as

$$H(X|E) = p_c[1 - h(Q) - h(\frac{5}{4}Q)], \quad (14)$$

where $H(X|E)$ is evaluated by using the QBER from the parameter estimation step.

In figure 1, we show the variation of the secret key rate r , with the number of signals N , which Alice sends to Bob when given a finite amount of resources. We found that the minimum number of signals required in order to extract a reasonable amount of secret key compares with

those of the BB84 protocol [14]. In figure 2, we show the variation of the secret key rate r , with the number of signals N for bit error $Q = 2.5\%$. In this figure, it can be observed that the post-selection technique gives optimal bounds when a finite amount of signals are used. Therefore, for a large number of signals the optimal attack is close to a collective attack. This shows that this technique is a powerful tool which can be considered in simplifying security proofs and also leads to improved key rates.

5. Conclusion

We have shown how one can apply the finite-key analysis formalism and the results of the post-selection technique in order to find the secret key rates for finite amount of resources. We have shown that the secret key rate largely depends on the number of signals sent. In particular, reasonable key rates are obtained for $N \approx 10^5$ - 10^6 signals. We have also shown that the post-selection technique leads to optimal security bounds for the PBC00 protocol. These results also appear in the longer version of our paper in Ref [26]. Therefore, in the absence of bounds for collective attacks one can appeal to the bounds given by the post-selection technique. These results can be applied to other protocols which are symmetric as well. This study has demonstrated the feasibility of applying the post-selection technique to the PBC00 QKD protocol, which is a specific and realistic protocol.

Acknowledgments

This work is based on research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation.

References

- [1] Scarani V, Bechmann-Pasquinucci H, Cerf N, Dušek M, Lütkenhaus N and Peev M 2009 *Reviews of Modern Physics* **81** 1301–1350 ISSN 1539-0756
- [2] Bennett C, Brassard G *et al.* 1984 *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* vol 175 (Bangalore, India)
- [3] Wiesner S 1983 *ACM Sigact News* **15** 78–88
- [4] Ekert A 1991 *Physical Review Letters* **67** 661–663
- [5] Bennett C H 1992 *Phys. Rev. Lett.* **68** 3121–3124
- [6] Bruß D 1998 *Phys. Rev. Lett.* **81** 3018–3021
- [7] Scarani V, Acín A, Ribordy G and Gisin N 2004 *Phys. Rev. Lett.* **92** 057901
- [8] Inoue K, Waks E and Yamamoto Y 2002 *Phys. Rev. Lett.* **89** 037902
- [9] Stucki D, Fasel S, Gisin N, Thoma Y and Zbinden H 2007 *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* vol 6583 p 18
- [10] Phoenix S J, Barnett S M and Cheffles A 2000 *Journal of Modern Optics* **47** 507–516
- [11] Mayers D 2001 *J. ACM* **48** 351–406 ISSN 0004-5411
- [12] Scarani V, Bechmann-Pasquinucci H, Cerf N J, Dušek M, Lütkenhaus N and Peev M 2009 *Rev. Mod. Phys.* **81** 1301–1350
- [13] Shor P and Preskill J 2000 *Physical Review Letters* **85** 441–444
- [14] Scarani V and Renner R 2008 *Phys. Rev. Lett.* **100**(20) 200501
- [15] Cai R and Scarani V 2009 *New Journal of Physics* **11** 045024
- [16] Sheridan L, Le T P and Scarani V 2010 *New Journal of Physics* **12** 123019
- [17] Abruzzo S, Kampermann H, Mertz M and Bruß D 2011 *Physical Review A* **84** 032321
- [18] Tomamichel M, Lim C C W, Gisin N and Renner R 2012 *Nature communications* **3** 634
- [19] Mafu M, Garapo K and Petruccione F 2013 *Physical Review A* **88** 062306
- [20] Boileau J, Tamaki K, Batuwantudawe J, Laflamme R and Renes J 2005 *Physical Review Letters* **94** 40503
- [21] Renner R and Cirac J 2009 *Phys. Rev. Lett.* **102** 110504
- [22] Christandl M, König R and Renner R 2009 *Physical Review Letters* **102** 20504
- [23] Renner R 2008 *International Journal of Quantum Information* **6** 1–127
- [24] Bennett C 1992 *Physical Review Letters* **68** 3121–3124
- [25] Kraus B, Gisin N and Renner R 2005 *Phys. Rev. Lett.* **95** 080501
- [26] Mafu M, Garapo K and Petruccione F 2014 *Phys. Rev. A* **90**(3) 032308

Quark-gluon plasma physics from string theory

Razieh Morad and William A Horowitz

Department of Physics, University of Cape Town, Private Bag X3, Rondebosch 7701, SA

E-mail: razieh.morad@uct.ac.za and wa.horowitz@uct.ac.za

Abstract. The goal of high-energy nuclear physics is to create and study quark-gluon plasma (QGP), the predicted deconfined state of QCD matter at energy densities greater than $1 \text{ GeV}/\text{fm}^3$ that permeated the universe a microsecond after the Big Bang. Contrary to original expectations, the properties of the QGP seem best described by the strong-coupling, phenomenological string theory methods of the AdS/CFT correspondence instead of the usual weak-coupling, Feynman diagram methods of perturbative QCD (pQCD). In particular, the AdS/CFT paradigm predicts a very small value for the viscosity to entropy ratio of the QGP, in remarkable agreement with data collected from the Relativistic Heavy Ion Collider (RHIC) and the Large Hadron Collider (LHC). In search of a consistent description for all observables related to QGP, we extend the AdS/CFT theory to that of high-momentum probes of the plasma and compare our results to data from LHC.

1. Introduction

Recent experiments performed at the Relativistic Heavy Ion Collider (RHIC) and the Large Hadron Collider (LHC) has provided spectacular evidence that suggests that a deconfined state of hadronic matter has been formed which is called quark gluon plasma (QGP) [1–4] with a small ratio of shear viscosity to entropy density [5; 6]. The experimental discovery that QGP is a strongly coupled plasma with low viscosity, poses a challenge to theorists. While lattice QCD is the proper tool for understanding the static equilibrium thermodynamics of such strongly coupled plasma, it does not allow us to calculate its dynamics evolution on heavy-ion collision.

Recently, a novel tool called "the AdS/CFT correspondence" [7–11] provide valuable insight into the strongly coupled plasma. In according to the original conjecture of Maldacena, the $\mathcal{N} = 4$ SYM theory in the large N_c and large 't Hooft coupling is dual to classical supergravity on ten-dimensional $AdS_5 \times S^5$ geometry [7]. In order to study the theory at finite temperature, one can add black hole (BH) to the geometry [8] which yields to the AdS-Sch metric. Fundamental quarks are described by open strings moving in the 10d geometry. In the large λ limit, the quantum fluctuation of string world sheet are suppressed and the dynamics of string are described by the classical string theory.

Jets are produced within the expanding fireball and probe the QGP. Analyzes the energy loss of these energetic partons as they travel throw QGP may reveal extremely valuable information about the dynamics of the plasma and exhibit distinctive properties such as jet-quenching which can clearly be observed at RHIC [1–4] and more recently LHC [12–14].

In this paper, we study the light quark jet energy loss in both gravitational dual to static and expanding plasma. We propose a new prescription of jets in string theory based on the separation of hard and soft sectors. We demonstrate that the light quark jet energy loss reveal

the “Bragg peak” at late times. Then we compute the nuclear modification factor of jets R_{AA}^{jet} , renormalise the quantity, and compare the results with the preliminary CMS data [15].

2. Light quark jet energy loss

According to the AdS/CFT correspondence [10], the $\mathcal{N} = 4$ SYM theory at finite temperature is dual to a $10d$ black hole geometry with the AdS-Schwarzschild (AdS-Sch) metric as follows,

$$ds^2 = \frac{L^2}{u^2} \left[-f(u) dt^2 + d\mathbf{x}^2 + \frac{du^2}{f(u)} \right], \quad (1)$$

where $f(u) \equiv 1 - (u/u_h)^4$ is the blackening factor and L is the AdS curvature radius. Four dimensional Minkowski coordinates are denoted by x_μ and the coordinate u is an inverse radial coordinate. So, the boundary of the AdS-Sch spacetime is at $u = 0$ and the event horizon is located at $u = u_h$. The temperature of the equilibrium SYM plasma relates to the event horizon as $T \equiv \frac{1}{(\pi u_h)}$. World sheet coordinates are as σ^a where $\tau \equiv \sigma^0$ is denoted as the timelike world sheet coordinate, while the spatial coordinate is $\sigma \equiv \sigma^1$.

Fundamental representation quarks added to the $\mathcal{N} = 4$ SYM theory are dual to open strings moving in the $10d$ geometry. Addition of a $\mathcal{N} = 2$ hypermultiplet to the $\mathcal{N} = 4$ SYM theory is performed by adding D7 branes to the $10d$ geometry [16]. These branes extend along the radial coordinate from the boundary at $u = 0$ down to maximal coordinate at $u = u_m$ as well as fill the whole 4d Minkowski space. Also, they wrap on S^3 from the S^5 sphere. The bare mass M of quark is proportional to $1/u_m$ [17], so for massless quarks the D7 brane should fill the whole radial direction. Open strings that are attached to the D7 brane are dual to the quark-anti quark pairs on the field theory side. In the $5d$ geometry these strings can fall unimpeded toward the event horizon until their end points reach the radial coordinate u_m where the D7 brane ends. Since for sufficiently light or massless quarks $u_m > u_h$, open string end points can fall into the horizon.

We are interested in studying the back-to-back jets so we consider the configurations in which the two endpoints of string move away from each other as the total spatial momentum of the string vanishes. By choosing the appropriate frame, one half of the string has a large spatial momentum in x direction, and the other half of the string carries a large negative spatial momentum. We will limit our attention to strings which move in one direction in the R^3 space, x direction, so the embedding function of string $X^\mu(\tau, \sigma)$ will be a map to $(t(\tau, \sigma), x(\tau, \sigma), u(\tau, \sigma))$. So, the profile of an open string which is created at a point in space at time $t = t_c$ is given by

$$t(0, \sigma) = t_c, \quad x(0, \sigma) = 0, \quad u(0, \sigma) = u_c. \quad (2)$$

By these conditions, the string created at time t_c and by time evolution, the string evolves from a point into an extended object and the string endpoints fall toward the horizon. Polyakov action for the string has the form

$$S_P = -\frac{T_0}{2} \int d^2\sigma \sqrt{-\eta} \eta^{ab} \partial_a X^\mu \partial_b X^\nu G_{\mu\nu}. \quad (3)$$

Variation of the Polyakov action with respect to the embedding functions X^μ lead to the equation of motion as

$$\partial_a \left[\sqrt{-\eta} \Pi_\mu^a \right] = \frac{1}{2} \sqrt{-\eta} \eta^{ab} \frac{\partial G_{\nu\rho}}{\partial X^\mu} \partial_a X^\nu \partial_b X^\rho, \quad (4)$$

where Π_μ^a are the canonical momentum densities associated to the string obtained from varying the action with respect to the derivatives of the embedding functions.

One then has the problem of finding the proper object in the dual string theory that corresponds to a jet, a slippery object even in field theory; jets are truly only defined by the algorithm used to measure them. Presumably the ideal way to compute jet observables in the dual theory is to compute the energy momentum tensor associated with a high-momentum probe and “run” a jet finding algorithm on the result. However, one can define a jet prescription in the AdS/CFT and calculate the rate of energy loss from the string profile itself.

The authors of [18] are motivated by the localization of the baryon density on the boundary which is of scale of order $\Delta x \sim 1/\pi T$ and defined jet as a part of string which is in the Δx spatial distance from the endpoint. We called this as the “ Δx – prescription” of jet [18].

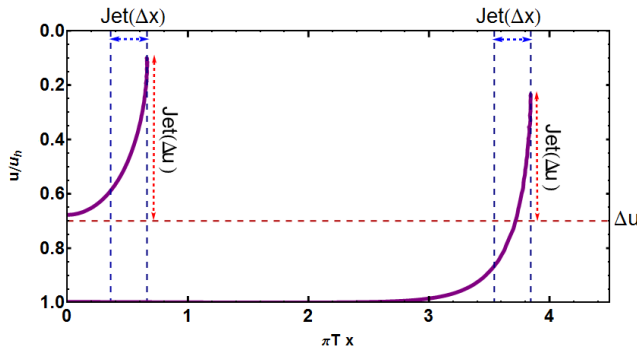


Figure 1. Illustration of the Δx and Δu prescriptions of a jet in the string theory; see text for details.

In this paper, motivated by the separation of energy scales in, e.g., thermal field theory, we propose rather a Δu prescription which we believe will ultimately provide a closer approximation to the result of a more complete calculation, figure 1. Since the radial coordinate in the string theory sets an energy scale in the field theory, in our Δu prescription the portion of the string above some cutoff $u = u^*$ in the radial direction is considered part of the jet; the portion of the string below the cutoff is considered part of the thermalized medium. By choosing any value of u above the black hole horizon as the cutoff, we regain the natural result that a jet that is thermalized no longer has detectable energy or momentum.

We evaluate the energy loss rate of jet in the radial $\sigma = u$ parametrization for both prescriptions of jet and plot in figure 2(a,b). In order to define jet using the Δx – prescription, we choose $\sigma_\kappa(t)$ as $\Delta x = 0.3/\pi T$. Our result on the energy loss rate of light quark using the Δu – prescription shows a Bragg peak at late times which means the explosive transfer of quark energy to the plasma at late times and is consistence with the previous works [18].

Since the quark-gluon plasma produced at ultra-relativistic heavy ion collision is an expanding and cooling medium, we study the light quark jet energy loss in a time dependent gravity dual to the boost invariant flow [19]. This geometry is similar to the static black hole geometry, but the location of the horizon moves in the bulk as $\tau^{1/3}$ where τ is the proper time. Also the temperature of plasma cools as $T(\tau) \tau^{-1/3}$.

We calculate the energy loss rate of light quark using both Δx – prescription and Δu – prescription of jet. We consider the initial temperature of plasma in JP metric T_c the same as the temperature of plasma in AdS-Sch metric. The result is shown in figure 2(c,d) which demonstrates that the behavior of light quark energy loss in the JP metric is same as the AdS-Sch metric, but the distance that quark traveled before thermalizing increases.

3. Jet Nuclear Modification Factor

To compare our toy model with experimental data, we calculate an approximation of the nuclear modification factor R_{AA} for jet in the following way. We consider the contributions of both quark and gluon jets. We assume that the produced parton with initial energy p_T^i loses a fraction of its

We evaluate the energy loss rate of jet in the radial $\sigma = u$ parametrization for both prescriptions of jet and plot in Fig. (2) (a,b). In order to define jet using the Δx -prescription, we choose $\sigma_\kappa(t)$ as $\Delta x = 0.3/\pi T$. Our result on the energy loss rate of light quark using the Δu -prescription shows a Bragg peak at late times which means the explosive transfer of quark energy to the plasma at late times and is consistence with the previous works [16].

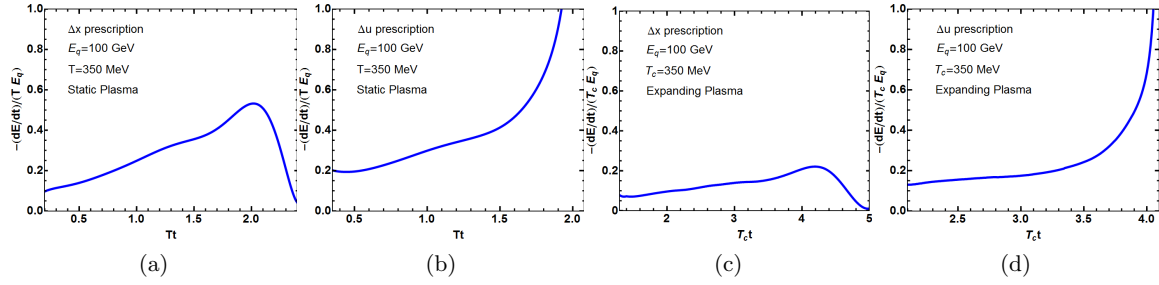


Figure 2: The instantaneous energy loss of light quark as a function of time in both AdS-BH metric using the (a) Δx -prescription of jet and (b) Δu -prescription of jet and JP metric using the (c) Δx -prescription of jet and (d) Δu -prescription of jet. The normalization constant $E_q = 100$ GeV is the initial energy of the jet which has a virtuality of 175 GeV^2 and $T = 350$ MeV is the temperature of the plasma. Our jet's definition reconstruct the Bragg peak in the energy loss rate of light quark in both static and expanding plasma.

energy ϵ with probability $P(\epsilon|p_T, L, T)$ as the final energy of parton is given by $p_T^f = (1 - \epsilon)p_T^i$. Since the quark-gluon plasma produced at ultra-relativistic heavy ion collision is an expanding and cooling medium, we study the light quark jet energy loss in a time dependent gravity dual [20], and the gluons loss their energy by the factor of 2 in the large N_c limit respect to the quarks, to the boost invariant flow [46]. This geometry is similar to the static black hole geometry. We suppose that the production spectrum can be approximately by a power law [20] with the location of the horizon moves in the bulk as $\tau \rightarrow \tau_0$ where τ is the proper time. Also the slowly varying with respect to τ then we may find a simple equation for the jet nuclear modification factor as follows,

$$R_{AA}^{R \rightarrow jet}(p_T) = \left\langle \int d\epsilon P(\epsilon|p_T, L, T) (1 - \epsilon^R)^{n_R(p_T)-1} \right\rangle. \quad (5)$$

For an absolutely uniform nucleus that is a 1D line, the geometric average is carried out as an integral over a line of production points with a parton that propagates through the line. In this case, $R_{AA}^{R \rightarrow jet}(p_T)$ is obtained from the below line integral [20]

$$R_{AA}^{R \rightarrow jet}(p_T) = \int_0^{L_{max}} \frac{dl}{L_{max}} (1 - \epsilon^R(p_T, l, T))^{n_R(p_T)-1}. \quad (6)$$

We calculate R_{AA} of jet by using our results of jet energy loss in both AdS-Sch and JP metric. Our results are shown in figure 3. The purple curve shows the R_{AA} obtained from the AdS-Sch metric, while the blue curve is the R_{AA} obtained from the JP metric. The AdS/CFT results of jet energy loss show an over suppression of jets in both static and expanding plasma.

The point-like initial condition falling string that we consider here is dual to creation of a pair of quark-antiquark which fly away each other in the strongly coupled plasma, interact and loss their energy. We expect that jets produced in the pp collision do not loss their energy. So, we consider the falling string with the same initial conditions in AdS_5 metric. Our results show that the string falls in the empty AdS_5 . So, jet loses its energy in the vacuum! We calculate the R_{AA} for a falling string in AdS_5 metric and show in figure 3 (a) in red curve.

In order to compare our results with the experimental data, we define a renormalized R_{AA} in AdS/CFT as

$$R_{AA}^{renorm} = \frac{R_{AA}^{medium}}{R_{AA}^{AdS_5}} \quad (7)$$

We plot the renormalized R_{AA}^{renorm} for jets in both AdS-Sch and JP metric in figure 3 (b) and compare with the CMS preliminary data for the most central Pb-Pb collision at $\sqrt{s_{NN}} = 2.76 \text{ GeV}$ [15]. A suppression factor of 0.5 for high p_T jets is observed in central Pb-Pb collision in comparison to the pp collisions. Our results also show surprisingly agreement with CMS preliminary results on jet R_{AA} .

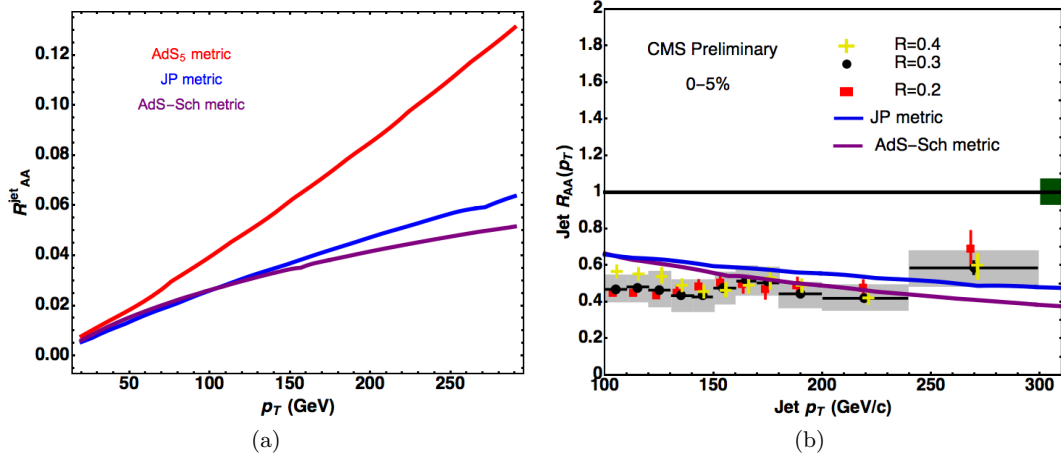


Figure 3: (a) Jet R_{AA} as a function of p_T in the most central Pb-Pb collision obtained via AdS/CFT in AdS_5 (red), JP (blue) and AdS-Sch (purple) metrics. (b) AdS/CFT jet R_{AA} as a function of p_T compared with the preliminary CMS data in different effective cone sizes for anti- k_T jets using the Bayesian unfolding method for most central Pb-Pb collision at the LHC with $\sqrt{s} = 2.76$ TeV per nucleon [43]. The results of our calculations in AdS-Sch and JP metrics are shown by the purple and blue curves, respectively. The vertical lines indicate uncorrelated statistical uncertainty, and the wide band the systematic uncertainty for Bayesian unfolding $R=0.3$. The green box above 300 GeV/c represents the overall combined uncertainty from T_{AA} and luminosities.

4. Conclusions

In this paper we have purposed a novel prescription of jet in the context of string theory and AdS/CFT correspondence. We have defined jet as a part of a falling string which lies above the radial coordinate u^* near the event horizon. Our motivation is defining the jets as hadrons. The initial energy loss of light quark is identified with the energy flux from the point at $u(\sigma, \partial_\sigma) = u^*$. We have shown that using the Δu prescription of jet the light quark energy loss exhibit the Bragg peak again at late times in both static and expanding plasmas. This late time behavior of jet energy loss implies that after traveling substantial distances through the plasma, the thermalization of light quark ends with a large amount of energy transferring to the plasma which is similar to the energy loss rate of a fast charge particle moving through ordinary matter.

We considered a brick of plasma and calculated the nuclear modification factor of jet in both AdS-Sch and JP metric. We assumed that the temperature of plasma around 350 MeV at AdS-Sch metric and at initial time in JP metric. Our results show an aver suppression of jet of order of ten respect to the data. We investigated that it is because of the falling string setup at AdS space. In fact, R_{AA} of jet using the falling string in AdS_5 which is dual to jets in vacuum is not one, even though it is less than one. We introduced a renormalized R_{AA}^{renorm} by dividing the R_{AA} in the medium to the R_{AA} in the vacuum. Surprisingly, our ratio shows good agreement with the experimental data on the R_{AA}^{jet} of most central Pb-Pb collision at LHC (fig. 3(a,b)). On the other hand, the light quark energy loss is highly depends on the initial conditions of falling string. The only way to determine the energy loss of a jet precisely in strong coupled regime is solving the gravitational bulk to boundary problem. One can solve Einstein equations for a perturbation on the 5d geometry map to the presence of the string and a conformal to the bulk to bound as the perturbation in the SYM energy behavior of the source by the presence of jet the perturbation in the SYM energy-momentum tensor by the presence of jet which will be left for the future work.

References

- [1] Adams J *et al.* (STAR Collaboration) 2005 *Nucl.Phys.* **A757** 102–183 (*Preprint nucl-ex/0501009*)
- [2] Adcox K *et al.* (PHENIX Collaboration) 2005 *Nucl.Phys.* **A757** 184–283 (*Preprint nucl-ex/0410003*)
- [3] Arsene I *et al.* (BRAHMS Collaboration) 2005 *Nucl.Phys.* **A757** 1–27 (*Preprint nucl-ex/0410020*)
- [4] Back B, Baker M, Ballintijn M, Barton D, Becker B *et al.* 2005 *Nucl.Phys.* **A757** 28–101 (*Preprint nucl-ex/0410022*)
- [5] Policastro G, Son D T and Starinets A O 2001 *Phys.Rev.Lett.* **87** 081601 (*Preprint hep-th/0104066*)
- [6] Kovtun P, Son D T and Starinets A O 2005 *Phys.Rev.Lett.* **94** 111601 (*Preprint hep-th/0405231*)
- [7] Maldacena J M 1998 *Adv.Theor.Math.Phys.* **2** 231–252 (*Preprint hep-th/9711200*)
- [8] Witten E 1998 *Adv.Theor.Math.Phys.* **2** 253–291 (*Preprint hep-th/9802150*)
- [9] Gubser S, Klebanov I R and Polyakov A M 1998 *Phys.Lett.* **B428** 105–114 (*Preprint hep-th/9802109*)
- [10] Aharony O, Gubser S S, Maldacena J M, Ooguri H and Oz Y 2000 *Phys.Rept.* **323** 183–386 (*Preprint hep-th/9905111*)
- [11] Casalderrey-Solana J, Liu H, Mateos D, Rajagopal K and Wiedemann U A 2011 (*Preprint 1101.0618*)
- [12] Yin Z B (ALICE Collaboration) 2013 *Acta Phys.Polon.Supp.* **6** 479–484
- [13] Aad G *et al.* (ATLAS Collaboration) 2010 *Phys.Rev.Lett.* **105** 252303 (*Preprint 1011.6182*)
- [14] Chatrchyan S *et al.* (CMS Collaboration) 2011 *Phys.Rev.* **C84** 024906 (*Preprint 1102.1957*)
- [15] Collaboration C (CMS Collaboration) 2012
- [16] Karch A and Katz E 2002 *JHEP* **0206** 043 (*Preprint hep-th/0205236*)
- [17] Herzog C, Karch A, Kovtun P, Kozcaz C and Yaffe L 2006 *JHEP* **0607** 013 (*Preprint hep-th/0605158*)
- [18] Chesler P M, Jensen K, Karch A and Yaffe L G 2009 *Phys.Rev.* **D79** 125015 (*Preprint 0810.1985*)
- [19] Janik R A and Peschanski R B 2006 *Phys.Rev.* **D73** 045013 (*Preprint hep-th/0512162*)
- [20] Horowitz W A 2010 (*Preprint 1011.4316*)

First principle study of structural, thermal and electronic properties of the chalcopyrites AlAgX_2 ($\text{X}=\text{S}, \text{Se}, \text{Te}$)

GM. Dongho Nguimdo, Daniel P. Joubert

School of Physics, University of the Witwatersrand, Private Bag 3, Johannesburg 2050, South Africa, tel.: +27 (0)11 717 6804, fax: +27 (0)11 717 6830.

E-mail: GuyMoise.DonghoNguimdo@students.wits.ac.za

Abstract. First principles density functional theory calculations of structural, thermal and electronic properties of the bulk chalcopyrites AlAgX_2 ($\text{X}=\text{S}, \text{Se}, \text{Te}$) were performed using the local density approximation and the modified Becke Johnson approximation. The optimized structure and lattice constants are obtained after a full relaxation of the structure while equilibrium volumes, bulk moduli and derivatives are extracted by fitting the Birch Murnaghan equation of state. For each compound, a systematic study of the density of states, the bandgaps and the band structure was carried out for the different approximations. We found that the modified Becke Johnson approximation gives the most accurate fundamental gap when compared to the experimental data. Phonon frequencies are used to predict the dynamical stability of the structure at the ground state. Thermal properties including free energy and heat capacity are also discussed.

1. Introduction

Increasing energy demand and the search for sustainable, environmental friendly and cheap sources of energy have been a challenge for the scientific community over the last decades. The ternary chalcopyrite semiconductors $\text{A}^{\text{I}}\text{B}^{\text{III}}\text{X}^{\text{VI}}$ ($\text{A}=\text{Ag}, \text{Cu}$; $\text{B}=\text{Al}, \text{Ga}, \text{In}$; $\text{X}=\text{O}, \text{S}, \text{Se}, \text{Te}$) are promising materials to investigate for meeting this challenge. They are also predicted to have potential uses in many other technology applications including low dimensional transistors and optoelectronic devices and photocatalytic splitting [1–3]. Several studies, both experimental and theoretical, have been carried out on these compounds. Hai *et al.* [4] showed from *ab initio* calculations that the bandgaps of CuInSe_2 and CuGaSe_2 are well predicted by the hybrid functional B3PW91 than any others hybrids functionals. Tsuyoshi *et al.* [5] studied the effect of Cu vacancies on the stability of CuXSe_2 ($\text{X}=\text{In}, \text{Ga}, \text{Al}$). Lekse *et al.* [6] investigated the effect of the solid state microwave synthesis methods on the purity of AgInSe_2 .

Our interest in the properties of the AlAgX_2 ($\text{X}=\text{S}, \text{Se}, \text{Te}$) stems from the fact that in these compounds, unlike to the other chalcopyrites families, properties such as vibrational and thermal behaviour have yet not been studied. Moreover, because of the well known problem of discontinuity of the functional derivative on the traditional functionals such as LDA and GGAs, previous studies led to a significant underestimation of the fundamental bandgap [7, 8].

To overcome this issue, more elaborate functionals have been developed including the LDA+U [9], the modified Becke-Johnson functional(MBJ) [10], the hybrid functional HSE [11],

and the GW approach [12]. In first principle calculations, a balance should be found between accuracy of the result and computational time. GW and HSE methods are usually accurate, but are computationally very expensive. Recently, Tran and Balha [10] have developed the MBJ potential which seems to be meeting both requirements and also for a large number of semiconductors and insulators. The MBJ potential is defined as:

$$v_{x,\sigma}^{\text{MBJ}}(\mathbf{r}) = cv_{x,\sigma}^{\text{BR}}(\mathbf{r}) + (3c - 2) \frac{1}{\pi} \sqrt{\frac{5}{12}} \sqrt{\frac{2t_{\sigma}(\mathbf{r})}{\rho_{\sigma}(\mathbf{r})}}, \quad (1)$$

where $v_{x,\sigma}^{\text{BR}}(\mathbf{r})$ is the Becke-Roussel potential, $\rho_{\sigma}(\mathbf{r})$ and $t_{\sigma}(\mathbf{r})$ are the electron and the kinetic energy density. The c parameter depends on the unit cell volume and some free parameters whose values are obtained according to a fit to experimental data.

In this paper, we use the MBJ potential to investigate the bandgaps and the density of states (DOS) of AlAgX₂. LDA bandgap is also included for comparison. Dynamical stability and thermal properties are also studied by means of phonon calculations. Our paper is organised as follows: In section 2, we outline the computational method. Section 3 is dedicated to results and discussion while the summary is given in Section 4.

2. Computational details

We employ the first principle DFT method as implemented in the VASP package [13] with inclusion of the plane wave Projector Augmented Wave PAW [14]. To determine the size of the plane wave basis set and \mathbf{k} -points mesh for sampling the Brillouin Zone (BZ), a series of convergence tests were performed in order to minimize the total energy. We choose a cut-off energy of 520 eV which leads to total energy tolerance of about 0.1 meV. For the integration over the BZ, a Γ -centered Monkhorst-pack grid [15] of $7 \times 7 \times 7$ was found sufficient. The LDA [16] and MBJ are used as exchange-correlation potential. The geometry optimization of the structure and lattice vibrational calculations are performed using LDA only. The linear response method within the density function perturbation theory (DFPT) [17] was used the calculation of phonon and thermal properties. The PHONOPY package [18] is used to extract information from DFPT calculations. Initial structures were obtained from Materials Project database [19, 20]. The high symmetry points for plotting the band structure and phonon dispersion curves were generated using the online version of the AFLOW software [21].

3. Results and discussions

3.1. Structural properties

As with other chalcopyrites, AlAgX₂ compounds crystallise in the zincblende structure with the space group $I\bar{4}2d$ (No. 122). In first principle calculation, the geometry optimization and ionic relaxation are performed through a self-consistent minimization of the forces and total energy. Usually, in the construction of the exchange-correlation potentials, the energy is approximated and the potential is then determined as a functional derivative of the energy functional, $V_{xc} = \left(\frac{\delta E_{xc}[n]}{\delta n} \right)$. The MBJ potential was originally constructed as an approximate potential itself [10]. In such cases, it is suggested to use the van Leeuwen-Baerends line integral [22] or the Levy-Perdew virial relation [23] to obtain the energy. Recently, Gaiduk *et al.* showed that the Becke-Jonsson potential does not satisfy those conditions and therefore is not a functional derivative [24]. Hence for the structural calculations, it is advised to use other functionals. We use the LDA functional for our calculations. Our results are summarized in Table 1 and compared with previous DFT calculations as well as experimental data. The lattice parameters are underestimated because of the well known problem of overbinding in LDA. Nevertheless, our results are within 5% of the experimental values. The slight difference

with other calculated results might come from the method used. Bulk modulus (B_0) and its derivative (B'_0) are extracted by fitting the energy per atom and the volume per atom to the Birch Murnaghan [25] equation. Results show that these compounds are not hard materials with AlAgTe₂ been the softest of the series. We could not find any previous theoretical or experimental values for B_0 and B'_0 . These results therefore may serve as a reference.

Compounds	Functional	a(Å)	c/a	V ₀ (Å ³)	B ₀ (GPa)	B' ₀
AlAgTe ₂	LDA	6.25	1.91	29.04	55.14	4.83
	LDA ^a	6.22	1.97	-	-	-
	Exp ^a	6.29	1.88	29.31	-	-
AlAgSe ₂	LDA	5.88	1.84	23.36	67.16	4.92
	LDA ^a	5.78	1.99	-	-	-
	Exp ^a	5.95	1.81	23.83	-	-
AlAgS ₂	LDA	5.61	1.82	20.07	80.5	5.65
	LDA ^a	5.48	1.99	-	-	-
	Exp ^a	5.72	1.77	20.86	-	-

^a = Ref. [7]

Table 1: Structural parameters of AlAgX₂(X=S,Se,Te) compared to the available experimental and theoretical data.

3.2. Electronic properties

Using the relaxed structures from previous calculations, we calculate the band structure and DOS using MBJ potential. The LDA calculations are also included for comparison. Bandstructures are depicted in Figure 1 and it can be noted that the overall features are similar. The main difference is on the size of the gap. The minimum of the conduction band (MCB) and the maximum of the valence band (MVB) both occur at the Γ point, indicating that these compounds are direct bandgap in the ground state. From Table 2, one can note that LDA dramatically underestimated the bandgap for all the three compounds. Such behaviour is common to LDA calculations. Previous studies have predicted bandgaps in the range of our LDA results (see Table 2). MBJ predicts bandgaps of 3.15 eV, 2.38 eV, and 2.14 eV for AlAgS₂, AlAgSe₂ and AlAgTe₂, respectively. These values are in good agreement with 3.14 eV, 2.55 eV, and 2.27 eV, respectively, obtained from experiment. It is expected that similar results would be obtained with more sophisticated techniques, namely, GW and HSE, but with a very long computational time. The advantage of MBJ is that it combines the accuracy and less computational time. Although the band structure gives information on the size and the nature of the bandgap, it is not obvious to state which orbital contribute to the formation of the bandgap. More details can be obtained from the Density of States (DOS).

Figure 2 displays the partial density of states (PDOS) of the compounds. The conduction band is mainly from Al-*p*, X-*p* and Ag-*s* states. It is worth to note that the small X-*p* peak which contributes to the formation of the edge of the MCB in AlAgS₂ tends to reduce in intensity and shifts to lower energy in AlAgSe₂ and does not exist on the AlAgTe₂ spectrum. It is rather the Al-*p* orbital which has the most significant contribution. This could be related to the bond lengths of the chalcogenide atoms with the transition metal atoms which tend to increase with atomic radius and hence, the volume per atom of AlAgTe₂ is the largest of the series. We believe such character to be the essence of the narrowing of the bandgap. The edge of the valence band is dominated by Ag-*d* and X-*p* orbitals and the strong hybridization between the two orbitals in AlAgS₂ reduces gradually in the two other compounds.

3.3. Dynamical stability and thermodynamic properties

Performing phonon calculation is usually a very delicate task because of the difficulty of calculation of the exact force constant acting on an atom. A reliable force constant is usually

Table 2: Bandgap E_g (eV) of AlAgX_2 compared to the available experimental and theoretical data. The calculations in Ref. [7] were performed using the tight- bonding linear muffin-tin orbital (TB-LTMO) method.

	AlAgTe_2				AlAgSe_2				AlAgS_2			
	LDA	LDA	MBJ	Exp.	LDA	LDA	MBJ	Exp.	LDA	LDA	MBJ	Exp.
E_g (eV)	1.12	1.36 ^a	2.14	2.27 ^a	1.07	1.59 ^a	2.38	2.55 ^a	1.83	1.98 ^a	3.15	3.13 ^a

^a = Ref. [7],

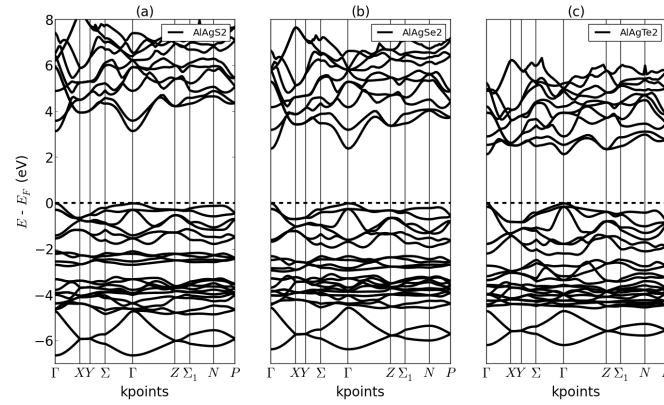


Figure 1: Band structure of AlAgX_2 calculated using MBJ. Both the minimum conduction and maximum valence band occur at the Γ point indicating a direct bandgap for all the compounds.

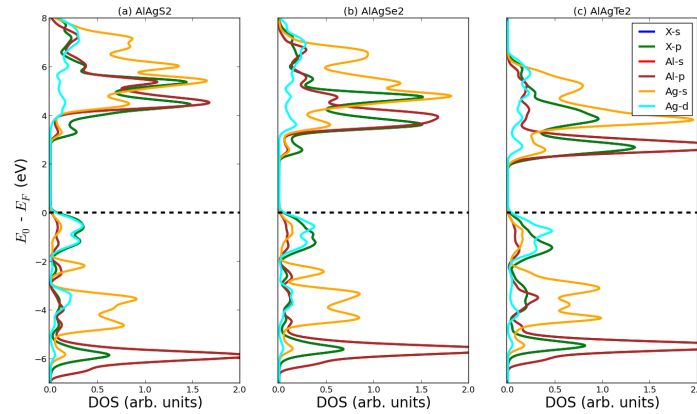


Figure 2: (Colour online) Density of states of AlAgX_2 . The Fermi level E_F is shifted to zero.

obtained by using a supercell since in the normal unit cell, interaction between the displaced atoms and their images is important. A supercell of 64 atoms and with a $2 \times 2 \times 2$ \mathbf{k} -points grid were used for our calculations. The calculations should also meet the total energy convergence criterion. Hence because of the non existence of MBJ exchange correlation energy as stated above, LDA is used as exchange correlation functional. A structure is predicted to be stable with respect to phonon when no vibration modes have imaginary real frequencies.. The phonon

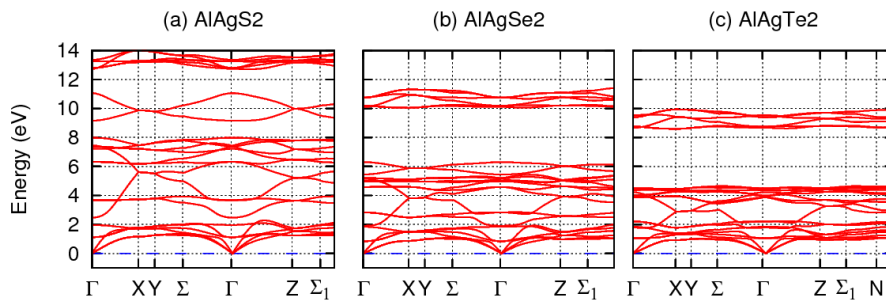


Figure 3: (Colour online) The phonon-dispersion curves of AlAgX_2 .

spectrum along the high symmetry points is depicted in Figure 3. A careful observation shows that apart from the acoustic modes going to zero frequencies around the Γ point, all the phonon-dispersion curves remain positive throughout the BZ attesting to the stability of the structures. Moreover, there is not a clear limit between the acoustic and the optical modes for all the three spectra with AlAgS_2 having the highest optical modes.

Phonons contribute to a range of thermodynamic properties including Helmholtz free energy $F(T)$, entropy $S(T)$ and heat capacity at constant volume $C_v(T)$. $F(T)$ and $C_v(T)$ are displayed in Fig.(4) up to 1000K. In this temperature range, it is assumed that phonon interactions are still small such that anharmonic interactions are not occurring yet. The Helmholtz free energy is defined as $F = U - TS$ with U the internal energy. Since entropy S is an increasing function of the temperature, F should decrease with increment of temperature. All the three compounds satisfy that condition. $F(T)$ does not go to zero at 0K indicating a zero point motion in the systems. $C_v(T)$ increases rapidly at lower energy [0 – 200K] as predicted in the *Debye model* [26]. From the room temperature $\sim 300\text{K}$, it converges to a limit close to 200J/K.mol . This asymptotic behaviour is in agreement with the *Dulong-Petit law* which states that $C_v(T)$ should tend to $3Rn$ at high temperature with $R = 8.31\text{J/K.mol}$ the gas constant and n the number of atoms in the unit cell [26]. Each of our unit cells has $n = 8$ atoms and leads to $C_v(T) \simeq 199, 44\text{J/K.mol}$ which corresponds to the limit obtained in Fig.(4-b).

Many factors contribute to the heat capacity and are of varying importance. For insulators and semiconductors, the principal contribution comes from lattice vibration at low temperature. By increasing the temperature, electrons get excited and their contribution cannot be neglected. The population of electrons in the conduction band starts increasing when the thermal energy $K_B T$ (K_B , the Boltzmann constant) approaches the energy gap ($K_B T \sim E_g$). The maximum temperature in our study was set at 500K which leads to a thermal energy of 0.042 eV. This value is lower than all the calculated bandgaps. We can conclude that up to 500K, the phonon contribution to C_v is still the most important for the compounds under investigation.

4. Conclusion

In this work, we have studied the structural, electronic and thermal properties of the chalcopyrites AlAgX_2 ($X=\text{S,Se,Te}$) using LDA and MBJ approximations. In spite of the overbonding in LDA, lattice parameters obtained from our calculations were in the range of experimental values. Band structures spectrum show that all the three compounds have a direct bandgap. By the mean of the MBJ potential, we obtained very accurate bandgaps compared to experimental data. The advantage of using the MBJ is that it is known to be as cheap as LDA and GGAs in term of computationally time. Similar agreement with experimental data is usually obtained with very computational expensive methods such as HSE and GW methods. The PDOS spectra show that the p orbitals from chalcogenide and Al atoms contribute significantly to the

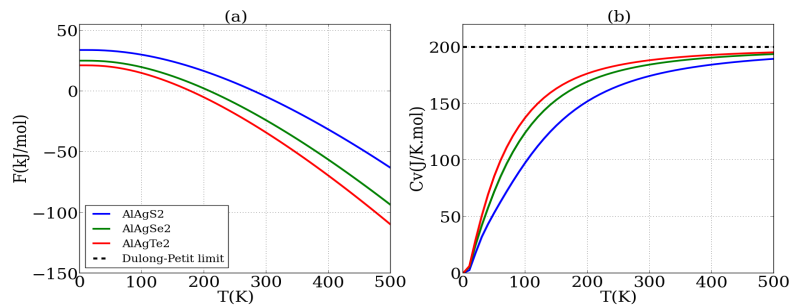


Figure 4: (Colour online) Free energy and Heat capacity of AlAgX_2 .

formation of the MCB and therefore are at the essence of the bandgap difference between the three compounds. Using the supercell methods with the DFPT, we study the dynamic stability of AlAgX_2 . Because of the absence of negative frequencies along the phonon-dispersion curves, AlAgX_2 are predicted to be dynamically stable at the ground state. Heat capacity follow the Debye model at low temperature and converges to the Dulong-Petit limit at high temperature.

Acknowledgements

This work was partially supported by DAAD-AIMS (African Institute for Mathematical Sciences) scholarship. The Centre for High Performance Computing (CHPC), Cape Town-South Africa is thanked for the computational resources.

References

- [1] Reshak A 2013 *Int. J. Electrochem. Sci.* **8** 9371–9383
- [2] Zunger A and Jaffe J 1983 *Phys. Rev. Lett.* **51** 662–665
- [3] Jaffe J and Zunger A 1984 *Phys. Rev. B* **29** 1882–1906
- [4] Xiao H, Tahir-Kheli J and Goddard III W A 2011 *J. Phys. Chem. Lett.* **2** 212–217
- [5] Maeda T and Wada T 2005 *J. Phys. Chem. Sol.* **66** 1924–1927
- [6] Lekse J W, Pischera A M and Aitken J A 2007 *Mater. Res. Bull.* **42** 395–403
- [7] Mishra S and Ganguli B 2011 *Solid State Comm.* **151** 523–528
- [8] Jayalakshmi V, Mageswari S and Palanivel B 2012 *Solid State Phys. Proc.* **1447** 1087–1088
- [9] Anisimov V I, Zaanen J and Andersen O K 1991 *Phys. Rev. B* **44** 943–954
- [10] Tran F and Blaha P 2009 *Phys. Rev. Lett.* **102** 226401
- [11] Heyd J, Scuseria G E and Ernzerhof M 2003 *J. Chem. Phys.* **118** 8207–8215
- [12] Shishkin M and Kresse G 2007 *Phys. Rev. B* **75** 235102
- [13] Kresse G and Furthmüller J 1996 *Phys. Rev. B* **54** 11169–11186
- [14] Kresse G and Joubert D 1999 *Phys. Rev. B* **59** 1758–1775
- [15] Monkhorst H J and Pack J D 1976 *Phys. Rev. B* **13** 5188–5192
- [16] Perdew J P and Wang Y 1992 *Phys. Rev. B* **45** 13244–13249
- [17] Baroni S, de Gironcoli S, Dal Corso A and Giannozzi P 2001 *Rev. Mod. Phys.* **73** 515
- [18] Togo A, Oba F and Tanaka I 2008 *Phys. Rev. B* **78** 134106
- [19] Belsky A, Hellenbrandt M, Karen V L and Luksch P 2002 *Acta Crystallogr., Sect. B: Struct. Sci* **58** 364–369
- [20] Jain A, Ong S P, Hautier G, Chen W, Richards W D, Dacek S, Cholia S, Gunter D, Skinner D, Ceder G and Persson K a 2013 *Appl. Phys. Lett. Mater.* **1** 011002
- [21] Curtarolo S, Setyawan W, Hart G L, Jahnatek M, Chepulskii R V, Taylor R H, Wang S, Xue J, Yang K, Levy O et al. 2012 *Comput. Mater. Sci.* **58** 218–226
- [22] van Leeuwen R and Baerends E J 1995 *Phys. Rev. A* **51** 170–178
- [23] Gross E K and Dreizler R M pp 11–31, 1995 *Density Functional Theory* (Springer)
- [24] Gaiduk A P and Staroverov V N 2009 *J. Chem. Phys.* **131** 044107
- [25] Birch F 1947 *Phys. Rev.* **71** 809–824
- [26] Kittel C and McEuen P 1986 *Introduction to solid state physics* vol 8 (Wiley New York)

Quantum Neural Networks - Our brain as a quantum computer?

Maria Schuld, Ilya Sinayskiy, Francesco Petruccione

Quantum Research Group, School of Chemistry and Physics, University of KwaZulu-Natal
Durban, KwaZulu-Natal, 4001, South Africa and National Institute for Theoretical Physics
(NITheP), KwaZulu-Natal, 4001, South Africa

E-mail: schuld@ukzn.ac.za

Abstract. Neuroscience and quantum physics have a central feature in common: both disciplines study objects that largely remain a mystery to scientists. While a century after the discovery of quantum theory, physicists still struggle to interpret their quantum objects' counterintuitive behaviour, biologists are not even close to understanding the mechanisms underlying the remarkable performance of our brain. The research field of quantum neural networks (QNNs) combines both 'mysteries' by investigating how established models of neural networks can be formulated in the language of quantum theory. Far away from the rather esoteric discourses of a 'quantum brain', QNN models first and foremost aim at developing efficient algorithms to run on future realisations of quantum computers. QNNs thereby promise to provide a substantial speed-up or increased memory capacity relative to classical neural networks. However, beyond questions of powerful computing technology, a success in the yet relatively small field of QNN research could give a first hint that our brain makes use of quantum mechanics to master its incredible tasks. In that sense, QNN research can be seen as a subfield of the 'dawn' of quantum biology which evaluates the question of how nature employs quantum effects in macroscopic (i.e. hot and dense) environments to optimise its processes.

1. Introduction

One of the most important scientific questions yet to answer is the problem of how the 'hardware' of our brain leads to its functioning in terms of thoughts, memory and consciousness. Some voices claim that an important ingredient of an explanation of 'how the mind emerges from matter' is quantum mechanics. On the more populist side, the discourse seems to be fuelled by the fascination of merging two scientific mysteries, namely the counterintuitive behaviour of the microscopic world and the black-box our most important organ still appears to be.¹ But also more established physicists argue in favour of a brain based on quantum mechanics as a potential avenue of solving the open problem of computation in the brain works. The most well-known is Sir Roger Penrose who, in collaboration with the anaesthesiologist Stuart Hameroff, located quantum computing in the microtubules or cytoskeleton of neural cells [1]. Another approach is the quantum brain model by Ricciardi and Umezawa [2] and further developed by Pessa and Vitiello [3] in which the states of neural networks are understood as collective modes using the formalism of Quantum Field Theory. The nonlocal properties of both quantum waves

¹ Popular science debates on a 'quantum brain' lead to journals of controversial scientific scope such as *QuantumNeurology*.

and consciousness gave rise to notions of a ‘quantum consciousness’ [4, 5]. However, solving the mind-matter problem through quantum physics is highly controversial [6, 7, 8].

Another meeting point between neuroscience and quantum physics are quantum neural networks (in short, QNNs). Conventional neural networks are simplified mathematical or computational models of the neural setup of our brain. QNNs are then models or devices that integrate quantum computing and neural networks as two promising paradigms of information processing in order to improve classical neural networks. Besides their potential computational power, QNNs contribute important arguments towards the debate on the role of quantum physics in the brain. If it can be shown that neural computation (which is widely assumed to be the basic mechanism of how the brain works) is dramatically improved by introducing quantum effects, we could further investigate if on a biological level, quantum neural computing can be observed. It turns out that QNN research feeds the arguments against quantum brain models: neural networks, characterised by nonlinear neural activation functions, and quantum systems based on a probabilistic description of linear operations, show very distinct mathematical structures. However, apart from their questionable biological explanation power, QNNs constitute an exciting research topic from a computational perspective. Increasing the performance of classical neural networks would have a significant effect on applications of machine learning and would extend the potential field in which future quantum computers could be useful.

This brief contribution intends to introduce into the theoretical foundations of QNNs (Section 2) and QNN research (Section 3) in order to reflect on the question if these have the potential to serve as realistic models of the brain. It comes to the conclusion that we have to see QNNs rather from a computational than a biological perspective, and that the challenges in merging quantum physics and neuroscience rather point towards the fact that quantum brain models are not that easy to derive (Section 4).

2. Information processing with neurons and quantum objects

To understand the basics of how neurons process information, we have to look at how they feed signals into one another through synaptic connections. Neurons are tube-like cells that transmit so called action potentials. An action potential is a localised depolarisation of the equilibrium membrane potential of usually $-70mV$ traveling along the neuron². A neuron transmitting an action potential is called ‘active’ or ‘firing’ as opposed to a ‘resting’ neuron.

Each of the approximately 10^{11} neurons in our brain has 1 to 10^4 synaptic connections to other neurons, making up for a total number of 10^{14} synapses [9]. When an action potential reaches a (chemical) synapse, neurotransmitters are released into synaptic cleft and open up ion channels in the membrane of the subsequent or post-synaptic neuron, so that a post-synaptic potential is created. In simple words, neurons produce signals in other neurons that depend on the synaptic strength with which they are connected. If all these incoming signals to a neuron exceed a certain threshold, the neuron produces an action potential and becomes active.

This synaptic activation mechanism can be translated into a simple mathematical model called *perceptron* and first introduced in [10]. It is based on binary neurons proposed by [11] in the 1940s. In a perceptron setup, N input neurons x_1, \dots, x_N with values of either 0 (resting) or 1 (active) feed into a neuron y with threshold θ_y , and the strength of the synaptic connections between x_i and y is simulated by a weight w_{iy} . Neuron y is activated (represented by setting it to 1) if the input signal from x_1, \dots, x_N multiplied by their respective weights exceeds the

² The resting potential of the membranes of neural cells is kept up by the relative ion concentration inside and outside the membrane permeable through voltage-guided ion channels.

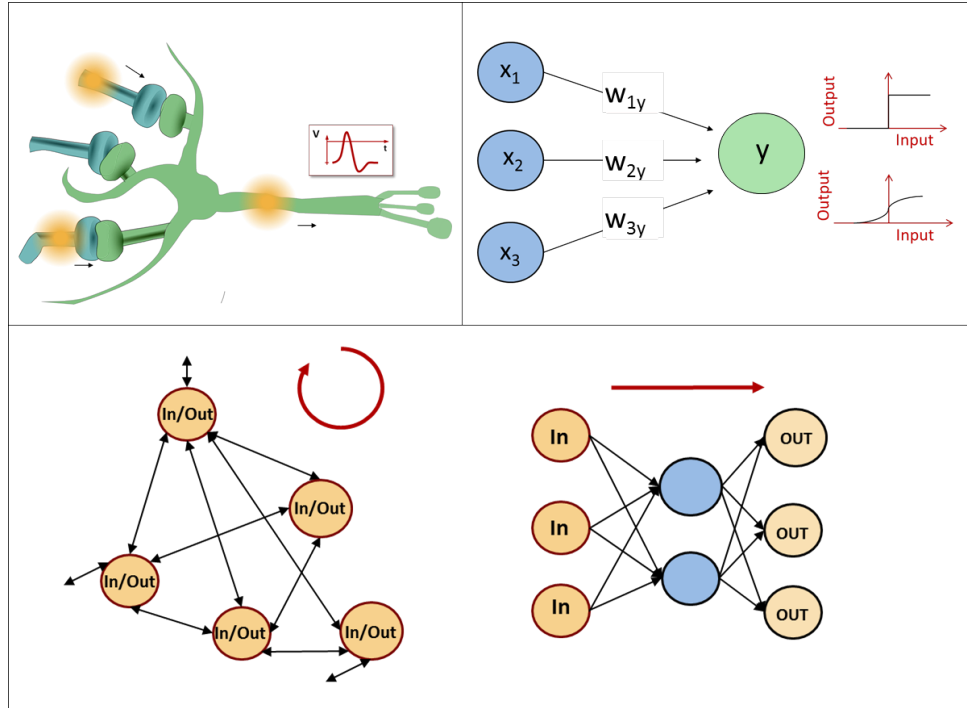


Figure 1. (Colour online) Neural computing derived from biological synapses. Top left: A neuron fires action potentials if the combined signals from neurons feeding into it through synaptic connections lead to an above threshold post-synaptic signal. Top right: Synaptic connections can be mathematically modelled by perceptrons with binary neurons $x_{1,2,3}$ connected to output neuron y through the connections weighted by $w_{1y,2y,3y}$ and a step or sigmoid activation function. Bottom: In neural networks, neurons are either recurrently connected (recurrent neural network on the left) or structured in subsequent layers (feed-forward neural network on the right).

threshold. The mathematical formulation is the activation or updating function

$$y = \begin{cases} 1, & \text{if } \sum_{i=1}^N w_{iy}x_i \geq \theta_y, \\ 0, & \text{else.} \end{cases}$$

This binary perceptron is often replaced by continuous versions, for example by choosing $x_i, y \in [-1, 1]$ and $y = \text{sgm}(\sum_{i=1}^N w_{iy}x_i + \theta_y)$, where sgm denotes the sigmoid function.

A neural network is a set of interconnected neurons whose dynamics are defined by the activation mechanism. After setting neurons to an initial value, they are successively updated in a given sequence and the output of the neural network can be read out of the final state of the neurons. In recurrent neural networks, all neurons take part in this process until the network states converges to a stable point, while in feed-forward neural networks the information processing goes through layers and the output is read out at a final set of neurons [12]. Neural networks are consequently like computers that convert an input signal into an output signal, and their success is based on the fact that through adjusting the weights, neural networks can learn a input-output mapping just as our brain does.

If we want to design a quantum neural network derived from biological foundations, we need to find a way to introduce quantum effects in the perceptron updating mechanism. What do

we mean by quantum effects? Quantum theory is a mathematical framework used by physicists to describe the behaviour of very small particles under isolated conditions. The physical laws on small scales are thereby very different from Newton's classical mechanics. Our description of atoms, electrons or photons for example assigns a wavefunction to a particle, that can be modelled by a vector with the notation $|\psi\rangle$ in a special vector space called Hilbert space. The wavefunction contains information on the probability of the particle to be in a certain state (e.g. position, momentum or energy). The central fact is that this probabilistic description does not correspond to our lack of knowledge on the system, but derives from the fact that the particle is in general in a so called superposition of different states, sometimes referred to as being in 'all states at the same time'. The interaction with a macroscopic environment such as a measurement then picks one of these states with the given probability and 'collapses' the particle superposition (which is why we could never observe quantum effects directly on our macroscopic scales).

This remarkable property of the microscopic world can be exploited for information processing, for example through a quantum system that encodes binary information like a computer, but is able to 'do calculations' on a superposition of all possible bit strings at the same time, and consequently retrieves a result by measuring the system. It is not easy to invent algorithms on such a quantum computer, but since two decades a number of powerful quantum routines are known to outperform classical computers, and many scientists think that it is only a question of time until quantum computers will become reality. The art of developing a quantum neural network which draws on neural computing is to use insights from quantum information theory in order to improve the performance of neural networks. We will briefly sketch the scope and results of QNN research before we discuss the central question of this article, namely what QNNs can tell us about the feasibility of a quantum brain.

3. Quantum neural network research

The development of a quantum neural network first and foremost aims at improving the computational efficiency of neural networks through the introduction of quantum effects. Neural networks are powerful devices with important application in tasks of machine learning [13, 14, 12], but they can be very costly in terms of computational resources. This is where the quantum speed-up is supposed to help.

The basic idea of introducing quantum properties into classical NNs is to replace the McCulloch-Pitts neuron $x = \{0, 1\}$ by a qubit $|x\rangle$ of the two-dimensional Hilbert space \mathcal{H}^2 with basis $\{|0\rangle, |1\rangle\}$. The state of a network with N neurons thus becomes a quantum product state of the 2^N -dimensional Hilbert space

$$|\psi\rangle = |x_1\rangle \otimes |x_2\rangle \otimes \dots \otimes |x_N\rangle = |x_1 x_2 \dots x_N\rangle \in \mathcal{H}^{2^N} = \underbrace{\mathcal{H}^2 \otimes \dots \otimes \mathcal{H}^2}_{N \text{ times}}.$$

Apart from the 'qubit neuron' (or 'quron'), proposals for QNN models vary strongly in their proximity to the basic idea of neural networks. Some try to directly translate the activation mechanism into quantum mechanics [15, 16], a task which is nontrivial because of the structural differences between the mathematical formalism of quantum computing and neural computing. In a more liberal approach, researchers introduce a hypothetical quantum evolution that corresponds to the nonlinear function, the so called dissipative operator D [17], but they fail to find a possibility to create such an operator. Others implement a quantum neural network through interacting quantum dots [18], a task that requires advanced technologies in controlling the interaction strengths, especially if the systems are scaled up from mere proof-of-principal examples.

An important approach is also the development of a quantum associative memory (QAM) which attempts to simulate the functioning of a neural network without considering the neuroscientific

basics [19, 20]. In our previous research, we followed this idea and used so called stochastic quantum walks in order to impose dynamics on an abstract quantum system that reproduces the basic features of an associative memory [21]. Quantum walks are equivalents to classical random walks which describe a stochastic process in which a walker jumps between nodes of a graph with predefined probabilities. We were able to show that a specific quantum version of these processes together with a certain construction of the underlying graph can lead the walker from a node representing an initial firing state of a neural network to a node representing the desired output. The quantum nature of the walk thereby led to a slight speed-up in a specific parameter range.

Altogether, QNN research is still in its infancy and a successful QNN model based on biological foundations is still outstanding. However, the efforts are part of the new emerging field of quantum machine learning, and might gain more importance if one day quantum computers are accessible for real applications.

4. Conclusion: Is the brain a quantum computer?

As mentioned above, it is difficult to combine quantum theory and neural computing. Neural networks rely on nonlinear activation functions, which in the case of the threshold function works as a switch to activate a neuron through an incoming signal. On the other hand, quantum theory is a probabilistic description and formulated through linear transformations of amplitudes that carry information on the probability of measuring a certain state. Even if we could understand neurons as quantum objects, neural quantum computing seems to fail with this incompatibility. As formulated in [17], we would require a dissipative quantum operator D that imitates the threshold activation function, but up to today there are no proposals known to the authors of how such a transformation could look like. Other ideas that translate neural networks into quantum mechanics while preserving the biologically observed activation mechanism fail to incorporate compatible (unitary) process of learning [15].

The difficulty in creating a powerful quantum version of neural computation supports arguments against simple notions of our brain to be a ‘quantum computer’ (as it was previously framed in a *Nature* contribution [22]). This adds to other important points made regarding the possibility of observing quantum effects in biological systems like the brain. First, the brain is a ‘hot and messy’ sphere of high temperatures and high particle densities. These properties destroy quantum coherence thereby rendering quantum effects irrelevant on observable timescales. Tegmark in fact estimated the decoherence time for collisions between the roughly 10^6 ions involved in the process of generating an action potential at a membrane site to be $\tau = 10^{-20}s$ [8]. Since this time scale is much smaller than the $10^{-3}s$ of a firing event in fast neurons, he concludes that “the computations in the brain appear to be of a classical rather than quantum nature,” [8, 12, italics left away]. Litt et al. add that the brain also lacks a mechanism for the complex task of quantum error correction [23]. Second, from what we know today, information processing in the brain is executed through the above described transmission of action potentials along neurons and their connections [24]. To explain how the brain performs quantum computing based on neuroscientific insights, we would have to introduce the entire (macroscopic) neuron as a quasi-particle and its complex firing process as a quantum state. This seems to be far away from quantum objects studied up to today.

As a conclusion, we can summarize that the challenge to express the neural updating function within the framework of quantum theory might be taken as an additional argument against the hypothesis that our brain is a quantum computer, adding towards the decoherence concerns put forward by various authors. Despite this fact, QNNs provide an exciting research field focusing on the potential of quantum computing to increase the performance of neural networks. These attempts are not confined to approaches preserving biological features of neural computing but can use the entire toolbox of quantum computing to develop a powerful QNN model that is

applicable for central machine learning tasks.

References

- [1] Hameroff S 1998 *Philos T Roy Soc A*, **356** 1743 1869
- [2] Ricciardi LM and Umezawa H 1967 *Kybernetik*, **4** 2 44
- [3] Eliano Pessa and Giuseppe Vitiello 1999 *Bioelectroch Bioener*, **48** 2 339
- [4] Loewenstein W 2013 *Physics in mind. A quantum view of the brain*, Basic Books
- [5] Stapp HP 2004 *Mind, matter, and quantum mechanics* Springer
- [6] Hagan S, Hameroff SR, and Tuszyński JA 2002 *Phys Rev E*, **65** 6 061901
- [7] McKemmish LK, Reimers JR, McKenzie RH, Mark AE, and Hush NS 2009 *Phys Rev E*, **80** 2 021912
- [8] Tegmark M 2000 *Phys Rev E*, **61** 4 4194
- [9] Purves D 2008 *Principles of cognitive neuroscience*. Sinauer Associates Inc
- [10] Rosenblatt F *Psychol Rev*, **65** 6 386
- [11] McCulloch WS and Pitts W 1943 *B Math Bio*, **5** 4 115
- [12] Hertz JA, Krogh AS, and Palmer RG 1991 *Introduction to the theory of neural computation* Westview press
- [13] Alpaydin E 2004 *Introduction to machine learning*, MIT press
- [14] Bishop CM 2006 *Pattern recognition and machine learning*, Springer
- [15] Altaisky MV 2001 *arXiv preprint quant-ph/0107012*
- [16] Fei L and Baoyu Z 2003 *ICNNSP '03 Proceedings* IEEE 539
- [17] Gupta S and Zia RKP 2001 *J Comput Syst Sci*, **63** 3 355
- [18] Behrman EC, Steck JE, and Skinner SR 1999 *IJCNN '99*. IEEE **2** 874
- [19] Ventura D and Martinez T 2000 *Inform Sciences*, **124** 273
- [20] Trugenberger CA 2001 *Phys Rev Let*, **87** 6 067901
- [21] Schuld M, Sinayskiy I, and Petruccione F 2014 *Phys Rev A*, **89** 3 032333
- [22] Koch C and Hepp K 2006 *Nature*, **440** 7084 611
- [23] Litt A, Eliasmith C, Kroon FW, Weinstein S, and Thagard P 2006 *Cognitive Sci*, **30** 3 593
- [24] Dayan P and Abbott LF 2001 *Theoretical neuroscience* MIT Press Cambridge

The stochastic Schrödinger equation approach to open quantum systems

I Semina¹ and F Petruccione^{1,2}

¹Quantum Research Group, School of Chemistry and Physics, University of KwaZulu-Natal, Durban, 4001, South Africa

²National Institute for Theoretical Physics, University of KwaZulu-Natal, Durban, 4001, South Africa

E-mail: iuliia.semina@mail.ru

Abstract. Under the system-environment interaction the dynamics of the open system can be described by a set of quantum trajectories satisfying some stochastic Schrödinger equation. Such an approach can be extended to the non-Markovian regime by replacing white noise with colored noise. This approach is efficient for simulations with the help of the time-discrete stochastic wave-function methods. As an illustrative example we consider the model of a dissipative qubit. The adaptive Platen method has been derived in order to introduce the colored noise to the iteration algorithm. Also, we test the validity of the approximation based on the Nakajima-Zwanzig method in the stochastic sense. This method is not well studied and needs further investigations.

1. Introduction

A first aim of the theory of open quantum systems is the description of the time evolution of a system S (the open system) interacting with an external environment E . One of the ways to describe the partial dynamics of such a system is to use the generalized master equation for the reduced density matrix $\eta(t)$. In this situation the simple approach is based on the absence of memory effects of the environment and is provided by the Markov approximation. Nevertheless, this approach is no more valid when the memory effects can not be excluded: strong coupling, correlation, and entanglement in the initial S - E state and the system at low temperature. This gives rise to the theory of non-Markovian quantum dynamics, for which a general theory does not exist, but only approaches, for example [1, 2, 3].

Under the system-environment interaction the dynamics of the system can be described by a set of quantum trajectories satisfying some stochastic Schrödinger equation. In this case, the density matrix of the system is recovered as an average over all possible number of trajectories of the state vector $\mathbb{E}[|\psi(t)\rangle\langle\psi(t)|] = \eta(t)$. Such an approach can be extended to the non-Markovian regime by replacing white noise with colored noise [3]. Specifically, the colored noise is represented by an Ornstein-Uhlenbeck process.

We describe the numerical investigation of the dynamics of a non-Markovian dissipative qubit, studied analytically in [3]. The stochastic simulations were done with the help of the extended Platen method.

The results of the simulations are compared with an analytical approximation firstly presented in [4] which is based on the Nakajima-Zwanzig projection method but conceptually is different

due to the appearance of the stochastic terms in the non-Markovian master equation. This method is not well studied and needs further investigations.

2. The linear SSE with colored noise and the closed stochastic master equation

In this section we will show two methods, presented recently in [3] and in [4]. Both the methods describe the non-Markovian dynamics. One method is based on the introduction of memory effects with the help of the colored noise and can be used for the simulations. Another method is an analytical approach, that is based on the Nakajima-Zwanzig method utilized to get a closed stochastic master equation. Furthermore, the simulation results will be compared with this analytical approach.

Let us start from the method, presented in [3]. A generic homogeneous linear stochastic differential equation for the non-normalized state $\phi(t)$ with $W = \{W_j(t), t \geq 0, j = 1, \dots, d\}$, a continuous d -dimensional Wiener process:

$$d\phi(t) = K(t)\phi(t) dt + \sum_{j=1}^d R_j(t)\phi(t) dW_j(t), \quad (1)$$

where $\phi(0) = \psi_0$, $\psi_0 \in \mathcal{H}$, the coefficients $R_j(t), K(t)$ are (non-random) linear operators on separable, complex Hilbert space \mathcal{H} . The normalized vector $\psi(t) = \phi(t) / \|\phi(t)\|$ corresponds to the conditional state of the system given the observed output up to time t and is often called the *a posteriori state*. For the case of measurement in continuous time the output is not discrete, but it is a whole trajectory of some observed quantity; this brings into play the stochastic processes. Apart from this complication, the linear stochastic Schrödinger equation is an evolution equation for the non-normalized vectors $\phi(t)$.

The stochastic differential equation (1) is to be intended in integral sense and the solution ϕ is the continuous, adapted Itô process [5] satisfying

$$\phi(t) = \psi_0 + \int_0^t K(s)\phi(s) ds + \sum_{j=1}^d \int_0^t R_j(s)\phi(s) dW_j(s).$$

The last term in the above equation is a stochastic Itô integral (see [5]). The coefficient in the drift part of (1) has the structure:

$$K(t) = -iH(t) - \frac{1}{2} \sum_{j=1}^d R_j(t)^\dagger R_j(t), \quad (2)$$

where H is the effective Hamiltonian of the system.

Finally, the *linear stochastic Schrödinger equation* (diffusive type) [6] is given by

$$d\phi(t) = \left(-iH(t) - \frac{1}{2} \sum_{j=1}^d R_j(t)^\dagger R_j(t) \right) \phi(t) dt + \sum_{j=1}^d R_j(t)\phi(t) dW_j(t), \quad (3)$$

$$\phi(0) = \psi_0, \quad \psi_0 \in \mathcal{H}, \quad \|\psi_0\| = 1, \quad H(t) = H(t)^\dagger. \quad (4)$$

The extension of the Markovian approach allows to describe the random dynamics for open quantum system with memory by introducing colored noise in the linear stochastic Schrödinger equation. In our case the model represents a dissipative evolution with memory, but without any observation of the quantum system.

The colored stationary Ornstein-Uhlenbeck (O-U) process $X(t)$ is defined by

$$X(t) = e^{-kt} \frac{Z}{\sqrt{2k}} + \int_0^t e^{-k(t-s)} dW(s), \quad k > 0, \quad 0 \leq s < t < +\infty, \quad (5)$$

and satisfies the stochastic differential equations:

$$dX(t) = -kX(t)dt + dW(t), \quad X(0) = Z/\sqrt{2k}, \quad (6)$$

where $W(t)$ is a one-dimensional Wiener process and Z is a standard normal random variable. We note, that Z is independent from the Wiener process. The non-Markovianity of the O-U process follows from the fact that its correlation function is no more a δ -function (the Markovian regime is recovered in the limit $k \downarrow 0$):

$$\mathbb{E}[\dot{X}(t)\dot{X}(s)] = \delta(t-s) - \frac{k}{2}e^{-k|t-s|}. \quad (7)$$

Let us consider a one-dimensional driving noise $X(t)$ and three non-random operators C , D and R on \mathcal{H} . The starting point is the basic linear stochastic Schrödinger equation

$$d\phi(t) = (A + BX(t))\phi(t)dt + R\phi(t)dX(t), \quad (8)$$

where $X(t)$ is the stationary O-U process. This can be rewritten by changing dX according to its definition:

$$d\phi(t) = (A + X(t)B - kX(t)R)\phi(t)dt + R\phi(t)dW(t), \quad (9)$$

and the initial condition is a wave function $\psi_0 \in \mathcal{H}$, such that $\|\psi_0\|^2 = 1$. To perform the normalisation condition for the probability $\mathbb{E}[\|\phi(t)\|^2] = 1$ we need to impose two self-adjoint operators K and H_0 such that

$$B = -iK + \frac{k}{2}(R + R^\dagger), \quad A = -iH_0 - \frac{1}{2}R^\dagger R. \quad (10)$$

As a consequence the initial Eq. (8) becomes

$$d\phi(t) = \left(-iH(t) - \frac{1}{2}R^\dagger R\right)\phi(t)dt + R\phi(t)dW(t), \quad (11a)$$

$$H(t) := H_0 + X(t)L, \quad L := K + \frac{ik}{2}(R^\dagger - R). \quad (11b)$$

Apart from the randomness introduced by the random Hamiltonian H with colored noise $X(t)$ we have the same situation of the linear (3).

For the case of a random unitary evolution we take $B = 0$ in Eq. (8), which gives also $K = 0$. Moreover, the conditions (10) become

$$R = -iV, \quad V^\dagger = V, \quad A = -iH_0 - \frac{1}{2}V^2.$$

Then, we get

$$L = -kV, \quad H(t) = H_0 - kX(t)V. \quad (12)$$

As it was demonstrated in [7], the linear stochastic Schrödinger equation reduces to :

$$d\phi(t) = -i\left[(H_0 - kX(t)V)dt + VdW(t)\right]\phi(t) - \frac{1}{2}V^2\phi(t)dt. \quad (13)$$

The evolution of the quantum system is completely determined by the time-dependent, random Hamiltonian $H(t)$ of an isolated and closed system incorporating a random environment characterized in terms of O-U process.

Sometimes, it is convenient to study the state corresponding to the statistical operator $\eta(t)$ which is the state we attribute to the system at time t , when the output is not known and this is the second method, derived in [4]. The *a priori* state that corresponds to the master equation (which is not closed) is:

$$\eta(t) = \mathbb{E}[\psi(t)\langle\psi(t)|], \quad (14)$$

where \mathbb{E} corresponds to the mean value and $\psi(t) = \phi(t)/\|\phi(t)\|$.

Using the definition (14) and (13) the stochastic master equation is obtained. The closed stochastic master equation has been derived by the adapting Nakajima-Zwanzig method (for full derivation see [4, 7]):

$$\dot{\eta}(t) \simeq \mathcal{L}_0[\eta(t)] + \frac{k}{2} \int_0^t \left[V, e^{(\mathcal{L}_0 - k)(t-s)} [[V, \eta(s)]] \right] ds. \quad (15)$$

The methods presented above will be considered for the model of the dissipative qubit. Also, in the next section it will be shown, how to introduce the colored noise to the simulation algorithm.

3. The simulation

The description of open quantum systems by using stochastic wave-function methods has recently received a great deal of attention. By using the wave function instead of the density matrix, one can significantly speed up computer simulations as the dimension of the system increases [8].

Let us consider the two presented methods on the concrete example, namely the dissipative qubit. For this model we have the following parameters:

$$H_0 = \frac{\omega_0}{2} \sigma_z, \quad \omega_0 > 0, \quad V = \sqrt{\frac{\gamma}{2}} \sigma_y, \quad \gamma > 0. \quad (16)$$

The linear stochastic Schrödinger equation (13) then becomes:

$$\begin{cases} d\phi_1(t) = -\frac{1}{2} \left(\frac{\gamma}{2} + i\omega_0 \right) \phi_1(t) dt - \sqrt{\frac{\gamma}{2}} \phi_2(t) dX(t), \\ d\phi_2(t) = -\frac{1}{2} \left(\frac{\gamma}{2} - i\omega_0 \right) \phi_2(t) dt + \sqrt{\frac{\gamma}{2}} \phi_1(t) dX(t), \end{cases} \quad (17)$$

and the operator \mathcal{L}_0 :

$$\mathcal{L}_0[\sigma] = -\frac{i\omega_0}{2} [\sigma_z, \sigma] - \frac{\gamma}{2} [\sigma_y, [\sigma_y, \sigma]].$$

We use the Platen scheme [9] for which it will be convenient to rewrite the system (17) in the matrix Itô form:

$$d\phi = Q\phi dt + C\phi dW, \quad (18)$$

where the corresponding support matrices are:

$$Q = \begin{pmatrix} -1/2(\gamma/2 + i\omega_0) & \sqrt{\gamma/2}kx & 0 \\ -\sqrt{\gamma/2}kx & -1/2(\gamma/2 - i\omega_0) & 0 \\ 0 & 0 & -k \end{pmatrix}, \quad C = \begin{pmatrix} 0 & -\sqrt{\gamma/2} & 0 \\ \sqrt{\gamma/2} & 0 & 0 \\ 0 & 0 & 1/x \end{pmatrix}, \quad (19)$$

and the vector ϕ is

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ x \end{pmatrix}. \quad (20)$$

Here the main difference from the classical Platen scheme is the presence of an O-U process with an extra variable x that have been added to the vector ϕ to provide an element of the colored noise.

The Platen scheme has the form:

$$\begin{aligned} \phi_{n+1} &= \phi_n + \frac{1}{2}(\mathbf{a}(\Upsilon) + \mathbf{a}) \Delta \\ &\quad + \frac{1}{4}(\mathbf{b}(R_+) + \mathbf{b}(R_-) + 2\mathbf{b})\Delta W \\ &\quad + \frac{1}{4}(\mathbf{b}(R_+) - \mathbf{b}(R_-))\{\Delta W^2 - \Delta\}, \end{aligned} \quad (21)$$

with corresponding:

$$\begin{aligned} \mathbf{a}(\phi) &= Q\phi, & \mathbf{b}(\phi) &= C\phi, \\ \Upsilon &= \phi_n + \mathbf{a}\Delta t + \mathbf{b}\Delta W, & R_{\pm} &= \phi_n + \mathbf{a}\Delta \pm \mathbf{b}\sqrt{\Delta}. \end{aligned}$$

Finally, we can extrapolate the results to provide a higher order approximation of the resulting functional. The order 4.0 weak extrapolation [5] has the form:

$$V_{g,4}^{\Delta}(T) = \frac{1}{21} \left[32E(g(Y_T^{\Delta})) - 12E(g(Y_T^{2\Delta})) + E(g(Y_T^{4\Delta})) \right],$$

where E denotes the mean value of the simulated function g with the time steps Δ , 2Δ and 4Δ correspondingly (for our case $\Delta = 0.05$).

The solutions of (15) for different values of k are shown in figure 1. One can see that the memory effects (increasing k) slow down the decay. In figure 2 and figure 3 the analytical approximation (15) is presented together with the simulation of (18). Each simulation was done for 10^4 realisations with the help of the Platen method adapted for the presence of O-U noise, with an additional term in (20), while the initial state was taken as:

$$\eta(0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \quad (22)$$

The choice of parameters for figure 2 and figure 3 has been optimized for the simplest case of equation (15).

4. Conclusion

One of the ways to describe the dynamics of an open quantum systems is to use the stochastic Schrödinger equations. The properties of this approach leads to effective computer simulations with the help of stochastic wave-function methods. These methods may also be used to describe systems with memory.

One of the strategies to describe memory effects in the system is to introduce the colored noise. In particular, the colored O-U noise has been added to the linear stochastic Schrödinger equation. This procedure was presented in [3] and was used to perform the simulations. To get the correct results, we adapt the Platen algorithm, with an additional term in (20), due to the presence of the colored noise.

Also, we have tested some approximation, derived in [4]. It is based on the Nakajima-Zwanzig technique but in connection to the stochastic equations. The final master equation was shown and its solution compared with the results of the simulations. The obtained curves show a good agreement for both methods.

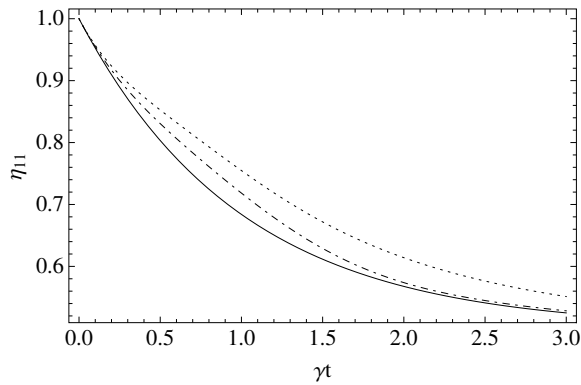


Figure 1. The solutions of (15) for different values of k : $k = 0$ (solid line), $k = 1$ (dot-dashed line), $k = 2$ (dotted line).

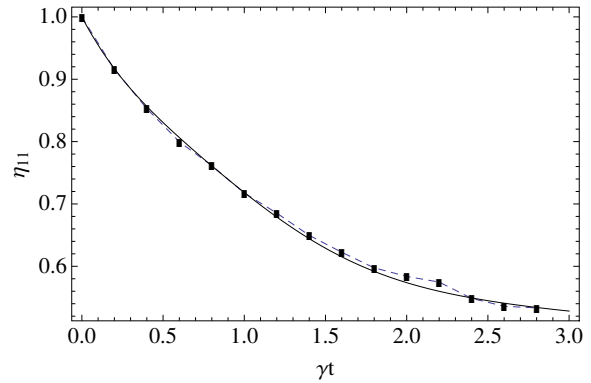
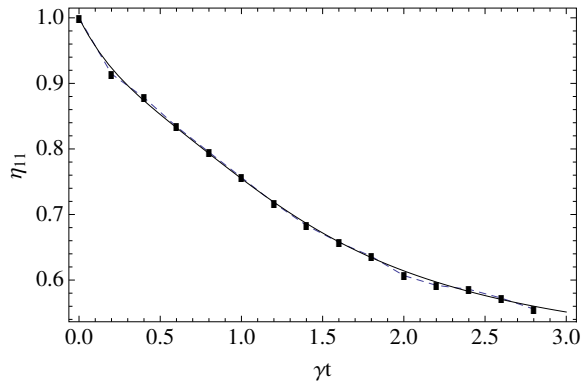


Figure 2. Plot of the mean occupation number of the excited state for the parameters $\gamma = 1, \omega_0 = \sqrt{37}/2, k = 1$. The solid line comes from the analytical approximation (15), while the dashed line and dots show the results of the stochastic simulation of (18).

Figure 3. Plot of the mean occupation number of the excited state for the parameters $\gamma = 1, \omega_0 = \sqrt{37}/2, k = 2$. The solid line comes from the analytical approximation (15), while the dashed line and dots show the results of the stochastic simulation of (18).

References

- [1] Barchielli A, Pellegrini C and Petruccione F 2012 Phys. Rev. A **86** 063814
- [2] W.T.Strunz, L. Diósi, N.Gisin, Open system dynamics with non-Markovian quantum trajectories, Phys. Rev. Lett. **82**, 1801 (1999).
- [3] Barchielli A, Pellegrini C and Petruccione F 2010 Euro Phys. Lett. **91** 24001.
- [4] Barchielli A, Di Tella P, Pellegrini C and Petruccione F 2011 *Stochastic Schrödinger equations and memory*, ed Rebolledo R and Orszag M, *Quantum Probability and Related Topics*, QP-PQ series: Quantum Probability and White Noise Analysis vol 27 (Singapore: World Scientific) pp. 52-67
- [5] Kloeden P and Platen E 1992 *Numerical Solution of Stochastic Differential Equations* (Heidelberg: Springer-Verlag)
- [6] Barchielli A and Gregoratti M 2009 *Quantum Trajectories and Measurements in Continuous Time: The Diffusive Case* (Berlin: Springer) Lect. Notes Phys. **782**
- [7] Semina I, Semin V, Petruccione F and Barchielli A 2014 Open Syst. Inf. Dyn. vol 21 nos. 1 & 2 1440008
- [8] H. Breuer, U. Dorner and F. Petruccione, Comput. Phys. Commun. **132**, (2000) 30.
- [9] Breuer H-P and Petruccione F 2002 *The Theory of Open Quantum Systems* (Oxford: Oxford University Press)

Molecular dynamics studies of some carbon nanotubes chiral structures

Moshibudi Shai¹, Thuto Mosuang¹, and Erasmus Rammutla¹

¹ Department of Physics & Geology, University of Limpopo, Private Bag x1106, Sovenga, 0727, Polokwane, South Africa

Corresponding author: thuto.mosuang@ul.ac.za

Abstract. Structural and equilibrium properties of armchair (cnt(12,12)) and two chiral (cnt(12,10) and cnt(10,12)) carbon nanotubes are studied using classical molecular dynamics. The formulation uses the Tersoff potential under the NVT ensemble to study these properties. Structural properties are studied using the radial distribution and structure factor functions. The equilibrium properties are studied using the total energy against lattice parameter variation. Similarities and differences in cnt(12,12), cnt(12,10), and cnt(10,12) symmetries are discussed.

1. Introduction

Carbon materials are widely used in various industries because of their exclusive mechanical properties [1]. It is now a known detail that the covalent bonding between carbon-carbon atoms contribute heavily to the strength of tetrahedral diamond bonds and to the in-plane hexagonal bonds. Therefore, carbon nanotubes, which are cylindrical hexagonal arrangements of carbon atoms along one given axis, represent the ideal carbon fibre, which should automatically hold the best mechanical properties [1-4]. This characteristic feature is crucial in the application of nanotubes, given the importance of strong light weight composites. Theoretical calculations have predicted the Bulk modulus B of single walled carbon nanotubes to be in the range of 38 to 191 GPa [5,6] disregarding the chirality. These calculations are based on the bulk modulus of graphite of which experimental measurements give 27-42 GPa [7,8]. Due to minute size effects, most of the experimental works have eluded measurements of the bulk modulus in multi-walled carbon nanotubes (mwcnt). But Young's modulus measurements, which can be described as the stiffness of the mwcnt in one dimension, were performed by Treacy *et al.* [9] to be in the range of ~2.0 TPa. Furthermore, atomic force microscopy techniques were used by Salvétat *et al.* [10] to obtain a Young's modulus of ~1.2 TPa. Measurements of single walled carbon nanotubes (swcnt) can pose a serious challenge due to their small diameters (~1.5 nm) and the fact that swcnt tend to form bundles. In so saying, measuring bundle nanotubes properties may not easily be projected to unit nanotubes and vice versa.

Various theoretical calculations of some small radii carbon nanotubes have focused on their electronic properties. First principles local density functional [11] and empirical tight binding approximations [12] predict that these nanoscale fibres will show conducting properties varying from metallic to semiconducting depending on the radii and chiral symmetry. Apart from these electronic and elastic properties, very little has been reported about their lattice and formation energies. It is such information which could facilitate optimal conditions essential for producing carbon nanotubes with high strength-to-weight ratios [13].

In this study, classical molecular dynamics (MD) is used to investigate the structural and equilibrium properties of swcnt. Structural configurations are extracted from the radial distribution functions and structure factors of the nanotubes. This is worked out at various temperatures in order to observe the flexibility of the potential used. Equilibrium properties are studied through structural optimizations along the a -axis using the Evans NVT ensemble at 300 K temperature and zero pressure. Here total energy is observed with little changes in the a -axis values at constant temperature. The implications of our results are then discussed with reference to earlier experiments and calculations and the accepted graphite and other nanotube results.

2. Computational aspects

Classical MD approach is used to solve the Newton's equations of motion for three types of swcnts discussed in this paper [14]. This MD simulation is performed under the Evans NVT ensemble for both structural and equilibrium properties. The Newtonian equations of motions are integrated using the leapfrog Verlet algorithm with the time step of 1.0×10^{-3} ps. Total energy is averaged after 2000 iterations with equilibration every 200 steps. The interaction between the carbon atoms is described by the well-known Tersoff two-body potentials [15,16]. Tersoff potentials have been used in many carbon and related materials calculations, and have proved that a wide range of structural properties can be determined using this formalism [15,16]. In addition, many studies relating bond length and bond order to total energy and force constants are interpreted correctly, suggesting that this method is a good start in predicting composite properties as discussed in this paper [17-19].

Swcnts modelled in this paper are an armchair (12,12) carbon nanotube, two chiral structures (12,10), and (10,12) nanotubes, thereafter all referred to as cnt(12,12), cnt(12,10), and cnt(10,12) respectively. The structure were chosen such that all have comparable number of carbon atoms; i.e cnt(12,12) has 312 carbon atoms, whilst cnt(12,10) and cnt(10,12) both have 264 and 260 carbon atoms respectively. Furthermore, cnt(12,12) and cnt(12,10) have equal values of a and b -axis, likewise, for cnt(12,12) and cnt(10,12) the c -axis values are equal. Loosely articulating, cnt(12,10) can be thought of as the projection of cnt(12,12) along the a and b -axis whereas cnt(10,12) can be thought of as the projection of cnt(12,12) along the c -axis. It need to be mentioned that all the literature studied works only with chiral structures of the $(2n,n)$ form [13,20,21]. The radial distribution functions ($g(\mathbf{r})$) and structure factors ($S(\mathbf{k})$) were calculated at 300, 3000, and 5000 K temperature, to observe the behaviour of atomic configurations and interatomic interactions with elevated temperatures. In order to obtain the equilibrium configurations of these models a -axis values were varied with total energy at constant temperature of 300 K until a convincing minimum energy is found. Based on the optimized structures and their energies, various equilibrium properties of these models had been extracted.

3. Results and discussion

3.1 Structural properties

In order to draw confidence on the empirical bond-order potential, radial distribution functions $g(\mathbf{r})$ for swcnt in two chiral forms have been calculated. When searching for the minimum of the three symmetry configurations; chiral cnt(10,12) configuration produced a maximum value instead of a minimum, so this symmetry was discontinued from the study. Even though $g(\mathbf{r})$ gives information only up to the second nearest neighbour, the insight upon the arrangement of carbon atoms in nanotubes can be extracted. This can be quantitatively compared with carbon atoms distribution in graphene and graphite formations. Using the peak positions of $g(\mathbf{r})$, the nearest neighbour and the second nearest neighbour carbon atoms in swcnt can be determined. The peak positions appear at 1.45 and 2.50 Å for both armchair cnt(12,12) and chiral cnt(12,10), in good agreement with other theoretical calculations and experiments [22-24]. The radial distribution function for cnt(12,12) at 300, 3000, and 5000 K is shown in figure 1. Detailed information from $g(\mathbf{r})$ along with experimental results of Boiko *et al.* [23] can be harvested in table 1. Figure 2 displays the structure factor functions ($S(\mathbf{k})$) for the same cnt(12,12) at 300, 3000, and 5000 K. $S(\mathbf{k})$ is the Fourier transform of the $g(\mathbf{r})$ from the real space to the reciprocal space.

Table 1 Quantities calculated from the radial distribution functions. r_1 and r_2 represent first and second nearest neighbour distances. n_1 and n_2 represent the coordination number of atoms in the first and second nearest neighbour distances.

	r_1 (Å)	r_2 (Å)	n_1	n_2
cnt(12,12)	1.45	2.50	4.25	2.70
cnt(12,10)	1.45	2.50	4.35	2.90
Graphite[23]	1.43	2.53	3.30	-

Using a pictorial scrutiny of the $g(r)$ in figures 1 and 2 and comparison of experimental results in table 1, it appears that the carbon nanotubes tubular structure is robust with cylindrically stable structure even at extremes of temperature. This toughness is comparable to that of diamond according to table 1 standards. Atomic distribution around the first nearest neighbour is uniform in all temperature ranges and it resembles that of diamond according to Kakinoki *et al.* [24] even though the peak positions are for graphite. At second nearest neighbour distance (2.50 Å) coordination is diamond-like at 300 K but becomes more graphitic at 5000 K. Wiggles at 3000 and 5000 K peaks on cnt(12,12) suggest that armchair (figure 3) is not as robust as chiral cnt(12,10).

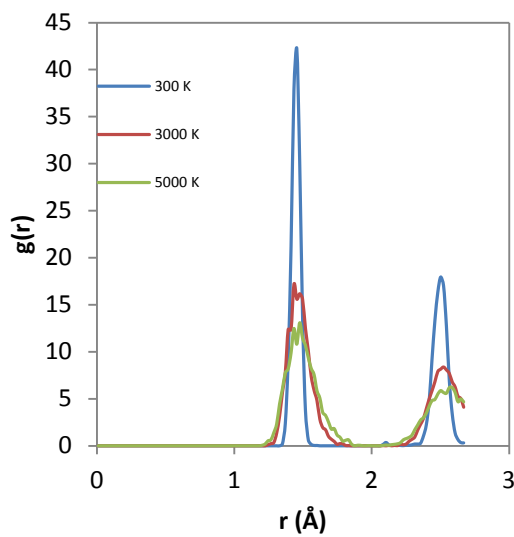


Figure 1. Radial distribution function for the armchair cnt(12,12) configuration at 300, 3000, and 5000 K.

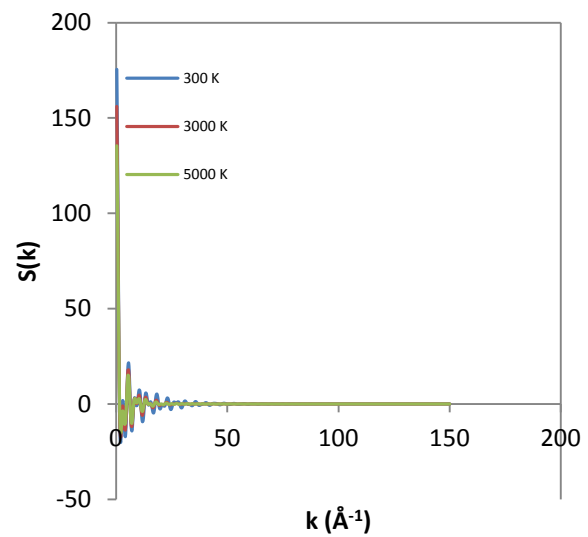


Figure 2. Structure factor functions for the armchair cnt(12,12) configuration at 300, 3000, and 5000 K.

3.2 Equilibrium properties

The structural optimization was performed under the Evans NVT ensemble at zero pressure and 300 K temperature. In this environment, equilibrium bundle lattice constant, bulk modulus, bulk modulus derivative, equilibrium energy, and equilibrium volume were calculated. Figures 3 and 4 show how the total energy per atom in nanotube bundles behaves with the varying a -axis in both the cnt(12,12) and cnt(12,10) tube bundles. The total energy against a -axis data was further least-squares fitted to Murnaghan's [25] equation of state in order to obtain the bulk modulus and its derivative. The results

obtained together with other carbon nanotube calculations [19] and graphite experimental results [7,8] are organized in table 2.

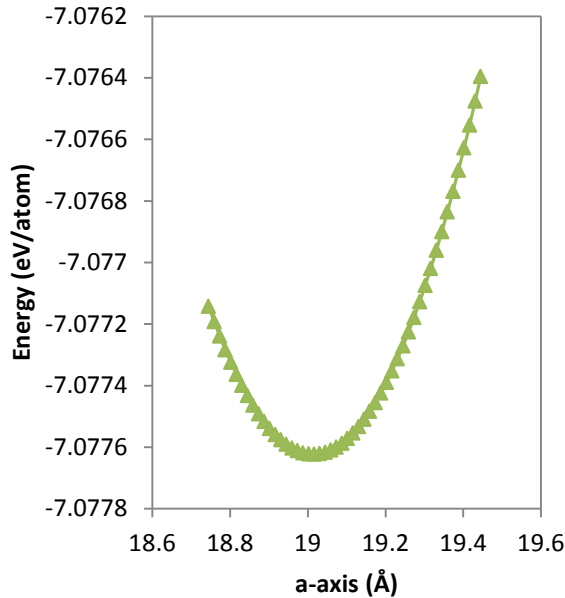


Figure 3. Equilibrium energy as a function of a -axis for the armchair cnt(12,12).

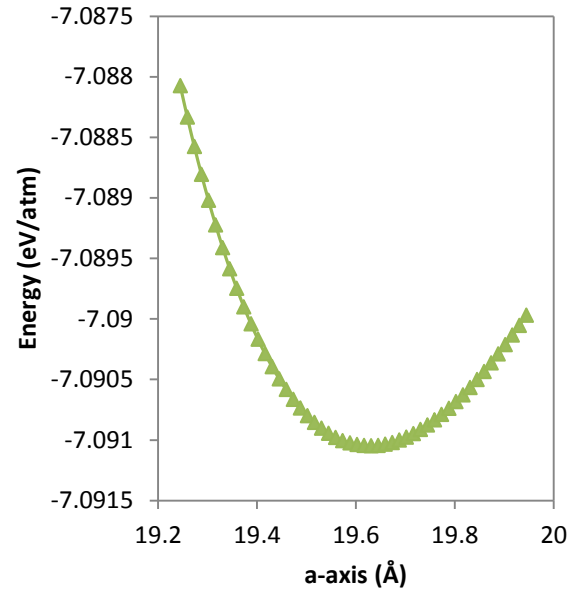


Figure 4. Equilibrium energy as a function of a -axis for the chiral cnt(12,10).

The equilibrium lattice constant a for the unit cells of the cnt(12,12) and cnt(12,10) bundles were found to be 19.014 and 19.628 Å for nanotube radii of 8.14 and 7.47 Å respectively. Reich *et al.* [19] demonstrated that the spacing between the tubes in a bundle cannot exceed the graphite equilibrium spacing of 3.35 Å. Applying this concept, tubes with bigger radii will have smaller inter-tubular spacing. To complement this, Tang *et al.* [26] merged x-ray measurements and elasticity theory calculations to show that the wall-to-wall distance between tubes within a bundle may not exceed the graphite equilibrium interlayer spacing. The bulk modulus for the armchair cnt(12,12) and chiral cnt(12,10) were computed to be 21.12 and 44.56 GPa respectively. These results are within the allowed limits of the graphite x-ray measurements of Hanfland *et al.* [7] and Zhao and Spain [8].

Table 2. Calculated lattice constants a , bulk modulus B_0 and its first derivative B' , equilibrium energy E_0 , and equilibrium volume V_0 of two carbon nanotubes symmetries compared with graphite experimental results [7,8] and calculated results of carbon nanotubes [19].

	cnt(12,12) this work	cnt(12,10) this work	Graphite experiment[7,8]	Nanotubes bundles[19]
a (Å)	19.01	19.63	2.603	11.43
B (GPa)	21.12	44.56	33.80	37.00
B'	1.000	18.37	8.900	11.00
E_0 (eV)	-7.078	-7.091	-	-
V_0 (Å ³)	27.21	28.93	35.12	-

3.3 Dynamical properties

Transport properties of carbon ions and even impurities within a tube can be extracted from the simulations using the time-dependent mean square displacement (msd), which is defined by:

$\langle r_i^2(t) \rangle = \frac{1}{N} \sum_{i=1}^N [r_i(t) - r_i(0)]^2$, where N is the total number of ions in the system, $r_i(0)$, and $r_i(t)$ are initial and advanced position of ions respectively. At elevated temperatures, the carbon ions gain momentum and become highly mobile and so the msd increases with time. The argument can be confirmed by the msd data plotted as function of time for armchair cnt(12,12) and chiral cnt(12,10) configurations as shown in figures 5 and 6.

As it can be seen, at room temperatures, both configurations demonstrate somewhat constant behaviour with time, which suggest that there is no notable carbon ion diffusion. A periodical rise and fall conduct of carbon ions with time at 3000 K is notable to both cnt(12,12) and cnt(12,10). At 5000 K, carbon function in cnt(12,10) increases with increasing time, indicating ion diffusion, whereas oscillatory behaviour is still observed in cnt(12,12). From the slope of carbon ions as function of time in cnt(12,10) at 5000 K, a diffusion coefficient of $0.108 \text{ \AA}^2/\text{ps}$ and a thermal factor of 1.029 \AA^2 were computed according to the equation: $\langle r_i^2(t) \rangle = 6Dt + A_0$, where D is the diffusion coefficient and A_0 is some small thermal factor due to atomic vibrations.

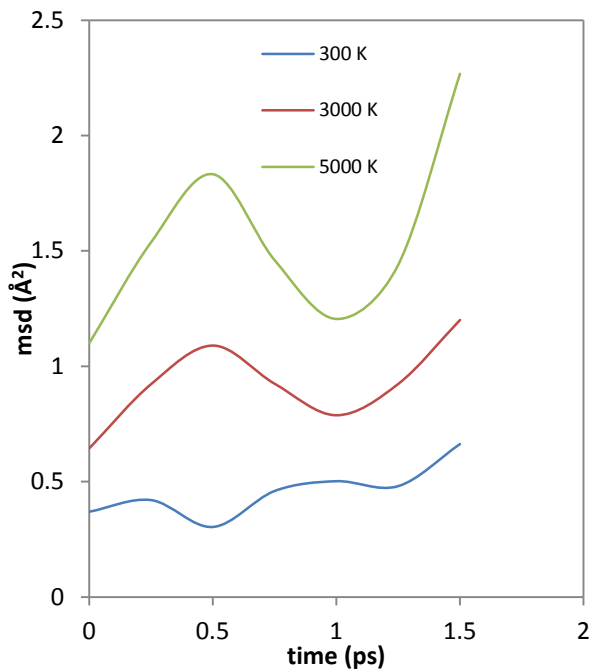


Figure 5. Mean square displacement (msd) of carbon ions in an armchair cnt(12,12) at 300, 3000, and 5000 K.

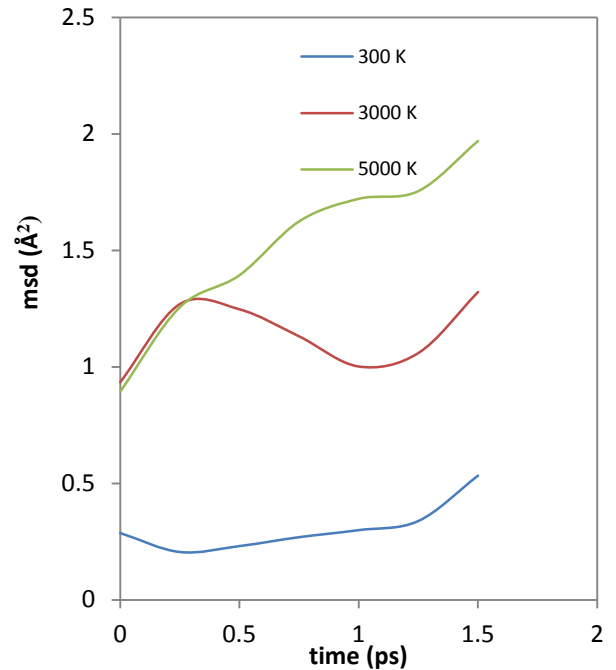


Figure 6. Mean square displacement (msd) of carbon ions in an achiral cnt(12,10) at 300, 3000, and 5000 K.

4. Conclusion

The structural configuration and equilibrium properties of some swcnts symmetry configurations were studied. $g(\mathbf{r})$ and $S(\mathbf{k})$ results demonstrate the strength and rigidity of atomic distributions even at extremes of temperature. This in turn validates the flexibility and transferability of the Tersoff potential to nanoscale level. The bulk modulus of the bundles is found to be within the graphite acceptable range. Specifically for cnt(12,12) and cnt(12,10), the larger radius implies smaller bulk modulus and the other way round. The equilibrium volumes are smaller than those of graphite.

Acknowledgements

The UL and IBSA are thanked for financial assistance. CHPC is thanked for computational facilities.

References

- [1] Dresselhaus MS, Dresselhaus G, Sugihara K, Spain IL, Goldberg HA 1988 *Graphite Fibers and Filaments* (Springer-Verlag, New York).
- [2] Dresselhaus MS, G. Dresselhaus G, and R. Saito R 1995 *Carbon* **33** 883.
- [3] Krishnan A, Dujardin E, Ebbesen TW, Yianilos TN, and Treacy MMJ 1998 *Phys. Rev. B* **58** 14013.
- [4] Hernández E, Goze C, Bernier P, and Rubio A 1998 *Phys. Rev. Lett.* **80** 4502.
- [5] Popov V, Doren VV, and Balkanski M 2000 *Solid State Commun.* **114** 395.
- [6] Lu J 1997 *Phys. Rev. Lett.* **79** 1297.
- [7] Hanfland M, Beister H, and Syassen K 1989 *Phys. Rev. B* **39** 12598.
- [8] Zhao YX and Spain IL 1989 *Phys. Rev. B* **40** 993.
- [9] Treacy MMJ, Ebbesen TW, and Gibson TM 1996 *Nature* **381** 680.
- [10] Salvétat JP, Bonard JM, Thomson NH, Kulik AJ, Farro L, Bennit W, and Zuppiroli L 1999 *J. Appl. Phys. A* **69** 255.
- [11] Mintmire JW, Dunlap BI, and White CT 1992 *Phys. Rev. Lett.* **68** 631.
- [12] Hamada N, Sawada S, and Oshiyama A 1992 *Phys. Rev. Lett.* **68** 1579.
- [13] Robertson DH, Brenner DW, Mintmire JW 1992 *Phys. Rev. B* **45** 12592.
- [14] Smith W, Forester TR and Todorov IT, 2009 *The DL_POLY 2 User Manual*, STFC Daresbury Laboratory, Daresbury, Warrington WA4 4AD Cheshire, UK.
- [15] Tersoff J 1988 *Phys. Rev. B* **75** 6991.
- [16] Tersoff J 1988 *Phys. Rev. Lett.* **61** 2879.
- [17] Tersoff J and Ruoff RS 1994 *Phys. Rev. Lett.* **73** 676.
- [18] Diao J, Srivastava D, and Menon M 2008 *J. Chem. Phys.* **128** 164708.
- [19] Reich S, Thomsen C, and Ordejón P 2002 *Phys. Rev. B* **65** 153407.
- [20] Pullen A, Zhao GL, Bagayoko D, and Yang L 2005 *Phys. Rev. B* **71** 205410.
- [21] Gao G, Çağın T, and Goddard WA III 1998 *Nanotechnology* **9** 184.
- [22] Beeman D, Silverman R, Lynds R, and Anderson MR 1984 *Phys. Rev. B* **30** 870.
- [23] Bioko BT, Palatnik LS, and Derevyanchenko AS 1968 *Dokl. Akad. Nauk SSSR* **179** 316 [1968 *Sov. Phys. – Dokl.* **13** 237].
- [24] Kakinoki J, Katada K, Hanawa T, and Ino T 1960 *Acta Crystallogr.* **13** 171.
- [25] Murnaghan F, 1944 *Proc. Nat. Acad. Sci. USA*, **30** 244.
- [26] Tang J, Qin L-C, Sasaki T, Yudasaka M, Matsushita A, Iijima S 2000 *Phys. Rev. Lett.* **85** 1887.

Parallel benchmarks for ARM processors in the high energy context

Joshua Wyatt Smith and Andrew Hamilton

University of Cape Town, Physics Department

E-mail: joshua.wyatt.smith@cern.ch

Abstract. High Performance Computing is relevant in many applications around the world, particularly high energy physics. Experiments such as ATLAS, CMS, ALICE and LHCb generate huge amounts of data which need to be stored and analyzed at server farms located on site at CERN and around the world. Apart from the initial cost of setting up an effective server farm the cost of power consumption and cooling are significant. The proposed solution to reduce costs without losing performance is to make use of ARM[®] processors found in nearly all smartphones and tablet computers. Their low power consumption, low cost and respectable processing speed makes them an interesting choice for future large scale parallel data processing centers. Benchmarks on the Cortex[™]-A series of ARM processors including the HPL and PMBW suites will be presented as well as preliminary results from the PROOF benchmark in the context of high energy physics will be analyzed.

1. Introduction

High energy physics (HEP) at the Large Hadron Collider (LHC) [1] creates an enormous amount of data that need to be stored for later analysis. Dedicated server farms have been built at CERN (Tier 0) as well as around the world (Tier 1's and Tier 2's) and are connected through The Worldwide LHC Computing Grid (WLCG). The initial setup costs of these server farms can be immense and the costs to maintain such a server farm can be even larger. Cooling and the power required to maintain such servers at their peak performance make server farms an expensive venture. In early 2013 Tier 0 had a power capacity of 3.5MW.

The proposed solution is to make use of ARM[®] processors [2] from here on referred to as ARM. These are found in smartphones and tablet computers where the combination of low power consumption and high performance is the top priority. Significant savings might be achieved if ARM processors are able to cope with the huge amount of data processed. This letter serves to explore the capabilities of ARM processors through parallel processing benchmarks.

2. Hardware and software

The ARM processor was designed to perform only a few instructions at once. This reduces the need for hardware such as transistors and thus minimizes power consumption. This letter addresses three system on chip (SoC) setups in the Cortex[™]-A range, namely the A7 MPCore[™], A9 MPCore[™] and A15 MPCore[™] [3, 4, 5], from here on referred to as the A7, A9 and A15 respectively. Table 1 summarizes the different processors used. An obvious advantage to the A9 is the number of cores, however this is irrelevant because the number of cores is an

intrinsic property of the SoC. FPU refers to *Floating Point Units*. The FPU generate results for the speed at which multiplication and addition operations are carried out. v3/4 refers to the respective FPU version. A traditional Intel[®] computer is also used (Hep405). This serves to provide some reference to the reader. Power measurements for the A7, A9 and A15 were taken using a Fluke 289 Digital Multimeter. Measurements on Hep405 were taken using the Intel[®] Power Gadget. However, if the power usage characteristics and temperatures vary from those used in Intel's calibration then there will be errors between estimated and actual power usage. For this reason, when referring to Hep405's results the reader must take care in remembering that these values serve as more of an estimate.

Table 1. The different setups with key features.

Setup	Processor	Cores	RAM	Cache	FPU	OS
Cubietruck	AllWinner A20, 1.2GHz	A7 dual core	2GiB DDR3	512 KiB L2	VFPv4	Archlinux, hard float
Wandboard-Quad	Freescall i.MX6 Quad, 996MHz	A9 quad core	2GiB DDR3	32KiB L1, 1 MiB L2	VFPv3	Archlinux, hard float
ArndaleBoard-K	Samsung Exynos 5250, 1.7GHz	A15 dual core	2GiB DDR3	32 KiB L1, 1MiB L2	VFPv4	Fedora 19, hard float
Hep405	Intel [®] Core i7-2600, 3.4GHz	quad core	16GiB DDR3	256KiB L1, 1MiB L2, 8MiB L3	-	Scientific Linux 6

3. Benchmarks and results

3.1. High Performance LINPACK suite

The LINPACK benchmark [6] is historically one of the most common tests in high performance computing (HPC) being used as early as the 1980's. The High-Performance LINPACK (HPL) benchmark is the parallel version of the LINPACK benchmark which is used to rank the world's TOP500 supercomputers. The user can specify how much memory to commit to solving the largest problem that the machine is capable of solving. It calculates the floating point operations per second or *flops* of a system by splitting the large matrix into blocks that are then solved on different cores or CPUs. This enables several blocks to be worked on in parallel. Ideally the increase in speed is scalable, i.e. four cores is four times quicker than one core, however, communication and latency between the cores hampers the performance and so speedup is never actually 100%. A list of block sizes and matrix sizes are specified at the beginning of the run (in ascending order). Then each matrix size is iterated over using the increasing block sizes. Thus, there is an overall increase in transferred matrix size from left to right.

Figure 1 shows the power consumption with the respective *Gflops* and *Gflops/watt*. Power measurements were taken at a 7 second resolution. The grey area is an "envelope function" which splits the data into blocks along the time axis and then finds the average power value for that window. It gives a representation of the average power consumption which is indicated by the blue line. A trend we see for the A7 and A9 is that as the size of the matrix block that is passed to the separate cores increases, the average power consumption decreases. This is because slower computation means that less work in the form of communication has to happen between the processors. The average power consumption remains fairly constant with the A15 because of the faster clock speed of 1.7GHz. The *Gflops/watt* is calculated by taking the average power consumption for the time window where there is a HPL measurement. We can see that the A9 is

the most efficient provided the block size is relatively large. The A7 does not have a fast enough processor and the A15 uses too much power for the output it delivers. The A9's *Gflops*/watt is comparable to the output of Hep405.

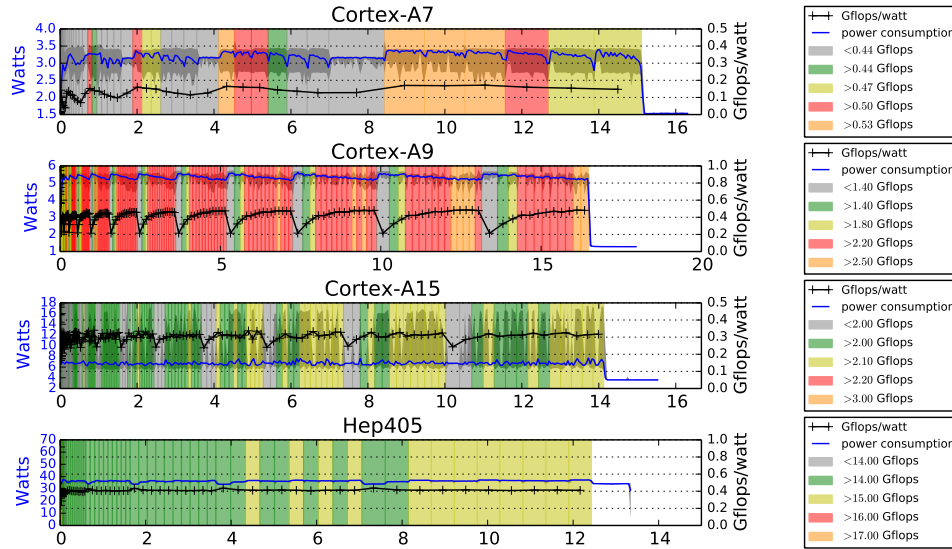


Figure 1. The relationship between matrix size, block size, *Gflops*, power consumption and *Gflops*/watt for the HPL benchmark. All *Gflops* values in each legend are present in each graph, they just have too small time scales to be seen.

3.2. Parallel Memory Bandwidth Benchmark suite

The Parallel Memory Bandwidth Benchmark (PMBW) [7] is a relatively new suite that measures bandwidth capabilities of a multi-core computer. This is an important test because more cores result in the floating point performance increasing in a linear fashion. However, if the memory bandwidth is not capable of processing the data fast enough those processors will stall. Unlike floating point units the memory bandwidth does not scale with the number of cores running in parallel. The code was developed in assembler language which means compile flags for the SoC become unimportant, thus there are no optimizations. The code uses two general synthetic access patterns, namely sequential scanning and pure random access. A real world application will fit somewhere in-between the two tests. The results for the benchmark are plotted in Figure 2.

For single threads, the A15 performs better than the A9 and A7 as expected due to the higher clock speed. If we look at the SoC with threads equal to the number of processors, for small cache memory transfers the peak bandwidth for the A9 and A15 are comparable but as the array size increases and bandwidth starts to stabilize the A9 doesn't perform as well as the A7 and A15 due to the lower clock speed at 996MHz. This figure shows only the results for the tests in which 32 bit message sizes are transferred for each clock cycle. It needs to be mentioned that the A15 has a higher peak performance when transferring 64 bit message sizes but it cannot maintain these speeds. As the total array size increases the bandwidth drops below the 32 bit values. For reference, on a single thread Hep405 is able to reach a peak bandwidth of 121 GiB/s for a 256 bit message size while reading and 60GiB/s while writing. It reaches 298GiB/s for reading and 163GiB/s for writing when running 4 threads. It is accurate to say that none of the ARM processors perform very well when it comes to memory bandwidth. A solution relies on using 64-bit architecture for the SoC instead of 32-bits. This will improve bandwidth speed

as it allows wider memory registers and access to more RAM. These chips have been released recently, however it will still be some time before they are put onto development boards.

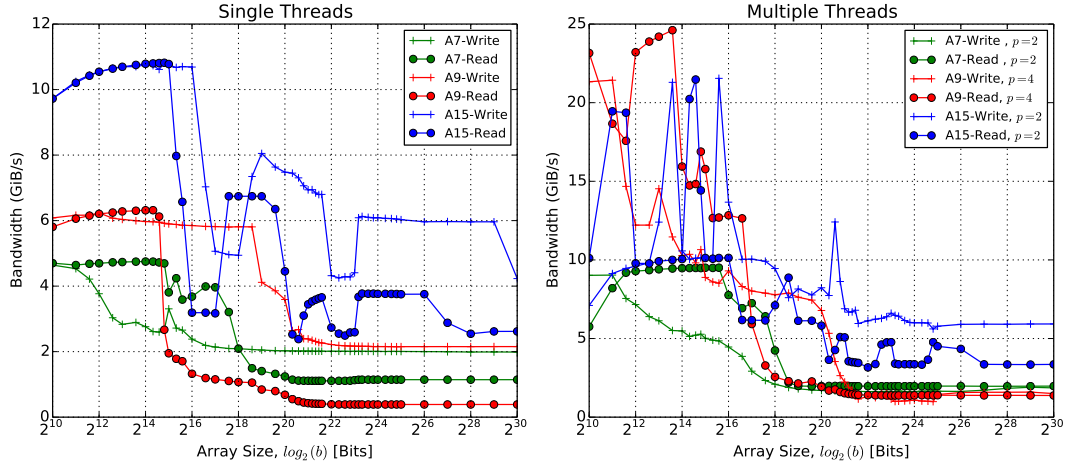


Figure 2. Memory bandwidth results for each processor. Read/Write refers to scanning operations (opposed to permutation operations) performed while doing read or write tasks. Each routine transfers 32-bits with increasing total array size.

3.3. PROOF benchmark suite

PROOF (Parallel ROOT Facility) [8], is a parallel extension to the well known data analysis tool used in HEP called ROOT [9]. It was designed as an extensive test suite in benchmarking potential configurations and performances of multicore computer clusters. PROOF exploits the fact that data analysis in the HEP context can be easily parallelizable. Tasks can therefore be neatly split up onto the individual cores (or *workers*) of each computer.¹ If the task size increases in proportion to the number of processors then the results are scalable. However, if data gets distributed or read sequentially by one core this may cause bottlenecks in the memory. Thus, different topologies can be tested such as having master nodes within smaller clusters which further facilitate communication. The CPU benchmark starts off by performing measurements with 1 active *worker* and enables additional *workers* at the start of each test. This benchmark is more appropriate for a setup with more nodes and cores, however it also returns a value for the SoC as well as a normalized value for a single *worker* which is good for comparing each of the ARM processors. The default setting creates 16 1d histograms filled with $1000000 \times \text{workers}$ random numbers. MRGPS is Mega Random Generations Per Second. One random generation produces 16 gaussian numbers and fills 16 histograms. The output is given for the SoC, as well as the normalized performance per worker. Power consumption was recorded with a resolution of 1 second. The results are shown in Table 2.

For the ARM processors the A9 is the most efficient returning the largest value for MRGPS and MRGPS/watt. However, it performs significantly worse than Hep405. This is a problem for the ARM processors as this is an easily parallelizable benchmark and the results should be very close to scaling linearly. We don't want to sacrifice speed and so it means that in order to reach computation speeds similar to that of Hep405 a minimum of 19 A9's must be present. This doesn't reduce power consumption.

¹ The number of *workers* isn't necessarily the same as the number of cores. For this letter, results showing *workers* equal to the number of cores were chosen.

Table 2. The PROOF CPU benchmark.

Setup	Power (W)	Workers	MRGPS	$\frac{MRGPS}{worker}$	$\frac{MRGPS}{watt}$
A7	1.997	2	0.108	0.055	0.054
A9	2.999	4	0.301	0.075	0.100
A15	7.254	2	0.296	0.118	0.041
Hep405	31.766	4	5.525	1.347	0.174

4. Conclusions

Table 3 contains a summary of the results. The A7, A9 and A15 use significantly less power than traditional processors. However, they are also significantly slower and so parallel computing for the right type of problem must be exploited. From a performance point of view the dual core A15 is only barely faster than the quad core A9. In terms of performance per watt the A9 performs better. The question arises as to whether or not it is beneficial to replace current computer processors with ARM processors. From our findings there is little advantage (if any) in using the 32-bit ARM processors presented in this letter. There is however, a quad core A15 available as well as the recently released CortexTM-A50 range with power efficient 64-bit architecture. Development boards with this newer technology should be available soon and may have a large impact on processing speeds and memory bandwidth.

Table 3. Summary of the important results for each benchmark.

	A7	A9	A15	Intel [®] i7
Cores	2	4	2	4
Idle (W)	1.518	1.282	3.573	4.356
HPL: Average <i>Gflops</i> /W	0.109	0.408	0.323	0.421
PMBW: Max read (GiB/s)	9.505	24.611	21.472	298.264
PMBW: Max write (GiB/s)	9.035	21.428	21.546	163.123
PROOF: $\frac{MRGPS}{watt}$	0.054	0.100	0.041	0.174

References

- [1] *Journal of Instrumentation*, vol. 3. IOP and SISSA, 2008.
- [2] ARM, *Procedure Call Standard for the ARM[®] Architecture*. ARM Ltd, November, 2012.
- [3] ARM Ltd, *CortexTM-A7 MPCoreTM Technical Reference Manual*, r0p5 ed., April, 2013.
- [4] ARM Ltd, *CortexTM-A9 MPCoreTM Technical Reference Manual*, r4p1 ed., June, 2012.
- [5] ARM Ltd, *CortexTM-A15 MPCoreTM Processor Technical Reference Manual*, r4p0 ed., June, 2013.
- [6] J. Dongarra, *Linpack benchmark*, Computer Science Technical Report CS - 89 - 85, University of Tennessee.
- [7] T. Bingmann, “pmbw - Parallel Memory Bandwidth Benchmark / Measurement.” website, 2013.
- [8] ACAT, *The PROOF benchmark suite measuring PROOF performance*, vol. Journal of Physics: Conference Series 368 (2012) 012020, IOP Publishing, 2011.
- [9] Antcheva, I. et al, ed., *ROOT - A C++ framework for petabyte data storage, statistical analysis and visualization*, vol. 180, 12; 2499-2512, Computer Physics Communications, 2009.

Methodology for the digital simulation of open quantum systems

R B Sweke¹, I Sinayskiy^{1,2} and F Petruccione^{1,2}

¹ Quantum Research Group, School of Physics and Chemistry, University of KwaZulu-Natal, Durban, 4001, South Africa

² National Institute for Theoretical Physics, University of KwaZulu-Natal, Durban, 4001, South Africa

E-mail: rsweke@gmail.com

Abstract. Quantum simulation of quantum systems is an extremely active field of contemporary research. In this paper a brief introduction is given to this field, with an emphasis on the distinction between analog and digital quantum simulations. Furthermore, conventional methods for the digital quantum simulation of Hamiltonian simulation are discussed and generalisations of these methods for the digital simulation of open quantum systems are presented.

1. Introduction

It has been known for some time that the simulation of quantum systems is a challenging problem [1]. In particular, one immediately apparent difficulty inherent in simulating quantum systems on conventional classical computers arises from the so called *exponential explosion* [2] - a description of the fact that a classical encoding of an N qubit quantum state (one may think of a qubit as a spin 1/2 particle [3]) typically requires the storage of 2^N probability amplitudes. As such, even writing the initial state of some desired simulation (before a discussion of the simulation itself can even take place) may require totally unrealistic memory resources. As an illustration, encoding the state of 40 spin 1/2 particles may require the storage of $2^{40} \approx 10^{12}$ complex numbers, which at standard precision requires approximately 4 terabytes of memory [2].

Due to such difficulties and the large number of potential applications for simulation of quantum systems in fields such as physics, biology and chemistry, the development of methods for the efficient simulation of quantum systems has become an extremely active area of contemporary research [2]-[4]. As we are concerned in this paper with a presentation of potential methodologies for the efficient quantum simulation of open quantum systems, it is useful to begin by defining what is meant by “efficient simulation of quantum systems”. Firstly, a simulation of a quantum system is understood as some process via which certain *pre-specified* properties of the the quantum system are accurately predicted [4]. It is important to note that we do not require our simulation to predict *all* properties of the quantum system of interest. For example, one may be interested in the expectation value of specific observables, scattering amplitudes within a quantum field theory or the state of the system at some particular time. The predictions of the simulation are considered accurate if the error is less than some pre-specified error tolerance

ϵ , and the simulation is considered successful if the specific properties of interest are accurately predicted, regardless of whether or not other properties of the quantum system are accurately predicted, or even accessible. Secondly, a simulation is considered efficient if the time and space resources required to obtain the required predictions can be determined as a polynomial function of the length of the bit string specifying the problem inputs and of $1/\epsilon$ [4, 5].

In the early 1980's Richard Feynman suggested that it may be possible to use quantum systems themselves as a means to efficiently simulate other quantum systems and simultaneously avoid difficulties such as the exponential explosion [1]. This suggestion has turned out to be remarkably prescient and while many approximation methods, such as Monte-Carlo techniques [6], have been developed for the efficient simulation of quantum systems on classical computers, we are concerned here with simulations in line with Feynman's original vision - so called "efficient *quantum* simulation of quantum systems".

2. Analog and digital quantum simulation

Quantum simulation, as envisaged by Feynman, requires a *quantum simulator*, which is understood as a *controllable quantum system used for the simulation of other quantum systems* [2]. For the purposes of illustration we will discuss in this section the simulation of closed quantum systems, undergoing Hamiltonian generated unitary evolution, but it is important to bear in mind that these principles may be applied to a broad class of quantum systems, such as quantum field theories and open quantum systems, as discussed in the review of Georgescu et. al. [2].

Consider a closed quantum system, initially in the state $|\phi(0)\rangle$ evolving via the Hamiltonian H_{sys} into the final state $|\phi(t)\rangle = U|\phi(0)\rangle$ where $U = \exp(-i\hbar H_{sys}t)$. As illustrated in Figure 1, any quantum simulation of such a system requires three distinct steps: The first step, or preparation phase, is an efficient preparation of the initial quantum state of the quantum simulator $|\psi(0)\rangle$, where there exists some well defined mapping between the states $|\psi(0)\rangle$ and $|\phi(0)\rangle$. The second step is a simulation of the system dynamics, i.e. an implementation of the unitary dynamics \hat{U} where \hat{U} approximates U , and the final step is an extraction of the desired system properties via measurements of the final state of the simulator $|\psi(t)\rangle = \hat{U}|\psi(0)\rangle$. It is important to note that *all steps* in the above process need to be efficient according to the criterion discussed in the introduction.

2.1. Digital quantum simulation

Digital quantum simulation refers to quantum simulation utilising the conventional quantum circuit model implemented on some universal quantum computer [3]. In order to perform a digital quantum simulation it is therefore necessary that the initial state of the simulator $|\psi(0)\rangle$ is some multi-qubit state. Furthermore, the unitary transformation \hat{U} is then implemented via a sequence of one and two qubit gates from some appropriate universal gate set [3]. It is interesting to note that in principle *any* unitary operation can be decomposed into a sequence of gates from any universal gate set (hence the name *universal* gate set) however not all such decompositions are efficient - i.e. not all unitary operators can be decomposed into a polynomial number of gates from some universal set. The question of which Hamiltonians can be efficiently simulated via digital quantum simulation has been exhaustively addressed [2, 4] and in Section 3 we provide an illustration of the prevailing methodology by using local Hamiltonians as an example. Digital quantum simulation is advantageous in that any universal quantum computer can be used as a digital quantum simulator. However, this advantage comes at the price of requiring a sufficiently sized universal quantum computer, which in itself is a significant theoretical and technological challenge, whose realisation has not yet been achieved.

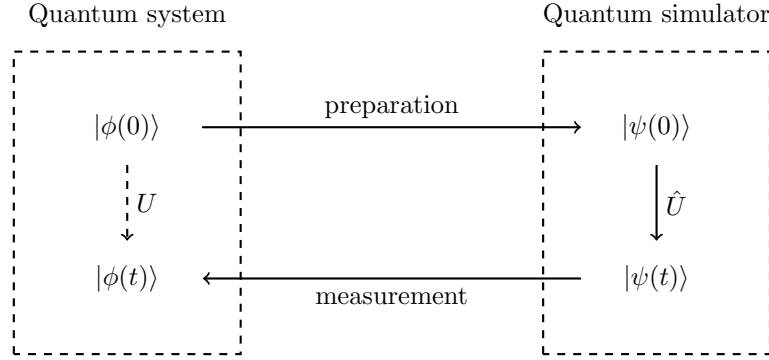


Figure 1. Schematic representation of a typical quantum simulation of a closed quantum system undergoing Hamiltonian generated unitary evolution $U = \exp(-i\hbar H_{sys}t)$. The first stage of the simulation is preparation of the initial state of the simulator $|\psi(0)\rangle$, which is in direct correspondence with $|\phi(0)\rangle$, the initial state of the system. The simulator then evolves via \hat{U} , an approximation to U , after which the desired system properties are obtained from measurements of the final simulator state $|\psi(t)\rangle$.

2.2. Analog quantum simulation

Analog quantum simulation, also known as quantum emulation, refers to quantum simulations in which one designs a quantum system (the quantum simulator) whose Hamiltonian H_{sim} is in direct correspondence with the Hamiltonian of the system which is to be simulated H_{sys} . The simulation then proceeds by letting the simulator evolve naturally via $\hat{U} = \exp(-i\hbar H_{sim}t)$. As an example [2] consider the Hamiltonian describing a gas of interacting bosonic atoms in a periodic potential

$$H_{sim} = -J \sum_{i,j} \hat{a}_i^\dagger \hat{a}_j + \sum_i \epsilon_i \hat{n}_i + \frac{1}{2} U \sum_i \hat{n}_i (\hat{n}_i - 1) \quad (1)$$

where \hat{a}_i^\dagger and \hat{a}_i are the bosonic creation and annihilation operators on the i 'th lattice site, $\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i$ is the atomic number operator and ϵ_i denotes the energy offset of the i 'th lattice site. The coefficient J provides a measure of the hopping strength between lattice sites and U quantifies the interaction strength between atoms at the same lattice site. The Hamiltonian H_{sim} is extremely similar to the Bose-Hubbard Hamiltonian

$$H_{BH} = -J \sum_{i,j} \hat{b}_i^\dagger \hat{b}_j - \mu \sum_i \hat{n}_i + \frac{1}{2} U \sum_i \hat{n}_i (\hat{n}_i - 1) \quad (2)$$

where J and U are the same as above and μ is the chemical potential. The correspondence between H_{sim} and H_{BH} is clear and as such it is evident that one may straightforwardly simulate the Bose-Hubbard model, via identification of appropriate parameters, using an analog quantum simulator consisting of atoms in an optical lattice. Analog quantum simulation is advantageous in that one does not require an entire universal quantum computer for such simulations, and as a result analog quantum simulations have already been successfully demonstrated [2].

3. Hamiltonian simulation

In 1996 Lloyd [7] demonstrated a method for the digital simulation of Hamiltonian systems, and since then this method has provided the fundamental methodology for more sophisticated digital Hamiltonian simulation algorithms [4]. In this section we briefly review these methods as they

provide the inspiration for the open quantum system digital simulation methods presented in Section 4. Lloyd considers a Hamiltonian which may be written as the sum of local Hamiltonians

$$H = \sum_{l=1}^M H_l \quad (3)$$

where each H_l may be efficiently simulated - i.e. there exists a known efficient gate decomposition for $U_l = \exp(-i\hbar H_l t)$. In the case that $[H_l, H_{l'}] = 0$ for all l and l' we then have that

$$U = \exp(-i\hbar H t) = \prod_l \exp(-i\hbar H_l t) = \prod_l U_l \quad (4)$$

and the efficient gate decomposition for U is easily achieved via the known gate decompositions for U_l . However in general it is not true that all constituent local Hamiltonians commute and for this case Lloyd recognized that the Hamiltonian H could be efficiently simulated via a discretization of time, where the errors resulting from such a discretization are bounded via the Lie-Trotter product formula [3]. In particular, if H_1 and H_2 can be efficiently simulated, then as a result of the Lie-Trotter formula

$$e^{-i\hbar(H_1+H_2)t} = \lim_{n \rightarrow \infty} \left(e^{-i\hbar H_1 t/n} e^{-i\hbar H_2 t/n} \right)^n, \quad (5)$$

one has that

$$\left\| \left(e^{-i\hbar H_1 t/n} e^{-i\hbar H_2 t/n} \right)^n - e^{-i\hbar(H_1+H_2)t} \right\| \leq \epsilon \quad (6)$$

provided $n = \mathcal{O}((\nu t)^2/\epsilon)$ where $\nu := \max\{\|H_1\|, \|H_2\|\}$. This result is easily extended to the sum in Eq. 3 and provides an effective method, via time discretization, for the simulation of any Hamiltonian which can be written as a sum of efficiently implementable Hamiltonians. Since this initial result new digital simulation algorithms have been developed which have broadened the class of Hamiltonians which may be simulated while simultaneously decreasing the algorithmic costs [4]. However all such methods rely on the following fundamental strategy:

- (i) Decompose the Hamiltonian H into a sum of Hamiltonians which can be efficiently simulated, as in Eq. 3.
- (ii) Simulate the Hamiltonian H through recombination of the constituent Hamiltonians, via a Lie-Trotter product formula or suitable generalisation.

This methodology may be described as the *Decomposition/Recombination* strategy and in the following section we illustrate how a similar strategy may be utilised for the digital simulation of open quantum systems.

4. Digital simulation of open quantum systems

The problem considered in this section, to which we would like to apply the methodology illustrated in Section 3, is that of simulating the dynamics of Markovian open quantum systems. In particular, given a continuous one parameter semi-group of quantum channels $\{T_t\}$ [8], where $T_t : \mathcal{B}(\mathcal{H}_s) \rightarrow \mathcal{B}(\mathcal{H}_s)$ is a completely positive trace preserving map from and onto the bounded operators on the Hilbert space of the system $\mathcal{H}_s \cong \mathcal{C}^d$, we would like a digital simulation algorithm which provides the state of the system at time t , given by $\rho(t) = T_t[\rho(0)]$, where $\rho(0) \in \mathcal{B}(\mathcal{H}_s)$ is the initial state of the system [3]. Every such continuous one parameter semigroup of quantum channels $\{T_t\}$ has a unique generator

$$\mathcal{L} : \mathcal{B}(\mathcal{H}_S) \rightarrow \mathcal{B}(\mathcal{H}_S) \quad (7)$$

such that

$$T_t = e^{t\mathcal{L}} = \sum_{k=0}^{\infty} \frac{t^k \mathcal{L}^k}{k!} \quad (8)$$

and \mathcal{L} satisfies the differential equation

$$\frac{d}{dt}\rho(t) = \mathcal{L}(\rho(t)), \quad (9)$$

known as a master equation. Furthermore, a linear super-operator $\mathcal{L} : \mathcal{B}(\mathcal{H}_S) \rightarrow \mathcal{B}(\mathcal{H}_S)$ is the generator of a continuous dynamical semigroup of quantum channels, if and only if it can be written in the form

$$\mathcal{L}(\rho) = i[\rho, H] + \sum_{k,l=1}^{d^2-1} A_{l,k} ([F_k, \rho F_l^\dagger] + [F_k \rho, F_l^\dagger]), \quad (10)$$

where $H = H^\dagger \in \mathcal{M}_d(\mathcal{C})$ is Hermitian, $A \in \mathcal{M}_{d^2-1}(\mathcal{C})$ is positive semidefinite and $\{F_i\}$ is a basis for the space of traceless matrices in $\mathcal{M}_d(\mathcal{C})$. Eq. 10 is known as the Gorini, Kossakowski, Sudarshan and Lindblad form of the quantum Markov master equation and we refer to A as the GKS matrix [8]. It is also crucial to note that for any quantum channel $T : \mathcal{B}(\mathcal{H}_s) \rightarrow \mathcal{B}(\mathcal{H}_s)$ it is always possible to introduce a dilation space \mathcal{H}_E with $\dim(\mathcal{H}_E) = [\dim(\mathcal{H}_S)]^2$ such that there exists a unitary matrix $U \in \mathcal{M}_{d^3}(\mathcal{C})$, known as the Stinespring dilation [8], where

$$T(\rho) = \text{tr}_E[U(|e_0\rangle\langle e_0| \otimes \rho)U^\dagger] \quad (11)$$

and $|e_0\rangle\langle e_0| \in \mathcal{H}_E$ is some initial state of the environment. In light of the above the Kliesch et. al. [9] have provided the following digital simulation algorithm for Markovian open quantum systems, which can be seen as a direct generalisation Lloyd's initial Hamiltonian simulation algorithm [7] discussed in the previous section:

- (i) Assume that the system consists of N subsystems of Hilbert space dimension d . Furthermore, assume that the generator \mathcal{L} is k -local in the sense that it can be written as the sum

$$\mathcal{L} = \sum_{\Lambda \subset [N]} \mathcal{L}_\Lambda \quad (12)$$

where $[N] := \{1, \dots, N\}$ and \mathcal{L}_Λ are *strictly* k -local, i.e. \mathcal{L}_Λ act non-trivially on at most k -subsystems.

- (ii) Simulate the total dynamics $T_t = \exp(t\mathcal{L})$ through recombination of the constituent channels $T_{t/m}^{(\Lambda)} = \exp(t/m\mathcal{L}_\Lambda)$. Each constituent channel may be simulated via the Stinespring dilation (whose unitary is guaranteed an efficient gate decomposition as a result of strict k -locality) and the error introduced by such time discretization is bounded by the Trotter decomposition theorem for Liouvillian dynamics [9], a direct generalisation of the Lie-Trotter theorem into the superoperator regime.

The algorithm described above can be seen as the direct analogue of Lloyd's initial Hamiltonian simulation algorithm in that it *assumes* the existence of a natural decomposition into constituent semi-groups which may be efficiently simulated, and then proves the efficiency of recombining these semi-groups. Given this fundamental platform and the success of similar strategies for digital simulation of Hamiltonian systems a general decomposition/recombination strategy is naturally suggested for Markovian evolutions, specified not by strictly k -local generators, but rather by generators with arbitrary GKS matrix A :

- (i) Decompose the generator \mathcal{L} into the sum of generators which may be efficiently simulated.
- (ii) Simulate the entire evolution through recombination of the constituent semi-groups via a Lie-Trotter theorem for Liouvillian dynamics or some suitable generalisation.

One such generalised decomposition/recombination strategy, for the simulation of arbitrary Markovian dynamics of a qubit, has recently been presented in [10] and as in Hamiltonian simulation it is expected that the class of evolutions which may be simulated using this strategy can still be greatly broadened while the algorithmic costs are simultaneously reduced.

5. Conclusion

We have provided a brief introduction to the field of quantum simulation. In particular we have described the difference between analog and digital quantum simulations and provided a description of how the decomposition/recombination strategy for the simulation of Hamiltonian evolution may be generalised for the simulation of Markovian open quantum systems.

Acknowledgments

This work is based upon research supported by the South African Research Chair Initiative of the Department of Science and Technology and National Research Foundation. Ryan Sweke acknowledges the financial support of the National Research Foundation SARChI program.

References

- [1] Feynman R 1982 *Int. J. Theor. Phys.* **21** 467-488
- [2] Georgescu I M, Ashhab S and Nori F 2014 *Rev. Mod. Phys.* **86** 153
- [3] Nielsen M A and Chuang I L 2000 *Quantum computation and quantum information* (Cambridge University press)
- [4] Sanders B C 2013 *Proc. Lecture Notes in Comp. Sci.* **7948** 1-10
- [5] Watrous J 2008 Quantum Computational Complexity *Preprint* arXiv:0804.3401 [quant-ph]
- [6] Suzuki M 1993 *Quantum Monte Carlo methods in condensed matter physics* (World Scientific)
- [7] Lloyd S 1996 *Science* **273** 1073
- [8] Wolf M M 2012 Quantum channels and operations: guided tour *Lecture note preprint* <http://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf>
- [9] Kliesch M, Barthel T, Gogolin C, Kastoryano M and Eisert J 2011 *Phys. Rev. Lett.* **107** 120501
- [10] Sweke R, Sinayskiy I and Petruccione F 2014 Simulation of single qubit open quantum systems *Preprint* arXiv:1405.6049 [quant-ph]

Entanglement of two distant nitrogen-vacancy-center ensembles under the action of squeezed microwave field

N Teper^{1,2}, L N Mbenza¹, I Sinayskiy^{1,2}, F Petruccione^{1,2}

¹ School of Chemistry and Physics, University of KwaZulu-Natal, Durban, South Africa

² National Institute for Theoretical Physics (NITheP), KwaZulu-Natal, South Africa

E-mail: teper@rambler.ru, sinayskiy@ukzn.ac.za

Abstract. We consider a circuit consisting of two distant nitrogen-vacancy-center ensembles coupled to separate transmission line resonators, which interact by means of a current biased Josephson junction. Our investigation is focused on transitions and dissipation in the Josephson junction leading to entanglement. In our approach the Josephson junction is regarded as a reservoir, whose variables are eliminated from the system dynamics. We include in this scheme also superconducting quantum interference devices, flux-driven Josephson parametric amplifiers, which are the sources of a squeezed microwave field. The entanglement was studied in terms of the logarithmic negativity. The logarithmic negativity was considered for different regimes: weak coupling and strong coupling of transitions of the Josephson junction, and under action of squeezed microwave fields. We show that different degrees, time and duration of entanglement can be reached for various parameters choices.

1. Introduction

We propose a new approach for the physical realization of thermal entanglement of a continuous variables system using spin ensembles. Nitrogen vacancy (NV) centers in diamond attract especial interest because the manipulation, storage, and readout of the quantum information encoded in the different sublevels can be implemented by means of laser and microwave fields [1, 2, 3, 4, 5]. Experimental confirmation of these properties led to engineering of various hybrid circuits, containing nitrogen-vacancy-center ensembles (NVEs), separate transmission line resonators (TRLs) and current biased Josephson junction (CBBJ) [6, 7, 8, 9, 10]. Application of NVE with N spins allows enhancement of the coupling strength by a factor \sqrt{N} , that is especially important for performing of measurement-based quantum computing. For the description of interaction of spin ensembles with external fields collective variables are used. In the low-excitation limit, collective variables of NVE can be described as bosonic modes or harmonic oscillators.

Recently, conditions of squeezing were investigated in the system of two distant NVEs coupled to separate TLRs, which are interconnected by a CBBJ [11]. Our investigation is focused on two-mode thermal entanglement, which can occur due to transitions and dissipation in the CBBJ and TLRs. Also we include in the scheme superconducting quantum interference devices, flux-driven Josephson parametric amplifiers (JPAs), which are the sources of a squeezed microwave field. Recently JPA was described theoretically [12] and also realized experimentally [13]. Squeezed

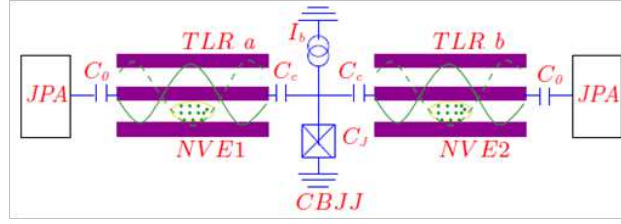


Figure 1. The coupled system of a current biased Josephson junction (CBJJ), two transmission line resonators (TLRs) and Josephson parametric amplifiers (JPAs). Current biased Josephson junction interconnects transmission line resonator a (TLR a) and transmission line resonator b (TLR b) by means of coupling capacitors C_c . I_b is bias current. C_J is the junction capacitor. Josephson parametric amplifiers enhance scheme on left and right sides via coupling capacitors C_0 . Nitrogen-vacancy-center ensemble 1 (NVE 1) and nitrogen-vacancy-center ensemble 2 (NVE 2) are shown inside transmission line resonator a and transmission line resonator b.

microwave field of JPA provides self-squeezing of each bosonic mode in separate TLR that can be important for applications of the continuous-variables approach. Logarithmic negativity was chosen as a measure of entanglement because the continuous-variables approach provides a simple way to find the covariance matrix, which is the basis for the calculation of the logarithmic negativity. In Section 2 we present the scheme, including NVEs, TLRs, CBJJ and JPA, and the model, which is proposed for the description of the interaction. In the section 3 we illustrate the results obtained for the entanglement investigation by means of a calculation of the logarithmic negativity for various parameters. In conclusion, we formulate the necessary conditions for the formation of entangled steady state.

2. Model description

We consider a circuit consisting of two distant nitrogen-vacancy-center ensembles coupled to separate transmission line resonators, which interact by means of a current biased Josephson junction, enhanced by JPAs symmetrically in both parts of system (figure 1).

Action of different types of fields, such as fields of two resonators, squeezed microwave field, classical microwave fields and constant magnetic field, is considered for the description of the dynamics of the scheme. A constant magnetic field removes the degeneracy between the levels of spin states, and induces energy splitting (figure 2). The experimental results [14] show that the resonant frequencies can be approximately equal for all NV centers under particular conditions. Sublevels of each NV center in the both ensembles are coupled by a corresponding mode of TLR with vacuum Rabi frequency g_a and g_b respectively. Simultaneously, NV centers of each ensemble are driven by two classical microwave fields, where Ω_1 and Ω_2 are their Rabi frequencies. In the scheme the logic states 0 and 1 are denoted by corresponding levels: $|0\rangle$ and $|1\rangle$, respectively. Due to the large detuning the coupling is considered perturbatively, using the second-order perturbation theory. The ground state is eliminated from the system dynamics. The interaction of the NV with two fields in such a system can be described by the corresponding Hamiltonian, where $\hbar = 1$

$$\begin{aligned}
 H_I = & \sum_{j=1}^{N_1} \left[\frac{\Omega_1^2}{4\Delta_1} |1\rangle_{1,j}^j \langle 1| + \frac{g_a^2}{\Delta_1} a^\dagger a |0\rangle_{1,j}^j \langle 0| + \left(\frac{\Omega_1 g_a a^\dagger}{2\Delta_1} |0\rangle_{1,j}^j \langle 1| + H.c. \right) \right] \\
 & + \sum_{j=1}^{N_2} \left[\frac{\Omega_2^2}{4\Delta_2} |1\rangle_{2,j}^j \langle 1| + \frac{g_b^2}{\Delta_2} b^\dagger b |0\rangle_{2,j}^j \langle 0| + \left(\frac{\Omega_2 g_b b^\dagger}{2\Delta_2} |0\rangle_{2,j}^j \langle 1| + H.c. \right) \right], \quad (1)
 \end{aligned}$$

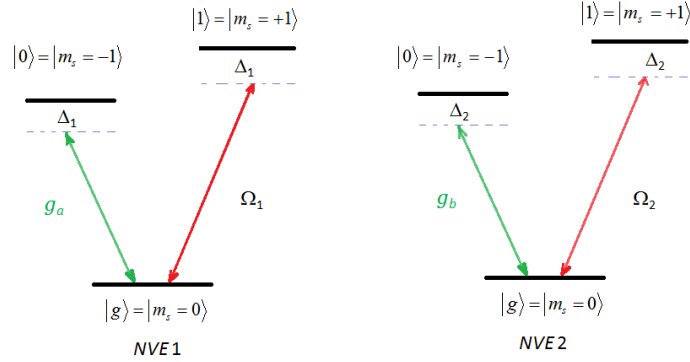


Figure 2. Level structure of single a NV center under the action of an external magnetic field.

where N_1, N_2 are the number of NV centers in the corresponding NVE. We use the collective spin operators [11] for each of two NVE ($i = 1, 2$ denotes ensemble)

$$S_i^- = \sum_{j=1}^{N_i} |0\rangle_{i,j}^j \langle 1|, \quad S_i^+ = \sum_{j=1}^{N_i} |1\rangle_{i,j}^j \langle 0|,$$

and map them into the boson operators $c_i(c_i^\dagger)$ by means of the Holstein-Primakoff transformation

$$\begin{aligned} S_i^- &= c_i \sqrt{N - c_i^\dagger c_i} = \sqrt{N_i} c_i, \\ S_i^+ &= c_i^\dagger \sqrt{N - c_i^\dagger c_i} = \sqrt{N_i} c_i^\dagger, \\ S_i^z &= \left(c_i^\dagger c_i - \frac{N_i}{2} \right). \end{aligned} \quad (2)$$

The effective Hamiltonian can be obtained by neglecting the constant energy terms in H_I

$$H_{\text{eff}} = \widetilde{\Omega}_1 (a^\dagger c_1 + a c_1^\dagger) + \widetilde{\Omega}_2 (b^\dagger c_2 + b c_2^\dagger), \quad (3)$$

where

$$\widetilde{\Omega}_1 = \sqrt{N_1} \frac{\Omega_1 g_a}{2\Delta_1}, \quad \widetilde{\Omega}_2 = \sqrt{N_2} \frac{\Omega_2 g_b}{2\Delta_2}.$$

We model the CBJJ as a two-level artificial atom, considering two lowest levels with frequency of transition $\omega_{10} = 12\text{GHz}$. Such distribution of energy levels is provided by the choice of CBJJ parameters [15, 16, 17]. We denote these two states by $|0\rangle_{\text{CBJJ}}, |1\rangle_{\text{CBJJ}}$. So we consider the two-level system driven by the quantized fields of TLRs, using the rotating-wave approximation. The Hamiltonian describing the interaction of CBJJ with TLRs fields is [15]

$$H_{\text{int}}^{\text{eff}} = \widetilde{g}_a (a \Sigma^+ e^{i\phi} + a^\dagger \Sigma^- e^{-i\phi}) + \widetilde{g}_b (b \Sigma^+ e^{i\theta} + b^\dagger \Sigma^- e^{-i\theta}), \quad (4)$$

where $\widetilde{g}_t = [2C_t(C_J + 2C_c)]^{-1/2} \omega_t C_c \cos \delta$ is the coupling factor, $t = a, b$; δ is a small phase shift on coupling capacitance C_c , which connects TLR and CBJJ; $\Sigma^+ = |1\rangle_{\text{CBJJ}} \langle 0|$, $\Sigma^- = |0\rangle_{\text{CBJJ}} \langle 1|$ are the raising and lowering operators of the CBJJ; $\Sigma_z = |1\rangle_{\text{CBJJ}} \langle 1| - |0\rangle_{\text{CBJJ}} \langle 0|$ is the inversion operator of the CBJJ; ϕ and θ represent phases. Including the expressions for dissipation in the

TRLs and in the CBJJ [15], we describe the transfer processes in the system by means of the master equation

$$\begin{aligned} \frac{\partial \rho}{\partial t} = & -i[H_{\text{eff}}, \rho] - i[H_{\text{int}}^{\text{eff}}, \rho] + k_a(a\rho a^\dagger - \frac{1}{2}a^\dagger a\rho - \frac{1}{2}\rho a^\dagger a) + \\ & + k_b(b\rho b^\dagger - \frac{1}{2}b^\dagger b\rho - \frac{1}{2}\rho b^\dagger b) + \frac{\gamma_\varphi}{2}(\Sigma_z \rho \Sigma_z - \rho) + \\ & + (\frac{\gamma_{10} + \Gamma_1}{4})(\Sigma^- \rho \Sigma^+ - \frac{1}{2}\Sigma^+ \Sigma^- \rho - \frac{1}{2}\rho \Sigma^+ \Sigma^-) - i[V_a, \rho] - i[V_b, \rho], \end{aligned} \quad (5)$$

where γ_φ is the dephasing factor for the considered transition, γ_{10} is the decay factor describing spontaneous emission, Γ_1 is the quantum tunnelling rate, that represents probability of transitions to continuum. The Γ_0 is not taken into account, because it is much smaller than other decay rates, ionization is mostly from the upper levels [17].

The terms $V_a = \beta(a^{\dagger 2}e^{-i\varphi} + a^2e^{i\varphi})$, $V_b = \xi(b^{\dagger 2}e^{-i\varphi} + b^2e^{i\varphi})$ describe squeezed fields of JPA; β, ξ are real amplitudes of squeezed fields, φ is phase of squeezed fields.

Based on the Master equation (5) set of differential equations for observables was obtained, using two different approaches: method of decorrelations and adiabatic elimination of fast variables of CBJJ and TLRs. Solution of this system leads to covariance matrix, which is necessary for the calculation of the logarithmic negativity [18] and for the analysis of the entanglement in our model.

3. Results

To analyse the entanglement between bosonic modes of two separate NVEs we use the logarithmic negativity [18]

$$E_N = -\sum_{i=1}^2 \log_2(\min(1, |\gamma_i|)), \quad (6)$$

where γ_i are symplectic eigenvalues of the partially transposed covariance matrix γ^{T_1} . γ^{T_1} is obtained from the covariance matrix γ by time reversal of the momentum operator of the first system by means of the transformation $\hat{p}_1 \rightarrow -\hat{p}_1$

$$\gamma^{T_1} = P\gamma P, \quad (7)$$

where

$$P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \oplus \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The elements of the covariance matrix γ are calculated, using the found values of NVEs observables $\langle c_1^\dagger c_1 \rangle$, $\langle c_2^\dagger c_2 \rangle$, $\langle c_1 c_2 \rangle$ and other. The symplectic eigenvalues γ_i are calculated as the positive square roots of the usual eigenvalues of $-\sigma\gamma^{T_1}\sigma\gamma^{T_1}$ [18], where

$$\sigma = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

If $\gamma_i \geq 1$ and $E_N = 0$, state is separable.

Figure 3 shows the dynamics of the entanglement for different values of the squeezed fields, which act on both sides of the circuit with the same amplitude. The phase of the squeezed fields and the phase in the term describing coupling transitions of the CBJJ are equal to zero. Thus we consider a symmetric system with equal parameters of dissipation, pumping and coupling for

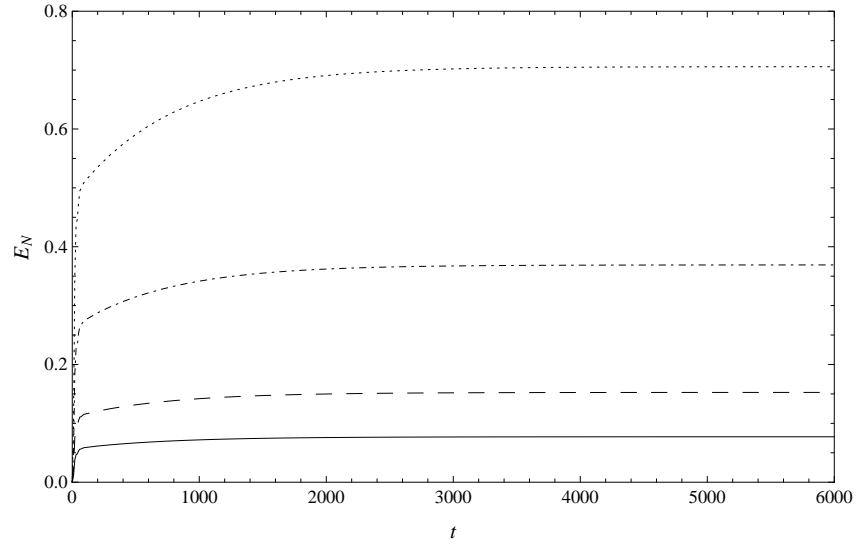


Figure 3. Logarithmic negativity as a function of dimensionless time for $\tilde{g}_a = \tilde{g}_b = 1\text{MHz}$, $\tilde{\Omega}_1 = \tilde{\Omega}_2 = 0.1\text{MHz}$, $\beta = \xi = 1\text{kHz}$ (solid line), 2 kHz (dashed line), 5 kHz (chain line), 10 kHz (dotted line), $\Gamma_1 = 0.1\text{MHz}$, $\gamma_{10} = 0.2\text{MHz}$, $\gamma_\varphi = 0.1\text{MHz}$ and $k_a = k_b = 0.1\text{MHz}$.

both sides of the scheme. One can see significant entanglement for various values of amplitude of the driving field. The growth of entanglement is finalised by stabilization after the steady state is reached. Such a steady state remains entangled. Increasing of the field amplitude leads to growth of entanglement, and direct dependence of the degree of entanglement from the field value is observed. But further amplification of pumping can cause loss of the stationary regime. In such a case the steady state is not formed.

Figure 4 confirms entanglement in the considered system for a range of parameters. One can see that the significant entanglement can be reached also for weak coupling of transitions in the CBJJ and NV centers. In this case dissipation in the CBJJ is the driving force for intermode entanglement. Weak coupling of transitions in the NV center causes an increase of entanglement time. Increasing of transitions coupling in CBJJ provides growth of entanglement, and very weak coupling induces delay in the growth of logarithmic negativity.

4. Conclusion

We investigated thermal entanglement in the system consisting of two distant nitrogen-vacancy-center ensembles coupled to separate transmission line resonators, which interact by means of a current biased Josephson junction. The case of equal parameters for both parts of the circuit was chosen. It was found, that without pumping of squeezed microwave field intermode entanglement is not observed. Entanglement in such a scheme is reached only under the action of the squeezed field. Increasing of field of JPAs leads to growth of the logarithmic negativity, but the amplitude of pumping must be much smaller than the decay rates for a steady state to be formed. If the fields of JPA are weak enough, the entangled steady state is formed. Entanglement can be observed for a large range of parameters. Values of the coupling factors for the transitions in the CBJJ play an important role for the entanglement, the increasing leads to growth of logarithmic negativity. But when these coupling factors are small, dissipation is a driving force, which provides intermode entanglement. The choice of coupling factors for transitions of NV centers allows to control the time of reaching the steady state.

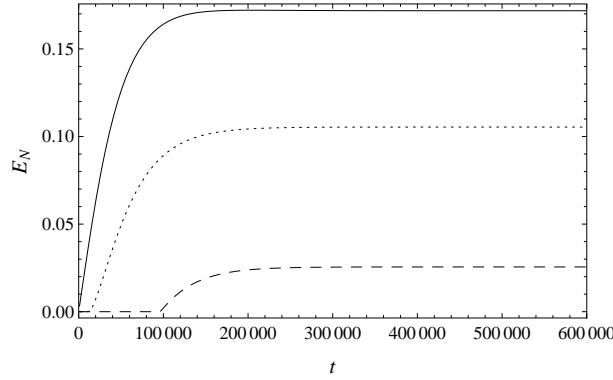


Figure 4. Logarithmic negativity as a function of dimensionless time for $\widetilde{\Omega}_1 = \widetilde{\Omega}_2 = 0.1\text{kHz}$, $\beta = \xi = 0.1\text{kHz}$, $\widetilde{g}_a = \widetilde{g}_b = 6\text{kHz}$ (solid line), 5 kHz (dotted line), 4 kHz (dashed line), $\Gamma_1 = 0.1\text{MHz}$, $\gamma_{10} = 0.2\text{MHz}$, $\gamma_\varphi = 0.1\text{MHz}$ and $k_a = k_b = 1\text{kHz}$.

References

- [1] Childress L, Gurudev Dutt M V, Taylor J M, Zibrov A S, Jelezko F, Wrachtrup J, Hemmer P R and Lukin M D 2006 *Science* **314** 281.
- [2] Gurudev Dutt M V, Jiang L, Togan E, Maze J, Jelezko F, Zibrov A S, Hemmer P R and Lukin M D 2007 *Science* **316** 1312.
- [3] Jiang L, Hodges J S, Maze J R, Maurer P, Taylor J M, Cory D G, Hemmer P R, Walsworth R L, Yacoby A, Zibrov A S and Lukin M D 2009 *Science* **326** 267.
- [4] Neumann P, Beck J, Steiner M, Rempp F, Fedder H, Hemmer P R, Wrachtrup J and Jelezko F 2010 *Science* **329** 542.
- [5] Buckley B, Fuchs G D, Bassett L C, Awschalom D D 2010 *Science* **330** 1212.
- [6] Twamley J and Barrett S D 2010 *Phys. Rev. B* **81** 241202(R).
- [7] Yang W L, Yin Z Q, Hu Y, Feng M, and Du J F 2011 *Phys. Rev. A* **84** 010301(R).
- [8] Yang W L, Hu Y, Yin Z Q, Deng Z J, and Feng M 2011 *Phys. Rev. A* **83** 022302.
- [9] Xiang Z-L, Ashhab S, You J Q, Nori F 2013 *Rev. Mod. Phys.* **85** 623.
- [10] Lü X-Y, Xiang Z-L, Cui W, You J Q, and Nori F 2013 *Phys. Rev. A* **88** 012329.
- [11] Yang W L, Yin Z Q, Chen Q, Chen C Y, and Feng M 2012 *Phys. Rev. A* **85** 022324.
- [12] Ojanen T, and Salo J 2007 *Phys. Rev. B* **75** 184508.
- [13] Zhong L, Menzel E P, Candia R Di, Eder P, Ihmig M, Baust A, Haeberlein M, Hoffmann E, Inomata K, Yamamoto T, Nakamura Y, Solano E, Deppe F, Marx A, and Gross R 2013 *New J. Phys.* **15** 125013.
- [14] Kubo Y, Diniz I, Dewes A, Jacques V, Dréau A, Roch J-F, Auffeves A, Vion D, Esteve D, and Bertet P 2012 *Phys. Rev. A* **85** 012333.
- [15] Hu Y, Xiao Y-F, Zhou Z-W, and Guo G-C 2007 *Phys. Rev. A* **75** 012314.
- [16] Martinis J M, Nam S, Aumentado J, and Urbina C 2002 *Phys. Rev. Lett.* **89** 117901.
- [17] Yu Y, Han S, Chu X, Chu S-I, Wang Z 2002 *Science* **296** 889.
- [18] Plenio M B, Hartley J and Eisert J 2004 *New J. Phys.* **6** 36.

Vacuum energies and frequency dependent interactions

Herbert Weigel

Physics Department, Stellenbosch University, Matieland 7602, South Africa

E-mail: weigel@sun.ac.za

Abstract. The vacuum polarization (or Casimir) energies of field configurations whose interaction with the quantum fluctuations is frequency independent are straightforwardly computed from scattering data since there is a simple relation between the frequency derivative of the scattering phase shift and a spatial integral of the Green's function at coincident points. In more complicated, but nevertheless typical frameworks, the interaction of the quantum fluctuations is frequency dependent and this relation must be modified. This modification may or may not additionally contribute to the vacuum polarization energy. Here we will consider three examples that naturally induce frequency dependent interactions. (I) Scalar electrodynamics with a static background potential. (II) An effective theory that emerges from integrating out a heavy degree of freedom. (III) Quantum electrodynamics coupled to a frequency dependent dielectric material. In cases (II) and (III) any omission of the frequency dependence violates the renormalizability of the theory. For case (III) an ambiguity arises because the introduction of a dielectric function comes at the expense of lacking a canonical Lagrangian formulation for the interaction of the photons with the constituents of the material. The physically motivated choice for the Hamiltonian leads to an attractive self-stress of a dielectric sphere.

1. Introduction

In quantum field theory the vacuum polarization (sometimes called Casimir[1]) energy refers to the change of the zero point energies due to the interaction of the quantum fluctuations with a (static) background $V(\mathbf{x})$. This background can be generated by a localized solution to the classical field equations, or it can model an interaction with some material, or it can mimic boundary conditions on the quantum fluctuations in some singular limit[2].

The energy eigenvalues emerge from a Schrödinger type wave-equation

$$[-\nabla^2 + V(\mathbf{x})] \psi_\nu(\mathbf{x}) = k_\nu^2 \psi_\nu(\mathbf{x}), \quad (1)$$

with $\omega_\nu = \frac{k_\nu^2}{2m}$ or $\omega_\nu = \sqrt{k_\nu^2 + m^2}$ for non-relativistic or relativistic dispersion relations, respectively. Here m is the mass of the quantum field and natural units ($\hbar = 1$, $c = 1$) are adapted. Let $\omega_\nu^{(0)}$ be the energy eigenvalues of the non-interacting problem, *i.e.* when $V \equiv 0$. Then the vacuum polarization energy is

$$E_{\text{vac}} = \frac{1}{2} \sum_\nu [\omega_\nu - \omega_\nu^{(0)}] + E_{\text{ct}}, \quad (2)$$

where E_{ct} is the counterterm contribution that cancels the ultra-violet divergence in the sum unambiguously. This equation is not suitable for calculations. First we have to distinguish between bound state and (continuum) scattering solutions to the wave-equation (1). The bound states energies are simply ω_n while the scattering solutions, whose energies are given by a continuous function $\omega(k)$, are characterized by the S -matrix or the phase shift $\delta(k)$ for a single mode¹. Properly identifying these contributions leads to the phase shift formula

$$E_{\text{vac}} = \frac{1}{2} \sum_{n=\text{b.s.}} |\omega_n| + \int \frac{dk}{2\pi} \omega(k) \frac{\partial}{\partial k} \delta(k) + E_{\text{ct}} \quad (3)$$

This formula can be derived/motivated in various ways. For example, it is popular to express the continuum contribution as the integral over frequencies weighted by change of the density of states induced by the background. This change is computed by the derivative of the phase shift according to the Krein–Friedel–Lloyd (KFL) formula; see Ref.[3]. However, there is a more rigorous approach based on the vacuum matrix element of the energy momentum tensor and Green’s function methods[4]. In that calculation the frequency derivative of the wave-equation (1)

$$\frac{\partial}{\partial k} \delta(k) = \int dx \operatorname{Im} \left\{ G(x, x, k) \frac{\partial}{\partial k} [k^2 - V(\mathbf{x})] - 2k G^{(0)}(x, x, k) \right\}, \quad (4)$$

proves essential. Here $G(x, y, k)$ is the Green’s function in the considered channel. It is hence very suggestive that a more complicated interaction with the background, that in particular is frequency dependent, will give rise to modifications of the phase shift formula (3). The main purpose of this study is to investigate possibilities for such modifications.

Before doing so, we must rephrase (3) in a form pertinent for computations. Individual contributions in (3) require regularization which can only be safely removed upon their combination. Hence it is necessary to synchronize the regularization procedure. This is achieved by noting that the Born series for scattering data and the Feynman series are both expansions in powers of the background $V(\mathbf{x})$. Thus we subtract the N leading terms of the Born from the phase shift and add back their contribution as Feynman diagrams:

$$E_{\text{vac}} = \frac{1}{2} \sum_{n=\text{b.s.}} |\omega_n| + \int_0^\infty \frac{dk}{2\pi} \omega(k) \left[\frac{\partial}{\partial k} \delta(k) \right]_N + E_{\text{FD}}^{(N)} + E_{\text{ct}}. \quad (5)$$

Since the Born series well approximates the phase shifts at large k , a small number N suffices to make the integral finite. The sum $E_{\text{FD}}^{(N)} + E_{\text{ct}}$ can then be straightforwardly treated by standard renormalization techniques of perturbative quantum field theory. This interesting topic will not central to the present study. But it is important to remember that there is a consistent procedure to render the frequency integrals finite.

2. Electric Potential in Scalar Electrodynamics

The prototype for a frequency interaction with a static background is the electric potential in scalar electrodynamics in one space dimension. In this case the gauge potential is $A_\mu = (V(\mathbf{x}), 0)_\mu$ and the local Lagrangian reads

$$\mathcal{L} = (D_\mu \Phi)^* (D^\mu \Phi) - M^2 \Phi^* \Phi \quad \text{with} \quad D_\mu = \partial_\mu + iA_\mu. \quad (6)$$

The canonical approach to this model is unconventional but straightforward; it can be formulated as an application of the random phase approximation [5]. After Fourier transforming the

¹ Because of the dispersion relation $\omega = \omega(k)$ we synonymously use ω and k for the frequency.

time-dependence, the field equations turn into wave-equations for bound states $\psi_n(x)$ with discrete energies $|\omega_n| < M$, and for scattering solutions $\psi_k^{(\pm)}(x)$ with energies $\pm\omega_k$ (where $\omega = \omega(k) = \sqrt{k^2 + M^2} > 0$),

$$\left[-\partial_x^2 - V(x)^2 + 2\omega_n V(x) + M^2 \right] \psi_n(x) = \omega_n^2 \psi_n(x) \quad (7)$$

$$\left[-\partial_x^2 - V(x)^2 \pm 2\omega V(x) + M^2 \right] \psi_k^{(\pm)}(x) = \omega^2 \psi_k^{(\pm)}(x). \quad (8)$$

Since the theory is not CP-invariant, the existence of a bound state with energy ω_n does not imply the existence of a corresponding bound state with energy $-\omega_n$, and likewise $\psi_k^{(+)}(x) \neq \psi_k^{(-)}(x)$ for the scattering states. Despite of additional time derivatives, canonical momenta are well defined and can be made subject to the conventional equal time commutation relations. It is also possible to construct a Fock space. We refer to Ref.[6] for details of the calculation. Here we just list the result for the vacuum matrix element of the energy density with the continuum contribution expressed in terms of the Green's function

$$u(x) = \sum_n \omega_n \psi_n(x) [\omega_n - V(x)] \psi_n(x) + \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \omega \operatorname{Im} \left\{ [\omega - V(x)] G(x, x, \omega) - \omega^2 G^{(0)}(x, x, \omega) \right\}. \quad (9)$$

If, as customary, we interpret the imaginary part of the Green's function as the density of states, we immediately recognize that the energy (density) is not given as the single particle energy times the density of states. Recall that this reasoning was fundamental in deriving the phase shift formula from the KFL relation. However, it is obvious that the operation in (4) produces the same coefficient of the Green's function as in (9) since, according to (8) we have to replace $V(\mathbf{x}) \rightarrow V(x)^2 - 2\omega V(x)$ before applying the frequency derivative to the wave-equation. As a consequence, the modification in the energy density is exactly compensated by the change in the relation between the derivative of the phase shift and the Green's function. It results in the validity of the phase shift formula for this model.

The interaction Hamiltonian contains time derivatives which spoils the simple relation that the interaction Hamiltonian is the negative interaction Lagrangian. Nevertheless the corresponding relation between the vacuum polarization energy and the quantum action, $S[A_\mu]$

$$E_{\text{vac}} = -\frac{1}{T} (S[A_\mu] - S[0]), \quad (10)$$

where T is some large time interval, holds.

3. Effective Model from a Heavy Particle

In this section, we consider a model of two scalar boson fields that are coupled via a static but space-dependent background $V(\mathbf{x})$,

$$\mathcal{L} = \frac{1}{2} \partial_\mu \Phi \partial^\mu \Phi - \frac{M^2}{2} \Phi^2 + \frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi - \frac{m^2}{2} \varphi^2 + V(\mathbf{x}) \Phi \varphi. \quad (11)$$

We assume that $M \gg m$ for the masses of the scalar fields. Eventually we want to consider an effective theory for φ which inherits a frequency dependent interaction with $V(\mathbf{x})$ from integrating out the heavier field Φ . From the Lagrangian, (11), we obtain the field equations $(\partial^2 + M^2) \Phi = V(\mathbf{x}) \varphi$ and $(\partial^2 + m^2) \varphi = V(\mathbf{x}) \Phi$ and the energy density

$$T^{00} = \frac{1}{2} \left(\dot{\Phi}^2 - \Phi \ddot{\Phi} + \dot{\varphi}^2 - \varphi \ddot{\varphi} \right) + \text{total space derivatives}, \quad (12)$$

where we have used the field equations to produce the total derivative terms.

Since the interaction is static we may consider frequency modes $\Phi_\omega(\mathbf{x}) = \int dt \Phi(t, \mathbf{x}) e^{i\omega t}$ and $\varphi_\omega(\mathbf{x}) = \int dt \varphi(t, \mathbf{x}) e^{i\omega t}$. Then the wave-equations become

$$\begin{aligned} -\nabla^2 \Phi_\omega(\mathbf{x}) &= (\omega^2 - M^2) \Phi_\omega(\mathbf{x}) + V(\mathbf{x}) \varphi_\omega(\mathbf{x}), \\ -\nabla^2 \varphi_\omega(\mathbf{x}) &= (\omega^2 - m^2) \varphi_\omega(\mathbf{x}) + V(\mathbf{x}) \Phi_\omega(\mathbf{x}). \end{aligned} \quad (13)$$

Assuming that the heavier field only varies slowly in space defines the local approximation $\Phi_\omega \approx V \varphi_\omega / (M^2 - \omega^2)$ in which the lighter field is subject to

$$-\nabla^2 \varphi_\omega \approx (\omega^2 - m^2) \varphi_\omega - \frac{V(\mathbf{x})^2}{\omega^2 - M^2 + i\epsilon} \varphi_\omega. \quad (14)$$

Here we have reinstalled the Feynman pole prescription from the defining functional integral. From this wave-equation we derive an orthogonality relation for modes of different frequencies

$$\int d^d x \varphi_\omega(x) \left[1 + \frac{V(\mathbf{x})^2}{(\omega^2 - M^2 + i\epsilon)(\omega'^2 - M^2 + i\epsilon)} \right] \varphi_{\omega'}(x) = 0 \quad \text{for} \quad \omega \neq \omega'. \quad (15)$$

In the same way, we can use the local approximation for Φ_ω directly in the energy density (12)

$$T^{00}(\mathbf{x}) = \int \frac{d\omega}{2\pi} T_\omega^{00}(\mathbf{x}) \equiv \int \frac{d\omega}{2\pi} \omega^2 \left[1 + \frac{V(\mathbf{x})^2}{(\omega^2 - M^2 + i\epsilon)^2} \right] \varphi_\omega(\mathbf{x})^2, \quad (16)$$

where we have omitted contributions that vanish upon spatial integration. Several remarks on this result are in order. First, the straightforward substitution of the frequency decomposition into (12) yields a double frequency integral over ω and ω' . It is only by the orthogonality condition, (15) that the off-diagonal parts, which are explicitly time dependent, vanish. Second, the energy $\int d\mathbf{x} T^{00}(\mathbf{x})$ is conserved by the wave-equation (14) in the local approximation and no further input from the full model is required. Third, the kernel of the frequency integral for $T^{00}(\mathbf{x})$ obviously relates to the frequency derivative of the wave-equation (14). It is exactly this derivative that is essential in the derivation, see (4), of the phase shift formula for $\omega^2 = k^2 + m^2$. In total this shows that even in a model in which the frequency dependence is not polynomial and canonical methods do not apply, the phase shift formula for the vacuum polarization energy remains valid. More details on this calculation as well as a numerical comparison of the vacuum polarization energies between the full model and its local approximation in one space dimension can be found in ref.[6].

4. Effective Model for a Dielectric Sphere

A particular interesting topic in the context of vacuum polarization energies is the Casimir self-stress of a dielectric sphere. Some time ago it has been claimed[7] to represent an exception from the rule[8] that Casimir forces are attractive. In this context the quantum fluctuations are photons and the background is a radially function that is strongly peaked at the radius of the sphere. The interaction between the photons and the background occurs via the frequency dependent dielectric function of the material.

The point of departure for this study are Maxwell's equations for a dielectric without sources

$$\nabla \cdot \mathbf{D} = 0, \quad \nabla \times \mathbf{E} + \partial_t \mathbf{B} = 0, \quad \nabla \cdot \mathbf{B} = 0 \quad \text{and} \quad \nabla \times \mathbf{B} - \partial_t \mathbf{D} = 0. \quad (17)$$

The dielectric nature enters through the convolution between the displacement and the electric field: $\mathbf{D}(t, \mathbf{x}) = \int dt' \epsilon(t', \mathbf{x}) \mathbf{E}(t - t', \mathbf{x})$. Introducing again frequency modes and assuming a

spherically symmetric dielectric $\epsilon_k(r) = \int dt e^{ikt} \epsilon(t, \mathbf{x})$ leads to the so-called Mie-model[9, 10]. Then Maxwell's equations (17) are transformed into decoupled second order differential equations for scalar fields $\varphi_k(\mathbf{x})$ and $\phi_k(\mathbf{x})$ by the ansätze

$$\begin{aligned} \text{TE :} \quad \mathbf{E}_k(\mathbf{x}) &= k \nabla \times [\varphi_k(\mathbf{x}) \mathbf{x}] \quad \text{and} \quad \mathbf{B}_k(\mathbf{x}) = i \nabla \times (\nabla \times [\varphi_k(\mathbf{x}) \mathbf{x}]) \\ \text{TM :} \quad \mathbf{B}_k(\mathbf{x}) &= -ik \nabla \times [\phi_k(\mathbf{x}) \mathbf{x}] \quad \text{and} \quad \mathbf{E}_k(\mathbf{x}) = \frac{-1}{\epsilon_k(r)} \nabla \times (\nabla \times [\phi_k(\mathbf{x}) \mathbf{x}]), \end{aligned} \quad (18)$$

for the transverse electric (TE) and transverse magnetic (TM) modes, respectively. The scalar functions obey simple differential equations that can be cast into

$$(k^2 + \nabla^2) \varphi_k(\mathbf{x}) = U_k^{(TE)} \varphi_k(\mathbf{x}) \quad \text{and} \quad (k^2 + \nabla^2) \phi_k(\mathbf{x}) = U_k^{(TM)} \phi_k(\mathbf{x}). \quad (19)$$

The explicit expressions for the potential in terms of the dielectric function are given in ref.[11].

Assuming that the dielectric function has a power expansion in k^2 , which in coordinates space is $-\partial_t^2$, it is possible to identify a conserved energy from Maxwell's equations (17). A single frequency mode contributes the spatial integral of

$$u_k(\mathbf{x}) = \frac{1}{2} \frac{\partial [k \epsilon_k(r)]}{\partial k} \mathbf{E}_k^2 + \frac{1}{2} \mathbf{B}_k^2 \quad (20)$$

to the spatial density of that energy. Expressing this energy density in terms of the scalar functions ϕ_k and φ_k and using the respective wave-equations to eliminate surface terms (in coordinate space) shows that the metric functions M_k in $u_k^{(TE)} = \frac{1}{2} M_k^{(TE)} \phi_k^2$ and $u_k^{(TM)} = \frac{1}{2} M_k^{(TM)} \varphi_k^2$ are indeed the frequency derivatives $M_k = \frac{k}{2} \frac{\partial}{\partial k} (k^2 - U_k)$ in both the *TE* and *TM* channels. These are the frequency derivatives of the wave-equations (19)[6]. Hence the phase shift formula does not get modified for this effective model.

However there are problems with this effective model. The assumption that ϵ_k is even in k is unphysical as it violates the Kramers–Kronig relations. The analog construction that allows for odd contributions requires ϵ_k to be a tensor with anti-symmetric components. However, by general arguments based on the Onsager relations in statistical mechanics, ϵ_k is established to be a symmetric tensor[12]. In that context it is more appropriate to study the free energy[13]. Then one considers the analytic continuation ($k = i\kappa$) of the standard energy density

$$\bar{u}_{i\kappa}(\mathbf{x}) = \frac{1}{2} \epsilon_{i\kappa}(r) \mathbf{E}_{i\kappa}^2 + \frac{1}{2} \mathbf{B}_{i\kappa}^2 \quad (21)$$

as the contribution of a single mode to the vacuum polarization energy. Formally it adds

$$\Delta u = \int_0^\infty \frac{d\kappa}{\pi} \sum_{\ell=1}^\infty (2\ell+1) \kappa^2 \int_0^\infty dr \left(\epsilon_{i\kappa}(r) - 1 - \frac{1}{2\kappa} \frac{dU_{i\kappa}(r)}{d\kappa} \right) |\psi_{i\kappa,\ell}(r)|^2 \quad (22)$$

to the unrenormalized energy. Note that Δu implies the sum over *TE* and *TM* modes and the $\psi_{i\kappa,\ell}$ are the analytic continuations of the radial parts of the angular momentum (ℓ) decompositions of ϕ_k and φ_k . The renormalization and the numerical computations have been performed in ref.[11]. The result is shown in figure 1.

While the naive application of the phase shift formula, *i.e.* adopting the energy functional from (20), indeed produces self-repulsion for the dielectric sphere as suggested some time ago[7] the additional term form (22) overcomes this contribution predicting an attraction in total.

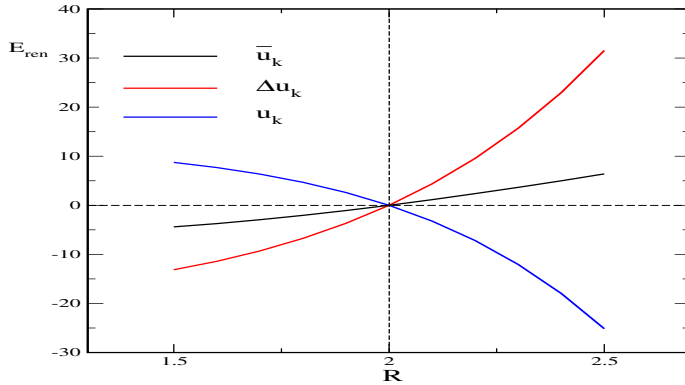


Figure 1. Renormalized vacuum energy of the dielectric sphere of radius R . The legend associates the kernels for the energy integrals: u_k from (20), Δu from (22) and \bar{u} from (21) is the total energy. Units are such that the plasma frequency of the dielectric is $\lambda = 2$.

5. Conclusion

In a rigorous field theory approach the phase shift formula for the vacuum polarization energy remains valid even when the interaction is frequency dependent. Here we have only considered a few examples. However, once the wave-equations produce a conserved energy functional (which may not always be the case) its kernel in frequency space is related to the frequency derivative of the wave-equations for the following consistency condition: This derivative features in the orthogonality condition for the solutions to the wave-equation and in this way ensures that the terms which are explicitly time dependent disappear from the total energy. On the other hand it is exactly this derivative that enters the derivation of the phase shift formula via the Wronskian between the regular and Jost solutions. Hence any additional term in that Wronskian identically appears in the energy functional thereby maintaining the phase shift formula for the vacuum polarization energy.

In particular cases effective theories require input from outside local quantum field theory. Then the phase shift formula is indeed modified. For the particular case of a dielectric shell this modification turns the self-stress from being repulsive to attractive.

Acknowledgments

This presentation is based on an ongoing collaboration with N. Graham and M. Quandt. Their input is highly appreciated. Support from the National Research Foundation (NRF), under Ref. No. IFR1202170025 is acknowledged.

References

- [1] Casimir H B G 1953 *Physica* **XIX** 846–849
- [2] Graham N, Quandt M and Weigel H 2009 *Lect.Notes Phys.* **777** 1
- [3] Faulkner J S 1977 *J. Phys. C: Solid State Phys.* **10** 4461
- [4] Graham N, Jaffe R L, Khemani V, Quandt M, Scandurra M and Weigel H 2002 *Nucl. Phys.* **B645** 49–84
- [5] Ring P and Schuck P 1980 *The Nuclear Many Body Problem* (Springer)
- [6] Graham N, Quandt M and Weigel H 2014 *Phys.Rev.* **D90** 085004 (*Preprint* 1406.0748)
- [7] Boyer T H 1968 *Phys. Rev.* **174** 1764–1774
- [8] Kenneth O and Klich I 2008 *Phys.Rev.* **B78** 014103 (*Preprint* 0707.4017)
- [9] Mie G 1908 *Ann. Phys.* **25** 377
- [10] Newton R G 1982 *Scattering Theory of Waves and Particles* (Springer, New York)
- [11] Graham N, Quandt M and Weigel H 2013 *Phys.Lett.* **B726** 846–849
- [12] Melrose D and McPhedran R 1991 *Electromagnetic Processes in Dispersive Media* (Cambridge University Press) chap 7
- [13] Rahi S J, Emig T, Graham N, Jaffe R L and Kardar M 2009 *Phys.Rev.* **D80** 085021