Proceedings of The 60th annual conference of The South African Institute of Physics (SAIP2015)

29 June - 3 July 2015

Boardwalk Convention Centre
Port Elizabeth

Editors:
Makaiko Chithambo (Rhodes University)
and
André Venter (Nelson Mandela Metropolitan University)

Nelson Mandela Metropolitan University

SOUTH AFRICAN INSTITUTE OF PHYSICS

RHODES UNIVERSITY
Where leaders learn

www.saip.org.za

PROCEEDINGS EDITORS:
Makaiko Chithambo and André Venter

PROCEEDINGS MANAGEMENT TEAM AT SAIP:
Juan Grey
Roelf Botha

PUBLISHER: The South African Institute of Physics (SAIP)

Tel. +27 (0)12 841 2627,
Fax +27 (0)86 648 8474,
E-mail secretary@saip.org.za.

ISBN: 978-0-620-70714-5

SAIP2015

The Proceedings of The 60th Annual Conference of The South African Institute of Physics (SAIP2015)

Jointly hosted by Nelson Mandela Metropolitan University and Rhodes University

29 June to 3 July 2015
Broadwalk Convention Centre
Port Elizabeth
South Africa

Editors:
Makaiko Chithambo and André Venter

Table of Contents

Conference Chairs and Committees . . . . . *ix*
Message from the Conference Chairpersons . . . . . *x*
Message from the Review Committee . . . . . *xii*
List of Reviewers . . . . . *xiii*

FULL RESEARCH PAPERS

Division A – Division for Physics of Condensed Matter and Materials

Electronic and Optical Properties of monolayer MX2(M= Zr, Hf; X= S, Se) from first principles calculations 2
Mahmud Abdulsalam, D. N. G. Moise and Daniel P. Joubert

Investigation of the introduction and annealing behaviour of the donor-vacancy complex in alpha-particle irradiated germanium. 8
Abraham W. Barnard, Walter E.Meyer, F. Danie Auret, Phuti N.M. Ngoepe and Sergio M.M. Coelho

Laser excitation studies and crystal-field analysis of ZnO:Tb3+ and ZnO:Eu3+ powders. 14
N A Bhebhe, M Mujaji, D Wamwangi, L F Koao and F B Dejene

Methane dissociation over Pd nanocluster: DFT study 20
M.H. Chuma, H.R. Chauke, G. Jones and P.E. Ngoepe

Effect of background gas and substrate temperature on ZnO:Zn PLD thin films 26
E Hasabeldaim, O M Ntwaeaborwa, R E Kroon and H C Swart

Structural and magnetic properties of NiFe2O4/NiFe bi-magnet and NiFe nano-alloy synthesized from thermal reduction of NiFe2O4. 32
Itegbeyogene P Ezekiel, Thomas Moyo and Hafiz M I Abdallah

Electronic and magnetic properties of the (Cr84Re16)100−yMny alloy system 38
BS Jacobs, AD Faceto, CJ Sheppard, ARE Prinsloo, PC de Camargo and AJA de Oliviera

Effect of pH on the properties of ZnO nanostructures prepared by chemical bath deposition method 43
L F Koao, F B Dejene and H C Swart

Phototransferred thermoluminescence andphosphorescence related to phototransfer in annealedsynthetic quartz 49
EFM Kombe-Atang and ML Chithambo

Structure prediction of manganese dioxide nanoclusters using computer simulation techniques. 55
R R Maphanga, P E Ngoepe, C R A Catlow and S M Woodley

Synthesis, structural and optical characterisation of cobalt and indium co-doped ZnO nanoparticles 61
MW Maswanganye, KE Rammutla, TE Mosuang, BW Mwakikunga, TS Bertrand and M Maaza

The modified interatomic potentials of FeS2 in atomistic simulations. 67
M AMehlape, P E Ngoepe and S C Parker

Néel temperature dependence of Cr + 1 at.% Al thin films on MgO(110), MgO(100) and fused silica 73
ZP Mudau, ARE Prinsloo, CJ Sheppard, AM Venter and EE Fullerton

Characterisation of traditional ceramic materials used in the Sotho culture (South Africa) for clay pot making 79
A F. Mulaba-Bafubiandi and P X D A Hlekane
Iron bearing minerals characterised with Mossbauer spectroscopy at the Mineral Processing and Technology Research Centre, University of Johannesburg, South Africa 84
A F Mulaba – Bafubiandi
Electronic and Magnetic Structure of Cr doped Rutile and Anatase TiO2; an ab-initio DFT+U study 91
Winfred M. Mulwa, Cecil N.M.Ouma, Martin O. Onani, Francis B. Dejene
Exciton energies of chalcopyrites AgAlX2(X=S,Se,Te) from GW and BSE calculations 97
GM Dongho Nguimdo and Daniel P Joubert
Collective Electronic Excitations in Ferromagnetic Metals 102
V Nolting
Dependence of the photoionization cross-section of alpa-Al2O3:C on the measurement temperature 108
A. Nyirenda, M.L. Chithambo
Simultaneous substitution of Ba, Mn and Co into Fe3O4 spinel structure: magnetic and electrochemical sensing properties of the synthesized nanoparticles. 114
Nadir S E Osman, Neeta Thapliyal, Thomas Moyo and Rajshekhar Karpoormath
An ab-initio study of the metastability of the boron-vacancy complex in silicon 119
Cecil N M Ouma and Walter E Meyer
Physical properties of Cr78Al22 thin films 125
ARE Prinsloo, CJ Sheppard, HA Derrett, N van den Berg and EE Fullerton
Electrical characterization of undoped and niobium-doped n-type silicon diodes. 131
M J Thebe, S J Moloi and M Msimanga
Electrical characterisation of 5.4 MeV alpha-particle irradiated; low doped, n-type gallium arsenide 136
S M Tunhuma ,F D Auret ,M J Legodi and M Diale
A quantitative evaluation of the depth resolution of AES depth profiling of Cu/Ni multilayer thin films using the MRI model. 141
X.L.Yan, Y. Liu, H.C. Swart, J.Y. Wang, J.J. Terblans
Study of the interdiffusion in Cu/Ni multilayer thin films by Auger electron spectroscopy depth profiling. 147
X.L.Yan, Y. Liu, H.C. Swart, J.Y. Wang, J.J. Terblans

Division B – Nuclear, Particle and Radiation Physics

High-momentum particle production at hadron colliders 154
D M Adamiak, W A Horowitz
Online energy reconstruction on ARM for the ATLAS TileCal sROD co-processing unit . . . 160
Mitchell A. Cox and Bruce Mellado
Measuring the orientation of the Ge Crystals of the iThemba LABS Segmented Clover Detector. 165
T. S. Dinoko, T. D. Bucher, E. A. Lawrie, J. Easton, N. Erasmus, S. P. Noncolela, O. Shirinda

Analysis and quantification of modelling errors introduced in the deterministic calculational path applied to a 2-node problem . . . 169
T.P. Gina, S.A. Groenewald , W.R. Joubert , S.H. Connell
Jet substructure: a discovery tool . . . 175
Deepak Kar
Ambient gamma dose rate measurements at Manyoni uranium deposits, Singida, Tanzania . . 180
K F Kasoga, D A Mwalongo, S F Sawe, M M Nyaruba and U Dammalapati
An electronics test-bench for the certification of the Tile Calorimeter of the ATLAS detector . . . 186
C O Kureba, X Ruan, M Spoor, M Govender, I Hofsajer, B Mellado and C Sandrock
The search for Dark Matter in association with the Higgs boson with the di-photon decay . . . 192
C. O. Kureba, X. Ruan and B. Mellado
Data processing at The South African Nuclear Energy Corporation SOC Ltd (Necsa) neutron diffraction facility . . . 198
D Marais, A M Venter and J Markgraaff
Status of the measurements of Higgs boson properties with the ATLAS detector . . . 204
Bruce Mellado
Probing new physics in the Higgs sector with effective field theories at the Large Hadron Collider . . . 210
Bruce Mellado, Luis March, Xifeng Ruan
Efficient processing of physics quantities for the Processing Unit for the upgrade of the TileCalorimeter of ATLAS . . . 216
D Ohene-Kwofie, E Otoo and B Mellado
A study of radiation damage in plastic scintillators using magnetic resonance techniques for the upgrade of the ATLAS detector . . . 222
C. Pelwan, H. Jivan, D. Joubert, J. Keartland, M. Madhuku, B. Mellado, G. Peters, K. Sekonya, E. Sideras-Haddad
A Di-Higgs Search in the gamma-gamma-beta-beta' Decay Channel Using the ATLAS Detector . . . 226
Robert G. Reed, Stefan Von Bruddenbrock, Deepak Kar, Bruce Mellado
A Framework Tool for Integrating the Back-End Infrastructure in the ATLAS Detector Control System . . . 232
Robert Graham Reed, Bruce Mellado
GPU-based Computation of Energy and Time for the Upgrade of the Tile Calorimeter of the ATLAS Detector . . . 238
Marc Sacks, Bruce Mellado

Division C – Photonics

Gene expression changes in diabetic wound healing as induced by photobiostimulation in vitro . . . 244
S M Ayuk, H Abrahamse and N N Houreld
Synchrotron modeling of the gamma-ray to optical afterglow of GRB 130427A and expected neutrino flux . . . 249
Jessymol K Thomas, Reetanjali Moharana and Soebur Razzaque

Novel zincate phosphors: A new red emitting phosphor for LED applications 255
Vijay Kumar, Sudipta Som, Somrita Dutta, Hendrik C Swart
Investigated charge carrier effects in silicon membranes using a femtosecond laser 261
WI Ndebeka, PH Neethling, CM Steenkamp, H Stafast, EG Rohwer
Fluorescence behaviour of Eu doped Gd2O3 nanosheets via CuO incorporation 267
Anurag Pandey, R. E. Kroon and H. C. Swart
A Nonlinear Optical loop Mirror enhanced three wavelength oscillations Erbium doped fiber laser source based on Fiber Bragg Grating reflectors 273
S. Qhumayo and R. Martinez

Division D1 – Astrophysics

Very-high energy emission from pulsars 278
M Breed, C Venter and A K Harding
Spectral studies of aring quasar PKS 1424-418 above 100 MeV with Fermi-LAT 284
Feraol F. Dirirsa, Richard J. Britto and Soebur Razzaque
The Vela Supercluster - does it provide the missing link to explain the local ow fields? 290
A Elagali, R C Kraan-Korteweg and M E Cluver
A 'road test' of ANOVA versus DFT and LS as a period-finding algorithm 296
C A Engelbrecht and F A M Frescura
Optical spectroscopic observations of unclassified Active Galactic Nuclei in the Fermi-2LAC catalogue 302
L Klindt, B van Soelen, PJ Meintjes and P Vaisanen
A Timing Noise Analysis Pipeline for HartRAO pulsars applied to PSR J1326-5859 308
Jacques Maritz, Pieter Meintjes, Sarah Buchner, Natalia Lewandowska
Possible extragalactic astrophysical counterparts of IceCube neutrino events 314
Reetanjali Moharana, Richard J. Britto, Soebur Razzaque
A comparative study of the three empirical solar models in North West Province, South Africa 320
T.S. Mulaudzi , N.E. Maluta , V Sankaran and F Nemangwele
A quasi-periodicity in the optical polarization of the blazar PKS 2155-304 326
N.W. Pekeur, R. Taylor, S.B. Potter and R.C. Kraan-Korteweg
Optical Observations of the Be/X-ray Binary A0538-66 332
A F Rajoelimanana, P A Charles, P J Meintjes and L J Townsend
Rotation Curves and Bars: Accounting for Non-circular Motions in Barred Spiral Galaxies . . 338
Toky H. Randriamampandry, Claude Carignan, Francoise Combes and Nathan Deg
Studying stellar populations of luminous red galaxies to probe the Hubble parameter H(z) . . . 344
A. L. Ratsimbazafy, C. M. Cress, S. M. Crawford, M. Smith
Implementation of a goodness-of-fit test for finding optimal concurrent radio and -ray pulsar light curves 350
A S Seyffert, C Venter, A K Harding, J Allison and W D Schutte
Numerical modelling of hydrodynamical astrophysical outflows: application using the PLUTO code 356
I P van der Westhuizen, B van Soelen, P J Meintjes, S J P K Riekert and J H Beall

Optical spectroscopy of PSR B1259-63 around the 2014 periastron passage 362
B van Soelen, P Vaisanen, I Sushch, L Klindt, PJ Meintjes, A. Odendaal and R. Armstrong
The unusually strong coronal emission lines of SDSS J1055+5637 366
Hartmut Winkler

Division D2 – Space Science

Modelling ground conductivity for computing the electric field associated with geomagnetically induced currents using the Finite Element Method. (A mid latitude case study) 373
E Matandirotya, P J Cilliers, and R R Van Zyl
An integrated software based analytical model for the signal path efficiency of the HartRAO lunar laser ranger optical system 379
S C Ndlovu, L Combrinck, N P Nkosi and R C Botha

Division E – Education

The light bulb effect: University students' problem solving cognitive processes in a physics problem solving skills test 386
C Albers, D Clerk and D Naidoo
Effect of guided inquiry laboratory activities on first-year physics students' views on the nature of science 392
V M Baloyi, W E Meyer and E Gaigher

Division F – Applied Physics

Performance Comparison between the Traditional Intensity Modulation Direct Detection and Coherent Detection in a High Speed Optical Fibre Communication System 399
T V Chabata, D Kiboi Boiyo, E K Rotich Kipnoo, R R G Gamatham, A W R Leitch, T B Gibbon
Monte Carlo based estimation of the effect of different aerosol classes on solar irradiance in African atmospheric conditions 404
Marie Chantal Cyulinyana and Hartmut Winkler
Characteristics and functions of the South African national measuring standard for force . . . 410
S T S Dlamini
Estimation of energy production decrease due to shading for the NamPower rooftop photovoltaic system 416
P Dobreva, E E van Dyk and F J Vorster
Measuring the optical thermometry properties of La2O2S:Eu phosphor material 422
L J B Erasmus, H C Swart, J J Terblans and R E Kroon
Acceleration parameters for fluid physics with accelerating bodies 428
I.M.A. Gledhill, H. Roohani, A. Biobaku and B. Skews
Unfolding the fast neutron energy distribution of a NE230 deuterated liquid scintillator detector using the MAXED code 434
M S Herbert
Fibre-to-the-Hut Technology: A solution for Cheap Access for High Speed-Optical Network in South Africa 440
G M Isoe, E K Rotich, R R G Gamatham, A W R Leitch and T B Gibbon

Time of Crossing in Pulsed Eddy Current Signals 445
B Kibirige
Electronic tracking system for quantum cryptography and radio telecommunication 451
M Mariola, Y Ismail, A Mirza, F Petruccione
Open-Source electronic board designed in South Africa for Africa 457
M Mariola, F Petruccione
Analysis of homogeneity in thin film photovoltaic modules using large area light beam induced current (LA-LBIC) measurements 463
M Okullo, F J Vorster, E E van Dyk and J L Crozier
Enhancing light absorption and lifetime stability of organic solar cells using pentacene encapsulation 469
F Otieno, K Kamalakannan, M Airo and D Wamwangi

Division G – Theoretical and Computational Physics

Density functional theory on a lattice: Particlenumber dependence of the exchange-correlationpotential 476
Kossi Amouzouvi and Daniel Joubert
Non-universality in a constrained period doubling route to chaos for Rössler's system 481
André E Botha and Wynand Dednam
Simulating cyclic loading of atomic-sized gold tips on gold surfaces via classical molecular dynamics and density functional theory transport calculations 487
W Dednam, C Sabater, M A Fernandez, C Untiedt, J J Palacios, M J Caturla
Quasi-normal Modes for Spin-3/2 Fields 493
Gerhard Erwin Harmsen
Thomas Rotation and Quantum Entanglement 499
JM Hartman, SH Connell and F Petruccione
Security of quantum key distribution 505
Mhlambululi Mafu
Security of the Bennett 1992 quantum key distribution protocol in the presence of noise 511
Mhlambululi Mafu, Makhamisa Senekane, Kevin Garapo and Francesco Petruccione
One and Two Dimensional Models of Dye Adsorption for application in Dye Sensitized Solar Cells 517
N.E. Maluta, N. Mphephu, V. Sankaran, T.S. Mulaudzi, F. Nemangwele
Molecular dynamics studies of Schottky and Frenkel defects in cerium dioxide 522
Thabiso Lekoko, Thuto Mosuang, and Erasmus Rammutla
A phenomenological description of the bath and its effect in photosynthetic light-harvesting systems 527
JA Nöthling, TPJ Krüger and T Mancal
Neutrino mass hierarchy and CP phase measurement using atmospheric neutrino flux 533
Soebur Razzaque
Projection operators in the theory of open quantum systems 539
V Semin, F Petruccione
Hypothesising the effects of Higgs portal dark matter in particle colliders 545
Stefan von Buddenbrock

Conference Chairs and Committees

SAIP2015 Conference Chairs

Makaiko Chithambo, Head: Department of Physics and Electronics, Rhodes University
André Venter, Head: Department of Physics, Nelson Mandela Metropolitan University

SAIP2015 Division Chairs

Division for Particle and Condensed Matter and Materials Physics:
Japie Engelbrecht (Nelson Mandela Metropolitan University)
Division for Nuclear and Particle Physics:
Simon Mullins (iThemba LABS)
Division for Photonics:
Erich Rohwer (University of Stellenbosch)
Division for Astrophysics:
Christian Engelbrecht (University of Johannesburg)
Division for Space Science:
John-Bosco Habarulema (South African National Space Agency)
Division for Physics Education:
Sam Ramaila (University of Johannesburg)
Division for Applied Physics :
Ernest van Dyk (Nelson Mandela Metropolitan University)
Division Theoretical and Computational Physics:
Kristian-Müller Nedebock (Stellenbosch University)

Proceedings Editorial Team

Editors: Makaiko Chithambo and André Venter (Rhodes University and NMMU)
Online Administration: Roelf Botha (SAIP and HartRAO)
Online Administration, Layout Reviewing and Proceedings Compilation:
Juan Grey (SA Institute of Physics)

Review Panel

Makaiko Chithambo (Rhodes University)
Deena Naidoo (WITS)
Azwinndini Muronga (NMMU)
André Venter (NMMU)
John-Bosco Habarulema (SANSA)
Simon Mullins (iThemba Labs)

Message from the Conference Chairpersons

The 60th annual South African Institute of Physics Conference, co-hosted by the physics departments of Rhodes University and the Nelson Mandela Metropolitan University, was held at the Boardwalk Convention Centre, Port Elizabeth, from 29 June to 3 July 2015.

On 20 December 2013, the United Nations General Assembly 68th Session proclaimed 2015 as the International Year of Light and Light-based Technologies (**IYL 2015**)**.** The SAIP consequently devoted a number of its 60th meeting activities to recognize the importance that light plays in modern society and the role it will play in advancing science and optoelectronic technologies in particular. The 2015 Winterschool was therefore aptly devoted to light-based science. Topics such as the trapping and cooling of single ionized atoms, THz spectroscopy, Luminescent materials, light manipulation and Photosynthesis received considerable attention.

This very successful conference was officially opened by the Dean of Science, NMMU, Prof Cedric McCleland. The conference was attended by 440 delegates from 8 countries, including, Austria, Botswana, Namibia, Swaziland, Sweden, USA and Tanzania. As part of the programme, the local organising committee initiated a public outreach programme consisting of 3 evening lectures. The first public lecture by Dr Don Mingay, entitled "The conflict in climate change", was presented on Tuesday 30 June at the Boardwalk Convention centre. Dr Mingay's rather controversial view on the cause of climate change was thoroughly enjoyed by over 300 members of the public and resulted in excited debate after his presentation. A most fascinating LASER show, organized by the Laser Research group from the University of Stellenbosch, preceded this lecture and truly captivated the mainly general public audience. On the evening of Thursday, 2 July, Dr Japie van Zyl from NASA held more than 500 members of the public spell-bound in his presentation "Exploring the Universe: The Search for Signs of Life".

As an outreach to senior science learners from Port Elizabeth and surrounding areas, Prof Andrew Forbes from the University of the Witwatersrand presented a lecture entitled "Light Fantastic" to teachers and senior science learners from this region. Almost 300 scholars from various schools as far as East London attended this fascinating presentation on the applications of light in modern society.

Plenary speakers provided much appreciated stimulus to the academic programme. These speakers included Prof Darrell Comins, (Optical techniques applied to material physics), Prof Michael Kosch (Space Weather
– why should we care?), Prof Liesel Folks (Status of women in STEM in the US), Prof Walter Kutschera (Exploring the world with Accelerator Mass Spectroscopy) and Prof Danie Auret (Characterization of electrically active defectsinsemiconductors).

A very enjoyable banquet was held at the Boardwalk Convention Centre. In addition to the customary student awards, two silver medals were awarded to Dr Angela Dudley from CSIR and Dr Shazrene Hohamed from SAAO for their outstanding contributions to Astronomy and Laser Physics respectively. Honorary membership of the SAIP was awarded to Prof David Wolfe from Institute of Physics (IOP).

Finally, this proceedings is the culmination of the commitment of a number of people: the administrative team in the SAIP office, Dr Roelf Botha, Mr Juan Grey, the Executive officer Mr Brian Masara, the SAIP Council, all division heads and especially to the reviewers for your patience and taking on more than could be reasonably expected. Thank you to you all.

SAIP 2016 will be hosted by UCT. We wish them well!

André Venter and Makaiko Chithambo
Chairpersons: SAIP2015 Conference

Message from the Review Committee

Papers submitted for the proceedings of SAIP2015 were handled using set procedures of the SAIP. All submissions were first checked for adherence to house style. Papers that failed to meet requirements were returned to authors to be suitably edited for style. The next step was peer-review and then, for accepted papers, publication. In theory, once submitted the publication of a paper should be swift but the reality is very different. Publication of proceedings requires as much the goodwill of reviewers as it does the patience of authors. Evaluating papers is onerous and many of our reviewers made commendable efforts to produce detailed and meaningful reports that no doubt many authors were grateful for. Indeed, a good number of reviewers read more than one paper. Ensuring that acceptable academic standards are upheld during peer-review is a necessary step to achieve credible proceedings. The editors spent much time sifting through reviewer reports to ensure that only well considered ones were used. Indeed, a review panel was constituted to work through some unusual and difficult cases. The cost of this is a delay between submission and publication and a test of patience for many authors who otherwise would expect a quick and objective review. We are thankful to the many authors who tolerated the delay as we worked behind the scenes to obtain objective and credible reviews. The editors are extremely grateful to reviewers who devoted their valuable time to these proceedings.

We received 158 papers of which 122 met the required criteria and were sent for evaluation by at least two reviewers. Of the 158, 63% were accepted for publication. The editors thank all authors for submitting their work to SAIP2015 proceedings.

Makaiko Chithambo (on behalf of the SAIP2015 Review team)

List of Reviewers

Dr. ACQUAVIVA, Giovanni	North West University
Dr. ANDALA, Dickson	Kenyatta University (Kenya)
Dr. ASANTE, Joseph	Tshwane University of Technology
Dr. BANGANAYI, Clever	Professional Engineering Technologist
Dr. BENECHA, Evans	University of South Africa
Dr. BERNARD, Claude	Institute of Nuclear Physics (France)
Dr. BLYTH, Sarah	University of Cape Town
Dr. BOKOV, Pavel	South African Nuclear Energy Corporation
Prof. BOTHA, André	University of South Africa
Prof. BOTTCHER, Markus	North-West University
Dr. BUTHELEZI, Zinhle	iThemba Labs
Prof. CHEN, Reuven	Tel Aviv University (Israel)
Prof. CHITHAMBO, Makaiko	Rhodes University
Dr. CHRUSCINSKA, Alicja	Nicholas Copernicus University (Poland)
Dr. CLARKE, M	Science and Technology Facilities Council
Prof. COE, Malcolm	University of Southampton
Prof. CORNELL, Alan	University of Cape Town
Dr. CROZIER, Jacqui	Nelson Mandela Metropolitan University
Dr. DACOSTA, Francis	University of Venda
Dr. DANSKIN, Donald	Cape Peninsula University of Technology
Dr. DAYDOV, Yuri	Joint Institute for Nuclear Research (Russia)
Mr. DE BEER, Frikkie	South African Nuclear Energy Corporation
Prof. DEANDREA, Aldo	Institute of Nuclear Physics (France)
Dr. DEANE, Roger	Rhodes University
Prof. DERRY, Trevor	University of the Witwatersrand
Dr. DIALE, Mmantsae	University of Pretoria
Dr. DINOKO, Tshepo	University of Dodoma (Tanzania)
Dr. DUDLEY, Angela	Council for Scientific and Industrial Research
Dr. ELIASSON, Peter	Swedish Defence Research Agency (Stockholm)
Dr. ENGELBRECHT, Christian	University of Johannesburg
Prof. ENGELBRECHT, Japie	Nelson Mandela Metropolitan University
Dr. ERASMUS, Rudolph	University of the Witwatersrand
Prof. FERREIRA, Stefan	North West University
Prof. FÖRTSCH, Siegfried	University of the Witwatersrand
Dr. GARRIGOUX, Tania	North West University
Dr. GIACOMO, Cacciapaglia	Institute of Nuclear Physics (France)
Prof. GIBBON, Timothy	Nelson Mandela Metropolitan University
Prof. HABARULEMA, John Bosco	South African National Space Agency
Dr. HAMILTON, Andrew	University of Cape Town
Prof. HANDLER, Gerald	Nicolaus Copernicus Astronomical Center (Poland)
Prof. HOROWITZ, William	University of Cape Town
Dr. USMAN, Iyabo	University of the Witwatersrand
Prof. JOUBERT, Daniel	University of the Witwatersrand
Dr. JÓZSA, Gyula	University of Cape Town
Dr. KACHA, Berhanu Tulu	University of the Free State
Dr. KAR, Deepak	University of Cape Town
Dr. KARRIEM, Zain	South African Nuclear Energy Corporation
Dr. KARSTEN, Aletta	National Metrology Institute of South Africa
Mr. KHAMLICH, Saleh	iThemba Labs
Dr. KIBIRIGE, Betty	University of Zululand
Dr. KIRUI, Joseph	University of Limpopo
Dr. KNOTT, Andy	National Physical Laboratory
Dr. KOEN, Chris	University of the Western Cape
Prof. KOLAHCHI, Mohammed	Institute for Advanced Studies in Basic Sciences (Iran)
Dr. KOMIN, Nukri	University of Witwatersrand
Prof. KOSCH, Michael	South African National Space Agency
Prof. KROON, Ted	University of the Free State
Dr. KRUGER, Paulus	North West University

Dr. KRÜGER, Tjaart	Stellenbosch University
Dr. KUGEL, Andreas	Deakin University (Australia)
Dr. KUMAR, Vinod	Indian Institute of Technology (India)
Dr. KUZMICH, Andrey	South African Nuclear Energy Corporation
Dr. LODYA, Lonzeche	SASOL
Dr. WU, Lorinda	CSIR-NLC
Prof. LOUBSER, Ilani	University of Cape Town
Prof. LOUW, Wynand	National Metrology Institute of South Africa
Prof. LOWTHER, John	University of the Witwatersrand
Dr. MACLEOD, Gordon	Hartebeesthoek Radio Astronomy Observatory
Dr. MAGHANA, Christopher	Kabarak University (Kenya)
Mr. SENEKANE, Makhamisa	University of KwaZulu-Natal
Dr. MALUTA, N Eric	University of Venda
Mr. MANYALI, George	University of Science and Technology (Kenya)
Dr. MAPASHA, Edwin	University of Pretoria
Dr. MARCHISONE, Massimiliano	University of the Witwatersrand
Dr. MARIOLA, Marco	University of KwaZulu-Natal
Dr. MATHEWS, Alan	University of KwaZulu-Natal
Dr. MCCRACKEN, John	Michigan State University (United States)
Dr. MCMAHON, Niall	Trinity College Dublin
Dr. MISHRA, Yogendra	University of Kiel (Germany)
Dr. MKHONTO, Donald	Tshwane University of Technology
Mr. MATEYISI, Mohau	Stellenbosch University
Dr. MOSUANG, Thuto	University of Limpopo
Dr. MOYO, Thomas	University of KwaZulu-Natal
Dr. MUJAJI, Marjorie	University of the Witwatersrand
Prof. MÜLLER-NEDEBOCK, Kristian	Nelson Mandela Metropolitan University
Dr. MWAKIKUNGA, Bonex	Council for Scientific and Industrial Research
Prof. NAIDOO, Deena	University of the Witwatersrand
Dr. NEL, Jacqueline	University of Pretoria
Mr. NGWIRA, Chigomezyo	CUA/NASA (USA)
Prof. NICHOL, Robert	Institute of Cosmology and Gravitation (UK)
Mr. NOTO, Luyanda	University of South Africa
Dr. NSENGIYUMVA, Schadrack	Rhodes University
Dr. NYAMHERE, Cloud	Midlands State University (Zimbabwe)
Mr. NYIRENDA, Angel	Rhodes University
Mr.OBODO, Kingsley	University of Pretoria
Dr. OGUNJOBI, Olakunle	North West University
Dr. OUMA, Cecil	Council for Scientific and Industrial Research
Mr. OYEDOKUN, David	University of Cape Town
Dr. OYEYEMI, Elijah	University of Lagos (Nigeria)
Prof. PAGONIS, Basile Vasilis	McDaniel College (USA)
Prof. PRINSLOO, Aletta	University of Johannesburg
Dr. PUCCI, Lorenzo	National Institute for Theoretical Physics
Dr. RAMAILA, Sam	University of the Witwatersrand
Prof. RAZZAQUE, Soebur	University of the Free State
Dr. REDDY, Leelakrishna	University of the Witwatersrand
Dr. RICHTER, Tobias	Science and Technology Facilities Council
Dr. RORO, Kittessa	Council for Scientific and Industrial Research
Dr. SOLANS SANCHEZ, Carlos	The European Organization for Nuclear Research (CERN)
Dr. SCHROEDER, Anja	South African Astronomical Observatory
Dr. SHEPPARD, Charles	University of the Free State
Dr. SITHOLE, Enoch	Southern Methodist University (United States)
Prof. SMITS, Derck	University of Johannesburg
Dr. SOM, Sudipta	National Taiwan University (Taiwan)
Mr. SPANIER, Felix	University of Johannesburg
Dr. SPEKKENS, Kristine	Queen's University (Canada)
Dr. STEIN, Matthias	Planck-Institute for Dynamics of Complex Technical Systems
Dr. SULEIMAN, Mohammed	Sudan University of Science and Technology
Prof. SWART, Hendrik	University of the Free State

Prof. TAYLOR, Dale	University of Cape Town
Dr. THORAT, Kshitij	Rhodes University
Dr. TORRES-GUZMAN, Jorge	Centro Nacional De Metrología (Spain)
Dr. TSHABALALA, George	University of the Witwatersrand
Dr. TYREL, Johnson	University of the Free State
Dr. UKEN, Daniel	University of KwaZulu-Natal
Dr. URGESSA, Zelalem	Nelson Mandela Metropolitan University
Dr. VAISANNEN, Petri	University of Nairobi (Kenya)
Dr. VAN DEN BERG, Jan-Albert	University of the Witwatersrand
Prof. VAN DYK, Ernest	Nelson Mandela Metropolitan University
Prof. VAN SOELEN, Brian	University of the Free State
Prof. VENTER, Christo	North West University
Dr. VORSTER, Frederik	Nelson Mandela Metropolitan University
Prof. WAANDERS, Frans	North West University
Dr. WAGENER, Magnus	Nelson Mandela Metropolitan University
Mr. WAKO, Ali	Meru University of Science and Technology (Kenya)
Dr. WAMWANGI, Daniel	University of the Witwatersrand
Dr. WARMBIER, Robert	University of the Witwatersrand
Prof. WEIGEL, Herbert	Stellenbosch University
Prof. WINKLER, Hartmut	University of Johannesburg
Prof. ZAMONSKY RIMOLDI, Oscar	South African Nuclear Energy Corporation
Ms. ZANDER, Claudia	University of Pretoria

Division A – Division for Physics of Condensed Matter and Materials

Electronic and Optical Properties of monolayer MX_2(M= Zr, Hf; X= S, Se) from first principles calculations

Mahmud Abdulsalam, D. N. G. Moise 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 117176804; Fax: +27 117176879.

E-mail: *mahmudaaa@yahoo.co.uk*

Abstract. Transition metal di-chalcogenide (TMDC) monolayers have potential applications in electronic and optical devices. Work on these atomically thin semiconductors is an exciting emerging field of research. In this research the electronic and photo-absorption properties of monolayer ZrS_2, $ZrSe_2$, HfS_2 and $HfSe_2$ have been investigated using density functional theory (DFT) and many body perturbation theory at the level of the partially self consistent GW_0 approximation and of the Bethe-Salpeter equation (BSE) in the Tamm-Dancoff approximation. The structures were found to be semiconductors with band gaps within the visible range of the spectrum with strong optical absorption in the visible range. Exciton binding energies were estimated from a comparison of the GW and BSE results.

1. Introduction

MX_2(M= Zr, Hf; X= S, Se) belong to the family of layered transition metal chalcogenides (TMC's). Layers in these materials are bonded to one another by weak van der Waals forces [1,2], making them a possible source of 2-D (two-dimensional) crystals [3–6]. Recently there has been a growing interest in 2-D materials due to their potential applications in the field of photonic and electronic applications [7–10]. Some of these materials have been used successfully in the fabrication of inverters, logic circuits, low powered field effect transistors and memory cells [5, 7, 10–14].

TMC's are expected to have a strong excitonic effects due to reduced screening, which is expected to affect their optical properties. There is a need to accurately calculate and understand the many-body electronic properties of these TMC's. One of the common approaches that can be used to investigate the electronic properties of these systems is to calculate the electronic excitation energies based on for example, Green's function for many body perturbation theory (MBPT) in the GW approximation with self energy [15]. This method has been proven to be accurate for a number of materials [16, 17]. Results from the GW-MBPT electronic excitation energies are then used to study neutral excitations by solving the two particle Bethe-Salpeter equation (BSE) [18–20]. The later method has been reported to give a reasonable description of optical properties of quite a number of systems [21, 22]. To the best of our knowledge, such methods have not been used to study the electronic and optical properties of monolayer MX_2(M= Zr, Hf; X= S, Se) structures. The interest of this paper, is therefore, to use the

many body GW and BSE equations to study the electronic and optical properties of monolayer MX_2(M= Zr, Hf; X= S, Se) structures.

In their bulk state, MX_2(M= Zr, Hf; X= S, Se) are iso-structural [23,24], they crystallize in a 1T-CdI_2 type structure with space group *164,* $P\bar{3}m1$ in which the metal atom is sandwiched between two sheets of the chalcogen atoms leading to X-M-X layers. Atoms within a layer are strongly held together by a covalent bond with weak ionic bonding within the layers. A monolayer of the structure is depicted in Figure 1.

2. Computational Method

Our first principle calculations were performed using the Projector Augmented Wave (PAW) [25] method as implemented in the Vienna ab initio Simulation Packege (VASP) [26, 27]. Perdew, Burke and Ernzerhof (PBE) [28, 29] parametrization of the generalised gradient approximation (GGA) [30–33] is used for exchange correlation. 17 x 17 x 1 Monkhorst-Pack meshes were used in sampling the Brillouin zones with an energy cut-off of 520 eV and were found to be sufficient enough for the energy convergence. The Brilloin zone sampling was chosen in such a way that the convergence of free energy is less than 1 meV/atom. The monolayer was generated by creating a space wide enough between the adjacent layers in the z-direction of the relaxed bulk structure. A vacuum region of 19 Å was used to isolate the layers along the c-axis and was found sufficient to eliminate interaction between the layers. The monolayer was generated by creating a space wide enough between the adjacent layers in the z-direction of the relaxed bulk structure.

Frequency dependent GW calculation at the level of partially self-consistent GW_0 approximation were performed, where the quasiparticle energies were updated while keeping the DFT wave functions. To calculate the optical spectra, BSE calculations using Tamm Dancoff (TD) approximation on the basis of a free qausielectron-quasihole were performed. Five highest valence bands and five lowest conduction bands were found to be enough to converged the peaks of the A and B energy peaks.

3. Results and discussions

We have systematically calculated the electronic and optical properties of the monolayer MX_2(M= Zr, Hf; X= S, Se) structures, the electronic and optical properties of the structures are discussed in subsection 3.1 and subsection 3.2 respectively.The unit cell of the structure used in the calculation is depicted in Figure 1.

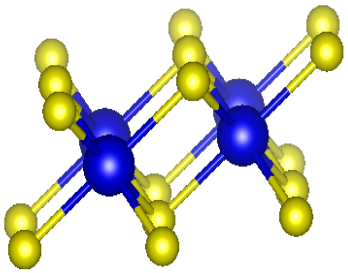


Figure 1: Monolayer MX_2(M= Zr, Hf; X= S, Se) relaxed structure: blue = M, yellow = X.

3.1. Electronic Properties

Our calculated fundamental electronic band gap of the monolayers MX_2(M= Zr, Hf; X= S, Se) structures are listed in Table 1. The PBE and GW calculated band gaps of ZrS_2, HfS_2 and $HfSe_2$ were found to be an indirect band gap semiconductors, their conduction band maximum lies at A and valence band maximum at Γ points of the high symmetry points in ZrS_2 and M in the case of HfS_2 and $HfSe_2$. $ZrSe_2$ was found to be a direct band gap material with both valence and conduction band at the Γ points of the high symmetry k-points. From the table we note that our calculated PBE electronic band gaps for the hafnium chalcogenides are in good agreement with those reported in [34]. The difference in our band gap in the zirconium compounds may be as a result of the choice of the lattice parameters used in the calculations. However, we have no experimental band gaps to compare our results with, the only available information we have is that of Li et. al [35], where the authors predicted the experimental band gap of monolayer ZrS_2 to be around 1.7 to 1.93 eV, which is in good agreement with our GW direct band results.

Table 1: Monolayer MX_2(M= Zr, Hf; X= S, Se) band gap in eV

	HfS_2	$HfSe_2$	ZrS_2	$ZrSe_2$
PBE	1.22	0.67	0.80	0.26
GW	2.02	0.93	1.95	2.53
Direct GW gap	3.07	2.31	2.83	2.53
Other PBE	1.27 [a]	0.61 [a]	1.10 [a]	0.45 [a]

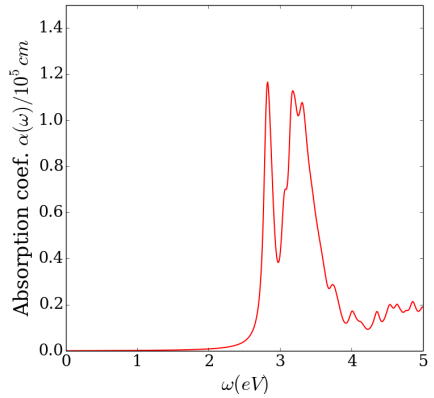
a = reference [34]

3.2. Optical Properties

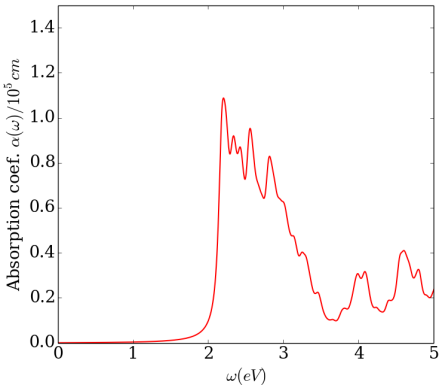
To determine the optical properties, absorption spectra of the ZrX_2 family structures were calculated at the BSE level of approximation. The results are plotted in Figures 2a, 2b, 2c and 2d. From Figure 2, we can note that HfS_2 and ZrS_2 start to absorb photons in the ultra-violet and infra red regions of the spectrum respectively and that there is downward shift in energy absorption energy in both zirconium and hafnium as we move from sulphur to selenium.

Table 2: Monolayer MX_2(M= Zr, Hf; X= S, Se) BSE calculated positions for A and B peaks in eV

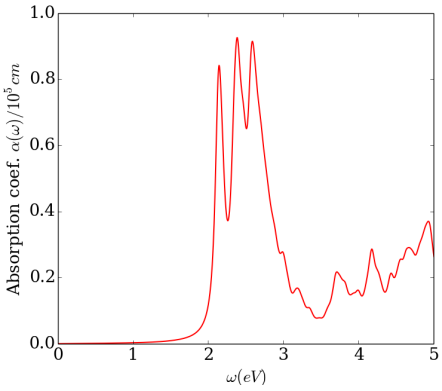
	HfS_2	$HfSe_2$	ZrS_2	$ZrSe_2$
E_A	2.84	2.20	2.15	1.57
E_B	3.17	2.60	2.40	1.71



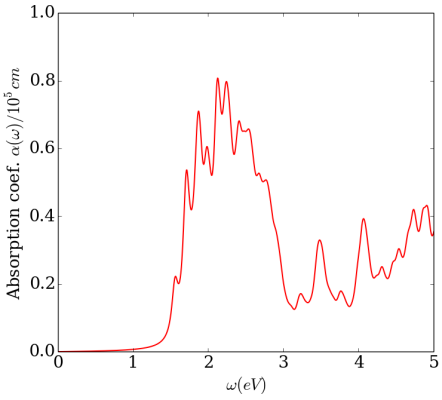
(a) HfS_2 optical absorption



(b) $HfSe_2$ optical absorption



(c) ZrS_2 optical absorption



(d) $ZrSe_2$ optical absorption

Figure 2: Monolayer MX_2(M= Zr, Hf; X= S, Se) optical absorptions

To determine the binding energies of the energetically lowest lying excitons, the energies of the peaks in the BSE calculated absorption coefficient were compared with the quasi-particle energies obtained from the GW calculations. The binding energies of the excitons can be estimated by subtracting the lowest BSE absorption peak energy listed in Table 2 from the GW direct band gap energy, i.e. E_A^a = GW direct gap -E_A and E_B^b = GW direct gap -E_B. The binding energies of the excitons are listed in Table 3. Negative binding energies suggest a resonant exciton at the corresponding energy level.

Table 3: Monolayer MX_2(M= Zr, Hf; X= S, Se) BSE calculated binding energy for A and B excitons in eV

	HfS_2	$HfSe_2$	ZrS_2	$ZrSe_2$
E_B^a	0.23	0.11	0.68	0.96
E_B^b	-0.1	-0.29	0.43	0.82

4. Conclusions

We have performed first principle calculations of the electronic and optical properties of monolayer MX_2(M= Zr, Hf; X= S, Se) structures. Three of structures were found to have indirect band gaps and $ZrSe_2$ was found to be a direct band gap material, with band gaps within the visible range of the spectrum. Strong optical absorption is predicted in the visible region. In a solar cell device, the monolayer structures will be mounted on a substrate which will affect the absorption properties. The results, however, suggest that it will be worthwhile investigating the MX_2(M= Zr, Hf; X= S, Se) further as potential components in second generation optical absorbers.

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] J. Wilson and A. Yoffe, "The transition metal dichalcogenides discussion and interpretation of the observed optical, electrical and structural properties," *Advances in Physics*, vol. 18, no. 73, pp. 193–335, 1969.

[2] A. Enyashin, S. Gemming, and G. Seifert, "Nanosized allotropes of molybdenum disulfide," *The European Physical Journal-Special Topics*, vol. 149, no. 1, pp. 103–125, 2007.

[3] J. N. Coleman, M. Lotya, A. ONeill, S. D. Bergin, P. J. King, U. Khan, K. Young, A. Gaucher, S. De, R. J. Smith, *et al.*, "Two-dimensional nanosheets produced by liquid exfoliation of layered materials," *Science*, vol. 331, no. 6017, pp. 568–571, 2011.

[4] Z. Zeng, Z. Yin, X. Huang, H. Li, Q. He, G. Lu, F. Boey, and H. Zhang, "Single-layer semiconducting nanosheets: High-yield preparation and device fabrication," *Angewandte Chemie International Edition*, vol. 50, no. 47, pp. 11093–11097, 2011.

[5] K. Novoselov, D. Jiang, F. Schedin, T. Booth, V. Khotkevich, S. Morozov, and A. Geim, "Two-dimensional atomic crystals," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 30, pp. 10451–10453, 2005.

[6] H. Ramakrishna Matte, A. Gomathi, A. K. Manna, D. J. Late, R. Datta, S. K. Pati, and C. Rao, "MoS_2 and WS_2 analogues of graphene," *Angewandte Chemie*, vol. 122, no. 24, pp. 4153–4156, 2010.

[7] B. Radisavljevic, A. Radenovic, J. Brivio, V. Giacometti, and A. Kis, "Single-layer MoS_2 transistors," *Nature nanotechnology*, vol. 6, no. 3, pp. 147–150, 2011.

[8] K. F. Mak, K. He, C. Lee, G. H. Lee, J. Hone, T. F. Heinz, and J. Shan, "Tightly bound trions in monolayer MoS_2," *Nature materials*, vol. 12, no. 3, pp. 207–211, 2013.

[9] D. Xiao, G.-B. Liu, W. Feng, X. Xu, and W. Yao, "Coupled spin and valley physics in monolayers of MoS_2 and other group-vi dichalcogenides," *Physical Review Letters*, vol. 108, no. 19, p. 196802, 2012.

[10] Z. Yin, H. Li, H. Li, L. Jiang, Y. Shi, Y. Sun, G. Lu, Q. Zhang, X. Chen, and H. Zhang, "Single-layer MoS_2 phototransistors," *ACS nano*, vol. 6, no. 1, pp. 74–80, 2011.

[11] B. Radisavljevic, M. B. Whitwick, and A. Kis, "Integrated circuits and logic operations based on single-layer MoS_2," *ACS nano*, vol. 5, no. 12, pp. 9934–9938, 2011.

[12] L. Britnell, R. Gorbachev, R. Jalil, B. Belle, F. Schedin, A. Mishchenko, T. Georgiou, M. Katsnelson, L. Eaves, S. Morozov, *et al.*, "Field-effect tunneling transistor based on vertical graphene heterostructures," *Science*, vol. 335, no. 6071, pp. 947–950, 2012.

[13] S. Bertolazzi, D. Krasnozhon, and A. Kis, "Nonvolatile memory cells based on MoS_2/graphene heterostructures," *ACS nano*, vol. 7, no. 4, pp. 3246–3252, 2013.

[14] Y. Zhang, Y.-W. Tan, H. L. Stormer, and P. Kim, "Experimental observation of the quantum hall effect and berry's phase in graphene," *Nature*, vol. 438, no. 7065, pp. 201–204, 2005.

[15] L. Hedin, "New method for calculating the one-particle green's function with application to the electron-gas problem," *Physical Review*, vol. 139, no. 3A, p. A796, 1965.

[16] W. G. Aulbur, M. Städele, and A. Görling, "Exact-exchange-based quasiparticle calculations," *Physical Review B*, vol. 62, no. 11, p. 7121, 2000.

[17] F. Aryasetiawan and O. Gunnarsson, "The gw method," *Reports on Progress in Physics*, vol. 61, no. 3, p. 237, 1998.

[18] M. S. Hybertsen and S. G. Louie, "Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies," *Physical Review B*, vol. 34, no. 8, p. 5390, 1986.
[19] S. Albrecht, L. Reining, R. Del Sole, and G. Onida, "Ab initio calculation of excitonic effects in the optical spectra of semiconductors," *Physical review letters*, vol. 80, no. 20, p. 4510, 1998.
[20] L. X. Benedict, E. L. Shirley, and R. B. Bohn, "Optical absorption of insulators and the electron-hole interaction: An ab initio calculation," *Physical review letters*, vol. 80, no. 20, p. 4514, 1998.
[21] G. Onida, L. Reining, and A. Rubio, "Electronic excitations: density-functional versus many-body greens-function approaches," *Reviews of Modern Physics*, vol. 74, no. 2, p. 601, 2002.
[22] M. Rohlfing and S. G. Louie, "Electron-hole excitations and optical spectra from first principles," *Physical Review B*, vol. 62, no. 8, p. 4927, 2000.
[23] H. Jiang, "Structural and electronic properties of ZrX_2 and HfX_2 (X= S and Se) from first principles calculations," *The Journal of chemical physics*, vol. 134, no. 20, p. 204705, 2011.
[24] M. Moustafa, T. Zandt, C. Janowitz, and R. Manzke, "Growth and band gap determination of the ZrS_xSe_{2-x} single crystal series," *Physical Review B*, vol. 80, no. 3, p. 035206, 2009.
[25] G. Kresse and D. Joubert, "From ultrasoft pseudopotentials to the projector augmented-wave method," *Physical Review B*, vol. 59, no. 3, pp. 1758–1775, 1999.
[26] G. Kresse and J. Hafner, "Ab initio molecular dynamics for liquid metals," *Physical Review B*, vol. 47, no. 1, pp. 558–561, 1993.
[27] G. Kresse and J. Hafner, "Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium," *Physical Review B*, vol. 49, no. 20, p. 14251, 1994.
[28] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Physical review letters*, vol. 77, no. 18, pp. 3865–3868, 1996.
[29] M. Ernzerhof and G. E. Scuseria, "Assessment of the perdew–burke–ernzerhof exchange-correlation functional," *The Journal of chemical physics*, vol. 110, no. 11, pp. 5029–5036, 1999.
[30] G. Kresse and J. Furthmüller, "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set," *Computational Materials Science*, vol. 6, no. 1, pp. 15–50, 1996.
[31] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Physical Review A*, vol. 38, no. 6, pp. 3098–3100, 1988.
[32] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, "**Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation**," *Physical Review A*, vol. 46, pp. 6671–6687, 1992.
[33] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, "Erratum: Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation," *Physical Review B*, vol. 48, pp. 4978–4978, 1993.
[34] H. Guo, N. Lu, L. Wang, X. Wu, and X. C. Zeng, "Tuning electronic and magnetic properties of early transition-metal dichalcogenides via tensile strain," *The Journal of Physical Chemistry C*, vol. 118, no. 13, pp. 7242–7249, 2014.
[35] Y. Li, J. Kang, and J. Li, "Indirect-to-direct band gap transition of the ZrS_2 monolayer by strain: first-principles calculations," *RSC Advances*, vol. 4, no. 15, pp. 7396–7401, 2014.

Investigation of the introduction and annealing behaviour of the donor-vacancy complex in alpha-particle irradiated germanium

Abraham W. Barnard, Walter E.Meyer, F. Danie Auret, Phuti N.M. Ngoepe and Sergio M.M. Coelho

Department of Physics, University of Pretoria, Private bag X20, Hatfield 0028, Pretoria

E-mail: u10688112@tuks.co.za

Abstract. The annealing behaviour of the donor vacancy complex (E-centre) in Ge has been investigated by high resolution (Laplace) deep level transient spectroscopy (L-DLTS). In this study Sb-doped Ge was used and the defect was introduced by irradiating the Ge sample with alpha particles from an Am-241 source. The Sb-vacancy complex has an activation energy of 0.37 eV ($E_{0.37}$) for electron emission as determined by L-DLTS. The E-centre in Ge has been observed to anneal out in a two stage process. In the first stage the defect concentration decreases rapidly when the sample is heated to approximately 320 K, and then remains relatively constant with annealing temperature. In the second final stage, at a temperature of approximately 370 K, the defect concentration decreases quite rapidly until the defect finally anneals out completely. A possible hypothesis is that the E-centre observed is in fact two different defects corresponding to the fast and slow annealing components. However, in this study, we found that both the slow and the fast annealing components of the E-centre have the same L-DLTS signatures (activation energy and apparent capture cross section) as well as the same true capture cross section. In effect, both the fast and the slowly annealing components of the E-centre seem to be the same defect. We investigated this phenomenon by investigating different irradiation and annealing procedures and suggest that the fast annealing component of the E-centre can be explained by Ge self-interstitials, released from other radiation induced defects recombining with the vacancy in the E-centre.

1. Introduction

Germanium shares unique properties with silicon in which both have 4 valence electrons that results in the formation of tetrahedral crystal lattice. This dramatically changes the electrical properties when substitution atoms are introduced [1]. It is mainly used in a highly pure form as a detector material and is not extensively used in devices due to the overwhelming success of silicon in applied fields. This material has however gained a lot of interest recently as a semiconductor material due to improved epitaxial growth techniques for electronic and opto-electronic applications. When compared to Si, Ge is a possible candidate for fast switching transistors or complimentary metal-oxide-semiconductor (CMOS) devices due to its higher hole and electron mobility [4].

The donor vacancy pairs (E-centre) formed by introducing vacancies into the antimony doped germanium crystal lattice through alpha or electron radiation has been identified to be trap with an

activation energy of 0.37 eV, which will be indicated as $E_{0.37}$. The rate of $E_{0.37}$ introduction depends on the concentration of the Sb in the material [3]. When high concentrations of Sb are present, the $E_{0.37}$ peak dominates and becomes the sole observed peak with other defects being secondary that grow over time. Defects within germanium are removed by low annealing temperatures between 625 K and 675 K with the E-centre annealing out by dissociation and diffusion of the vacancy. Annealing with a reverse bias delays or prevents annealing of the E-centre in Ge which is a contrast to the annealing of the E centre in Si which has a bias enhancement [2].

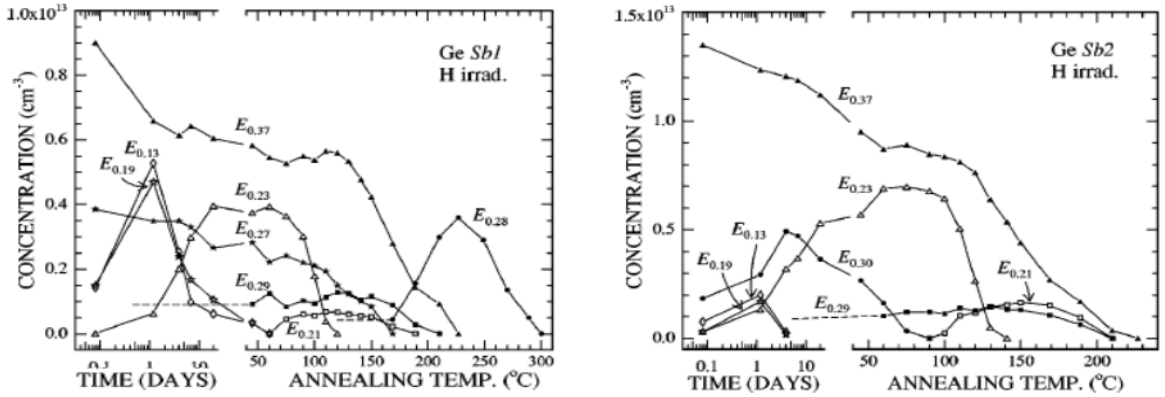


Figure 1. Isothermal annealing at room temperature followed by isochronal annealing at intervals of 15 minutes for two germanium samples Sb1 and Sb2 after exposure to proton radiation [2].

The $E_{0.37}$ observed by J. Fage-Pedersen and A. Nylandsted Larsen in figure 1 showed a reduction in concentration when exposed to room temperature and a further reduction once exposed to temperatures greater than 400 K. It was found that a major fraction of $E_{0.37}$ anneals at approximately 450 K. Over a large temperature span, thermally activated diffusion or association would not proceed. Thus if the $E_{0.37}$ peaks don't contain large contributions from other defects there has to be some kind of unstable source created during the irradiation that release mobile species at room temperature that consumes the E-centre [3].

2. Experimental details

The Ge supplied by Umicore used in this study was bulk grown (100) n-type germanium. Multiple samples of approximately 6 mm by 3 mm were first degreased with trichloroethylene, isopropanol and methanol for 5 minutes each then etched in a mixture of H_2O_2:H_2O with a ratio of 1:5 for 1 minute and dried with nitrogen. Immediately after cleaning a layer of 80 nm AuSb was deposited on the backside forming the ohmic contacts through resistive evaporation. After formation of the ohmic contacts, the samples were rinsed of with isopropanol for 5 minutes and dried off with nitrogen before being annealed at 650 K for 10 minutes in an environment flushed with argon at a rate of 0.1 litres per minute. The samples were cleaned with isopropanol for 5 minutes again and etched with the same mixture of H_2O_2:H_2O for 1 minute before being dried off with nitrogen. Gold circular dots, 100 nm thick, with a diameter of approximately 0.6 mm were grown at a rate of 0.1 nm per second through resistive evaporation on the front of the samples.

Three doping densities were deployed and are denoted as GeSb1, GeSb2 and GeSb3 with the first two approximated to contain 1×10^{15} cm^{-3} Sb and the third measured to have 2.6×10^{15} cm^{-3}. One contact on GeSb1 was exposed to Am-241 in intervals of 30 minutes and annealed at 330 K for 60 minutes after each exposure. L-DLTS was performed at 195 K after each exposure and annealing, up to a resultant exposure of 180 minutes. The second sample (GeSb2) underwent the same procedure but was exposed in intervals of 40 minutes with the third exposure being 50 minutes. Three other contacts on GeSb1 were exposed for the full length of 180 minutes then annealed at 330K for 60 minutes with a L-DLTS

spectrum taken at 195K. Another contact on GeSb2 and a contact on GeSb3 were exposed at intervals of 30 minutes with a Laplace transient recorded after each exposure.

In this paper we have studied the introduction and annealing kinetics of two defects that have a similar energy level which has previously been identified as the E-centre.

3. **Experimental results**

The energy level $E_{0.37}$ may represent at least two defects with very similar energy levels that cannot be distinguished through conventional DLTS methods due to the resolution not being high enough. For convenience sake we will denote the first reduction as the removal of the $E'_{0.37}$ and the second reduction as the removal of the E-centre ($E_{0.37}$) that was observed by many. This peak was identified to lie approximately at 195 K on the conventional DLTS spectrum, thus we used L-DLTS at this temperature for higher resolution [5]. However even using this high resolution technique, it was not possible to distinguish between the two traps.

3.1 Introduction kinetics of $E_{0.37}$

The introduction rate of the $E_{0.37}$ was determined by introducing both defects through alpha radiation as seen in figure 2 and removing only the $E'_{0.37}$ from the two germanium samples by means of low temperature annealing. The introduction of $E_{0.37}$ was tested by introducing both defects for the full exposure time period on multiple points on the GeSb1 sample. After removing the $E'_{0.37}$ defect it was found that the introduction of the $E_{0.37}$ was linear and did not depend on whether the $E'_{0.37}$ was annealed out during or after the irradiation. This is shown clearly in figure 2 where the concentration after annealing (red dots) showed a linear relationship with irradiation time, irrespective of the annealing occurring in one step after irradiation or in shorter steps between irradiations. The introduction kinetics for $E_{0.37}$ was observed to be linear which suggests that is was of a first order introduction which is consistent with vacancies captured by Sb.

3.2 Introduction kinetics of $E'_{0.37}$

The combined introduction rate of $E'_{0.37}$ and $E_{0.37}$ through alpha particle radiation was determined to be that of a quadratic function. This was determined by introducing both defects into the crystals at intervals of 30 minutes and then modelling their combined introduction rate against the equation

$$Nt = At^B$$

where *Nt* is the trap concentration and t is the total exposure time to alpha radiation. In figure 3 the best fit with the equation, GeSb2 was found to have $A = 6.67 \times 10^9$ with $B = 2.04$ and GeSb3 was found to have $A = 1.02 \times 10^{10}$ with $B = 2.02$. This suggests an introduction rate for the simultaneous introduction of both defects to be that of a quadratic nature. Since it was earlier determined that the introduction kinetics of the $E_{0.37}$ was that of a linear nature, it can easily be seen that the introduction kinetics of the $E'_{0.37}$ is that of a second order process.

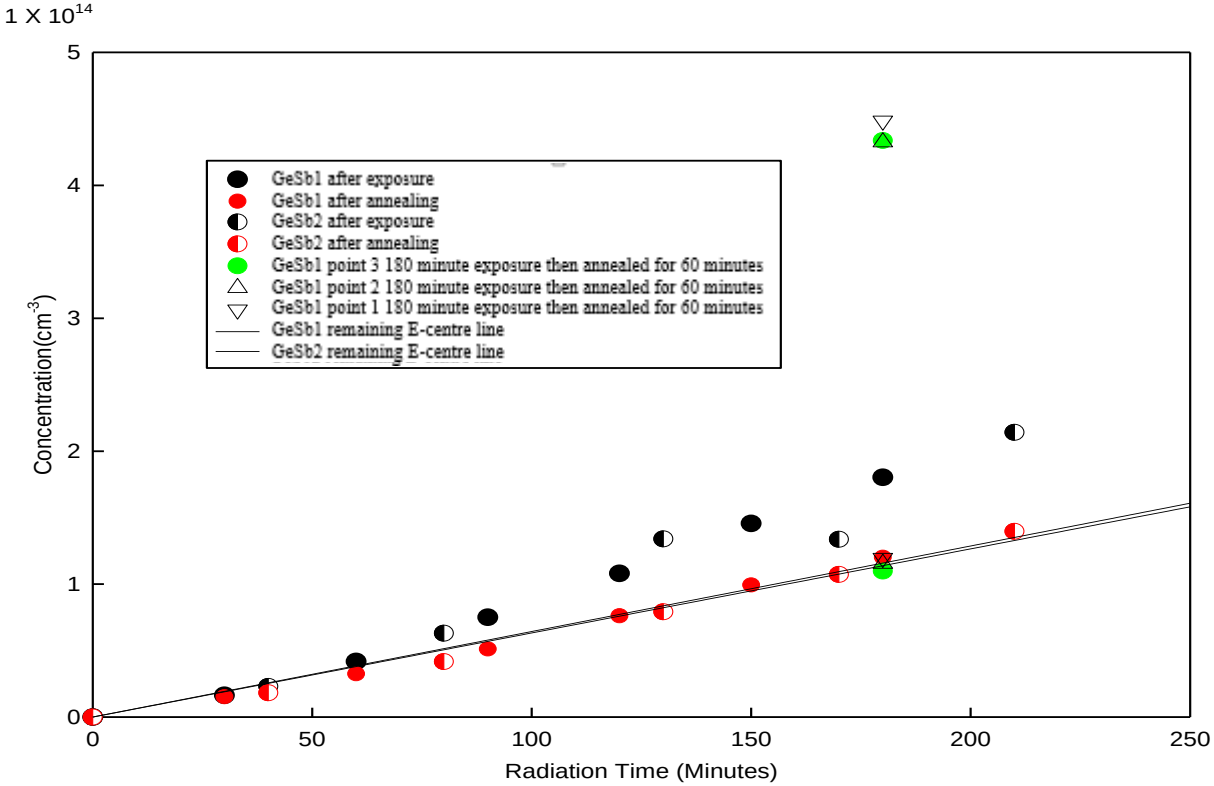


Figure 2. GeSb1 exposed to alpha radiation at 30 minute intervals with 60 minutes 330 K annealing after each exposure. GeSb2 exposed to alpha radiation at 40 minute intervals with same annealing procedure with the third exposure being 50 minutes. Three GeSb1 points exposed to alphas for 180 minutes then annealed at 330 K for 60 minutes.

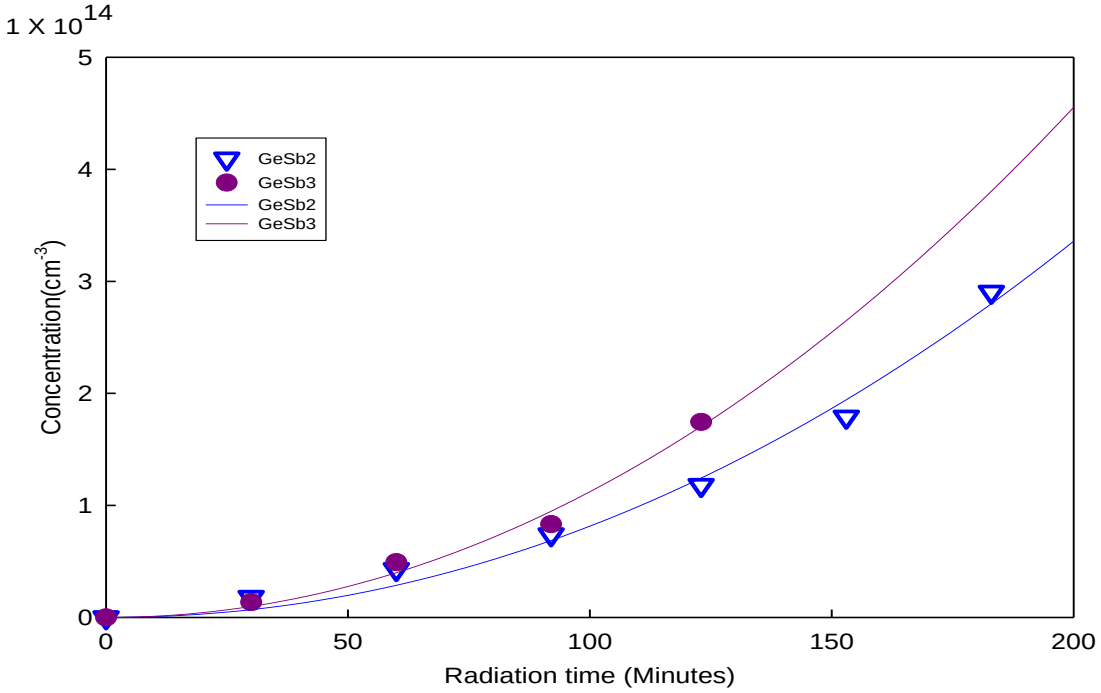


Figure 3. The concentration of the simultaneous introduction of $E'_{0.37}$ and $E_{0.37}$ in the samples GeSb2 and GeSb3 which were exposed to AM-241 at intervals of 30 minutes.

3.3 Annealing behaviour of $E_{0.37}$ and $E'_{0.37}$

This annealing behaviour of the $E'_{0.37}$ peak that is observed through Laplace DLTS was reconfirmed by exposing one of the germanium samples to alpha radiation and immediately doing isothermal annealing measurements on it. It was found that the annealing behaviour at room temperature (300K) was described as exponential decay to a constant concentration. This is consistent with first order decay of the first component with a second component remaining. This annealing behaviour was also confirmed in figure 4 at temperatures of 315 K and 330 K where exponential decay with greater decay constants were observed. We will refer to the component that anneals out in this first stage as the $E'_{0.37}$ defect. Annealing at 330 K for 60 minutes has been found to guarantee the removal of the $E'_{0.37}$ for experimental purposes. Only once the sample was exposed to temperatures greater than 400 K which is seen in isochronal annealing's in figure 5, did the rest of the $E_{0.37}$ peak that is observed in Laplace DLTS anneal away.

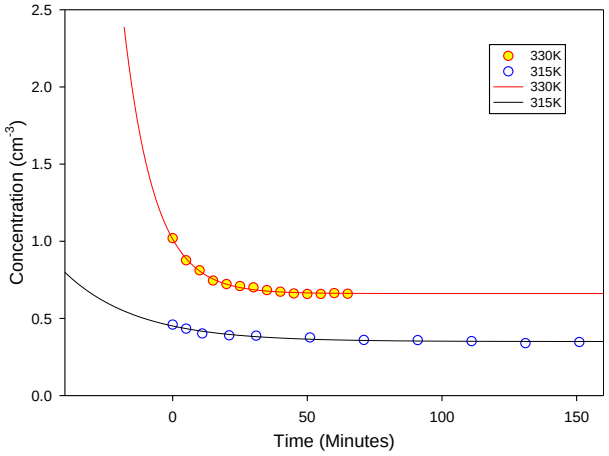


Figure 4. Isothermal annealing of the $E'_{0.37}$ defect at 315 K and 330 K.

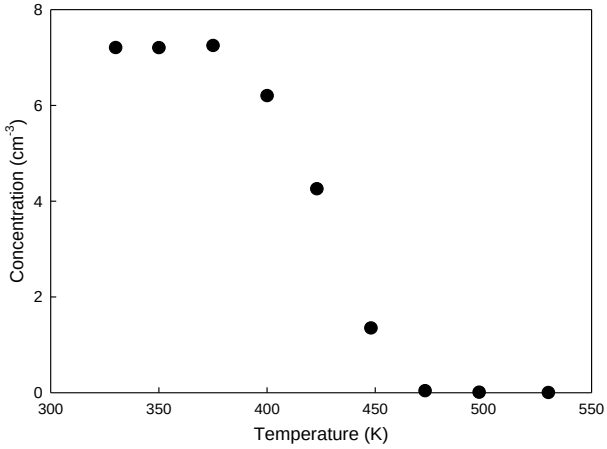


Figure 5. Isochronal annealing of the $E_{0.37}$ defect at time intervals of 15 minutes at 25 K increments.

4. Conclusion

It was found that the two defects $E'_{0.37}$ and $E_{0.37}$ cannot be distinguished through DLTS or through high resolution techniques such as L-DLTS. The introduction kinetics of the $E_{0.37}$ was found to be linear and the introduction kinetics of the $E'_{0.37}$ was found to be quadratic. The annealing of both the $E_{0.37}$ and $E'_{0.37}$ through isochronal annealing was found to exhibit exponential decay. The $E'_{0.37}$ anneals out first at temperatures as low as room temperature with the $E_{0.37}$ which experiences annealing at temperatures greater than 400K. Since the introduction rate of the $E_{0.37}$ is linear it is consistent with the theory of vacancies captured by Sb. The quadratic nature of the $E'_{0.37}$ suggests the reaction of newly introduced defects with previously introduced defects through radiation. A possible theory that will still need to be investigated through density functional theory would be that there may be a self-interstitial captured by the E centre defect which is then seen as the $E'_{0.37}$ defect.

5. References

[1] S.M. Sze, Physics of Semiconductor devices, Wiley & Sons, Canada, 1981.

[2] J.Fage-Pederson, A. Nylandsted Larsen, Irradiation-induced defects in Ge studied by transient spectroscopies, Physical Review B, **62** 15 (2000) 10116.

[3] A. Mesli, L. Dobaczewski, K. Bonde Nielsen, V.L. Kolkovsky, M. Christian Petersen, and A. Nylandsted Larsen, Low-temperature irradiation-induced defects in germanium: In situ analysis, Physical Review B, **78** (2008) 165202.

[4] R. Hull, J.C. Bean (Eds.), Germanium Silicon, Physics Materials, Semiconductor and Semi-metals, vol. 56, Academic, San Diego, 1999.
[5] A.R. Peaker, Laplace deep level transient spectroscopy: Embodiment and evolution, Pysica B: Condensed Matter, **407** 15 (2012) 3026

Laser excitation studies and crystal-field analysis of ZnO:Tb^{3+} and ZnO:Eu^{3+} powders

N A Bhebhe[1], M Mujaji[1], D Wamwangi[1], L F Koao[2] and F B Dejene[2]

[1] School of Physics and DST-NRF Centre of Excellence in Strong Materials, University of the Witwatersrand, Johannesburg, Wits 2050, South Africa
[2] Department of Physics, University of the Free State (Qwaqwa Campus), Private Bag X13, Phuthaditjhaba ZA 9866, South Africa

E-mail: Nkosiphile.Bhebhe@students.wits.ac.za

Abstract. Results from laser excitation studies of Tb^{3+} ions and Eu^{3+} ions in zinc oxide (ZnO) powders are presented; rare-earth doped ZnO (ZnO:RE^{3+}) is currently of great interest as a prospective solid-state laser matrix and for optoelectronic device applications. The chemical bath deposition technique was utilized for synthesizing the ZnO:Tb^{3+} and ZnO:Eu^{3+} powders. Photoluminescence spectra of the pellet samples were obtained in the 460 – 900 nm range. The spectra exhibit sharp emission lines superimposed on a broad emission background with 457.9 nm, 476.5 nm and 488.0 nm argon laser-line excitation. The sharp peaks are attributed to the $^5D_4 \rightarrow {}^7F_J$ (J = 0, 1, 2, 3, 4, 5, 6) and the $^5D_0 \rightarrow {}^7F_J$ (J = 0, 1, 2, 3, 4) electronic transitions of Tb^{3+} and Eu^{3+}, respectively. Crystal-field energy levels for the Tb^{3+} ion and the Eu^{3+} ion occupying a C_{3v} symmetry site were deduced.

1. Introduction

Semiconductors such as zinc oxide (ZnO) and gallium nitride (GaN) are of great importance to the optoelectronic industry because of their inherent optical and electronic properties [1,2]. ZnO has drawn considerable interest because of its superior properties which include a wide band gap (3.36 eV) [3] and a large exciton binding energy of 60 meV [1]. Furthermore, ZnO is a cheap, chemically and thermally stable material [4].

Optical studies have shown ZnO to exhibit a broad emission in the visible region of the electromagnetic spectrum [1,4,5]. The broad emission has been attributed to deep level transitions within the band gap due to intrinsic ZnO defects such as zinc vacancies and oxygen vacancies [3,4]. To improve on the visible emission, studies on rare earth (RE^{3+}) doped ZnO continue to be pursued with successful incorporation of RE^{3+} ions such as trivalent terbium (Tb^{3+}) [1,6] and trivalent europium (Eu^{3+}) [7]. In the literature, it is suggested that the RE^{3+} ion resides at either a Zn^{2+} lattice site or an interstitial position [6,7,8] with a neighboring zinc vacancy for charge compensation [6]. However, there has been no experimental verification of the RE^{3+} location. Intense green and red emission bands have been observed for ZnO:Tb^{3+} [1,6] and ZnO:Eu^{3+} [7], respectively, with ultra-violet excitation. RE^{3+} doping therefore significantly enhances visible luminescence in ZnO and could be useful for solid-state laser development and for optoelectronic device applications. The relatively intense peaks are attributed to the RE^{3+} intra-4f radiative transitions [1,7]. The narrow line widths are

a result of the free ion like behaviour of the RE^{3+} ion due to 5s and 5p shielding of the 4f orbital from the ligand environment [1].

Among the many different techniques of synthesizing RE^{3+} doped semiconductors [3,5] interest has shifted towards semiconductor powders [1,7]. Powder synthesis provides a simple and low-cost procedure for the incorporation of RE^{3+} dopants into semiconductors [2]. In particular, the chemical bath deposition method is a low temperature (80 °C) technique.

In the literature, studies of ZnO:Tb^{3+} and ZnO:Eu^{3+} have focused on ultra-violet excitation and visible emission for white light production [1,7]. In this work, visible-region excitation and visible-region emission studies of ZnO:Tb^{3+} and ZnO:Eu^{3+} are presented. Crystal-field analyses of Tb^{3+} ions and Eu^{3+} ions occupying a C_{3v} symmetry site have been performed and crystal-field energy levels deduced. This study of the visible-region emission characteristics of Tb^{3+} ions and Eu^{3+} ions in ZnO is important for development of solid-state lasers and for other optoelectronic device applications. Furthermore, the effect of the incorporation of RE^{3+} ions in the ZnO structure has been investigated.

2. Experimental aspects

The chemical bath deposition method was utilized to synthesize Tb^{3+} and Eu^{3+} doped as well as undoped ZnO powders. The precursors were analytical grade 0.03 mol of zinc acetate ($Zn(CH_3COO)_2 \cdot 2H_2O$) dissolved in 60 ml of deionised water, 0.09 mol of thiourea ($(NH_2)_2CS$) dissolved in 500 ml of deionised water, 123.5 ml of ammonia (25% NH_3) dissolved in 500 ml of deionised water; these were mixed in volume ratios of 60 ml : 60 ml : 60 ml. To obtain a 1.0 mol% RE^{3+} dopant concentration 0.0003 mol of either terbium nitrate ($Tb(NO_3)_3 \cdot 6H_2O$) or of europium acetate ($Eu(CH_3COO)_3 \cdot 2H_2O$), both of 99.9 % purity, were added. The detailed preparation procedure was as presented by Koao et al. [7] and the resulting powder was pressed using a hydraulic press at 80 kN to obtain a pellet of diameter 8 mm and a thickness of 2 mm. The pellets were annealed at 700 °C for two hours in air.

Room temperature photoluminescence measurements were carried out on the pellets, using 457.9 nm (21 838.8 cm^{-1}), 476.5 nm (20 986.4 cm^{-1}) and 488.0 nm (20 491.8 cm^{-1}) excitation lines of the Spectra Physics 2080 argon ion laser. The emission was detected using a Burle C31034-A02 photomultiplier tube coupled to a McPherson 2062DP monochromator together with an SRS preamplifier and photon counter. Typically, laser powers of 35 mW and monochromator slit-widths of 250 μm were used. Spectra were recorded in the 460 – 900 nm range. The 515 nm and 495 nm long pass filters were used with the 488.0 nm and the 476.5 nm laser lines, respectively. Grazing incidence x-ray diffraction (GIXRD) measurements were also carried out on all the samples using the 0.154 nm Cu K_α wavelength of the Bruker D8 Discover at $\omega = 1°$. The measurements were carried out in the range $20° \leq 2\theta \leq 80°$.

3. Results and discussion

3.1. XRD studies

The grazing incidence XRD patterns obtained at room temperature for pristine ZnO, 1.0 mol% Tb^{3+} doped and 1.0 mol% Eu^{3+} doped ZnO powders are presented in figure 1 for the as-prepared and the annealed samples. The annealing treatment was carried out so as to remove precursor impurity phases that were observed in the as-prepared samples. GIXRD results show the presence of a wurtzite ZnO phase ((100), (002), (101), (102), (110), (103), (112) and (202)) in all the annealed samples; there are no phase changes arising from the addition of the Eu^{3+} ions and the Tb^{3+} ions. This supports the suggestion that the incorporation of the Tb^{3+} ions and Eu^{3+} ions into the host matrix is substitutional in nature [6,7,8]. In this study, it was not possible to experimentally identify the lattice location of the RE^{3+} ions. However in the literature, the Tb^{3+} ion (0.92 Å) and the Eu^{3+} ion (0.95 Å) have been proposed to occupy the Zn^{2+} (0.74 Å) lattice site in the ZnO structure [4,7].

The observed broadening of the XRD peaks has been attributed to the small crystallite sizes of the powders [9]. As such the average crystallite sizes of the samples were calculated from the Debye

Scherrer formula [4]. The lattice constants 'a' and 'c' are presented in table 1 together with the crystallite sizes 'D'.

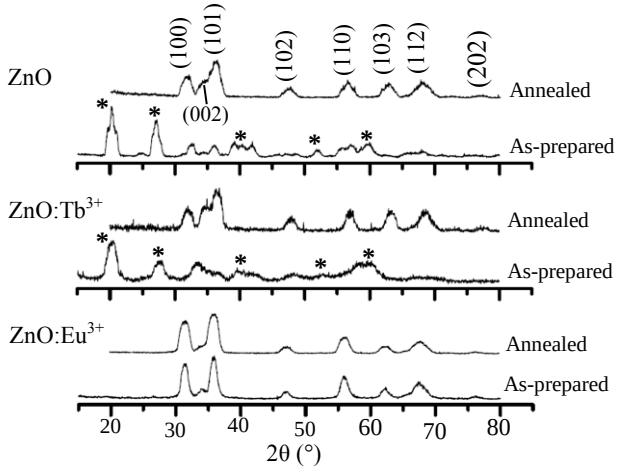


Figure 1. Room temperature grazing incidence XRD patterns for as-prepared and annealed pristine ZnO, 1.0 mol% Tb^{3+} doped and 1.0 mol% Eu^{3+} doped ZnO powders. (*) Zinc acetate precursor impurity.

The XRD patterns show a hexagonal wurtzite ZnO structure for all the samples. Therefore, the incorporation of the Tb^{3+} ions and the Eu^{3+} ions does not result in an effective change of the ZnO host structure.

Table 1. Lattice constants and crystallite sizes for undoped, Tb^{3+} and Eu^{3+} doped ZnO powders.

Sample	Lattice constants (±0.006 nm)		Crystallite size (±1 nm)
	a (nm)	c (nm)	D (nm)
ZnO	0.326	0.519	9
ZnO:Eu^{3+}	0.327	0.523	9
ZnO:Tb^{3+}	0.322	0.517	10

3.2. Laser excited photoluminescence

Photoluminescence spectra of undoped as well as 1.0 mol% Tb^{3+} and 1.0 mol% Eu^{3+} doped ZnO pellets were obtained at room temperature with argon laser excitation. The 457.9 nm, 476.5 nm and 488.0 nm excitation wavelengths were used. Although there are spectral variations in the doped samples arising from the different excitation lines, the same spectrum was obtained for the undoped sample with all three excitation lines. The 488.0 nm spectrum included in figures 2(a) and 3(a) was of slightly better intensity. The photoluminescence spectra of the annealed samples showed no significant changes in the Tb^{3+} ion and Eu^{3+} ion emissions compared to the un-annealed (as-prepared) samples. The spectra of the as-prepared samples are presented in figures 2(a) and 3(a).

The spectrum for undoped ZnO (figures 2 and 3) shows a broad emission band with almost uniform intensity from 550 nm to about 700 nm, followed by a gradual decrease in intensity with increase in wavelength thereafter. This emission is attributed to deep level transitions within the ZnO band gap due to intrinsic defects [3,4].

For ZnO:1.0 mol% Tb^{3+}, electronic transitions from the $^5D_4$ multiplet to all the ^7F_J (J = 0, 1, 2, 3, 4, 5, 6) multiplets were observed (figure 2(a)). Energy transfer from the intrinsic defects in the host lattice to the RE^{3+} ion results in the population of the Tb^{3+} $^5D_4$ multiplet [1,4]. Transitions to the $^7F_2$, $^7F_1$ and $^7F_0$ multiplets were quite weak by comparison (figure 2(a) inset). With either 457.9 nm or 476.5 nm excitation the sample exhibits a broad emission band spanning the visible region of the electromagnetic spectrum on which relatively intense peaks are superimposed. As the excitation wavelength is increased from 457.9 nm to 488.0 nm the broad emission band is increasingly suppressed while the Tb^{3+} emission bands increase in intensity. The $^5D_4 \rightarrow {}^7F_{6,5,4,3}$ emission bands are

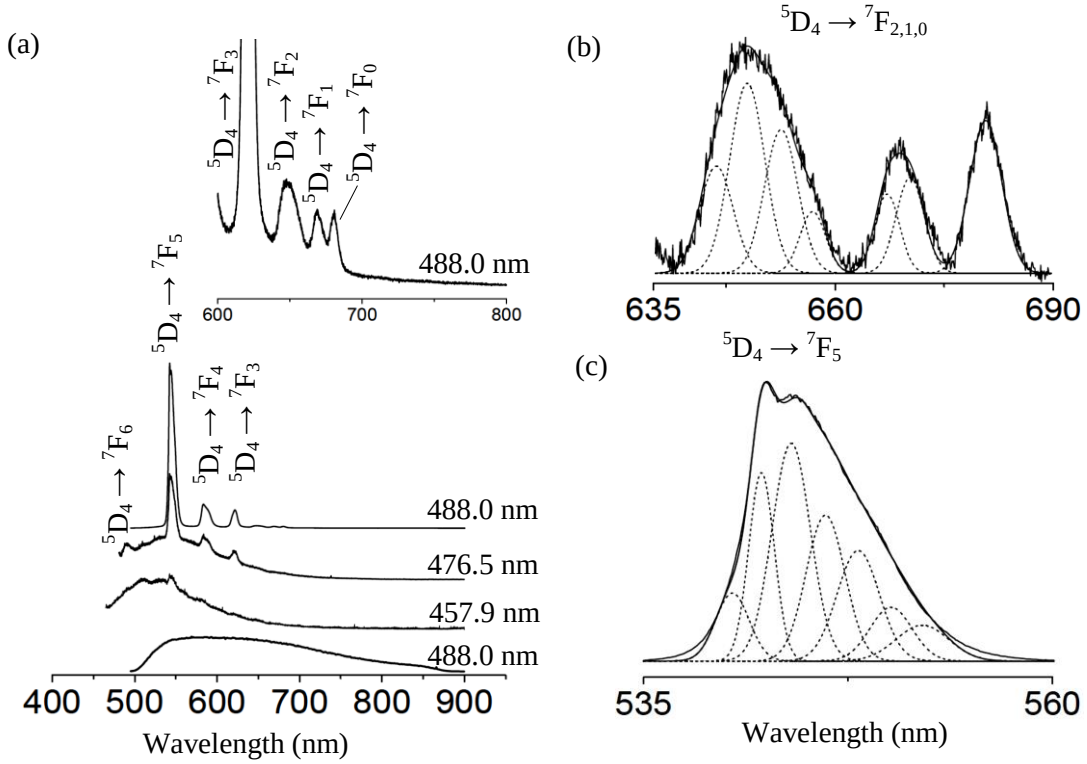


Figure 2. (a) Room temperature photoluminescence spectra of ZnO:Tb^{3+} and undoped ZnO pellets. (b) and (c) Gaussian peak fitted emission bands. The excitation wavelengths are as indicated while the peak labels identify the multiplets associated with each transition band.

as reported in the literature from ultra-violet excitation [1,4] while the $^5D_4 \rightarrow {}^7F_{2,1,0}$ emission bands have not been reported before. The 488.0 nm excitation wavelength is in resonance with the $^5D_4$ multiplet and as a result, the emission is significantly stronger rendering the weaker transitions observable.

Knowledge of the RE^{3+} ion site symmetry in the host matrix is important for determining the RE^{3+} ion energy levels resulting from crystal-field effects [10]. It is from the site symmetry and the selection rules that RE^{3+} ion transitions can be assigned to specific originating and terminating crystal-field energy levels, thereby determining the energy level scheme of the RE^{3+} ion in a particular host lattice [10]. Pereira et al. [8] proposed a C_{3v} site symmetry for the RE^{3+} ion at a Zn^{2+} lattice site and this configuration is adopted for the analyses presented here. In this work, the number of crystal-field energy levels expected for a C_{3v}-symmetry site were determined from group theory decomposition tables [11]. However, electronic transitions between some of the crystal-field energy levels are forbidden resulting in fewer experimentally-determined levels for some of the multiplets.

In order to deduce the crystal-field energy levels from the observed photoluminescence spectra, Gaussian peak profiles were fitted to the observed Tb^{3+} emission bands as demonstrated in figure 2(b) and (c). Governed by the selection rules, fittings were carried out starting with the $^7F_0$ multiplet as it has only one energy level and therefore a single transition. Each Gaussian peak in the emission band represented an electronic transition from a crystal-field energy level in the $^5D_4$ multiplet to a crystal-field energy level in each of the ^7F_J multiplets. From the results, two crystal-field energy levels of the excited $^5D_4$ multiplet were deduced to be at 20 483 cm^{-1} (488.2 nm) and 20 602 cm^{-1} (485.5 nm). At room temperature electrons have sufficient thermal energy to occupy the second level of the $^5D_4$ multiplet resulting in transitions from the higher level (20 602 cm^{-1}) being present as well and mixed in with those from the lower level at 20 483 cm^{-1}. The crystal-field energy levels determined here for the

$^5D_4$ and the ^7F_J (J = 0, 1, 2, 3, 4, 5, 6) multiplets of Tb^{3+} in ZnO are presented in table 2; the number in brackets is the total number of levels for each multiplet as obtained from the decomposition tables.

Table 2. Crystal-field energy levels of the Tb^{3+} ion in ZnO in cm^{-1} (± 2 cm^{-1}).

Multiplet	$^7F_6$ (9)	$^7F_5$ (7)	$^7F_4$ (6)	$^7F_3$ (5)	$^7F_2$ (3)	$^7F_1$ (2)	$^7F_0$ (1)	$^5D_4$ (6)
Levels	0	2 046	3 287	4 330	5 065	5 495	5 794	20 483
	95	2 106	3 340	4 359	5 162	5 571		20 602
	199	2 178	3 398	4 399	5 267			
	408	2 245	3 465	4 439				
		2 301	3 531	4 486				
		2 377	3 597					

For ZnO:1.0 mol% Eu^{3+}, electronic transitions from the $^5D_0$ multiplet to the ^7F_J (J = 0, 1, 2, 3, 4) multiplets were observed (figure 3(a)). In addition, transitions from the $^5D_1$ multiplet to the $^7F_3$ and $^7F_4$ multiplets overlapped with the $^5D_0 \rightarrow {}^7F_{1,2}$ transitions, respectively. All three excitation wavelengths show Eu^{3+} ion related emission bands also superimposed on a broad emission background. The emission spectrum obtained with 457.9 nm excitation was quite noisy and had the least intense Eu^{3+} emission bands in comparison to those for the 476.5 nm and 488.0 nm excitation lines hence it is not included in figure 3(a). These emission bands have been reported in the literature from ultra-violet excitation studies but with no further analysis [7]. Increasing the excitation wavelength from 457.9 nm to 488.0 nm resulted in a relative increase in the intensity of the Eu^{3+} emission bands.
The Eu^{3+} crystal-field energy levels were deduced using the same deconvolution procedure as for Tb^{3+}, as in figure 3(b) and (c). The position of the $^5D_1$ multiplet was deduced from the $^5D_1 \rightarrow {}^7F_{3,4}$ transitions although it was not possible to identify the $^5D_1 \rightarrow {}^7F_{0,1,2}$ transitions. From the results, the crystal-field energy level of the $^5D_0$ multiplet was deduced to be at 17 259 cm^{-1} (579.4 nm) while the lower-lying crystal-field energy level of the $^5D_1$ multiplet was placed at 18 847 cm^{-1} (530.6 nm). The crystal-field energy levels deduced for the $^5D_{0,1}$ and the ^7F_J (J = 0 to 4) multiplets of Eu^{3+} in ZnO are presented in table 3.

The photoluminescence spectra (figures 2 and 3) show strong green emission for ZnO:Tb^{3+} and strong red emission for ZnO:Eu^{3+}. In general, the Eu^{3+} ion emission was weaker than the Tb^{3+} ion

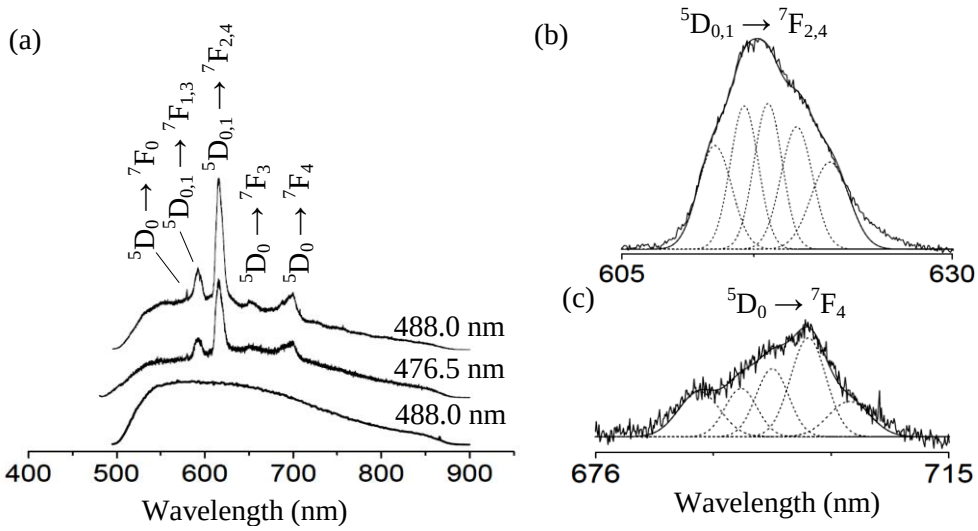


Figure 3. (a) Room temperature photoluminescence spectra of ZnO:Eu^{3+} and the undoped ZnO pellets. (b) and (c) Gaussian peak fitted emission bands.

Table 3. Crystal-field energy levels of the Eu^{3+} ion in ZnO in cm^{-1} (± 2 cm^{-1}).

Multiplet	$^7F_0$ (1)	$^7F_1$ (2)	$^7F_2$ (3)	$^7F_3$ (5)	$^7F_4$ (6)	$^5D_0$ (1)	$^5D_1$ (2)
Levels	0	373	919	1 849	2 568	17 259	18 847
			1 025	1 904	2 734		
			1 088	1 964	2 812		
				2 034	2 885		
				2 091	2 967		
					3 058		

emission for the same excitation conditions. The 457.9 nm excitation proved to be inefficient for populating the emitting multiplets while the 488.0 nm wavelength was the most suitable excitation line.

4. Conclusions

The XRD results showed no changes in the ZnO phase with the introduction of Tb^{3+} or Eu^{3+} ions. Room temperature Tb^{3+} and Eu^{3+} transitions in the visible region of the electromagnetic spectrum were observed in ZnO powders with excitation from the 457.9 nm, 476.5 nm and 488.0 nm argon laser lines. A total of 29 crystal-field energy levels for Tb^{3+} and 18 for Eu^{3+} were determined. The ZnO:Tb^{3+} and ZnO:Eu^{3+} emission transitions and energy level schemes are of significance for solid-state laser development and for optoelectronic device applications such as green and red light emitting diodes.

Acknowledgements

Financial support from the DST-NRF Centre of Excellence in Strong Materials (CoE-SM) and the University of the Witwatersrand, Johannesburg is gratefully acknowledged. We are also grateful to J. Augustine, C. Sandrock, V. Govender, L. Mafemba and G. Peters of the School of Physics, University of the Witwatersrand for their technical support.

References

[1] Kumar V, Som S, Kumar V, Kumar V, Ntwaeaborwa O M, Coetsee E and Swart H C 2014 *Chem. Eng. J.* **255** 541-552

[2] Podhorodecki A, Nyk M, Misiewicz J, and Strek W 2007 *J. Lumin.* **126** 219-224

[3] Ziani A, Davesnne C, Labbe C, Cardin J, Marie P, Frilay C, Boudin S and Portier X 2014 *Thin Solid Films* **553** 52-57

[4] Pal P P and Manam J 2013 *Mat. Sci. Eng. B* **178** 400-408

[5] Zhang J, Feng H, Hao W, and Wang T 2006 *Mat. Sci. Eng. A* **425** 346-348

[6] Yang L, Tang Y, Hu A, Chen X, Liang K and Zhang L 2008 *Phys. B* **403** 2230-2234

[7] Koao L F, Dejene F B, Kroon R E, and Swart H C 2014 *J. Lumin.* **147** 85-89

[8] Pereira A S, Peres M, Soares M J, Alves E, Neves A, Monteiro T and Trindade T 2006 *Nanotechnology* **17** 834-839

[9] Mote V D, Purushotham Y and Dole B N 2012 *J. Theoretical and Appl. Phys.* **6** 1-8

[10] Hufner S 1978 *Optical spectra of transparent rare earth compounds*. (New York: Academic Press) pp 1-13

[11] Koster G, Dimmock J O, Wheeler R G, and Statz H 1963 *Properties of the thirty-two point groups*. (Cambridge: M.I.T press) pp 55-57

Methane dissociation over Pd nanocluster: DFT study

M.H. Chuma[1], H.R. Chauke[1], G. Jones[2] and P.E. Ngoepe[1]

[1]Materials Modelling Centre, University of Limpopo, Private Bag x1106, Sovenga, 0727
[2]Johnson Matthey Technology Centre, Building 22 CSIR, Meiring Naude Road, Brummeria, Pretoria

E-mail: moyahaboc@gmail.com

Abstract. Palladium is often used as a catalyst for many processes in emissions control technologies [1, 2]. This is due to its potential of becoming a novel catalyst for low temperature methane combustion [3]. Palladium nanoclusters have a number of surface features and various adsorption sites, their relative activity play important role if predictions are to be made for improved materials properties. The dissociation of methane over palladium nanoclusters and high index surfaces was investigated using density functional theory (DFT) as implemented in grid based projector augmented wave (GPAW) code [4]. Methane has been adsorbed on the Pd_{13} nanoclusters at various active sites, and the results shows that CH_4 dissociate into CH_3 and H. It is found that CH_3 bind strongly to the top site forming a Pd-C bond of 2.050 Å and H on bridge with a Pd-H bond of 1.716 Å. Methane dissociation was found to be more favorable on the nanocluster than on Pd surfaces.

1. Introduction

Precious metal catalysts, which include the platinum group metals (PGMs) for example Pt, Pd, Rh and noble metals such as Ag and Au, find their way into processes of emissions control technology for application in methane oxidation and catalytic converters [5, 6]. The mostly used catalyst is Pt and Pt-based. Recently, Pd-based catalysts are being developed since they have similar electronic configurations and lattice constants as Pt but cheaper [7].

In recent years, metal nanoparticles have been of interest in several areas including physics and material sciences. The properties of these nanoparticles were found to differ from those of bulk materials [8-10]. In addition, nanoparticles have large surface area and are rich in defect such as steps and kinks. This enhances the activity of most reactions in the emission control technology [11].

Methane dissociative adsorption is generally assumed as the rate determining step for methane oxidation [12]. In the current study, we use the DFT calculations to understand the methane dissociation reaction process over Pd_{13} nanocluster and compare its catalytic activity to Pd surface. This was achieved by carrying out adsorption calculations of molecules and intermediates that form part of the methane dissociation reaction on possible adsorption sites of Pd_{13} nanocluster and Pd(100) surface. The uniqueness of the Pd nanocluster and surface under consideration is their similarity of active adsorption sites and allow comparison to be made precisely; that is the top, bridge and 4fold hollow adsorption sites.

2. Methodology

Calculations were performed using DFT as implemented in the GPAW code [4]. GPAW is an implementation of the projector augmented wave (PAW) method [13] that uses real space grids to represent the wave functions, densities and potentials. The revised version of the Perdew–Burke–Ernzerhof (RPBE) exchange correlation function was used with a grid-spacing of 0.18 Å and k-points mesh of 4x4x4 for the bulk Pd. The Pd_{13} nanocluster was set to have a cell volume of 18x18x18. A five layered Pd(100) slab surface with a vacuum thickness of 10 Å was used in the calculations. The surface was set to be periodic along the x- and y-directions, whereas non-periodic along the z-direction. A k-mesh of 4x4x1 was used for the surface relaxation and molecule adsorption calculations.

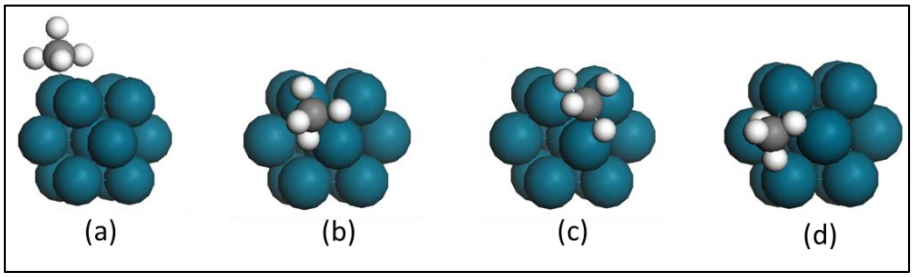


Figure 1. Pd_{13} nanoclusters showing adsorption of CH_4 at various sites namely (a) top, (b) bridge, (c) 3fold hollow and (d) 4fold hollow.

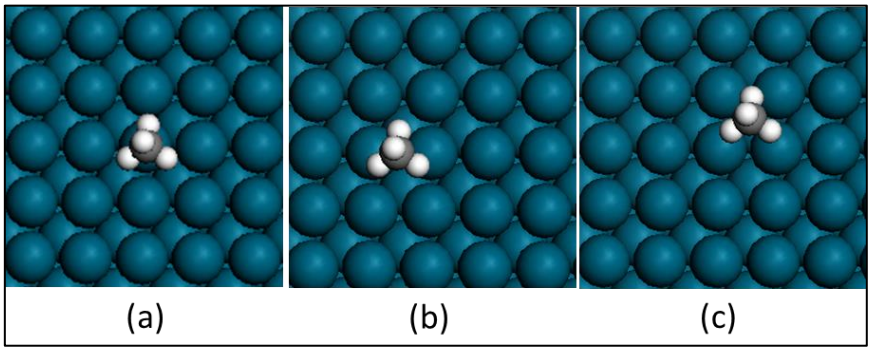


Figure 2. Top view of Pd(100) surfaces showing adsorption of CH_4 at various sites namely (a) top, (b) bridge, (c) 4fold hollow.

Pd_{13} nanocluster consists of four adsorption sites whereas Pd(100) surface exhibits three high symmetry adsorption sites as shown in Figure 1 and 2. Different adsorbates can possibly be adsorbed on these sites and the site with the lowest adsorption energy is considered as the most preferred adsorption site. The adsorption energies for the adsorbates are calculated using the equation,

$$E_{ads} = E_{a/m} - E_a - E_m \quad (1)$$

where *a* represents the adsorbate and *m* the Pd_{13} nanocluster or Pd(100) surface.

3. Results and Discussion

3.1. Validation of the method

The bulk Pd structure with a space group of Fm-3m was used to cleave the Pd(100) surface and create the Pd_{13} nanocluster. The Pd nanocluster has a radius of 5 Å and consists of 13 atoms with Pd-Pd average bond length of 2.751 Å Pd(100) was set to have a 2x2 supercell with five atomic layers, where the bottom two layers were constrained while allowing the remaining three layers to relax.

The DFT method used to perform the calculations was tested and validated by determining the lattice parameter for bulk Pd structure, bond length for CH_4 molecule, Pd(100) surface and Pd_{13} nanocluster. Note that, before the properties could be calculated, all the structures were subjected to full geometry optimization allowing cell and volume to change, and the equilibrium parameters were recorded when the structures reached their ground state energy and converged to 1 meV. We found that the lattice parameter and bond length are in reasonable agreement to the experimental results with bulk Pd value of a = 3.983 Å as compared to 3.891 Å and the C-H bond length of 1.097 Å for CH_4, as shown in table 1. Lattice parameters and bond length of bulk Pd and Pd(100) surface agrees to within 3% with the available experimental data. The Pd-Pd and C-H average bond length for Pd_{13} and CH_4 are also in good agreement to within 1% of the experimental results.

Table 1. Lattice parameters of bulk Pd, the average bond length (C-H) for CH_4 molecule, (Pd-Pd) for Pd(100) surface and Pd_{13} nanocluster.

Structure	Lattice parameters (Å)		Average bond length (Å)	
	DFT	Experimental	DFT	Experimental
Pd	3.983	3.891[a]	2.816	2.751[a]
Pd(100)	2.821	2.744[b]	2.850	-
Pd_{13}	-	-	2.751	2.726[c]
CH_4	-	-	1.097	1.087[d]

[a] The experimental lattice parameter for bulk Pd were taken from ref [14]
[b] The experimental lattice parameter were for Pd(100) were obtained from ref [15]
[c] The Pd_{13} experimental bond length was taken from ref [16]
[d] The CH_4 experimental bond length was obtained from ref [17]

3.2. Dissociation of methane over Pd_{13} nanocluster and various Pd surfaces

Methane dissociation is determined by examining the adsorption of various molecules and intermediates that form part of the dissociation reaction. In this work CH_4, CH_3, CH_2, CH, C and H are adsorbed on various sites of Pd_{13} nanocluster and Pd(100) surface. The preferred adsorption sites and energies of these adsorbates are listed in Table 2We note that CH_3 prefers to be adsorbed on the top site for both Pd_{13} nanocluster and Pd(100) surface, with a Pd-C bond length of 2.050 Å and 2.059 Å, respectively. CH_2 prefers the bridge site whereas CH and C prefer various hollow sites depending on the adsorbents. The H prefers bridge adsorption site on the nanoclusters with Pd-H bond 1.716 Å, whereas on the Pd surface, H will likely be adsorbed on the 4fold hollow adsorption site. Interestingly, adsorption of CH_4 on Pd_{13} nanocluster gives -0.51 eV, suggesting that the nanocluster slightly binds on the 4fold hollow site. The adsorption energy of -0.01 eV has been reported for Pd(100) by Trinchero *et al.* [11], which is less as compared to the calculated value obtained for Pd_{13} nanocluster. However, this has been recorded for different adsorption sites.

Table 2. Preferred adsorption sites and adsorption energy (E_{ads}) of various molecules and intermediates that form part of the methane dissociation reaction (listed under adsorbate) over Pd_{13} nanocluster and Pd(100) surface.

Adsorbate	Pd_{13}		Pd(100)	
	site	E_{ads} (eV)	site	E_{ads} (eV)
CH_4	4fold hollow	-0.51	-	-
CH_3	top	0.41	top	1.01
CH_2	bridge	1.04	bridge	1.64
CH	4fold hollow	0.69	4fold hollow	1.55
C	4fold hollow	0.47	4fold hollow	1.85
H	bridge	-0.59	4fold hollow	-0.32

The superimposed reaction profile for methane dissociation on Pd_{13} nanocluster and Pd(100) surface are shown in Figure 3. The dissociation of methane is denoted as a process of continuously removing hydrogen from CH_4 during the interaction of methane with the Pd nanoclusters and surfaces. The dissociation of methane is then expressed by equation (2) – (5):

$$CH_4\,(g) + 2^* = CH_3^* + H^* \quad (2)$$

$$CH_3^* + {}^* = CH_2^* + H^* \quad (3)$$

$$CH_2^* + {}^* = CH^* + H^* \quad (4)$$

$$CH^* + {}^* = C^* + H^* \quad (5)$$

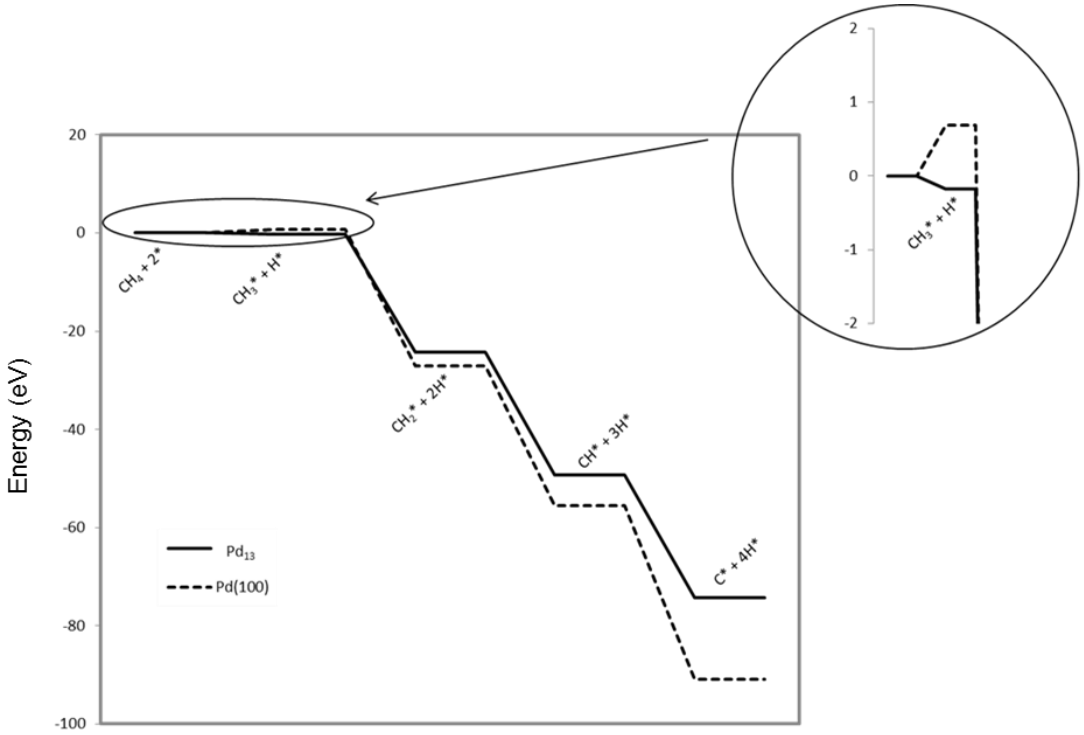


Figure 3. The superimposed reaction profile for methane dissociation on Pd_{13} nanocluster and Pd(100) surface.

The reaction starts with methane in a gas phase and two identical Pd nanoclusters or Pd surfaces denoted by an asterisk (*) and ends with C* and 4H*. Four possible dissociation processes are indicated. The first dissociation stage, is the production of CH_3* and H*. This process is exothermic on Pd_{13} nanocluster whereas endothermic on Pd(100) surface. This suggests that Pd_{13} is the most preferred adsorbent for methane dissociation at standard conditions.

Now, considering methane dissociation for the other three stages, that is, the production of CH_2* + 2H*, CH* + 3H* and C* + 4H*. The processes are exothermic on both the adsorbents. The production of C* + 4H* dominate the reaction process, showing the lowest energy. We also note that methane dissociation process is spontaneous for the Pd_{13} nanocluster, whereas non-spontaneous for Pd(100) surface. This suggest that Pd_{13} is more favourable as compared to Pd(100) surface.

4. Conclusion

DFT results predicted good agreement of the lattice parameters and bond lengths with experiments to within 3% and 1%, respectively. Furthermore, adsorption of various adsorbates on three and four possible adsorption sites reflect lower adsorption energies for Pd surface and nanocluster, respectively. It was obtained that CH_3 prefers the top site on Pd_{13} nanocluster whereas CH_2 and H prefer the bridge. Furthermore, both CH and C prefers 4fold hollow adsorption site. The first dissociation stage of methane, is observed when CH_4 dissociates into CH_3* forming a Pd-C bond of 2.050 Å and H* with a Pd-H bond of 1.716 Å on Pd_{13} nanocluster. Methane dissociation over Pd_{13} nanocluster is favorable for all the dissociation stages under standard condition, and thus considered the most preferred adsorbent as compared to Pd(100) surface.

Acknowledgments

The research was conducted at the Materials Modelling Centre (MMC), University of Limpopo; and has benefited from computing resources at the Centre for High Performance Computing (CHPC). We acknowledge the financial support of Johnson Matthey Technology Centre and the National Research Foundation.

References

[1] Mateos-Pedrero C, González-Carrazán S R, Soria M A and Ruíz P 2013 *Catal. Today* **203** 158-62
[2] Lyubovsky M, Roychoudhury S and LaPierre R 2005 *Catal. Lett.* **99** 113-17
[3] Lang S M, Fleicher I, Bernhardt T M, Barnett R N and Landman U 2014 *J. Phys. Chem.* A **118** 8572-82
[4] Enkovaara J, Rostgaard C, Mortensen J J, Chen J, Dułak M, Ferrighi L, Gavnholt J, Glinsvad C, Haikola V, Hansen H A, Kristoffersen H H, Kuisma M, Larsen A H, Lehtovaara L, Ljungberg M, Lopez-Acevedo O, Moses P G, Ojanen J, Olsen T, Petzold V, Romero N A, Stausholm-Møller J, Strange M, Tritsaris G A, Vanin M, Walter M, Hammer B, Häkkinen H, Madsen G K H, Nieminen R M, Nørskov J K, Puska M, Rantala T T, Schiøtz J, Thygesen K S and Jacobsen K W 2010 *J. Phys.: Condens. Matter* **22** 253202-24
[5] Rao C R K and Trivedi D C 2005 *Coord. Chem. Rev.* **249** 613-31
[6] Shao Y, Liu J, Wang Y and Lin Y J 2009 *J. Mater. Chem.* **19** 46-59
[7] Wang M, Guo D and Li H 2005 *J. Solid State Chem.* **178** 1996-2000
[8] Wu B, Kuang Y, Zhang X and Chen J 2011 *Nano Today* **6** 75-90
[9] Piccolo L and Henry C R 2000 *Surf. Sci.* **452** 198-206
[10] Thomann A L, Rozenbaum J P, Brault P, Andreazza-Vignolle C and Andreazza P 2000 *Appl. Surf. Sci.* **158** 172-83
[11] Trinchero A, Hellman A, Grönbeck H 2013 *Surf. Sci.* **616** 206-13
[12] Butch R, Urbano F J, Loader P K 1995 *Appl. Catal.,* A **123** 173-84
[13] Blöchl P E 1994 *Phys. Rev.* B **50** 17953-79

[14] Stevens K J, Ingham B, Toney M F, Brown S A and Lassesson A 2008 *Curr. Appl. Phys.* **8** 443-6
[15] Lide D R (Editor) 1995 CRC handbook of chemistry and physics: a ready-reference book of chemical and physical data, CRC Press **76**
[16] Schwerdtfeger P (Editor) 2004 Elsevier, Theoretical and computational chemistry **14** 691 ISBN: 0-444-51299-3
[17] Hirota E 1979 *J. Mol. Spectrosc.* **77** 213-21

Effect of background gas and substrate temperature on ZnO:Zn PLD thin films

E Hasabeldaim, O M Ntwaeaborwa, R E Kroon and H C Swart
Department of Physics, University of the Free State, Bloemfontein ZA-9300, South Africa

E-mail: swarthc@ufs.ac.za

Abstract. The dependence of the structural and optical properties of ZnO:Zn thin films deposited by Pulsed Laser Deposition at different preparation conditions have been investigated. All the films showed highly preferential c-axis orientation, and their crystallinities were improved with increasing the substrate temperature. The stress in the thin films has varied from -12.7 to -7.30 GPa depending on the background gas. The minimum value was obtained in the case of the oxygen background gas. The optical bandgap varied from 3.12 to 3.20 eV with the deposition background gas for Argon, Vacuum and Oxygen. The bandgap was also affected by the substrate temperature for the thin films deposited in oxygen gas. All the thin films exhibited ultraviolet emission. The thin films deposited in the oxygen atmosphere showed strong photoluminescence emission around the green and orange region (500 – 627 nm). These results suggest that the ZnO:Zn thin films may be used in the design, simulation and fabrication of optoelectronics devices such as white light emitting diodes.

1. Introduction

Recently, zinc oxide (ZnO) thin films have been studied extensively because of their excellent properties and various applications. The inherent defects in the ZnO matrix such as oxygen vacancies (V_o), oxygen interstitials (O_i), oxygen antisites (Zn_o), Zn vacancies (V_{Zn}), Zn interstitial (Zn_i) and Zn antisites (O_{zn}) play important roles in determining its optical and structural properties [1]. Various studies have predicted that oxygen vacancies, interstitial oxygen, zinc vacancies, and zinc interstitials are responsible for the luminescence in ZnO [2].The green emission from ZnO has been known at least since the turn of the century. The addition of dopants influences the ZnO defects as well as enhancing its optical properties [3].

A variety of techniques have been utilized to fabricate ZnO thin films, such as thermal deposition, sputtering, chemical vapour deposition, spray pyrolysis and pulsed laser deposition (PLD) [4,5]. In this study commercially available ZnO:Zn was used to fabricate thin films on Si(111) substrates by using the PLD technique. Photoluminescence (PL) using a He-Cd laser at 325 nm and UV-Vis spectroscopy were employed to elucidate the origin of the luminescence emission as well as determine the bandgap. It was found that orange and green emissions of ZnO were due to defects. The effect of background gas and the substrate temperature on the luminescence of ZnO:Zn PLD thin films were explored.

2. Experimental details

A commercial ZnO:Zn phosphor powder, obtained from Phosphor Technology, was pressed without binders using an in-house built sample holder to create an ablation target. The ablation target was

annealed at 250 ^{0}C for 24 h in air to remove all adventitious water containing species that might be present. The target was then placed inside the chamber of the PLD system on a rotating target holder that can also move linearly (up and down). Si (111) was used as a substrate, and was cleaned in an ultrasonic bath using acetone, then ethanol and finally deionized water, and then dried with N_2 gas. The 266 nm fourth harmonic of a Nd:YAG laser (40 mJ/pulse and 10 Hz pulse repetition rate) was used for the preparation of the ZnO:Zn thin films with PLD under different experimental conditions. The substrate temperature was fixed at 300 ^{o}C during the deposition of the thin films in vacuum and with argon background gas atmosphere, but varied for the thin films deposited in an oxygen atmosphere (50, 200, 300 and 400 ^{o}C). The laser deposition time was fixed at 25 min. The chamber was pumped down to a base pressure of 5×10^{-5} mbar and then backfilled with O_2 and Ar gases to a partial pressure of 5×10^{-2} mbar.

The structure of the thin films samples were characterized by X-ray diffraction (XRD) using a Bruker D8 Advance Diffractometer (40 kV, 40 mA) with Cu K_α x-rays (λ = 0.154 nm). Atomic force microscopy (AFM) using a Shimadzu SPM-9600 instrument was used to study the surface and roughness morphology. UV-Vis diffuse reflection spectra were measured using a Perkin Elmer Lambda 950 spectrophotometer. PL spectra were collected with a He-Cd laser PL system with a 325 nm excitation wavelength.

3. Results and discussion

3.1. XRD and structure analysis

Figure 1 shows the XRD patterns of the ZnO:Zn thin films deposited on the silicon substrates for (a) different background gasses and (b) different substrate temperatures, respectively. ZnO has a tetrahedral coordinated (wurtzite) structure in which the (0 0 2) plane has the lowest surface energy [6]. Hence, all the samples exhibited preferential (0 0 2) plane orientation, which is in line with the characteristics of the hexagonal wurtzite structure of ZnO where the c-axis lies perpendicular to the substrate surface. Crystallite size, stress and lattice parameters (a, c) of the thin films were calculated using the full width at half maxima of the (0 0 2) peak by using the Scherrer equation, biaxial strain model and the Miller indices formulas [7] and are tabulated in tables 1 and 2. The crystallite size was altered by increasing the substrate temperature or changing the chamber atmosphere (vacuum, oxygen and argon). The stress on the thin film varied as a function of lattice parameters, which were affected by the background gas. The crystallite size also increased when increasing the substrate temperature (table 2), this is probably due to the agglomeration of the grains or particles on the film. It can also be seen from figure 1 that the intensity of the (0 0 2) diffraction peak was affected by the background gas and the substrate temperature. The highest intensity was obtained from the thin film deposited in oxygen gas at a substrate temperature of 300 ^{o}C.

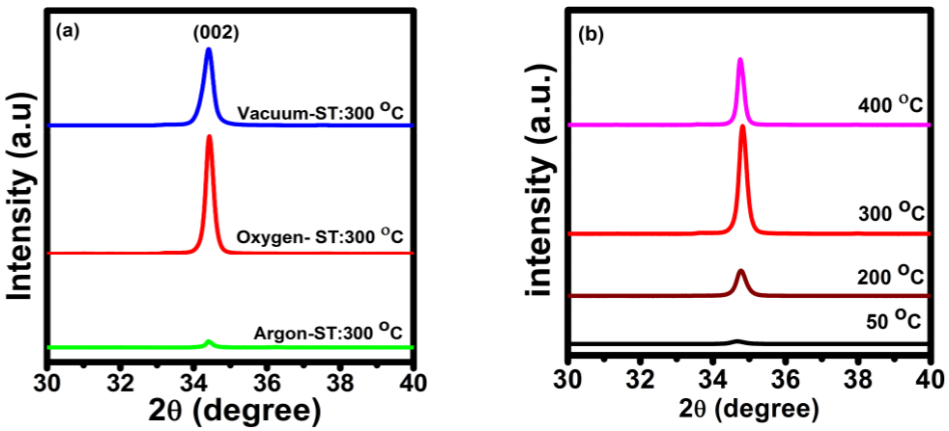


Figure 1. XRD patterns of ZnO:Zn thin films deposited (a) in different growth atmosphere and (b) for different substrate temperatures.

Atmosphere	Lattice constant (Å)		Stress (GPa)	Crystallite size (nm)	Surface roughness (nm)	Bandgap (eV)
	c	a				
Vacuum	5.123	2.958	-7.305	40	2.048	3.19
Argon	5.077	2.931	-11.481	46	7.085	3.12
Oxygen	5.063	2.923	-12.767	38	2.016	3.20

Table 1. Effect of growth atmosphere on ZnO:Zn thin films produced with a substrate temperature of 300 $^{\circ}$C.

Table 2. Effect of different substrate temperature on ZnO:Zn thin films produced with oxygen background gas.

Substrate temperature ($^{\circ}$C)	Lattice constant (Å)		Stress (GPa)	Crystallite size (nm)	Surface roughness (nm)	Bandgap (eV)
	c	a				
50	5.077	2.931	-11.481	24	3.023	3.16
200	5.063	2.923	-12.767	31	3.461	3.16
300	5.063	2.923	-12.767	38	2.016	3.20
400	5.063	2.923	-12.767	48	7.525	3.14

3.2. Surface morphology and roughness

The surface morphology of the ZnO:Zn films are shown as two and three dimensional AFM micrographs (3.00 × 3.00) μm^2 in figure 2. All the thin films showed distinguishable particles with a uniform distribution. The clear particle images were obtained from the thin films deposited in an argon atmosphere and oxygen (400 ^{o}C substrate temperature). The roughness of the thin films varied from 2.016 to 2.048 to 7.085 nm with the variation in background gas from oxygen, vacuum and argon, respectively (table 1). In reality, the roughness for the samples in vacuum or oxygen are almost the same, while that for argon is rougher. The roughness of the thin films obtained in an oxygen background gas also varied with increasing the substrate temperature as shown in table 2.

3.3. UV-Vis reflectance and photoluminescence properties

Figure 3 shows the UV-Vis reflectance curves of the ZnO:Zn thin films as deposited in different atmospheres, and at different substrate temperatures for an oxygen background gas. A sharp band edge observed between 353 and 376 nm (figure 3(a,b)) was assigned to the bandgaps absorption of ZnO due to electron transitions from the valence band to the conduction band ($O_{2p} – Zn_{3d}$) [7].

The optical bandgap was estimated by the extrapolation of the linear portion of the $[F(R_o)h\nu]^2$ versus $h\nu$ plots and was calculated by the Kubelka – Munk equation using the Tauc method:

$$[F(R_o)h\nu]^2 = C_2(h\nu\text{-}E_g). \quad (1)$$

C_2 is a constant, E_g is the optical bandgap, h is Plank's constant and $F(R_o)$ is the absorption coefficient.

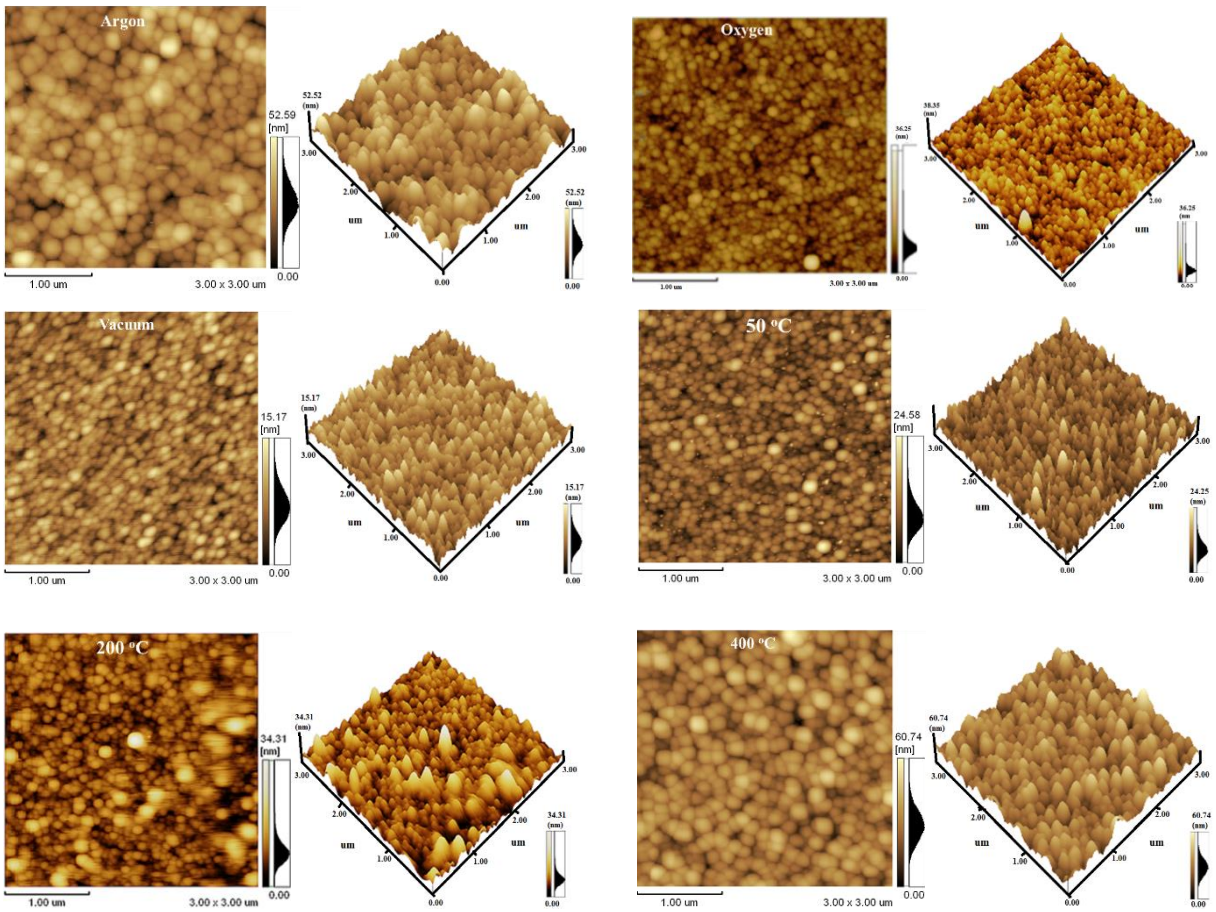


Figure 2. AFM images of ZnO:Zn thin films obtained in different growth atmospheres (300 ^{o}C substrate temperature) and for different substrate temperatures for an oxygen background gas.

The calculated optical band-gaps for the different background gas and substrate temperature for the thin films obtained at different background gas and temperature are tabulated in tables 1 and 2, respectively. The bandgaps of the thin films were strongly affected by the type of background gas and substrate temperature. In the case of the thin film obtained in an oxygen background gas shown in figure 3(c, d), the band marked with (*) is corresponding to the orange and green band in PL spectra figure 4(c).

The PL spectra are shown in figure 4. All the samples exhibited ultraviolet emission around 378 nm characteristic of band-to-band emission from ZnO, and their intensities were altered by the chamber background gas (fig. 4(a)). The weak emission band in some spectra near about 760 nm (e.g. for argon in fig. 4(a)) occurs when this ultraviolet emission is intense and is its second order. The thin films deposited in an oxygen atmosphere, which is deconvoluted in fig 4(c) showed green and orange bands. The green band was attributed to recombination of carriers bound at V_o and O_i defects, while the orange band was attributed to the recombination between carriers bound at Zn_i and O_i defects [8]. The oxygen vacancies were formed due to the excess of zinc atoms in the original ZnO:Zn PLD target and the oxygen interstitial defects were formed from oxygen in the oxygen atmosphere, respectively. Worth noting is that the green and orange bands correspond to the second absorption peak in the UV-

Vis curve (figure. 2(c)), which is labeled with an asterisk (*). The green and orange emission that appeared in the oxygen atmosphere due to the defects was completely diminished in the thin film deposited in the vacuum and argon gas atmospheres (figure 4(a)). This confirmed that there was a low defect concentration in the samples prepared using these two atmospheres.

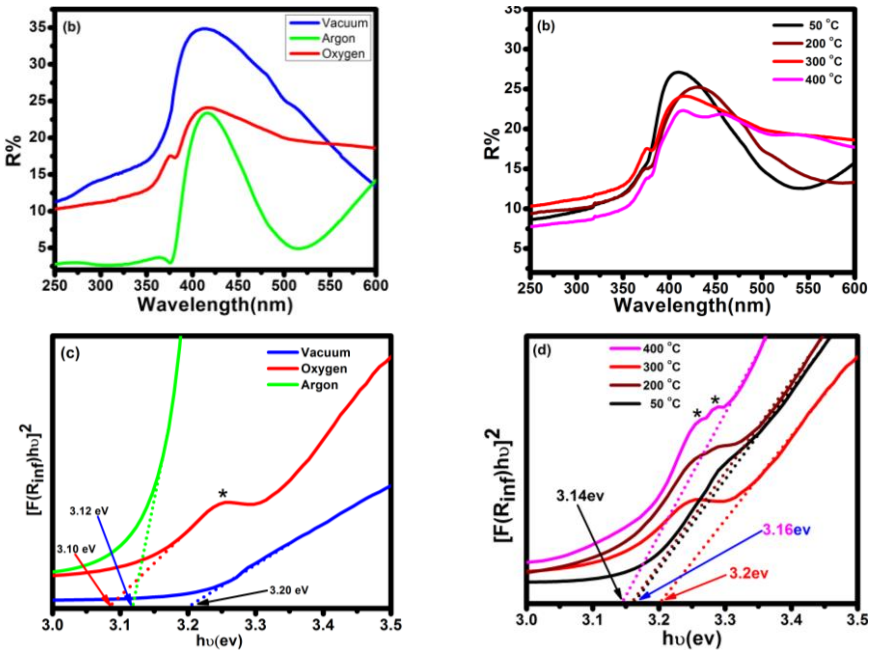


Figure 3. Reflectance curves (a,b) and Tauc plots (c,d) used to calculate the band-gaps for ZnO:Zn thin films deposited in different growth atmospheres (a,c) and different substrate temperatures (b,d).

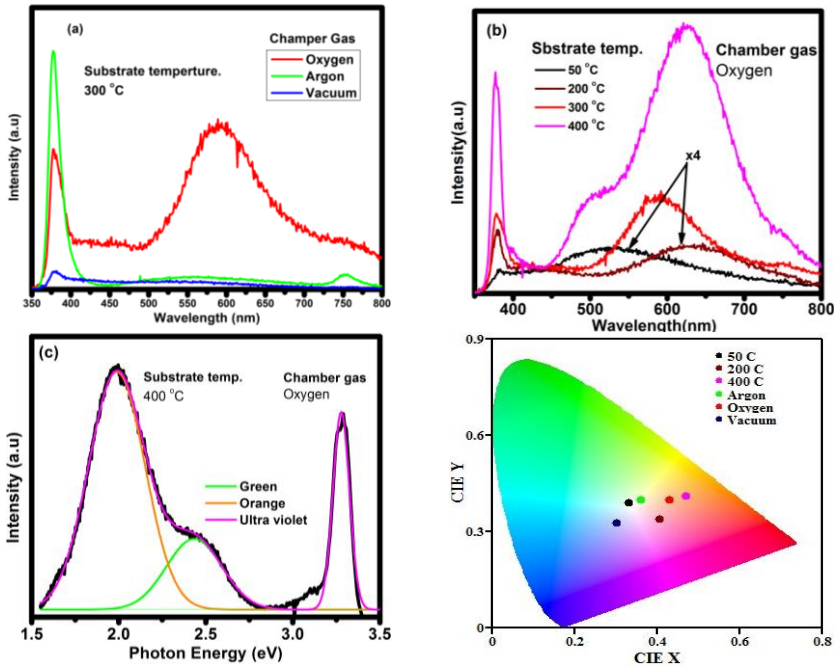


Figure 4. PL spectra of ZnO:Zn thin films deposited for (a) different growth atmospheres (300 °C substrate temperature), (b) for different substrate

temperatures for an oxygen background gas, (c) Deconvoluted peak of oxygen atmosphere at 400 ^{o}C substrate temperature and (d) CIE diagram of all samples. Data was averaged over three scans and spikes caused by the photomultiplier tube detector, detected in only one of the three scans, were removed.

4. Conclusion

C-axis oriented ZnO:Zn thin films were deposited on Si (111) substrates in different growth atmospheres by the PLD technique. The PL results showed a broad emission around 500 and 600 nm, which confirmed the presence of oxygen related defects in the thin films deposited in the oxygen atmosphere. The crystallite size and PL emission intensity depended on the growth atmosphere and substrate temperature. The results suggested that ZnO:Zn thin films may be a suitable candidate material for white light emitting diode applications.

Acknowledgment

This work is based on the research supported by the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation of South Africa (84415). The University of the Free State Cluster program is thanked for financial support. The PL system used is supported by the rental pool program of the National Laser Centre (NLC).

Reference

[1] Anderson J and Chris G V 2009 *Rep. Prog. Phys*. **72** 126501
[2] Kumar V, Swart H C, Ntwaeaborwa O M, Kroon R E, Terblans J J, Shaat S K K, Yousif A, Duvenhage M M 2013 *Mater. Lett.* **101** 57
[3] Wei L, Dongsheng M, Fumin Z, Xi W, Xianghuai L, Shichang Z, Yukun Z, Qiong L, Jingfang X 2000 *Nucl. Instrum. Meth. Phys. Res.* B. **169** 59
[4] Ohyam M, Kozka H, Yoko T 1997 *Thin Solid Films* **306** 78
[5] Kim Y, Weon B, Su-Jeong Shu A 2004 *Thin Solid Films* **491** 153
[6] Nakanishi Y, Miyake Y, Kominami Y , Aoki T, Hatanaka Y, Shimaoka G 1999 *Applied Surface Science*. **142** 233
[7] Kumar V, Swart H C, Som S, Kumar V, Yousif A, Pandey A, Shaat S K and Ntwaeaborwa O M 2014 *Laser Phys*. **24** 105704
[8] Kumar V, Som S, Kumar Vijay, Kumar Vinay, Ntwaeaborwa O M, Coetsee E, Swart H C 2014 *Chem. Eng. J*. **255** 541

Structural and magnetic properties of $NiFe_2O_4$/NiFe bi-magnet and NiFe nano-alloy synthesized from thermal reduction of $NiFe_2O_4$

Itegbeyogene P Ezekiel [1], **Thomas Moyo and Hafiz M I Abdallah**
School of Chemistry and Physics, University of KwaZulu-Natal, P/Bag X5400, Durban 4000, South Africa

E-mail: itegbeyogene@gmail.com

Abstract. $NiFe_2O_4$/NiFe nanocomposites, NiFe nano-alloy were synthesized by reduction of $NiFe_2O_4$ nano-ferrite with activated charcoal under Ar gas atmosphere at 900 °C for 3 hours. The $NiFe_2O_4$ was synthesized by a glycol-thermal. Partial and complete reduction yielded $NiFe_2O_4$/NiFe and NiFe, respectively. NiFe was formed at an optimum amount of activated charcoal of n_c = 5. Phase identification, morphology and magnetic properties were performed by X-ray diffraction (XRD), high-resolution scanning electron microscope, ^{57}Fe Mössbauer spectroscope and mini cryogenic-free system. $NiFe_2O_4$ has an average crystallite size of 10 nm, an XRD density of 5.3 g/cm^3 and an average lattice parameter of a = 8.362 ± 0.007 Å. NiFe exhibited the martensite bcc (α-Fe) and austenite fcc (γ-Fe) phases in coexistence with average a of 2.867±0.001 Å and 3.580±0.001 Å, respectively. Fitted Mössbauer analysis for n_c = 5 and 6 show high hyperfine magnetic fields associated with the bcc phase while the lower field component is associated with the fcc phase of NiFe. The saturation magnetization increased significantly from 57 emu/g to 141 emu/g at room temperature. The saturation magnetization is enhanced at low temperatures with a maximum of 161 emu/g at ≥ 30 K. The coercivity showed no significant increase at low temperatures.

1. Introduction

Research efforts for scalable and efficient synthesis method of nano-materials have increased in recent times. This has happened mainly because of targeted properties of these materials for specific applications in areas such as catalysis, ferrofluids, drug delivery, microwave devices, laser cavity, gas-sensor, magnetic recording and dimension stabilizers [1, 2]. Nano-ferrites and their nanocomposites are promising and attractive materials for most of these applications. Nickel nano-ferrite ($NiFe_2O_4$) is a typical example of such material. $NiFe_2O_4$ is a soft magnetic material [3] with Ni having much lower anisotropy effect compared to cobalt Co. $NiFe_2O_4$ has an inverse spinel structure where the Ni^{2+} ions are located in octahedral sites and the Fe^{3+} ions located both in octahedral (B) and tetrahedral (A) sites [4]. The location of the divalent cations has a relationship with the magnetic properties. The nanocomposite of $NiFe_2O_4$/NiFe is expected to exhibit an enhanced magnetization due to the synergetic saturation magnetization arising from the ferrite and alloy phases. The alloy phase within an appropriate composition shows some interesting anomalous thermal and magnetic properties [5] for

[1] To whom any correspondence should be addressed.

suitable applications. The improvement of these soft magnetic materials to maintain high saturation, permeability, resistivity and relative stability in air is paramount. Hu *et al* [6] reported the synthesis and characterization of $NiFe_2O_4$/NiFe nanocomposites synthesized by colloidal chemical method combined with hydrogen (H_2) reduction with a relatively low magnetization. In our research, we report the synthesis of our parent sample to be $NiFe_2O_4$ nano-ferrite, NiFe as the nano-alloy and $NiFe_2O_4$/NiFe bi-magnetic nanocomposites. The as-prepared sample was synthesized by a glycol-thermal method and reduced with activated charcoal. The use of activated charcoal is cheap and safer compared to the use of H_2. The nanocomposite synthesis of $NiFe_2O_4$/NiFe by reduction using activated charcoal on $NiFe_2O_4$ nano-ferrite has not been reported. A strong interest in investigating the improved magnetic and structural properties arising from these synthesis methods informed this research. These properties were investigated using X-ray powder diffraction, high-resolution scanning electron microscope, ^{57}Fe Mössbauer spectroscope and a mini cryogenic-free system.

2. Experimental details

$NiFe_2O_4$ was synthesized by the glycol-thermal method [7]. $NiFe_2O_4$ was reduced with specific amounts of carbon n_c = 1, 2, 4, 5, 6, 8 and 10 to produce $NiFe_2O_4$/NiFe and /NiFe based on this equation; $NiFe_2O_4 + n_cC \rightarrow NiFe + n_cCo_2$, for 0.5g fixed mass of NiFe2O4. The reductions were performed at 900 °C for 3 hours in argon atmosphere. XRD, high resolution scanning electron microscope (HRSEM), ^{57}Fe Mössbauer spectroscope and a mini cryogen free system was used to characterize the samples.

3. Results and discussion

The X-ray diffraction (XRD) patterns for $NiFe_2O_4$ nano-ferrite, $NiFe_2O_4$/NiFe nanocomposites and NiFe nano-alloy formed with various amounts of activated charcoal (n_c) are shown in Figure 1. The pure nano-ferrite at n_c = 0 conforms to the cubic spinel structure of bulk $NiFe_2O_4$ (JCPDS Card No. 10-0325) [1]. All the peaks are indexed with respect to Miller indices (111), (220), (311), (400), (422), (511) and (440) without evidence of any impurity peaks. The crystallite size (G) was calculated using the Scherrer's formula with the full-width at half-maximum (w_{hkl}) of the line broadening diffraction peak derived from the most prominent (311) peak of the $NiFe_2O_4$. The shape factor K is equal to 1, λ is the wavelength of a CoKα source monochromatic X-ray beam and θ is the Bragg's angle. The X-ray density of the ferrite calculated using; $\rho_{XRD} = 8M/N_A a^3$ where M is the molecular weight, N_A is the Avogadro's number and a is the lattice parameter [7]. The crystallite size and the X-ray density of $NiFe_2O_4$ was estimated to be G = 10.48±0.08 nm and 5.3±0.1 g/cm^3 respectively. The value of the lattice parameter of this compound was calculated from the (311) peak to be a = 8.362±0.007 Å using $a = \lambda/2\sin\theta\sqrt{(h^2 + k^2 + l^2)}$ where h, k, and l are the Miller indices. This is close to a reported value of 8.339 Å [1]. The XRD diffraction results show that partial reduction occurs for n_c = 1 to 4 for activated charcoal to yield $NiFe_2O_4$/NiFe nanocomposites. The small unidentified peak close to the (311) peak when n_c = 1 is attributed to a possible transient impure phase which eventually disappears for n_c > 4. Complete reduction of the ferrite phase to NiFe nano-alloy to form fcc and bcc phases [8] occurs for at least n_c > 4. The presence of FeO for n_c = 1 to 4 is attributed to the possibility of still unreacted iron ions in the sample crystallite interior with the activated charcoal atoms reacting most effectively with surface nano-ferrites. In all reduced samples, NiFe nano-alloy phase shows the coexistence of the austenite γ-Fe fcc (Ni-rich) indexed to the (111) peak and the martensite α-Fe bcc (Fe-rich) lattice structures indexed to the (110) peak [9]. The fcc and bcc crystallite sizes G were calculated from the (111) and (110) peaks, respectively. The abrupt growth in the crystallite size at n_c = 2 could be influenced by the most prominent presence of FeO phase at n_c = 2. The (200) peak has a lattice parameter associated with the fcc lattice structure with a value of a = 3.585 Å. The values of the lattice parameter suggest the nano-alloy composition to be Fe-Ni at 42 % Ni [10] because around this composition the lattice parameter expands from 3.572 Å to 3.588 Å at room temperature. The (111) γ-Fe peak has the most predominant crystalline phase with no obvious change in peak growth compared to the bcc α-Fe phase. Figure 2 shows that crystallite size of the bcc phase is smaller with a continuous

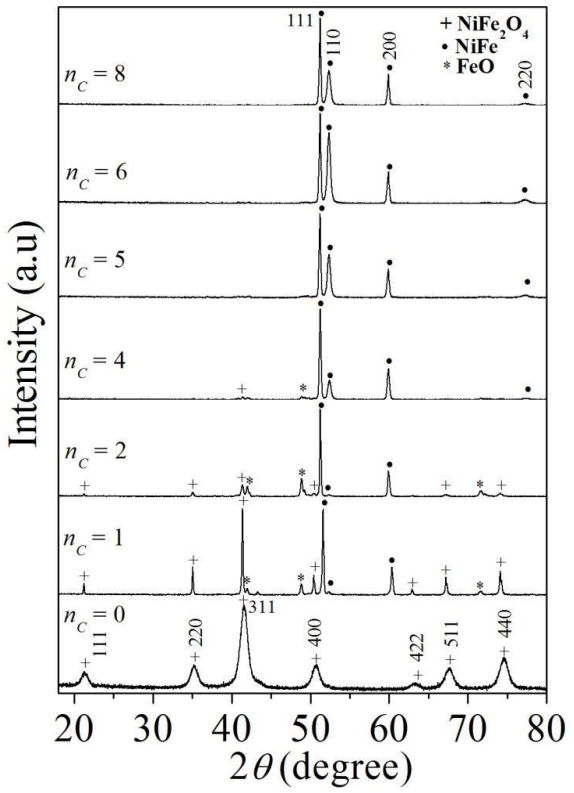


Figure 1. XRD patterns of $NiFe_2O_4$ reduced with n_c = 0, 1, 2, 4, 5, 6 and 8.

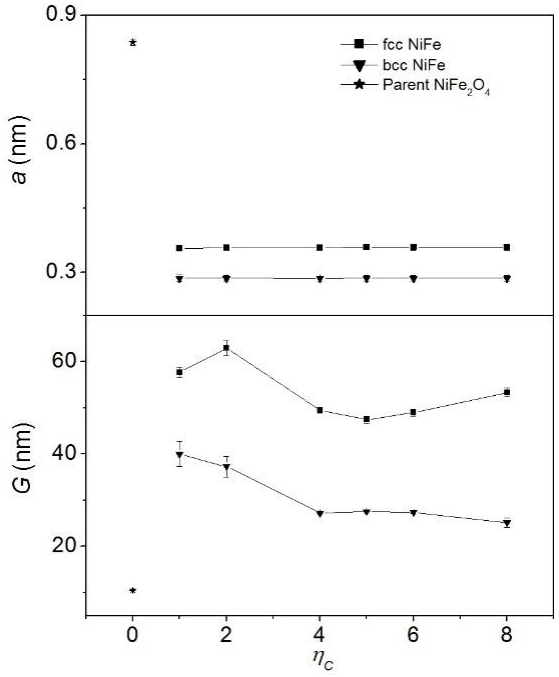


Figure 2. Variations of crystallite size G and lattice parameter a of the parent $NiFe_2O_4$, the bcc and fcc phases of the reduced samples at different amounts of activated charcoal n_c.

while that of the fcc phase is larger. The coexistence of the bcc and fcc phases give these nanocomposites and nano-alloys a plasticity and lattice resistance to deformation advantage because fcc structure favors plasticity while bcc structure favors lattice resistance to deformation [11]. Such NiFe particles have a higher resistance to oxidation than separate Fe and Ni particles [9]. The average values are close to the reported values of 2.87 Å and 3.52 Å respectively [5, 9].

The surfaces of the as-prepared nano-ferrite, nanocomposites formed at n_c = 2 and nano-alloys formed at n_c = 5 and 6 were examined by high resolution scanning electron microscope HRSEM as shown in Figure 3. The as-prepared sample is observed to consist of rough spherical particles which are less compact compared to the reduced samples. The nanocomposites are more compact with significant growth in the size of the particles. For n_c = 5 and 6 the particles are more closely packed which we attribute to the presence of the NiFe nano-alloy phase with higher density due to the presence of the fcc phase [10]. Similar images but less compacted were observed for FeNi alloy synthesized by chemical co-precipitation method [8].

^{57}Fe Mössbauer spectra measurements were performed at room temperature. The fitted Mössbauer spectra for n_c = 0, 5 and 6 are presented in Figure 4. Table 1 gives the values of the fitted parameters. The error function χ^2 provides an estimation of the fitting. Using the covariance matrix model in the Recoil fitting program, values of χ^2 of about one or less indicate better fits. The doublets in the reduced samples account for the coordination of Fe^{3+} and Fe^{2+} at both sites. The as-prepared $NiFe_2O$ Mössbauer spectrum was best fitted with two sextets corresponding to tetrahedral A- and octahedral B-sites. The hyperfine magnetic field B_{hf} values of 239 kOe and 420 kOe are consistent with α-Fe with comparable isomer shifts δ of 0.339 mm/s and 0.272 mm/s associated with Fe^{3+} ions [8, 12, 13]. The slightly smaller value of the isomer shifts at B-sites is due to larger covalency. The low quadrupole splitting values at both sites is indicative of the cubic symmetry around the Fe^{3+} ions [14]. The added doublet D accounts for the superparamagnetic relaxation effects due to the nano-size of the sample. Three sextets

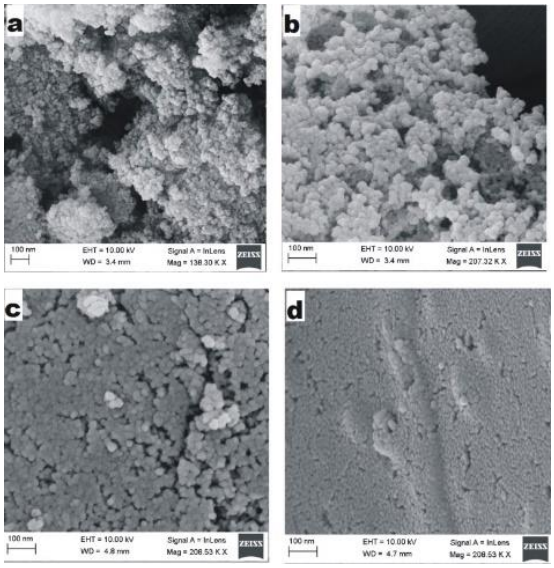


Figure 3. HRSEM images of: (a) the as-prepared $NiFe_2O_4$, (b) $NiFe_2O_4$/NiFe at n_c = 2, (c) NiFe at n_c = 5 and (d) NiFe at n_c = 6.

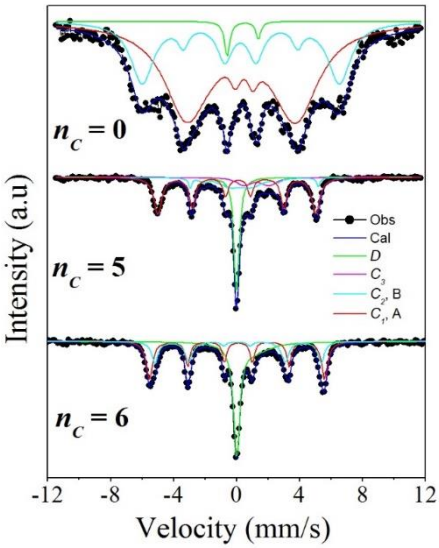


Figure 4. Mössbauer spectra of $NiFe_2O_4$ reduced with different amounts of activated charcoal n_c = 0, 5 and 6.

Table 1 Mössbauer parameters include hyperfine field (B_{hf}), isomer shifts (δ), quadrupole splitting (Δ_{EQ}), line widths (Γ), fraction populations (f) and reduced chi^2 (χ^2) of Fe ions of the as-prepared sample and alloy composite at n_c = 5 and 6 of activated charcoal.

n_C	Sub-pattern	B_{hf} (kOe) ±3	δ (mm/s) ±0.009	Δ_{EQ}(mm/s) ±0.002	Γ(mm/s) ±0.01	f (%) ±0.6	χ^2
0	A	293	0.339	0.005	0.58	68.0	0.7825
	B	420	0.272	0.014	0.49	30.0	
	D	-	0.420	2.150	0.23	1.9	
5	C_1	339	0.049	-0.048	0.19	39.5	1.3207
	C_2	325	0.900	-0.737	0.44	11.3	
	C_3	274	0.786	0.455	1.16	24.1	
	D	-	-0.017	0.139	0.22	25.0	
6	C_1	345	0.067	-0.042	0.16	35.1	0.9815
	C_2	327	0.053	0.032	0.19	24.5	
	D	-	0.540	1.080	-	40.4	

C_1, C_2 and C_3 are required to fit the spectra for n_c = 5 and two sextets C_1 and C_2 for n_c = 6 which we associate with the bcc and fcc phases. The high field components C_1 has a B_{hf} value of 339 kOe and 345 kOe which is attributed to the bcc FeNi and the lower field components C_2 has a value of 325 kOe and 327 kOe attributed to the fcc phase [10, 15] for n_c = 5 and 6, respectively. Sextet C_3 accounts for the small mixed phase denominated by fcc phase with a B_{hf} value of 274 kOe. These results are close to reported values of 343 kOe, 320 kOe and 273 kOe for FeNi alloy [8, 10, 16]. The isomer shift values obtained for the fully reduced samples are significantly different and appear to confirm change from Fe^{3+} to Fe^{2+} [17].

Room temperature magnetic hysteresis loops of the parent sample, nanocomposites and NiFe nano-alloys recorded in applied magnetic fields of up to 50 kOe are presented in Figure 5. Magnetic properties such as coercivity (H_C) and saturation magnetization (M_S) were calculated from the hysteresis loops. The samples exhibit soft magnetic hysteresis loops and the maximum magnetization

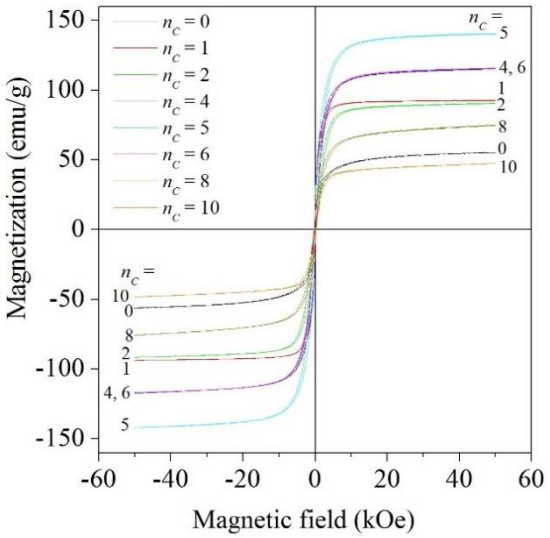


Figure 5. Magnetic hysteresis loops for the as-prepared $NiFe_2O_4$, $NiFe_2O_4$/NiFe nanocomposites and NiFe nano-alloy samples measured at room temperature.

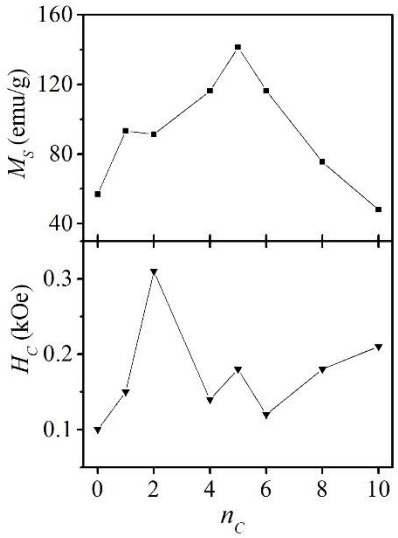


Figure 6. Variations of saturation magnetization M_S and coercivity H_C with n_c.

at a magnetic field of 50 kOe is herewith regarded as the estimate for saturation magnetization at a particular temperature. The as-prepared nano-ferrite has a coercivity of 0.10 kOe and saturation magnetization of 57 emu/g. The variations of the magnetization and coercivity with ratios of activated charcoal n_c are presented in Figure 6. The reduction of $NiFe_2O_4$ nano-ferrite appears to enhance the magnetizations and coercive fields. The fluctuating variation of the coercivity is not quite clear. The magnetization attains a maximum value of 141 emu/g at n_c = 5 while the coercive field has a maximum of 0.31 kOe at n_c = 2. For n_c > 5 magnetization decreases which we associate with the onset of the NiFe bcc and fcc phases with weaker internal hyperfine fields as deduced from Mössbauer spectroscopy. Higher magnetization and coercive fields are therefore obtained in the $NiFe_2O_4$/NiFe nanocomposites which begin to decline as inter metallic NiFe phase begins to dominate more and more. Low temperature measurements were also performed on a mini cryogen free measurement system. Typical hysteresis loops obtained at 4 K are presented in Figure 7. Figure 8 shows the coercivity and saturation magnetization dependence on temperature. The magnetizations are observed to increase gradually with decreasing temperature. A more significant increase in coercive fields is observed in the $NiFe_2O_4$ nano-ferrite which appears to follow the typical behavior for the spinel ferrites [4]. The nanocomposite sample $NiFe_2O_4$/NiFe at n_c = 4 had the highest magnetization below about 100 K compared to the parent nano-ferrite or fully reduced sample. The coercivity of the NiFe nano-alloy seems to be slightly affected by temperature change compared to the nanocomposite or the $NiFe_2O_4$.

4. Conclusions

Nanocomposites of $NiFe_2O_4$/NiFe and NiFe nano-alloy were synthesized by the reduction of $NiFe_2O_4$ using activated charcoal n_c. XRD measurements showed NiFe nano-alloy for n_c = 5 and 6. The peaks for the reduced samples were indexed to the bcc α-Fe and fcc γ-Fe lattices for the NiFe nano-alloy. Fitted Mössbauer analysis for n_c = 5 and 6 show high hyperfine magnetic fields associated with the bcc phase while the lower field component is associated with the fcc phase of NiFe. The isomer shift values indicate change to Fe^{2+} from Fe^{3+} in the spinel phase. The surface morphology for n_c = 5 and 6 have their particles closely packed and compacted which we attribute to the presence of NiFe alloy phase with high density of fcc packing. Room temperature measurements of magnetization showed enhanced saturation magnetizations. The coercivity of NiFe is less dependent on temperature.

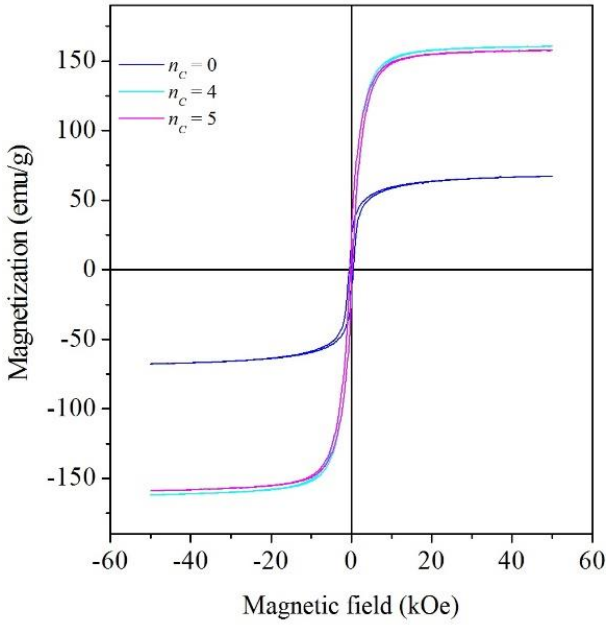


Figure 7. Magnetic hysteresis loops for as-prepared $NiFe_2O_4$ nano-ferrite, $NiFe_2O_4$/NiFe nanocomposite and NiFe nano-alloy samples measured at 4 K.

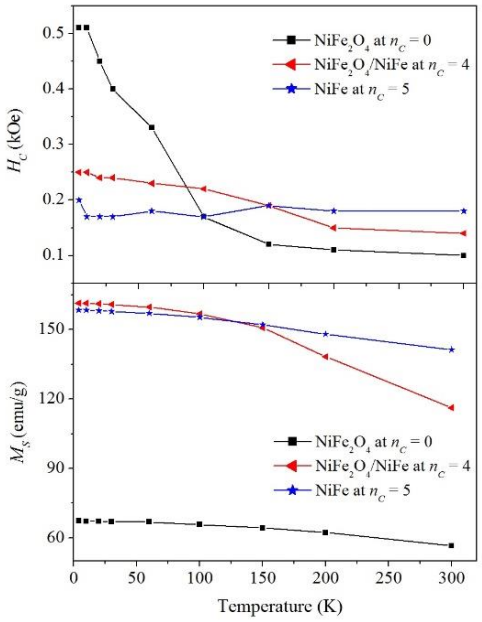


Figure 8. Coercivity H_C and saturation magnetization M_S dependence on temperature T for as-prepared $NiFe_2O_4$ nano-ferrite, $NiFe_2O_4$/NiFe nanocomposite and NiFe nano-alloy.

Acknowledgments

We express our gratitude to the National Research Foundation (NRF), South Africa for the NNEP grant for the VSM equipment, EM unit (UKZN, WC) for HRTEM and HRSEM.

References

[1] Sivakumar P, Ramesh R, Ramanand A, Ponnusamy S and Muthamizhchelvan C 2011 *J. Mater. Res*. **46** 2204
[2] Liu X G, Geng D Y, Choi C J and Kim J C 2009 *J. Nanopart. Res*. **11** 2097–2104
[3] Zhang Y, Zuo T, Cheng Y and Liaw P K 2013 *J. Sci. Rep*. **38** 1455
[4] Mazz K, Mumtaz A, Hasanain S K and Bertino M F 2010 *J. Magn. Magn. Mater*. **322** 2199–2202
[5] Rawwagah F H, Lehlooh A F, Mahmood S H, Mahmoud S, El-Ali A R, Said M R, Odeh I and Abu-Aljarayesh I 2012 *Jord. J. Phys*. **5** 9–14
[6] Hu R J and Chen A 2012 *J. Advn. Mater. Res*. **486** 65–69
[7] Abdallah H M I and Moyo T 2013 *J. Alloys Comp*. **562** 156–163
[8] Dong X L, Zhang Z D, Zhao X G and Chuang Y C 1999 *J. Mater. Res*. **14** 398–406
[9] Valderruten J F, Pérez A G A and Grenéche J M 2008 *J. Appl. Phys. Condens. Mater* **20** 485304
[10] Zhang Y, Zuo T, Cheng Y and Liaw P K 2013 *J. Sci. Rep*. **38** 455
[11] Leite G C P, Chagas E F, Pereira R, Prado R J, Terezo A J, Alzamora M and Saitovitch E B 2012 *J. Magn. Magn. Mater*. **324** 2711–2716
[12] Guittoum A, Layadi A, Tafat H, Bourzam A, Souam N and Lenoble O 2009 *J. Phil. Magn.* **88** 1085–1098
[13] Umare S S, Ningthoujam R S, Sharma S J, Kurian S S S and Gajbhiye N S 2008 *Hyperfine Interact*. **184** 235–243
[14] Rodriguez R R, Valenzuela J L and Tabares J *A Hyperfine Interact*. DOI 10.1007/s10751-013-0834-5.
[15] Lehlooh D A and Mahmood H S 2002 *Hyperfine Interact*. **139-140** 387–392
[16] Hassan M Y, El-Desoky M M, Masuda H, Kubuki S and Nishida T *Hyperfine Interact*. DIO 10.1007/s10751-011-0478-2
[17] Chunhui D, Gaoxue W, Dangwei G, Changjun J and Desheng X 2013 *Nanoscale Res. Lett*. **8** 196

Electronic and magnetic properties of the $(Cr_{84}Re_{16})_{100-y}Mn_y$ alloy system

BS Jacobs[1], AD Faceto[2,3], CJ Sheppard[1], ARE Prinsloo[1], PC de Camargo[2] and AJA de Oliviera[2]

[1] Department of Physics, University of Johannesburg, PO Box 524, Auckland Park, 2006, South Africa
[2] Universidade Federal de São Carlos, Rodovia Washington Luís, km 235 - SP-310 São Carlos, São Paulo, Brazil
[3] Universidade Federal dos Vales do Jequitinhonha e Mucuri, Avenida Vereador João Narciso, 1380, Unaí, Minas Gerais, Brazil

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

Abstract. This paper reports the results of the temperature dependence of the electrical resistivity (ρ) and magnetic susceptibility (χ) of $(Cr_{84}Re_{16})_{100-x}Mn_x$ alloys with x = 0.3, 0.4, 0.6, 0.8 and 3.1. Anomalies are observed in the $\rho(T)$ curves corresponding to the Néel temperatures (T_N) of the samples. χ (T) curves of the $(Cr_{84}Re_{16})_{100-x}Mn_x$ alloy system in a constant applied field of 100 Oe was obtained on increasing T after cooling in zero magnetic field (ZFC) and after cooling in a field of 100 Oe (FC). At low temperatures (T < 10 K), a sharp increase in χ is observed on increasing T after ZFC. A prominent sharp peak is observed close to 30 K beyond which χ rapidly decreases to lower values. The behaviour of χ is indicative of possible spin glass state at the boundaries of the itinerant antiferromagnetic order established by Mn frozen local moments.

1. Introduction

The magnetic properties of alloys of Cr have been a topic of research for many years, however, without much attention recently. Magnetism of Cr alloys still excites the experimentalist as the unique antiferromagnetic properties of Cr [1] are significantly altered by adding small quantities of other elements. Lomer [2] used the two-band model to explain the Fermi surface of Cr. This model describes the two important features on the Fermi surface of paramagnetic Cr namely the electron "jack" and a slightly larger hole "octahedron". The shapes of these two Fermi surface features are almost identical and thus shifting of one in reciprocal space through a wave vector (***Q***) results in the surfaces "fitting" or "nesting" nearly perfectly against each other [3]. These "connected" electron and hole states condense via Coulombic interaction into a spin-density-wave (SDW) of this wave vector [3]. This leads to the incommensurate (I) SDW magnetic state. The nesting is sensitive to the degree of matching between the two sheets and thus it is expected that the magnetic properties of Cr rely on the topology of the Fermi surface. The Fermi surface topology can be altered by varying the electron distribution [4]. The electron to atom ratio (e/a) of Cr is 6. This can be increased or decreased by alloying with elements to its right or left respectively in the periodic table. Alloying with Mn (e/a =7) results in the increase in size of the electron sheet and decrease in the size of the hole sheet. This makes Cr more commensurate (C) and eventually a CSDW phase is formed. The CSDW state is more

stable than the ISDW phase and the Néel temperature T_N is expected to increase with increase in Mn concentration [4].

The spin glass (SG) state is observed in cases where magnetic impurities (such as Mn or Fe) are hosted in non-magnetic metals [5]. In these systems, below a "freezing (or pinning) transition" at T_P, there is no long range magnetic order but the spins are locked in a frustrated low energy state [6]. The occurrence of the spin glass state in alloys containing Cr and Mn have been reported before [7,8,9,10]. In $Cr_{1-x}Mn_x$ alloys, characteristics typical of the spin-glass state, namely hysteresis of $M(H)$, peak in $\chi(T)$ in the ZFC state and relaxation of M as the logarithm of time when H is changed were observed [8]. However, it differs from a conventional SG since the magnetic susceptibility $\chi(T)$ is essentially independent of temperature between T_P and T_N, and that the pinning temperature (T_P) is essentially independent of Mn concentration. In the $(Cr + 1.5\ \% Fe)_{1-x}Mn_x$ system, the SG state was reported to coexist with Curie-Weiss paramagnetism and at low Mn concentrations ($x \leq 0.3\%$ Mn) the Fe impurity suppresses the spin-glass behaviour [9]. The $Cr_{80-x}Fe_{20}Mn_x$ system displays a sequence of SG-ferromagnetic-antiferromagnetic-paramagnetic magnetic phase transitions on heating from low temperatures [10].

The $Cr_{84}Re_{16}$ alloy is in the CSDW state at room temperature [11]. Doping with V (element to the left of Cr in the periodic table) drives the system towards incommensurability and reduces T_N due to the lower electron/atom ratio (*e*/*a)* of V [11]. Results of the effect of V doping on this alloy system has been reported [12]. The present study was conducted to understand the nature of the magnetic state when Mn (element to the right of Cr in the periodic table) is used as a dopant. Preliminary investigation confirms the increase of T_N with Mn doping and suggests the presence of the SG state in the $(Cr_{84}Re_{16})_{100-x}Mn_x$ alloy system.

2. Experimental

Ternary $(Cr_{84}Re_{16})_{100-y}Mn_y$ alloys with y = 0.3, 0.4 and 0.6, 0.8 and 3.1 were prepared by repeated arc melting in a purified argon atmosphere from Cr, Re and Mn each having mass fractional purity of 99.99 %. The alloys were separately sealed into quartz ampoules and 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. The crystal structure was confirmed using X-ray diffraction (XRD). Electrical resistivity (ρ) was measured above 300 K 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 using a SQUID-type vibrating sample magnetometer (VSM) MPMS3 of Quantum Design [13]. χ (T) curves of the $(Cr_{84}Re_{16})_{100-y}Mn_y$ alloy system in a constant applied field (100 Oe) was obtained on increasing T after zero field cooling (ZFC). Measurements were then also done on heating after field cooling (FC) at 100 Oe.

3. Results and discussion

Figures 1(a) and 1(b) show the representative XRD patterns for the $(Cr_{84}Re_{16})_{99.7}Mn_{0.3}$ and $(Cr_{84}Re_{16})_{96.9}Mn_{3.1}$ alloys using with Cu radiation (0.154 nm wavelength). The four characteristic peaks displayed in the XRD patterns confirm the expected body centred cubic (bcc) crystal structure for these alloys.

Measuring the temperature dependence of resistivity (ρ) is a well-established method in locating T_N [11]. The increase in ρ below T_N is due to electron-hole condensation in the nesting octahedra [11] leading to a change in the number of charge carriers available for conduction. As the temperature decreases, the condensation process progresses and there is a decrease in the final density of states. This leads to a decrease in intraband (between nesting octahedra) and interband (between octahedra and reservoir) scattering by impurity atoms which in turn decrease the contribution to resistivity [11]. Addition of Re to Cr, like Mn, increases the electron concentration and T_N [11]. As the spin-glass state is observed as a result of competing magnetic orders, it was necessary to first identify T_N to ensure that

the trends observed in the magnetic susceptibility curves correspond to the antiferromagnetic state of the alloys.

Representative $\rho(T)$ curves for the $(Cr_{84}Re_{16})_{100-y}Mn_y$ alloys with y = 0.3, 0.4 and 3.1, are shown in figure 2. The temperature associated with the minimum in $d\rho(T)/dT$ accompanying the $\rho(T)$ magnetic anomaly is defined as T_N [11]. In general, the size of the anomaly decreases but T_N increases with increase in Mn concentration as expected.

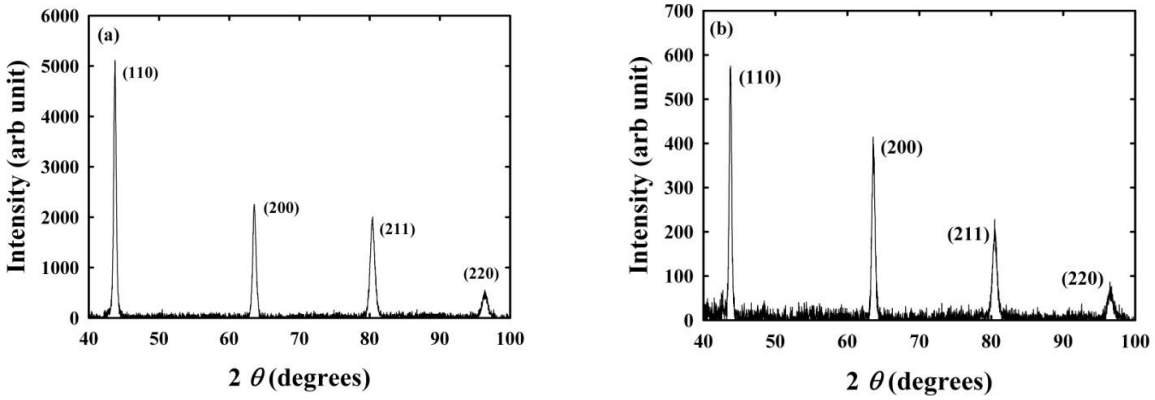


Figure 1: The XRD patterns for (a) $(Cr_{84}Re_{16})_{99.7}Mn_{0.3}$ sample and (b) $(Cr_{84}Re_{16})_{96.9}Mn_{3.1}$with the Bragg peaks indexed.

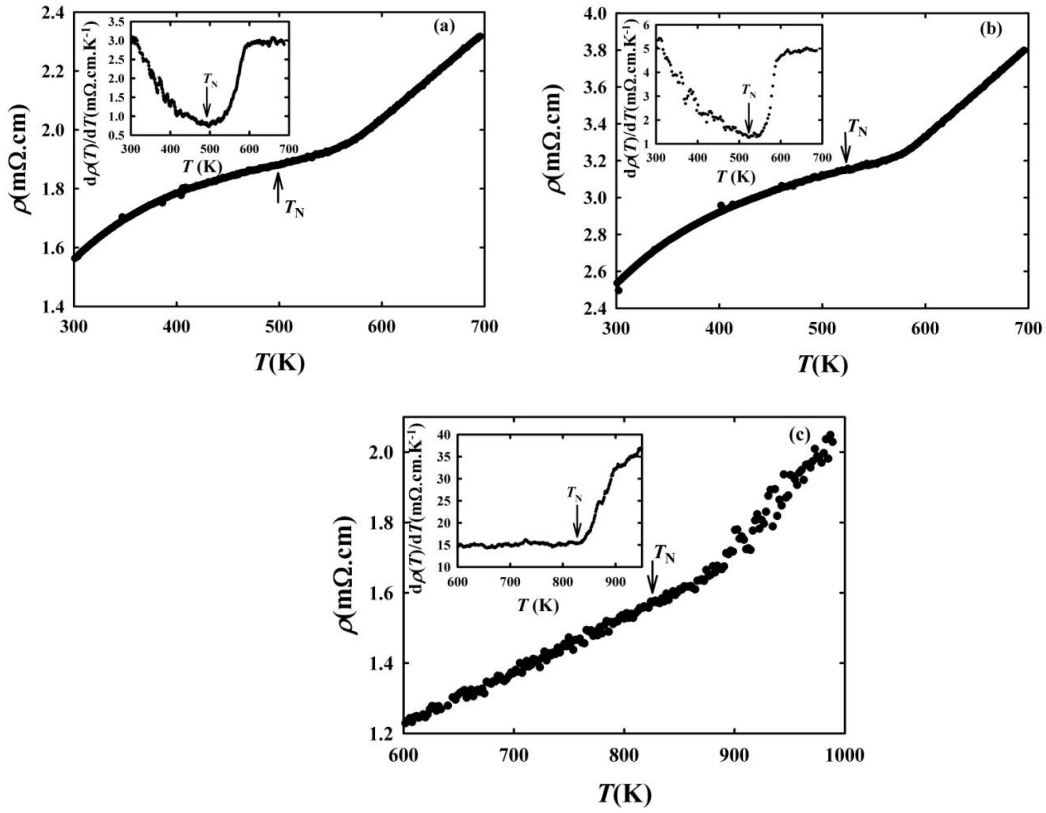


Figure 2: The $\rho(T)$ curves for the $(Cr_{84}Re_{16})_{100-y}Mn_y$ alloy system with, with (a) y = 0.3, (b) y = 0.4 and (c) y = 3.1. The arrow indicates the position of T_N for each alloy. Inset shows the temperature dependence of $d\rho(T)/dT$ with T_N marked at the minimum.

The temperature dependence of magnetic susceptibility (χ) for the alloys with y = 0.4, 0.6, 0.8 and 3.1 at.% Mn are shown in figure 3. At low temperatures (T < 10 K), a sharp increase in χ is observed on increasing T after ZFC. A prominent peak is observed close to 30 K beyond which χ rapidly decreases to lower values. In samples with concentrations 0.6, 0.8 and 3.1 at.% Mn, χ approaches zero above 40K. In the FC state, there is a slower decrease in χ on increasing T up to around 30 K beyond which the behaviour is identical to that observed in the ZFC state except for the alloy with 3.1 at. % Mn. In this case, the χ values obtained in both the FC and ZFC state first increases to a maximum value resulting in a peak before rapidly decreasing to low χ values. These results are similar to those observed in $Cr_{1-x}Mn_x$ alloys [8]. Here too, it is possible to define clearly a "pinning temperature" (T_p) to characterize the onset of the peak. It is important to note that T_p as well as the temperature of the peak show no variation with y. Also, above T_p, χ shows no dependence on temperature up to 400 K. The peak in the temperature dependence of χ with irreversibility with respect to FC and ZFC, the invariance of T_P with y and the invariance of χ with T above T_P are all indicators of this peculiar spin-glass state [8].

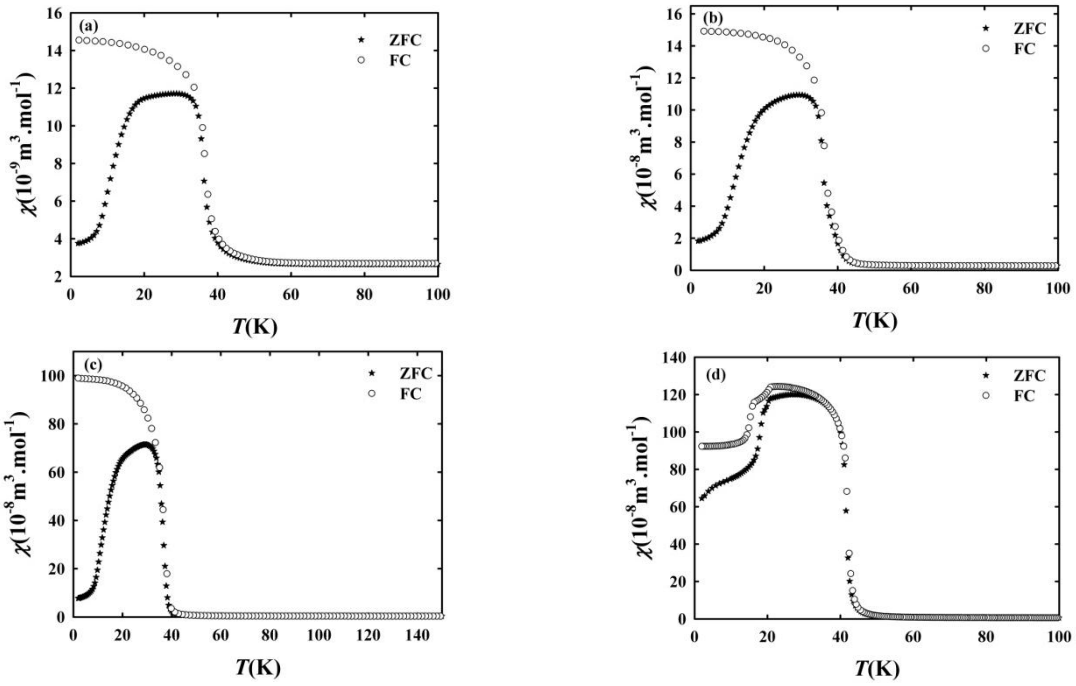


Figure 3. The temperature dependence of χ for the (a) y = 0.4 (b) y = 0.6 (c) y = 0.8 and (d) y = 3.1alloys.

The new type of spin-glass effect found previously in $Cr_{1-x}Mn_x$ and $(Cr + 1.3\% Si)_{1-x}Mn_x$ [7] is also found in $(Cr_{84}Re_{16})_{100-y}Mn_y$ alloys. The essence of this new type of spin-glass is the frustration of itinerant spins at the interfaces between AFM domains established around each Mn local magnetic moment [11]. Above T_p of the Mn magnetic moments, they follow the AFM domains of the AFM matrix. However, below T_p, the Mn moment is frozen and the AFM order has to follow it, becoming frustrated whenever it finds another fixed moment not pointing in the expected state. This approach supposes that below certain concentration of Mn, the interaction that determines the spin-glass state is at the boundaries of the AF SDW domains, resulting in a peculiar spin-glass that requires the existence of itinerant AFM matrix, contrasting with the canonical spin-glass of CuMn [14]. This is similar to the results reported by Li *et al.* [10] where the sharp downturn of the ZFC *M-T* curves is an indication of entry into the spin glass or spin frustration state having demagnetizing ferromagnetic domains that

are frozen below T_p and free to rotate above T_p. Figure 3 presents evidence of this behaviour where the pinning temperature does not depend linearly on Mn concentration below the expected incommensurate-commensurate transition for CrMn and CrRe. There is a strong increase of the magnetisation as a function of temperature for commensurate magnetic order.

4. Conclusion

The present study was a preliminary investigation of the electronic and magnetic properties of the $(Cr_{84}Re_{16})_{100-y}Mn_y$ alloy system. Increasing the concentration of Mn leads to increase in T_N as expected. Magnetic susceptibility curves indicate the presence of a possible spin-glass state in the $(Cr_{84}Re_{16})_{100-y}Mn_y$ system requiring the existence of itinerant AFM matrix which is different from that in the canonical CuMn alloy. Results need to be extended to include the tests for spin glass in order to confirm its presence.

Acknowledgements

This work was financially supported by the NRF of South Africa under Grants 80928, 80631, 93551, 80626 and the Faculty of Science at the University of Johannesburg.

References

[1] Fawcett E 1988 *Rev. Mod. Phys*. **60** 209-283
[2] Lomer WH 1962 *Proc. Phys. Soc.* **80** 489-496
[3] Alberts HL 1988 South African Journal of Science **84** 32-34
[4] Alberts HL and Lourens JAJ 1989 South African Journal of Science **85** 53-59
[5] Mydosh JA 1993 "Spin Glasses : An Experimental Introduction" ISBN 0-7484-0038-9 Taylor & Francis London, Washington DC
[6] Binder K and Young AP 1986 *Rev. Mod. Phys.* **58** 801-976
[7] Galkin VY, de Camargo PC, Ali N, Schaf J and Fawcett E 1995 *J. Phys. Condens Matter* **7** L649
[8] Galkin VY, de Camargo PC, Ali N an Fawcett E 1996 *J. Phys. Condens Matter* **8** 7925
[9] Galkin VY, de Camargo PC, Ali N and Fawcett E 1997 *Physica B* **237-238** 443
[10] Li B, Alberts HL, Wu BM, Prinsloo ARE and ZJ Chen 2009 *J. Magn. Magn. Mater.* **321** 61-73
[11] Fawcett E, Alberts HL, Galkin VY, Noakes DR and Yakhmi JV 1994 *Rev. Mod. Phys*. **66** 25-127
[12] Jacobs BS, Prinsloo ARE, Sheppard CJ and Strydom AM 2013 *J. Appl. Phys.* **113** 17E126
[13] Quantum Design Inc, 6325 Lusk Boulevard, San Diego, USA
[14] Retat I and Schwink Ch. 1983 *J. Magn. Magn. Mat.* **38** 123-132

Effect of pH on the properties of ZnO nanostructures prepared by chemical bath deposition method

L F Koao[1], F B Dejene[1] and H C Swart[2]

[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.

E-mail: koaolf@ufs.ac.za

Abstract. ZnO powders were prepared by chemical bath deposition method by varying the pH of the precursor. Different amount of ammonia hydroxide were added to change the pH of the solution from 5 to 12 values. The effect of pH of the precursor on the structure, morphology, optical and luminescence properties of ZnO nanostructures were investigated. The X-ray diffraction (XRD) patterns of the ZnO nanostructures correspond to the various planes of a single hexagonal ZnO phase. It was observed that the diffraction peaks increased in intensity with an increase in pH. The estimated average grain sizes calculated using the XRD spectra were found to be in the order of 38 ± 1 nm. It was observed that the estimated average grain sizes increased slightly with an increase in pH. The surface morphology study revealed that the grains were flakes-likes at low pH (< 6) but flower-like for high pH (12). The UV-Vis spectra showed a red shift with an increase in pH. The band gap energy of ZnO was found to decrease but the luminescence intensities increased with an increase in pH values. PL showed that the ultra-violet emission intensity of the nanostructures decreased with an increase in pH but the visible emission increased with an increase in pH.

1. Introduction

Nowadays, zinc oxide (ZnO) is a very attractive material in numerous promising applications such as gas sensors, photocatalyst, biosensnsres, solar cells, electrical and optical devices due to its non-toxic material with a wide direct band gap (3.37 eV) and a large exciton binding energy (60 meV) [1]. ZnO practical application is decided by its properties which can be modulated by varying its morphology [2-3]. A number of ZnO nanostructures such as nano wires [4], nanorods [5], nanotube [6], nanobelts [7], nanoflower-like [3], nanoplates [8], tetrapod [9] and nanoparticles [10] have been reported. Numerous methods have been employed for the production of ZnO nanostructures such as vapor liquid solid (VLS) [11], sol–gel method [12], Ball milling (BM) [10], and combustion method [13]. Some of these synthetic processes require long reaction time, high synthesis and annealing temperatures. To avoid some of these challenges the chemical bath deposition (CBD) method was used, is a simple, cheap and convenient process to prepare semiconducting materials. The more recent interest in all things 'nano' has provided a boost for CBD, since it is a low temperature, solution (almost always aqueous) technique and it gives better homogeneity. The aim of this paper was to investigate the effect of pH on the structure, morphology, optical and luminescence properties of the ZnO. Nesakumar et al. [14] prepared ZnO nanoparticles at different pH values under harsh conditions. Change in shape from nanosphere to nanorods was observed by varying pH from 11 to 10.7. Other researchers [15] showed that as pH increases the luminescence intensity in the UV region increases. From our work, we were able to observed the change of X-ray diffraction (XRD) structure from $Zn(OH)_2$ to hexagonal ZnO by increasing the pH.

2. Experimental

The pH of the bath was controlled by varying different volumes of ammonia. The ZnO powders were prepared by holding volumes of zinc acetate and thiourea constant and varying the volumes of ammonia in the precursor. Ammonia was used as a complexing agent. The 0.46 M of zinc acetate was dissolved in 500 mL of deionized water. 0.18 M thiourea was dissolved in 500 mL of deionized water and 123.5 mL of ammonia were dissolved in 500 mL of deionized water. Then a magnetic stirrer was used to stir each of the mixtures for overnight at room temperature to ensure homogenous distribution of the solution reagents. The amount of each solution was taken in the following order: 60 mL quantity of zinc acetate was first added to the beaker which was placed in the water bath, followed by addition of 60 mL of thiourea solution and the mixture was stirred for 30 s, following that, 5 mL of ammonia solution was then added slowly in the mixture, while continuously stirring for 5 min. The pH measured was 5. Water bath was maintained to be at a constant desired temperature of 80 ºC. The precipitates were then left overnight and filtered thereafter. The precipitates were later washed with 60 mL of ethanol and secondly with 60 mL of acetone. The obtained precipitates were dried at ambient conditions for 3 days. After that the powders were ready to be characterized. The synthesis of ZnO nanostructures in the presence of different volumes of was performed similarly. The varied amounts of ammonia present in the precursor solution are 24 mL, 60 mL, and 75 mL, the pH measured for each mixture were 6, 8 and 12, respectively. The particle size and morphology and the structural and luminescent properties of the as-synthesized phosphors were examined by means of Scanning Auger electron microscopy (SAM) using a PHI 700 Scanning Auger Nanoprobe, X-ray diffraction (XRD), Uv-vis spectroscopy and Photoluminescence (PL).

3. Results and Discussion

3.1 Structural analysis and Composition analysis

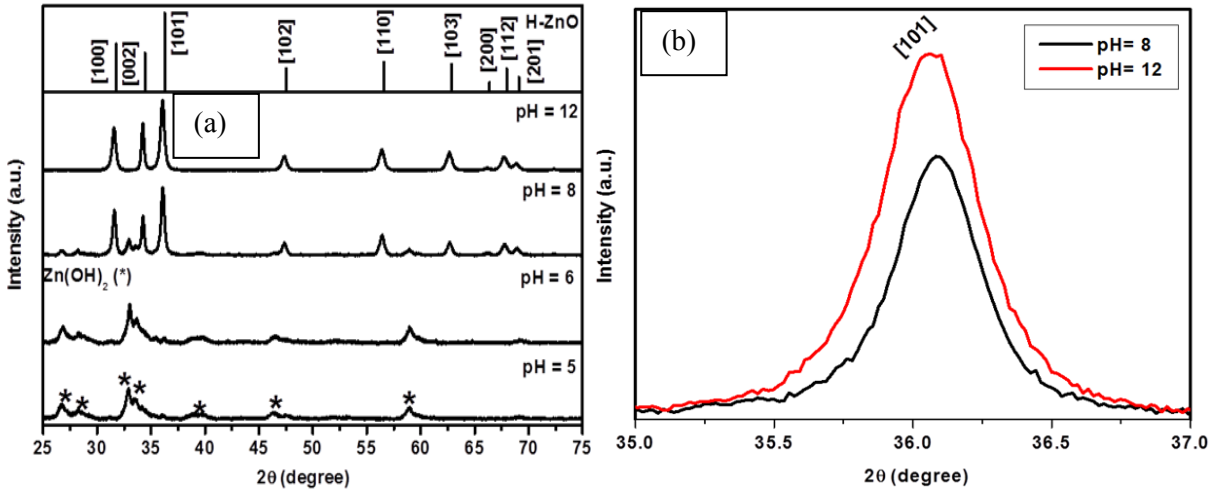


Figure 1. X-ray powder diffraction patterns for $Zn(OH)_2$ and ZnO structures prepared at (a) various pH and (b) XRD patterns at (103).

Figure 1 (a) shows the XRD patterns of the $Zn(OH)_2$ and ZnO nanostructures synthesised at various pH. One can see that at low pH the XRD structure was for $Zn(OH)_2$ [16], when the pH was increased further the structure changed to hexagonal ZnO. By increasing the pH using ammonia the $Zn(OH)_2$ peaks diminished. It is suggested that those peaks that are marked with star (*) are related to the $Zn(OH)_2$ as confirmed by the JCPDS no: 38-0356 [17]. The hexagonal ZnO match perfectly with the data of Koao et al. [3]. The average crystallite size of the as-prepared crystals can be estimated from the Full Width Half Maximum (FWHM) of the diffraction peaks using the Debye formula [18]. The

average crystallite sizes estimated using the XRD spectra were found to be 50 and 43 nm for samples synthesized at pH of 12 and 8, respectively. The estimated average grain size increased with an increase in the pH. The increase in estimated crystallite size may be due the zinc hydroxide that dissolves as the PH increases which are the primary precursor for growth of ZnO that lead to the formation of bigger crystallite. The average estimated values of the cell constants of *a* and *c* are 3.26979 and 5.23373 Å and 3.26727 and 5.23092 Å for pH= 8 and pH= 12, respectively. The estimated lattice parameters match perfectly with standard data available in JCPDS card no. (36-1451, a = 3.24982 and c= 5.20661 Å). It is clear that the estimated lattice parameters decreased with an increase in pH. This may be due to the increase in estimated crystallite sizes. In Figure 1(b), it was observed that the XRD diffraction intensities increased with an increase in pH, implying that the crystallite sizes increased and the improvement of the crystallinity.

3.2 Surface morphological analysis

Figure 2 (a), (b) and (c) show SEM images of the as prepared $Zn(OH)_2$ and ZnO nanostructures, respectively prepared at various pH but synthesized at constant time and bath temperature. At low pH of 5 the surface aspects of the SEM image was composed of flakes-like structures. It is clear that at the pH of 8, the morphology consists of mixtures of flakes-like and flower-like structures. However, by further increasing the pH to 12, morphology consists of flower-like structures surrounded by few spherical nanoparticles. By comparing the SEM micrographs it is clearly seen that by increasing the pH using ammonia the flake-like changed to the flower-like structures. The change in morphology may be due to the increasing pH inducing the nucleation growth yielded the flower-like structures.

AES measurements were carried out to check the elements and to evaluate their purity. The Auger spectrum of the $Zn(OH)_2$ and ZnO are shown in Figure 3. The typical spectrum of the nanoflowers confirmed the presence of Zn and O. The signals at energy of 276 eV are ascribed to the presence of adventitious carbon, which come from the atmosphere and due to handling. The Zn and O concentration were relatively very low for the sample synthesized at pH of 5 compared to the sample prepared at pH of 8 and 12. Moreover, no other impurities were found on the surface of the ZnO, suggesting that the as-synthesized ZnO were relatively pure.

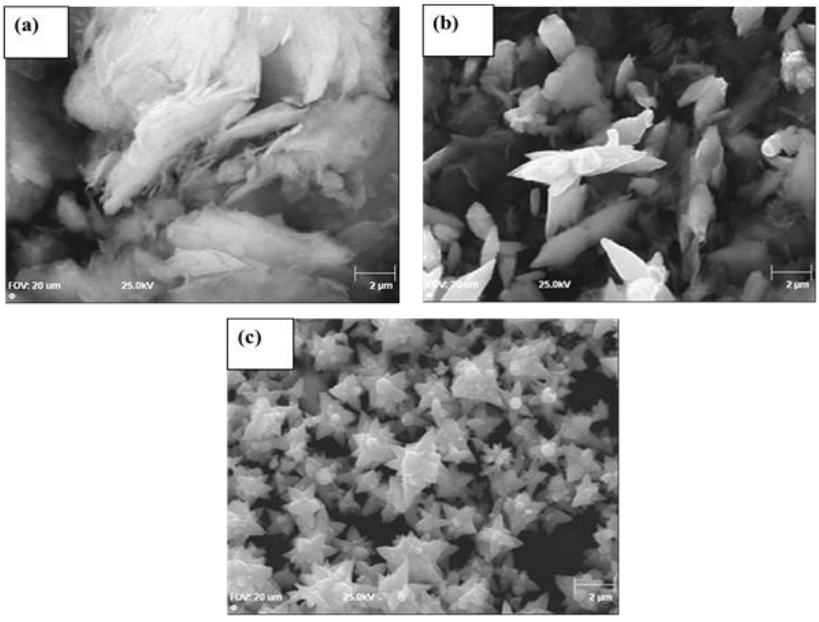


Figure 2. SEM images of $Zn(OH)_2$ and ZnO nanostructures for pH of (a) 5, (b) 8 and (c) 12 synthesized by CBD at bath temperature of 80 ºC.

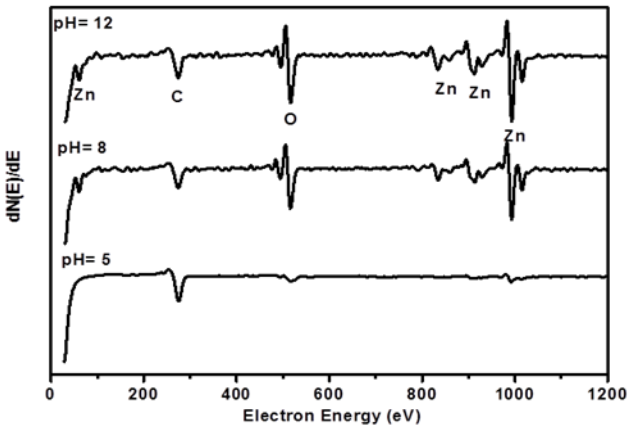


Figure 3. Auger spectra of the $Zn(OH)_2$ and ZnO powders prepared at pH= 5, pH= 8 and pH= 12 using the CBD method.

3.3 Optical properties

The UV-visible reflectance spectra of the as prepared samples are illustrated in Figure 4(a). It clearly indicates that by increasing pH, the optical absorption edge shifted to a higher wavelength. At low pH the absorption edge is at around 261 nm, which may be due to the presence of $Zn(OH)_2$ as confirmed by XRD analysis. The absorption edges are homogeneous, no other band was observed in the high pH of 12, which confirmed that the synthesized sample was pure ZnO.

The energy band gap of these materials was estimated using the Kubelka-Munk remission function **[19].** In Figure 4(b) it can be seen clearly that the band gap energy of the ZnO nanostructures decreased slightly with an increase in the pH. The estimated band gap energy for the synthesis time were 3.36 ± 0.01, 3.31 ± 0.01 eV for flower-like ZnO synthesized for pH of 8 and 12, respectively. All the calculated band gap energies were below the theoretical band gap value of 3.37 eV at room temperature. This may be due to an increase in grain size as confirmed by XRD analyses [20].

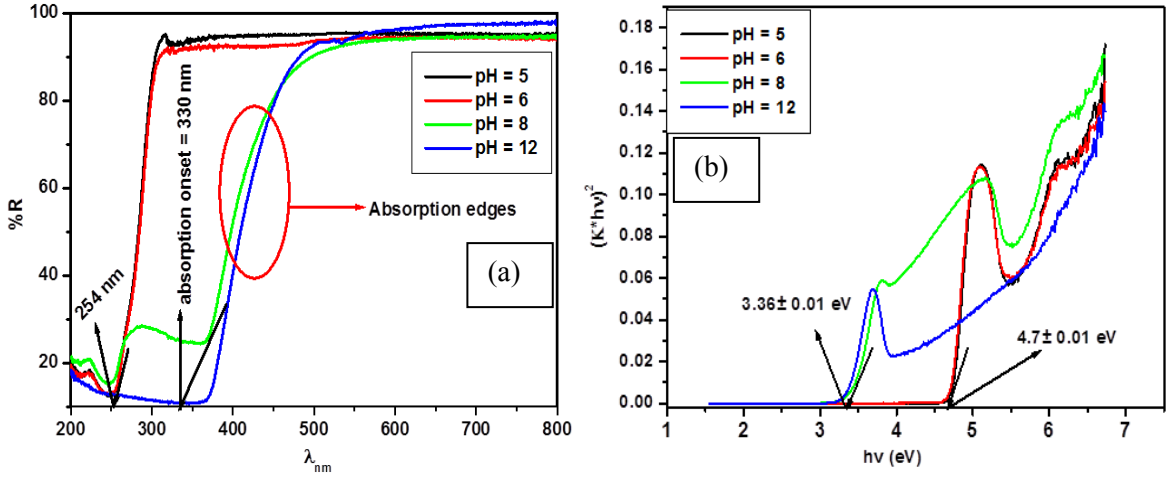


Figure 4. The absorbance spectra (a) and the band gap energy (b) of the $Zn(OH)_2$ and ZnO structures prepared at different pH.

It is known that as the grain size increased, the electronic states are not discrete and results in the reducing of the band gap and decreased the oscillator strength [21]. The highest band gap energy was

obtained for the sample synthesized at low pH, with its estimated band gap energy of 4.70 ± 0.01 eV. The band gap energy may be due to the $Zn(OH)_2$.

3.4 Photoluminescence

Figure 5 shows the photoluminescence (PL) spectra of the $Zn(OH)_2$ and ZnO nanostructures synthesized at different pH with the excitation wavelength at 300 nm. The excitation wavelength was taken from absorption onset on reflectance results. Four obvious luminescence peaks in the UV region were observed at around 364, 379, 414 and 450 nm from both low pH of 5 and 6. Those luminescence peaks at around 364 and 379 nm are usually ascribed to the radiative recombination of ZnO free excitons [22]. They are both attributed to the transition from valence band to conduction band of ZnO semiconductors. A violet peak at around 414 nm was ascribed as the emission from zinc vacancies present in the ZnO lattice [23]. The blue emission centred at 450 nm may be due to the electron transition from Zn_i to V_{zn} [22]. By increasing the pH further, those peaks decreased in luminescence intensity and a known yellow emission peak started to emerge at around 570 nm. This yellow emission was due to the presence of oxygen interstitial in the synthesized material [17]. The emission peak in the visible region increased in luminescence intensity with an increase in pH. The decreased in luminescence intensity in UV region may be due to an increase in estimated crystallite size as confirmed by the XRD [20]. The increase in the visible emission as the pH increased may be due to the reduction of $Zn(OH)_2$ as confirmed by the XRD analysis. The low pH provided enough ionization energy and increased the concentration of $Zn(OH)_2$, strengthening the UV to blue emissions. The high pH induced the outward diffusion of $Zn(OH)_2$ and quenching the UV to blue emissions. It was very interesting to observe that the XRD, SEM, XPS, UV and PL results support each other.

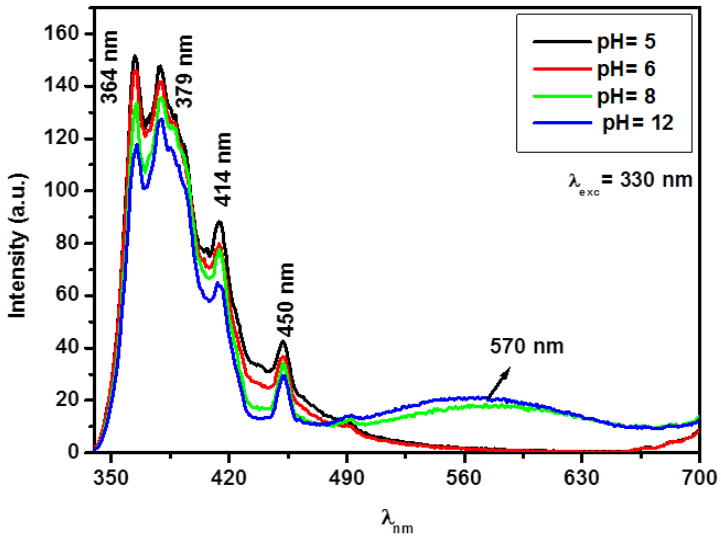


Figure 5. The PL spectra of ZnO nanostructures, prepared at different pH of pH= 5, pH= 8 and pH= 12 using the CBD method.

4. Conclusion

The ZnO nanostructures have been successfully synthesized by the chemical bath deposition technique at 80°C by varying the pH using ammonia. XRD showed that the structure of the material changed from $Zn(OH)_2$ to the hexagonal wurtzite ZnO. XPS showed that the amount of oxygen and Zn increased with an increase in pH. UV spectroscopy showed that the band gap energy of the ZnO flower-like decreased with an increase in the pH. PL showed that the visible emission intensity of the nanostructures increased with an increase in pH.

Acknowledgement
The author would like to acknowledge the National Research Foundation and the University of the Free State for financial support.

Reference

[1] Heo S N, Park K Y, Seo Y J, Ahmed F, Anwar M S, Koo B H 2013 *Electro. Mater. Lett.* **9** 261-265.
[2] Park W I, 2008 *Met. Mater. Int.* **18** 659.
[3] Koao L F, Dejene B F, Swart H C, Botha J R 2013 *J. Lumin* **143** 463-468.
[4] Amiruddin R, Kumarn M C S, 2014 *Ceram. Int.* **40** 11283–11290.
[5] Srivatsa K M K, Chhikara D, Kumar M S, 2011 *J. Mater. Sci. Technol*. **27** 701–706.
[6] Angwafor N G N, Riler D J, 2008 *Phys. Status Solid A.* **205** 2351-2354.
[7] Gao P X, Ding Y, Mai W J, Hughes W L, Lao C S, Wang Z L, 2005. *Sci*. **309** 1700-1704.
[8] Molefe F V, Koao L F, Dolo J J, Dejene B F, 2014, *Physica B*. **439** 185-188.
[9] Feng L, Liu A, Liu M, Ma Y, Wei J, Man B, 2010, *Mater. Character*. **61** 128-133.
[10] Amirkhanlou S, Ketabchi M, Parvin N, 2012, *Mater. Lett*. **86** 122-124.
[11] Huang M H, Wu Y Y, Feick H, Tran N, Weber E, Yang P H, 2001, *Adv. Mater*. **12** 113-116.
[12] Zhang C, 2010, *J. Phys. Chem. Solids*. **71** 364-369.
[13] Lamas D G, Lascalea G E, Walsoc N E 1998 *J. Eur. Ceram. Soc*. **18** 1217–1221.
[14] Nesakumar N, Rayappan J B B, Jeyaprakash B G, Krishnan U M, 2012, *J. Appl. Sci*. **12** 1758-1761.
[15] Yuan H, Xu M, Huang Q Z, 2014, *J. Alloy. Comp*. **616** 401-407.
[16] Jia W, Dang S, Liu H, Zhang Z, Yu C, Liu X, Xu B, 2012, *Mater. Lett*, **82** 99-101.
[17] Peng W Q, Qu S C, Cong G W, Wang Z G, 2006, *Mater. Sci. Semi. Process*, **9** 156-159.
[18] Cullity B D1978, 1956*Elements of X-ray Diffraction (2nd Ed)*, (Addison Wesley) 285-284.
[19] Molefe F V, Koao L F, Dejene B F, Swart H C,2015, *Optical. Mater*. **46** 292–298.
[20] Farhadi-Khouzani M, Fereshteh Z, Loghman-Estarki M R, Razavi R. S, 2012, *J. Sol-Gel Sci. Technol*, **64** 193-199.
[21] Sahu S N, Nanda K K, 2001, Proc. *Indian Natl. Sci. Acad*. **67 (A)** 103-130.
[22] Wang Y F, Yao J H, Jia G, Lei H, 2011, *Acya Physica Polonica A*, **119 (3)** 451-454.
[23] Lin B, Fu Z, Jia Y, 2001, *Appl. Phys. Lett*. **79** 943-945.

Phototransferred thermoluminescence and phosphorescence related to phototransfer in annealed synthetic quartz

EFM Kombe-Atang and ML Chithambo [a]

Department of Physics and Electronics, Rhodes University, Grahamstown, 6139, South Africa

E-mail: [a] m.chithambo@ru.ac.za

Abstract. Phototransferred thermoluminescence and phosphorescence related to phototransfer in synthetic quartz is reported. The glow curve shows a number of unstable electron traps only one of which is involved in phototransfer. Kinetic analysis of both types of phototransferred signal gives an activation energy of ~ 0.55 eV for this electron trap. The PTTL intensity goes through a peak with illumination time and this is ascribed to the relative concentration of phototransferred electrons and, holes at a recombination centre. On the other hand, the dose dependence of the phototransferred thermoluminescence is superlinear.

1. Introduction

Quartz is a common natural mineral used in retrospective dosimetry using luminescence based methods [1]. Luminescence is typically emitted from previously irradiated insulators as a result of optical or thermal stimulation [2]. The ionizing radiation creates free electrons some of which get localized at pre-existing point defects termed electron traps. Thermoluminescence (TL) is observed when an irradiated material is heated at a controlled rate. The signal appears as a set of peaks known as a glow curve. Each peak is associated with an electron trap.

When an irradiated material is partially heated to remove lower temperature peaks, it is found that exposing the same material to light of certain wavelengths reproduces these peaks. This phenomenon, otherwise called phototransferred thermoluminescence (PTTL), is observed in natural and synthetic quartz [3–5]. The general principle of PTTL is that electrons are moved, by light, from thermally inaccessible traps (otherwise called deep traps) to unstable shallow traps.

The study of PTTL in this report was augmented by use of phosphorescence. This is the decay of thermoluminescence as a function of time at a constant temperature. Isothermal decay curves from phosphorescence can be analysed using various kinetics methods to determine parameters such as activation energy associated with the TL process.

This report is concerned with features of PTTL and, phosphorescence resulting from phototransfer in annealed synthetic quartz. In particular, we report the dependence of the PTTL on illumination time, and the influence of dose on the phototransferred phosphorescence as well as kinetic analysis of the phosphorescence. To the best of our knowledge kinetic analysis of phototransferred phosphorescence in synthetic quartz has never been reported.

2. Experimental details

Synthetic quartz of grain size 90-500 μm (Sawyer Research Products, Ohio, USA) was used. A few milligrams of sample was placed on stainless steel discs of 1 mm thickness and 10 mm diameter. Luminescence was measured using a RISØ TL/OSL-DA-20 Luminescence Reader. Samples were irradiated using an inbuilt ^{90}Sr/^{90}Y beta source at a dose rate of 0.1 Gy/s. Luminescence was detected by an EMI9235QB photomultiplier tube through a 7.5 mm thick Hoya U-340 filter (transmission band 280-390 nm). All TL measurements were made in a nitrogen atmosphere to avoid spurious luminescence from air and to improve thermal contact between the sample disc and the heater plate. Samples were annealed at 500^{o}C for 10 minutes before use in order to remove any residual signal. The quartz is subject to the pre-dose effect and the heating also improved its sensitivity by way of thermal activation [6].

3. Results

3.1. General features

Figure 1 shows a glow curve obtained following heating to 500^{o}C at 1^{o}C/s after irradiation to 100 Gy. The plot shows six peaks at 76, 104, 170, 210, 285 and 340^{o}C labelled as I to VI, respectively.

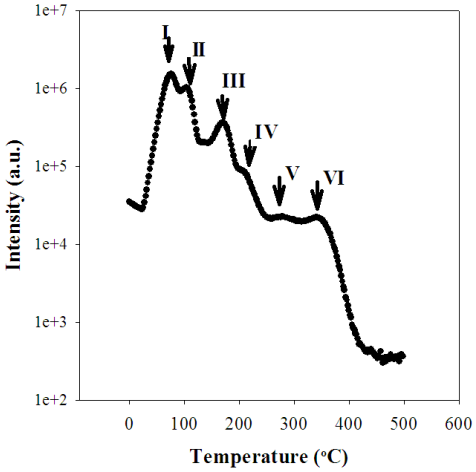


Figure 1. A glow curve measured at 1^{o}C/s following irradiation to 100 Gy.

3.2. Phototransferred thermoluminescence

PTTL was produced in a number of steps. The sample was first irradiated to 100 Gy and then preheated to 200^{o}C in order to empty the shallow traps at 76, 104 and 170^{o}C. These three peaks were removed to ease the search for phototransferred signal. The sample was thereafter exposed to 470 nm blue LED light for 20 s to cause the transfer of electrons from deep to the shallow traps. Finally, the sample was heated to 500^{o}C at 1^{o}C/s to obtain the complete glow curve.

Figure 2 shows the glow curve obtained after phototransfer. The phototransferred peak is seen at 76^{o}C, the position of the original peak I. Peaks II and III were not reproduced by phototransfer.

To observe the effect of illumination time on the PTTL, similar experiments were conducted except that the sample was illuminated for durations between 1 and 1000 seconds each time. Figure 3 shows the change of peak intensity with illumination time. The intensity of the photransferred peak increases with time up to 20 seconds and decreases thereafter to some stable value. The solid line through data will be explained later in the text.

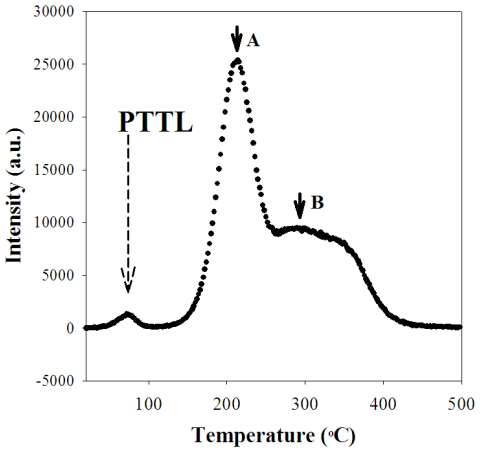


Figure 2. A TL glow curve after phototransfer. The arrow indicates the phototransferred peak. The other peaks A and B, are peaks resulting from the residual TL signal in the sample.

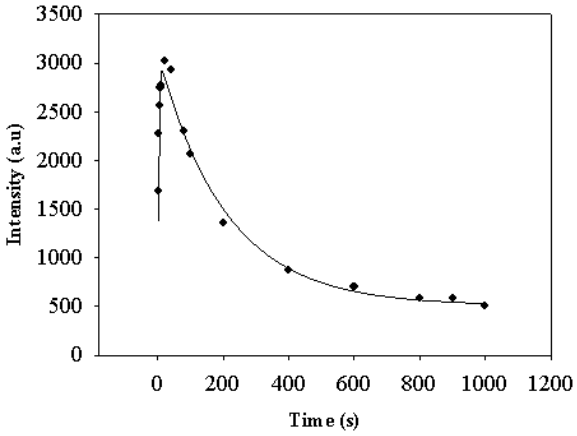


Figure 3. Variation of PTTL intensity with illumination time.

3.3. Phosphorescence

Since shallow traps are unstable, the expected loss of charge at ambient temperature may be monitored and analysed using phosphorescence. Here, the sample was beta irradiated to a certain dose and the phototransfer induced as before. The phosphorescence related to phototransfer was then measured at various temperatures between 52 and 60^{o}C. These temperatures correspond to the rising edge of peak I and measurements were made for doses between 100-500 Gy.

Figure 4 shows a phosphorescence decay curve measured at 55^{o}C following a dose of 200 Gy. The time dependence of the phosphorescence intensity $I(t)$ is exponential, that is,

$$I(t) = I_o \exp(-p\,t) \tag{1}$$

where I_o is initial intensity and p is the decay constant given by

$$p = s \exp(-E\,/kT) \tag{2}$$

where E is the activation energy, s the frequency factor, k Boltzmann's constant and T the measurement temperature [2]. Equation (1) implies that the phosphorescence follows first order kinetics. The linear form of the semi-logarithmic plot of $I(t)$ and t (figure 4, inset) suggests that this is indeed the case. The decay constant p, which is just the slope in a semi-logarithmic plot of $I(t)$ against t, provides a means to calculate the activation energy E of the associated electron trap. As can be deduced from equation 2, if values of p are obtained at different temperatures, the value of E can be found from a plot of p against t on a semi-log scale.

Figure 5 shows a plot of $\ln(p)$ against $1/kT$ corresponding to 200 Gy. These measurements were made for doses from 100-500 Gy. The mean activation energy was 0.55±0.04 eV. The mean frequency factor was 6×10^6 s^{-1}. The kinetic parameters E and s for the phototransferred peak were also evaluated using the various heating rates method [2] where the peak temperature T_m depends on heating rate β as

$$\ln\left(\frac{T_m^2}{\beta}\right) = \frac{E}{kT_m} + \ln\left(\frac{E}{sk}\right) \tag{3}$$

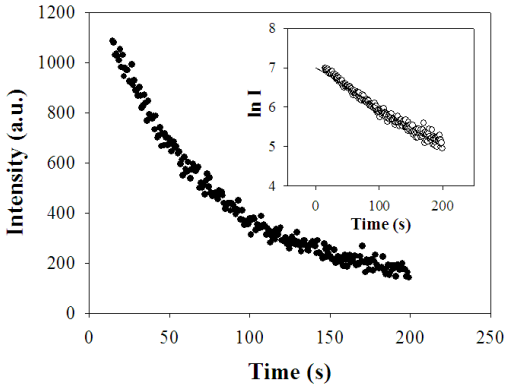


Figure 4. Phosphorescence measured at 55^{o}C following a 200 Gy dose. The inset shows a semi-log plot of the intensity against time.

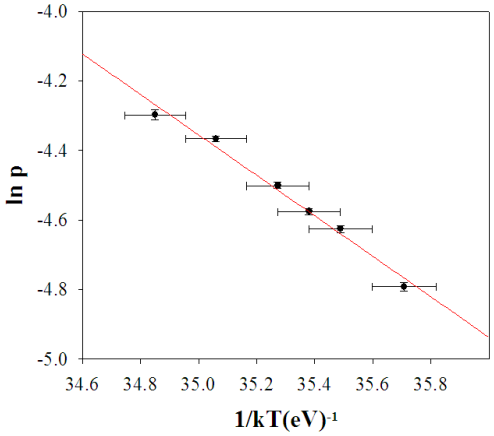


Figure 5. The relationship between ln (p) and $1/kT$ for phosphorescence.

Figure 6 shows $\ln(T_m^2/\beta)$ against $1/kT_m$ for heating rates from 0.1 to 5^{o}C/s. The activation energy E, worked out from the slope of best fit was 0.56±0.04 eV and frequency factor s was 4×10^6 s^{-1}. These values are consistent with those obtained from figure 5 for phosphorescence related to phototransfer. The reasons for the anomalous low values for s are under consideration. The quartz studied here sometimes shows a weak peak collocated with the main one [7]. The

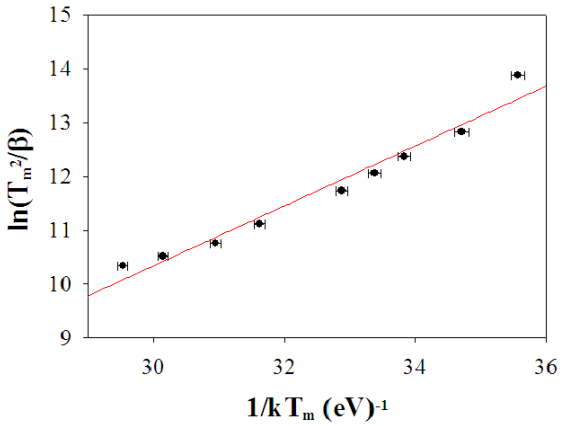


Figure 6. Application of the variable heating rate method on the PTTL.

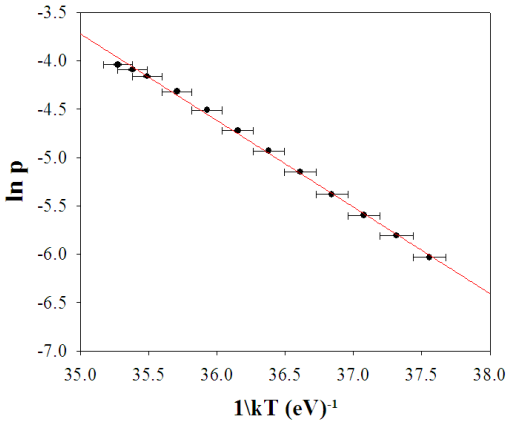


Figure 7. The plot is ln p against $1/kT$ for original TL.

activation energy of this secondary peak was determined as 0.53±0.01 eV using the area method of Chithambo [7]. The consistency of E values from figures 5 and 6 and ref. [7] suggest that this is the peak that is phototransferred. To verify this assumption, measurements of phosphorescence were made on a separate sample, preheated to remove peak I. The activation energy of the next peak in the series was determined using the temperature dependence of phosphorescence (figure 7) as 0.89±0.01 eV and its s as 10×10^{11} s^{-1}.These values indeed differ from those of peak I.

3.4. Dose response of photransferred phosphorescence

The dependence of the phosphorescence intensity on irradiation dose was studied for doses between 5 and 500 Gy. Measurements were made at an arbitrarily chosen temperature of 56^{o}C. Intensities were determined as areas under each phosphorescence decay curve.

Figure 8 shows the growth curve where the intensity changes with dose D as

$$S(D) = aD^2 + bD + c \tag{4}$$

where $S(D)$ is the analytical form of the intensity and a (in Gy^{-2}), b (in Gy^{-1}), c (in arbitrary units) are constants. The non-linear dose response of figure 8, was quantified using the superlinearity index $g(D)$ which gives a measure of the change in slope of the growth curve and is given by

$$g(D) = \left[\frac{DS''(D)}{S'(D)}\right] + 1 \tag{5}$$

where $S'(D)$ and $S''(D)$ are respectively the first and second derivatives of $S(D)$ [6]. A value of $g(D) > 1$ indicates superlinearity, $g(D) = 1$ linearity and $g(D) < 1$ signifies sublinearity. The inset to figure 8, a plot of $g(D)$ against D where $g(D) > 1$ despite the dose shows that the dose response is superlinear.

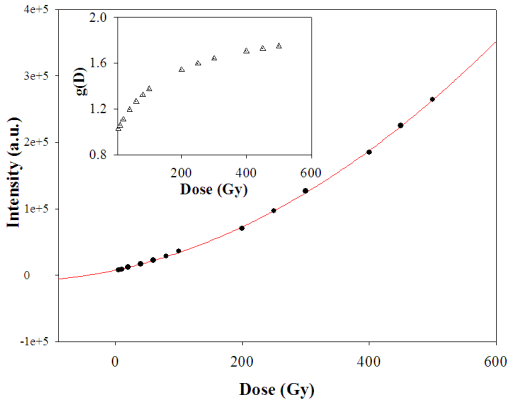


Figure 8. The phototransfered phosphorescence growth curve at 56^{o}C. The inset shows the superlinearity.

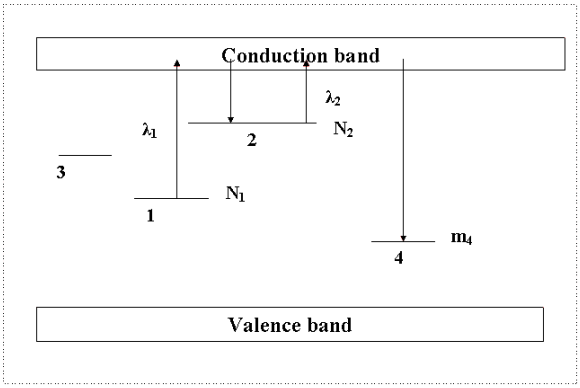


Figure 9. Energy band diagram to explain PTTL.

4. Discussion

The change of PTTL intensity with illumination (figure 3) can be explained with reference to the energy band diagram of figure 9. Some electrons optically excited out of the deep trap (level 1), with electron concentration N_1 at a rate λ_1, are retrapped, via the conduction band, at the shallow trap (level 2). The concentration of electrons in the shallow trap at the end of the illumination is N_2. Intermediate energy electron traps not involved in phototransfer are shown as level 3. Phototransferred luminescence is produced when electrons from the shallow trap recombine with holes at the recombination centre (level 4). The holes, as electrons at traps, are produced following irradiation. For short illumination, the number of electrons moved from the deep to the shallow trap is less than the concentration of holes m_4 at the recombination centre. Assuming that the resultant thermoluminescence intensity is proportional to N_2, the PTTL intensity would be expected to increase with N_2 as long as $N_2 < m_4$. At extended illumination,

the number N_2 may exceed m_4 and so the PTTL intensity will tend to decrease to some constant value with time as long as $m_4 \neq 0$. This qualitatively explains the result in figure 3.

The charge movement between the deep and shallow trap may alternatively be expressed as

$$dN_1/dt = -\lambda_1 N_1 \tag{6}$$

$$dN_2/dt = -\lambda_2 N_2 + \alpha \lambda_1 N_1 \tag{7}$$

where λ_2 is the rate of optical transfer of charge out of the shallow traps [3]. The general solution of coupled equations 6 and 7 is

$$N_2(t) = A(e^{-\lambda_1 t} - e^{-\lambda_2 t}) \tag{8}$$

where A is a constant. Assuming that the PTTL intensity I is proportional to N_2, the data in figure 3 can be described as $I \propto (e^{-\lambda_1 t} - e^{-\lambda_2 t})$. The solid line through data in figure 3 is the best fit of an equation of the form $(e^{-\lambda_1 t} - e^{-\lambda_2 t}) + d$ where d is an experimental offset. This simple model assumes no retrapping into the deep or shallow traps, but as evident, satisfactorily describes the data.

Concerning the dose response (figure 8), $N_2 < m_4$ for any dose. The superlinear response occurs due to the increase in phototransferred electrons and so does PTTL with dose.

5. Summary

Synthetic quartz has a number of shallow electron traps one of which was reproduced by phototransfer. The PTTL of this trap goes through a peak with illumination time. This is due to the relative concentrations of the photransferred electrons and, holes at the recombination centre. The unstable nature of the shallow trap is illustrated by the presence of phosphorescence emitted even at ambient temperature. The activation energy of this trap was calculated as 0.55±0.04 eV by kinetic analysis of phosphorescence and verified by the variable heating rate method as 0.56±0.04 eV. The PTTL intensity depends on irradiation and its increase with dose is superlinear.

Acknowledgments

We gratefully acknowledge financial support from the National Research Foundation and Rhodes University.

References

[1] Preusser F, Chithambo M L, Götte T, Martini M, Ramseyer K, Sendezera E J, Susino G J and Wintle A G 2009 *Earth-Science Reviews* **97** 184-214
[2] McKeever S W S 1985 *Thermoluminescence of Solids* (Cambridge, Cambridge University Press)
[3] Wintle A G and Murray A S 1997 *Rad. Meas.* **27** No 4 611-24
[4] Alexander C S, Morris M F and McKeever S W S 1997 *Rad. Meas.* **27** No 2 153-9
[5] Santos A J J, de Lima J F and Valerio M E G 2001 *Rad. Meas.* **33** 427-30
[6] Chen R and McKeever S W S 1997 *Theory of Thermoluminescence and Related Phenomena* (Singapore, World Scientific)
[7] Chithambo M L 2014 *J. Lumin.* **151** 235-43

Structure prediction of manganese dioxide nanoclusters using computer simulation techniques

[1]R R Maphanga, [1]P E Ngoepe, [2]C R A Catlow and [2]S M Woodley

Materials Modelling Centre, School of Physical and Mineral Sciences, University of Limpopo, Private bag x 1106, Sovenga, 0727, South Africa

Department of Chemistry, Materials Chemistry, University College London, Kathleen Lonsdale Building, Gover Street, London, WC1E 6BT, United Kingdom

E-mail: rapela.maphanga@ul.ac.za

Abstract. The characteristics of nanoclusters are linked to the high value of their surface/volume ratio, and therefore the structure of nanoclusters plays an important role in determining their physical properties. In order to enhance the properties of MnO_2 for various applications, some new or modified MnO_2 compounds are developed recently. One of the major demands for developing these materials is to modify and strengthen the structural stability in order to prevent the rapid capacity fading during the process of charge/discharge cycling. The interest in synthesis and characterisation of nanoclusters is driven by a wide range of applications of nanoparticle materials in catalysis, electronics and energy conversion. The lowest energy configurations for $(MnO_2)_n$ clusters, n =1 to 4 are predicted, employing the interatomic potential technique and electronic structure density functional theory method at the PBEso10 level. The application of an evolutionary algorithm to different energy landscapes, as defined by interatomic potentials, for each cluster size was used to generate the plausible structures for refinement using DFT. The geometrical properties of different sizes of nanoclusters are investigated. The DFT based, MnO_2 global minima configurations clusters were found to be similar to those predicted for isostructural TiO_2.

1. Introduction

MnO_2 as a well-known transition-metal oxide is one of the most attractive inorganic materials because of its structural flexibility and wide range applications in many chemical processes such as ion exchange, separation, catalysis, molecular adsorption, biosensors and energy storage in batteries and supercapacitors. MnO_2 exists in several crystallographic forms, such as α-, β-, γ-, λ-, δ-, ε–type, where the basic unit (MnO_6) octahedron links in different ways. Most of these systems have been tested for the application in various batteries, such as alkaline battery, lithium metal battery and lithium-ion battery. One of the major demands for developing new materials is to modify and strengthen the structural stability in order to prevent the rapid capacity fading during the process of charge/discharge cycling [1, 2].

Materials at the nanometer scale display different chemical and physical properties from bulk systems. Recently, special attention has been focused on the fabrication of the hierarchically superstructures based on low-dimensional nanocrystals [3], which are expected to play a crucial role in fabricating the next generation of nanodevices because they can be used as both building units and interconnections and will offer opportunities to explore their novel collective optical, mechanical,

magnetic and electronic properties. There has been an increased interest in manganese dioxide nanostructures for their potential applications in catalysis, rechargeable batteries, ion-sieves and supercapacitors. Nanowires, nanorods, nanobelts, nanotubes and nanosheets of MnO_2 have been widely reported. Several experimental methods (such as liquid-phase, thermal decomposition, electrochemical deposition, sol-gel, molten salt) have been developed for the synthesis of MnO_2 nanostructures, including, nanorods, nanowires, nanobelts and nanoneedles, nanosheets. However, experimental organisation of MnO_2 nanosheet clusters into hierarchical super structures still remains a challenge. Feng *et al* [4] explored the preparation of the hierarchical MnO_2 sheets via electrochemical deposition. Recently, low temperature hydrothermal method was used to synthesise uniform nanoclusters of γ–MnO_2. Li ions were treated with MnO_2 nanoclusters without any intercalation to investigate the effect of lithium ions on the morphology, particle size and charge/discharge behaviours of the synthesised MnO_2 nanocusters [5]. It has been found that the solid state treatment can change the morphology and particle size of γ–MnO_2 nanoclusters, without considerable change in the phase composition. In addition, low-dimensional structures of MnO_2 have been synthesised as nanorods, nanosheet [6] and modelled as nanoparticle [7]. In addition, MnO_2 structures have been simulated and studied in varying shapes and architectures, that is, nanosheet, nanorod and mesoporous [8]. The control of morphologies, sizes, architectures, and patterns of nanocrystals have become a dominant theme in material field since these parameters are the key elements that determine physical properties. It is well known that battery performance is highly dependent on MnO_2 morphologies and crystallographic forms.

Low-lying local minima on the energy of formation landscape defined by interatomic potentials (IPs) were found by employing one of a number of global optimisation algorithms [9, 10]. The use of IPs approaches reduces the computational cost of locating plausible structures, which can be refined at a later stage. Searching landscapes based on IPs has been successfully applied to finding a wide range of bulk metal oxide phases [11-15] and therefore such techniques have been applied to clusters, such as, MgO [16], ZnO [17], ZnS [18, 19], TiO_2 [19], SiO_2 [20], ZrO_2 [21], Al_2O_3 [13] and In_2O_3 [22]. The IPs are typically fitted so as to reproduce the structures and properties of the bulk phases, which may be one of the causes of a change in the ranking in terms of stability order for the clusters. When the plausible structures are found, are refined using a DFT approach. The unprecuresedented efforts have focused on the synthesis [23, 24] structure [25, 26] properties [27] and applications [28-30] of nanomaterials. However, structural features that exist at the nanoscale are difficult to characterise experimentally, and therefore, computer simulation methods offer a unique window of exploration into nanomaterials. In this study we develop simulation strategies for generating MnO_2 nanoclusters. To understand the behaviour of MnO_2 small clusters at nanoscale, a combination of global search techniques and density functional theory method are employed to elucidate the energy landscapes and geometrical configurations of MnO_2 clusters.

2. Methodology

Energies of nanoclusters can be calculated using methods based either on interatomic potentials or electronic structure techniques. The interatomic potential (IP) method is computationally inexpensive and allows large numbers and sizes of clusters to be explored. However, they have a number of limitations which may results in giving inaccurate results. Combining the two approaches (interatomic potential and density functional theory) proved to be insightful, where IP methods are being used in the first stage to explore the range of cluster sizes and structures, while DFT methods are used in the second stage to refine the energy ordering and structures of a selected subset of clusters. Thus, those found to be low in energy from the first stage. In addition, the DFT methods can be employed to test the reliability of the first stage calculations by probing cluster structures, which are likely to be sensitive to the parameterisation of interatomic potential used. Interatomic potentials are commonly parameterised by reference to bulk crystalline properties. This may cause difficulties when they are applied to clusters where the coordination number and bonding may differ. It is vital that the potential based calculations are validated by DFT methods.

All IP calculations are based on Born model of ionic solids, in which the Mn and O ions interact via a long range Coulombic interactions and short-range parameterised interactions. The potential parameters were previously developed by Matsui and were refined and used to model bulk, surfaces and structures of MnO_2 [7, 8, 31]. For the shell model, potentials developed by Ammundsen [32] were employed and fitted to reproduce the lattice parameters of MnO_2 bulk structures. The potentials gave good agreement with the crystal structures of both pyrouliste and ramsdellite polymorphs of MnO_2. The evolutionary algorithm implemented within the GULP [33] code was used to perform IP calculations. The IP calculations are based on two potential models, namely, the partially charged rigid ion model and the fully charged shell model. The energy is given by the sum of the two-body interactions and its analytical expression is given below:

$$V(r_{ij}) = \frac{q_i q_j}{r_{ij}} + B_0 + B_1 r + B_2 r^2 + B_3 r^3 + B_4 r^4 + B_5 r^5 \quad (1)$$

where the first term is the Coulomb contribution to the energy between two point charges q_i and q_j at distance r_{ij} apart. The analytical expression for Buckingham potential employed in this work, describes the interaction between two ions, *i* and *j*, of charge q separated by a distance r and is given in equation below:

$$Ur_{ij} = \left[\frac{q_i q_j}{r_{ij}} + A_{ij} e^{\left(-r_{ij}/\rho_{ij}\right)} - \frac{c_{ij}}{r_{ij}^6}\right] \quad (2)$$

An evolutionary algorithm (EA) method for global optimisation in stage 1 was employed to search for lower Local Minima (LM) on the energy hypersurface to predict stable and low-energy metastable atomic configurations. Different relaxed configurations, initially random stationary points on the energy hypersurface, from a population in which competition to survive and procreate is simulated. The probability of success on any current configuration is based on its energy relative to the other configurations in the population. In the current study, only the best 30 unique structures are kept, at the end of each cycle. Within a population, if the difference in the energy of formation between the two configurations is within E_d, then the higher energy configuration is removed from the population. The EA method was applied at least twice to each hypersurface, when E_d = 0.01 eV and E_d = 0.004 eV. On each of these standard runs, 200 EA cycles were performed. The atoms of each isolated cluster are constrained to be within a spherical container with a radius of 8 Å. The constraint prevents any initial fragmentation of a cluster and helps to speed up convergence of the local optimiser. Each LM corresponds to a configuration of atoms of the cluster, referred to as plausible structure.

As a second stage, each set of plausible structures for the IP are refined using standard local optimisation techniques. In the final stage the global minima structures and a number of metastable local minima configurations as defined by the IPs, were refined using a DFT method. The DFT (PBEsol) all-electron method with a local numerical orbitals basis set and local optimisation techniques as implemented in the FHI-AIMS [34] was employed. The FHI-AIMS two-tiered basis set was employed for the oxygen atoms and only the first-tier basis set for the manganese atoms and scalar-relativistic effects treated at the scaled ZORA level. To improve the energies and the respective electronic structures, the hybrid PBEsol0 functional and the full ZORA level of the theory were employed for recalculation. Configurations in the final population of the evolutionary algorithm are refined using the shell model. Then 30 lowest energy candidate structures, as measured by the energy of formation using the rigid ion and the shell models, were then subjected to geometry optimisation using DFT.

3. Results and Discussions

The lowest energy local minima configurations for $(MnO_2)_n$ n=1-4, clusters have been published previously, where the order of stability as measured by IP and DFT PBEsol0 methods were presented [35]. In addition, the saddle point configurations for MnO_2 in comparison with global minima for TiO_2

clusters were reported. Figure 1 shows the most stable $(MnO_2)_n$ n=1-4 nanoclusters in comparison to literature results.

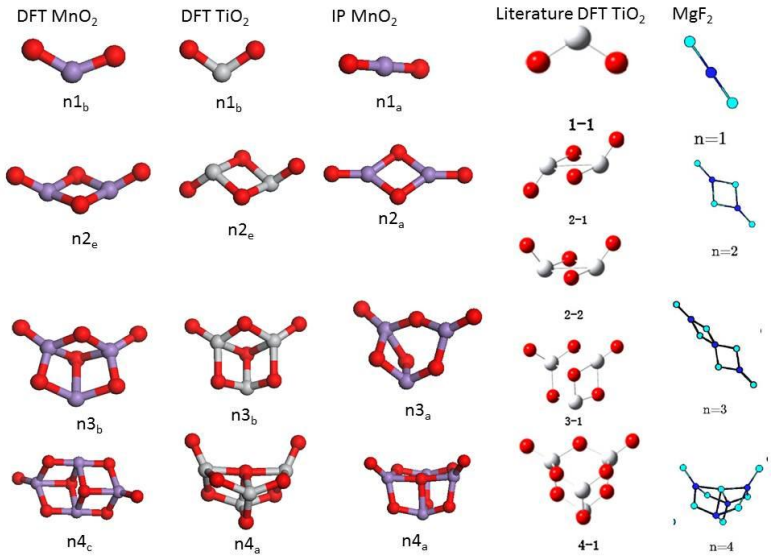


Figure 1: Geometrical configuration of MnO_2 nanoclusters in comparison with nanoclusters from the literature, obtained using IP and DFT methods.

In similar approach used to determine MO_2 (M=Ti, Zr, Hf, Si) clusters, to that discussed above for obtaining low-energy structures, the one key point is that the influence of the d-electrons is not usually accounted for in models based on interatomic potentials (IPs), i.e. when searching for plausible structures. In this study, the d-electrons are accounted for when searching for plausible structures. Thus, the spin configuration is incorporated within IP calculations. Typically, this influence has a greater effect on the bond angles about under-coordinated atoms on the surface, which can always be corrected for in a later stage, e.g. further refinements of the plausible structures using DFT approaches.

For n = 1, interatomic potential predicts a linear configuration with an angle of 180° during stage 1. This is in agreement with previous studies on MgF_2 clusters [36]. The structure bends into a boomerang (n1a) during relaxation in stages 3 and 4 when the DFT methods are employed. During DFT relaxation, the bond angle tends to decrease from 180° to 129° because polarisation is explicitly included in DFT models. Previously, different basis sets (ECP and DZV) have been employed to calculate geometries for isostructural TiO_2 and they gave similar results except small differences in bond lengths and angles [37]. They found that IP predicted a linear configuration with an angle of 180° while B3LYP/6 predicted an angle of 110.41°. Thus, n=1 TiO_2 nanocluster, the geometry was predicted to be of angular configuration. As it has been reported for isostructural TiO_2, our results predict the same configuration. Currently, there are no literature results on MnO_2 nanoclusters to compare with. Previous studies reported that d-orbitals have influence on the nanocluster arrangement. Thus, smallest global minima configurations for germania, silica and stania have a linear arrangement whilst titania, hafnia and zirconia have a boomerang arrangement. Similar local minima configurations for DFT calculations have been found for MnO_2 nanoclusters even though bond angles associated with any singly coordinated atoms are typically different.

The interatomic potential predicts $n2_a$ to be the global minimum, which has planar arrangement with bond distance of 1.687 Å for Mn-O1 and 1.822 Å for Mn-O2. The configuration depicts that oxygen atoms form a square on the mirror plane between the two manganese atoms. The IP predicts the GM structure to be the same as those previously reported for silica, germania, stania [21] and MgF_2 smallest cluster configurations. In stage 3 where the global minima found in stage 1 is relaxed, $n2_e$ $n2_b$ and $n2_c$ configurations are found to be the local minima. When the DFT method is employed to relax plausible structures, $n2_b$ and $n2_a$ configurations are obtained as the global minima for MnO_2 and TiO_2 clusters, respectively. The models predict the bond distances of the two Mn-O1 pair to be

smaller than that of four Mn-O2. For n = 2 global minima clusters, IP geometries predict bond distance results that are closely comparable to DFT results. Both systems predict a configuration with two-terminal Mn-O bond as the global minima in agreement with previous studies on TiO_2 DFT nanoclusters calculations. The first and second local minima are predicted to be the configurations with two terminal Ti-O and one terminal and three bridging O atoms respectively. Comparing the IP configurations for $(MnO_2)_2$ nanoclusters with those obtained using DFT, the atomic arrangement is similar with DFT configurations bending to form a quasi-like isomer for the first two local minima, i.e., $n2_e$ and $n2_b$.

The cluster n3A was predicted to be the global minima for IP calculations. The bond lengths and angles for this cluster are also listed in table 1, where O1 and O2 indicate that oxygen atoms is coordinated to 1 and 2 manganese atoms respectively. The lowest energy configuration, $n3_b$ for $(MnO_2)_3$ is composed of two terminal oxygen atoms and one three-fold coordinated oxygen atoms in agreement with previous DFT calculations. The DFT predicts that the GM configuration is composed of three tetragonal sides of a cuboid with two singly coordinated oxygen atoms. The cluster n3c has the same topology, except that one bond is missing to complete a three-fold coordinated oxygen atoms. The third lowest-energy configuration depicts a quasi-like isomer with two orthogonal MnO_2. The quasi-linear arrangement was suggested to be the global minimum in the Hatree-Fock [38] and pair-potential [39] calculations, whilst n3c was suggested as the global minimum in recent pair potential studies [37]. The local minima configurations are found to be of the same arrangement with the global minima, in which one Mn-O bond is broken. Similarly, DFT calculations for the nanocluster n=3, predicts the configuration of the lowest GM to be identical to those found for TiO_2, ZrO_2 and HfO_2 in previous studies. Thus, $n3_b$ is the same GM as reported for the aforementioned nanoclusters. The GM configuration for DFT-PBEsol0, can be constructed from smaller clusters, in particular, $n1_b$ and $n2_b$. Thus, the smallest lowest energy configurations can be used as building blocks for the larger clusters.

The configuration n4e is found to be the global minima for $(MnO_2)_4$, which is composed of two terminal oxygen atoms and six bridging oxygen atoms. The GM configurations found for n=4 can be constructed from smaller clusters. Numerous DFT studies on $(TiO_2)_n$ showed that DFT calculations predict the lowest energy global minima better than the other methods, such as Hatree-Fock (HF) and pair-potential. The latter two methods neglect the partial covalent nature of Ti-O bonds. The comparison of the lowest three LM MnO_2 cluster structures and those predicted for titania and zirconia is noted that the same configurations and rank is found for n = 2. The clusters are ranked according to PBEsol0 functional because it is one of the least biased, among complementary functionals. The small clusters are isostructural to those predicted for titania, silica and hafnia [21] and characterised by Mn ions concentrating in the core region of the cluster decorated by dangling, or singly coordinated, oxygen ions at the periphery. However, when the energy of formation is calculated using a DFT approach, after relaxation, the dangling oxygen atoms for MnO_2 do not remain in the plane containing the nearest Mn_2O_2 tetragon.

4. Conclusion

Interatomic potential and density functional theory methods were employed successfully to predict the global and local minima configurations of small $(MnO_2)_n$ n=1-4 nanoclusters. The IP method allows a large number of configurations to be explored. The plausible structures are then refined using more accurate DFT-PBEsol0 method. There is a good agreement between our generated MnO_2 nanocluster results and literature results for other metal oxides, in particular, isostructural TiO_2. The IP and DFT methods predict LM configurations for $(MnO_2)_n$ n=1-4 to be similar to those reported for TiO_2 obtained using Hybrid Simulated Annealing-Monte Carlo Basin Hopping (SA-MCBH) and DFT (B3LYP/DZVP) respectively. For these small clusters, the two methods predict similar global minima configurations with slight difference in bond angles. It is noted that all the lowest-energy clusters always possess one or two terminal Mn-O bonds of almost a length of approximately 1.6Å. The results are of great importance in understanding the formation and nucleation of MnO_2 nanostructures. In summary, there are similarities in the configurations found in simulating metal oxides, namely, MnO_2, TiO_2, ZrO_2 and HfO_2 nanoclusters.

References

[1] Fernandes J B, Desai B D and Kamat Dalal V N 1984 *Electrochim. Acta* **29** 181
[2] Ghaemi M, Ghavami R K, Khosravi-Fard L and Kassaee M Z 2004 *J. Power Sources* **125** 256
[3] Wang D, Luo H, Kou R, Gill M P, Xiao S, Golub V O, Yang Z, Brinke C J and Lu Y 2004 *Angew. Chem. Int. Ed.* 436169
[4] Feng Z -P, Li G -R, Zhong J -H, Wang Z -L, Ou Y -N and Tong Y -X, *Electrochem. Commun.* 2009 **11** 706
[5] Karami H, Ramandi-Ghamoshi M, Moeini S and Salehi F 2010 *J. Clust. Sci.* **21** 21
[6] Jiao F and Bruce P G, *Adv. Mat.* 2007 **1**9 657
[7] Sayle T X T, Catlow C R A, Maphanga R R, Ngoepe P E and Sayle D C 2005 *J. Am. Chem. Soc.* **127** 12828
[8] Sayle T X T, Maphanga R R, Ngoepe P E and Sayle D C 2009 *J. Am. Chem. Soc.* **131** 6161
[9] Woodley S M and Catlow C R A 2008 *Nat. Mater.* **7** 937.
[10] Woodley S M, Applications of Evolutionary Computation in Chemistry, Berlin: Springer-Verlag Berlin, 2004
[11] Woodley S M, Battle P D, Gale J D and Catlow C R A 1999 *Phys. Chem. Chem. Phys.* **1** 2535
[12] Woodley S M, Battle P D, Gale J D and Catlow C R A, 2004 *Phys. Chem. Chem. Phys.* **6** 1815
[13] Woodley S M 2007 *Phys. Chem. Chem. Phys.* **9** 1070
[14] Woodley S M and Catlow C R A 2009 *Comput. Mater. Sci.* **45** 84
[15] Schön J C 2004 *Z. Anorg. Allg. Chem.* **630** 2354
[16] Roberts C and Johnston R L 2001 *Phys. Chem. Chem.* Phys. **3** 5024
[17] Al-Sunaidi A A, Sokol A A, Catlow C R A and Woodley S M 2008 *J. Phys. Chem. C* **112** 18860
[18] Woodley S M, Sokol A Aand Catlow C R A 2004 *Z. Anorg. Allg. Chem.* **630** 2343
[19] Hamad S, Woodley S M and Catlow C R A 2009 *Mol. Simul.* **35** 1015
[20] Flikkema E and Bromley S T 2004 *J. Phys. Chem. B* **108** 9638
[21] Woodley S M, Hamad S and Catlow C R A 2010 *Phys. Chem. Chem. Phys.* **12** 8454
[22] Walsh A and Woodley S M 2010 *Phys. Chem. Chem. Phys.***12** 8446
[23] Pinna N and Niederberger M 2008 *Angew. Chem. Int. Ed.* 47 5292
[24] Lai M and Riley D J 2008 *J. Colloid Interface Sci.* **323** 203
[25] Thomas J M and Midgley P A 2004 *Chem. Commun.* **11** 1253
[26] Möbus G and Inkson B J 2007 *Mater. Today* **10** 18
[27] Lai M, Kulak A N, Law D, Zhang Z B, Meldrum F C and Riley D J M. Lai 2007 *Chem. Commun.* **34** 3547
[28] Feng X D, Sayle D C, Wang Z L, Paras M S, Santora B, Sutorik A C, Sayle T X T, Yang Y, Ding Y, Wang X D and Her Y S 2006 *Science* **312** 1504
[29] Berry C C 2005 *J. Mater. Chem.* **15** 543
[30] Tarnuzzer R W, Colon J, Patil S and Seal S, *Nano Lett.* **5** 2573
[31] Maphanga R R, Parker S C and Ngoepe P E 2009 *Surf. Sci.* **603** 3184
[32] Ammundsen B, Islam M S, Jones D J and Rozière J 1999 *J. Power Sources* **81-82** 500
[33] Gale J D and Rohl A L 2003 *Mol. Simul.* **29** 291
[34] Havu V, Blum V, Havu P and Scheffler M 2009 *J. Comput. Phys.* **228** 8367
[35] Ngoepe P E, Maphanga R R and Sayle D C 2013 Computational Approaches to Energy Materials, Wiley & Sons, UK
[36] Francisco E, Martín Pendás A and Blanco M A 2005 *J. Chem. Phys.* **123** 234305
[37] Hamad S, Catlow C R A, Woodley S M, Lago S and Mejias J A 2005 *J. Phys. Chem. B* **109** 15741
[38] Hagfeldt A, Bergstroöm R, Siegbahn H O G and Lunell S 1993 *J. Phys. Chem.* **97** 12725
[39] Yu W and Freas R B 1990 *J. Am. Chem. Soc.* **112** 7126

Synthesis, structural and optical characterisation of cobalt and indium co-doped ZnO nanoparticles

MW Maswanganye [a1], KE Rammutla [a], TE Mosuang [a], BW Mwakikunga [b], TS Bertrand [c] and M Maaza [c]

[a] Department of Physics, University of Limpopo, Private Bag x1106, Sovenga, 0727, RSA
[b] DST/CSIR National Centre for Nano-structured Materials, P O Box 325, Pretoria 0001, RSA
[c] Materials Research Department, iThemba labs, P O Box 722, Somerset West, 7129, RSA

Email address: mwmaswanganye@gmail.com or Thuto.Mosuang@ul.ac.za

Abstract. The undoped, 5% Co, In single doped and 5% In-Co double-doped ZnO nanoparticles were successfully prepared using sol-gel method. The structural and optical properties of the samples were investigated using XRD, UV-vis, TEM, EDS and Raman spectroscopy. There were no peaks associated with In or Co detected in the XRD patterns indicating that In^{3+} and Co^{2+} substituted for Zn^{2+} ions in the ZnO wurtzite structure this was also confirmed by the EDS and Raman results. TEM results showed that the prepared ZnO nanoparticles were spherically shaped. Single doping reduced the grain size and the energy band gap of the ZnO nanoparticles while combinational doping reduced them even further.

1. Introduction

Nano-materials of transition metal oxides have been found to be very interesting and currently they are investigated for their different properties compared to those of bulk materials [1]. For example, ZnO is an n-type semiconductor material having a wurtzite hexagonal crystal structure. It has a wide energy band gap of 3.36 eV and a large excitation binding energy of 60 meV [2]. It has a wide range of resistivity, high transparency at room temperature [3], high electron hall mobility, good chemical and thermal stability under operation conditions [3]. It's widely being used in a variety of applications such as opto-electronic devices [4], [5], light emitting diodes (LEDs) [6], flat panel displays [7], surface acoustic wave devices [8], chemical sensors, gas sensors [9] and solar cells [10]. Previous studies have shown that the physical properties of ZnO nanostructured materials depend on their size, shape and structural aspects and those factors are mostly influenced by the method of preparation [11][12] [13]. It has been reported that single doping ZnO nanoparticles with a controlled amount of impurities such as Cu, Al, In, Sb, Ga, Fe, Co, Ni, etc. can improve its structural, electrical and optical properties [14] while combinational doping is found to enhance its properties even further [15] [16]. Different methods such as chemical vapour deposition (CVD) [17], physical vapour deposition (PVD) [18], hydro thermal and sol-gel method [19] have been used to synthesize the undoped and doped ZnO nanoparticles. In this work we will be reporting on the structural and optical properties of 5% Co, In single doped and 5% In-Co co-doped ZnO nanoparticles prepared using sol-gel method.

[1] To whom any correspondence should be addressed.

2. Experimental

2.1. Sample preparation and characterization

The undoped, single doped and double-doped ZnO nanoparticle samples were synthesised using chemicals zinc acetate dihydrate ($C_4H_6O_4Zn{\cdot}2H_2O$), ethanol (C_2H_6O), ethanolamine (C_2H_7NO), cobalt (II) nitrate hexahydrate ($CoN_2O_6{\cdot}6H_2O$) and indium (III) nitrate hydrate ($InN_3O_9{\cdot}xH_2O$). The water soluble salts; cobalt (II) nitrate hexahydrate and indium (III) nitrate hydrate were used as sources of Co and In metal dopants. All of these chemicals were purchased from Sigma-Aldrich. Single doped ZnO nanoparticles with Co and In were prepared at 5 wt. % while the double-doped ZnO nanoparticles sample were prepared at 2.5 wt. % of each dopant in order to obtain 5 wt. %.

When preparing the undoped ZnO nanoparticles 0.2 M of zinc acetate dihydrate solution was mixed with 0.2 M of ethanol in order to obtain homogeneity, this was done with continuous stirring. For the single doped ZnO nanoparticles, cobalt (II) nitrate hexahydrate or indium (III) nitrate hydrate were added into 0.2 M of zinc acetate dihydrate solution and for the co-doped ZnO nanoparticles the cobalt (II) nitrate hexahydrate and indium (III) nitrate hydrate were added simultaneously into 0.2 M of zinc acetate dihydrate solution then mixed with 0.2 M of ethanol in order to obtain homogeneity. This was also done while continuously stirring. Proportionate amounts of 0.2 M ethanolamine were added in all the solutions as a stabilizer. Thereafter the solutions were stirred for 2 hours at 70℃. The precipitates formed were collected using filter paper and annealed at 400 ℃ for 1 hour. After annealing all the samples were characterised using X-ray diffraction (XRD), Transmission Electron Microscope (TEM) and Ultraviolet-visible spectroscopy (UV-vis), Energy Dispersive Spectroscopy (EDS) and Raman spectroscopy.

3. Results and discussions

3.1. *X-ray diffraction*

From fig.1. the XRD patterns show that all the prepared samples have a hexagonal wurtzite structure with ZnO peaks observed at $2\Theta = 31.77°$, $34.43°$, $36.27°$, $47.56°$ and $56.63°$ which belongs to the (100), (002), (101), (102) and (110) planes respectively (JCPDS No.36-1451). No peaks associated with Co or In were detected in the XRD pattern indicating that Co^{2+} and In^{3+} ions have substitutionally replaced the Zn^{2+} ions. Additional peaks (indicated with *) were detected in the sample belonging to Al from the XRD sample holder. The lattice parameters a = b and c for the undoped and doped ZnO nanoparticles were determined using the (100) and (002) planes and were found to be similar to the reported values of the bulk ZnO (JCPDS No.36-1451); the results are tabulated in table 1.

The average grain sizes of the undoped and doped ZnO nanoparticles were estimated using Debye-Scherrer's equation and are also presented in table 1. The values indicate that the grain sizes of the doped ZnO nanoparticles are smaller compared to the grain size of the undoped ZnO nanoparticles and co-doping ZnO nanoparticles significantly reduces the grain size of ZnO nanoparticles. The strains of all the undoped and doped ZnO nanoparticles were determined and found to be inversely proportional to the grain size. The values of the strain can also be viewed in table 1.

Table 1. Lattice parameters, average grain size and strain.

Sample	Lattice parameters		D(nm)	ε ($\times \mathbf{10^{-3}}$)
	a (Å)	c (Å)		
Undoped-ZnO	3.2490	5.2049	42.476	2.3982
5% Co-ZnO	3.2582	5.2171	19.242	5.3116
5% In-ZnO	3.2429	5.1893	22.179	4.5884
5% In-Co-ZnO	3.2394	5.1836	11.858	8.6131

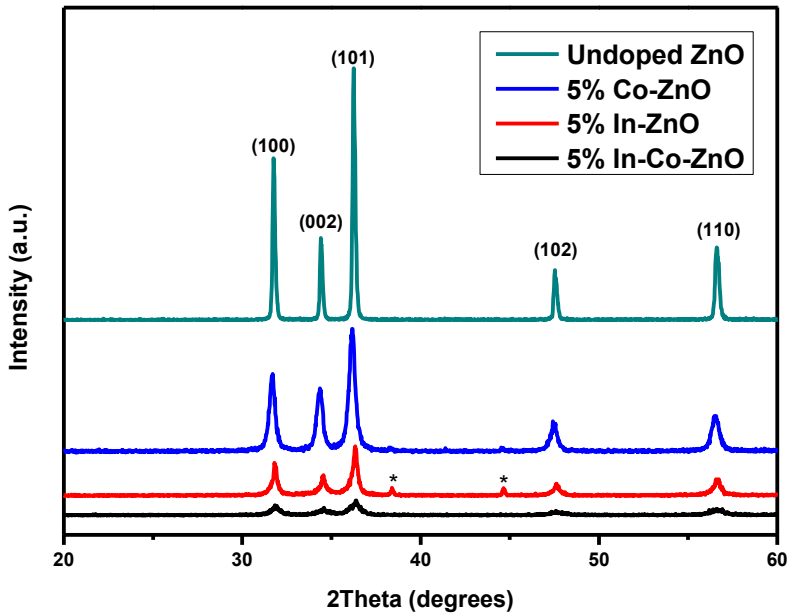


Figure 1. XRD patterns of undoped ZnO, 5% Co-doped ZnO, 5% In-doped ZnO and 5% In-Co co-doped ZnO nanoparticles.

3.2. *TEM*

Fig.2. shows the TEM images of (a) Undoped ZnO nanoparticles and (b) 5% In-Co-ZnO nanoparticles. Both the undoped and 5% In-Co co-doped ZnO nanoparticles are found to be spherical. In Fig.2 (a) big rods can be seen alongside the spherical nanoparticles and these rods are basically created when spherical nanoparticles come together. Consequently smaller particles can be seen within the rods; similar results were reported by L. Xu et al. [20]. Fig.3. shows the EDS images of 5% In-Co co-doped ZnO nanoparticles. The elements Zn, O, In and Co were observed in the image which confirms that Co and In were successfully incorporated into the ZnO nanoparticle structure. While C, Cu, Si elements may be from the TEM sample grid.

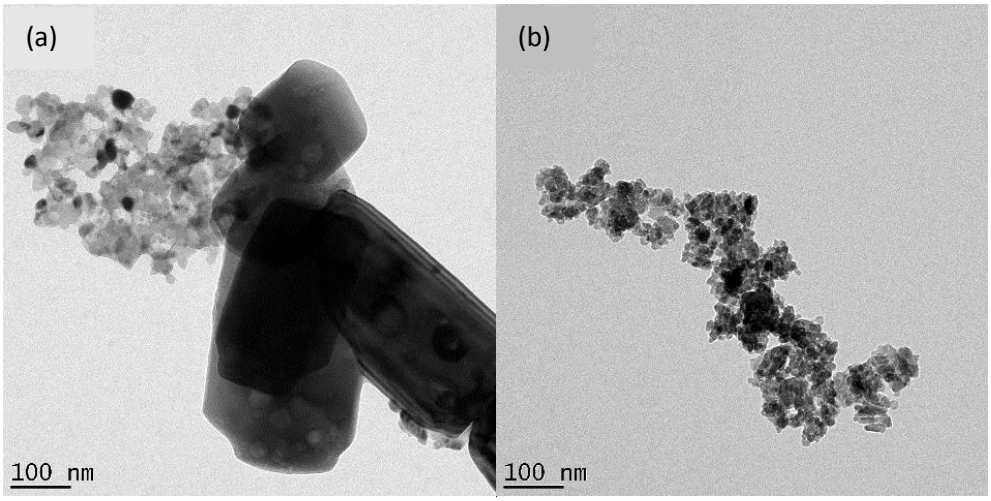


Figure 2. TEM images of (a) undoped ZnO (b) 5% In-Co-ZnO

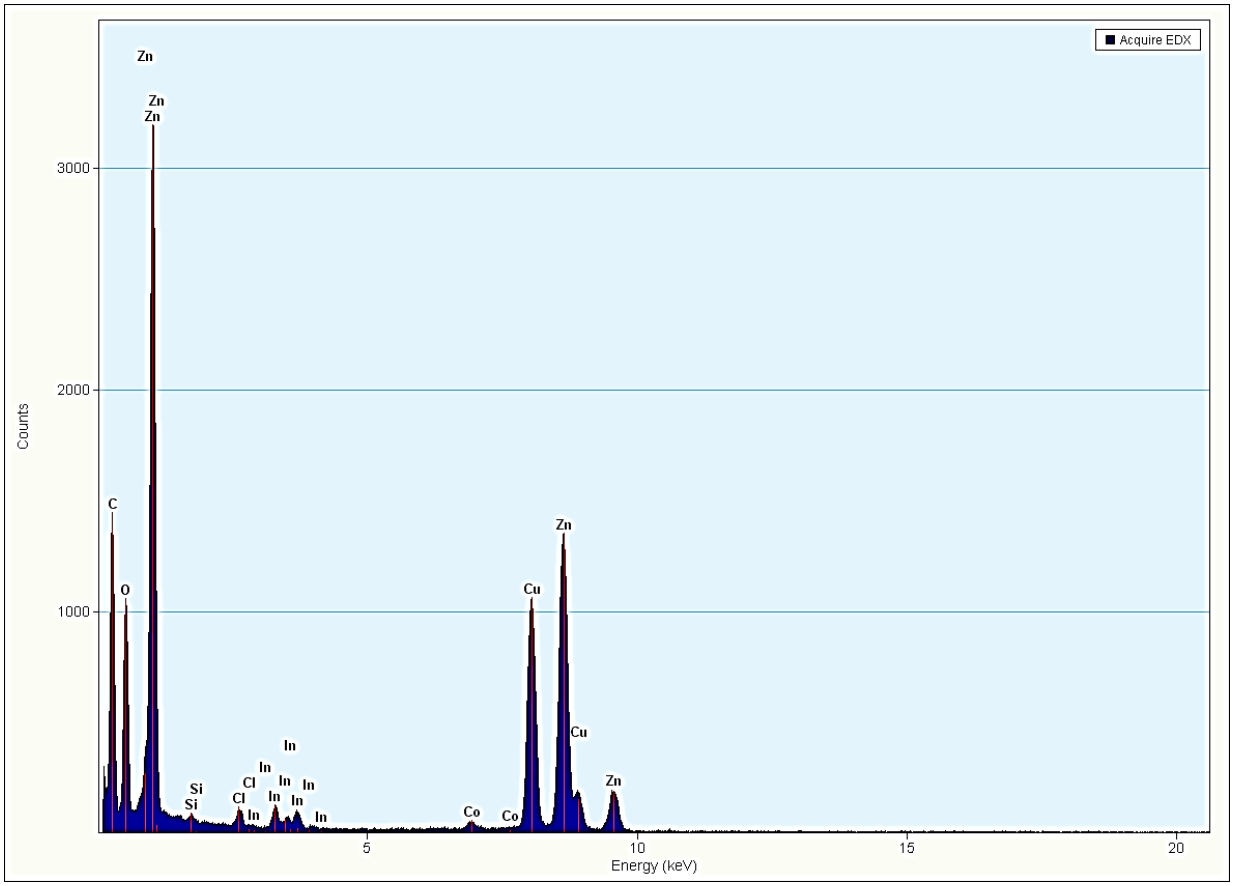


Figure 3. EDS image of 5%In-Co-ZnO nanoparticles.

3.3. UV-vis spectroscopy

The optical absorption spectra of the undoped and doped ZnO nanoparticles were recorded as shown in Fig. 4 in order to report the optical properties of these nanoparticles. For the nanoparticles doped with Co three additional absorption peaks are observed between 550-700 nm. These peaks are attributed to $^4A_2(F) \rightarrow {}^2A_1(G)$, $^4A_2(F) \rightarrow {}^4T_1(P)$ and $^4A_2(F) \rightarrow {}^2E(G)$ transitions which suggest that the tetrahedrally coordinated Co^{2+} ions substituted for Zn^{2+} ions in the hexagonal ZnO wurtzite structure [21] [22].

The energy band gap values have been determined and found to be 3.10, 2.48, 2.74, 2.37 eV for the undoped ZnO, 5% Co-ZnO, 5% In-ZnO and 5% In-Co-ZnO nanoparticles, respectively. It can be seen that the value of the undoped ZnO nanoparticles is smaller compared to the reported value of the bulk ZnO [23] and this could be due to the fact that nanostructured materials have different properties from the bulk materials. A red shift is observed between the undoped ZnO nanoparticles and the doped ZnO nanoparticles. It has been reported by K. Ando et al. [23] that this red shift could be attributed to the fact that Co^{2+} and In^{3+} ions substituted for the Zn^{3+} ions in the crystal lattice**.**

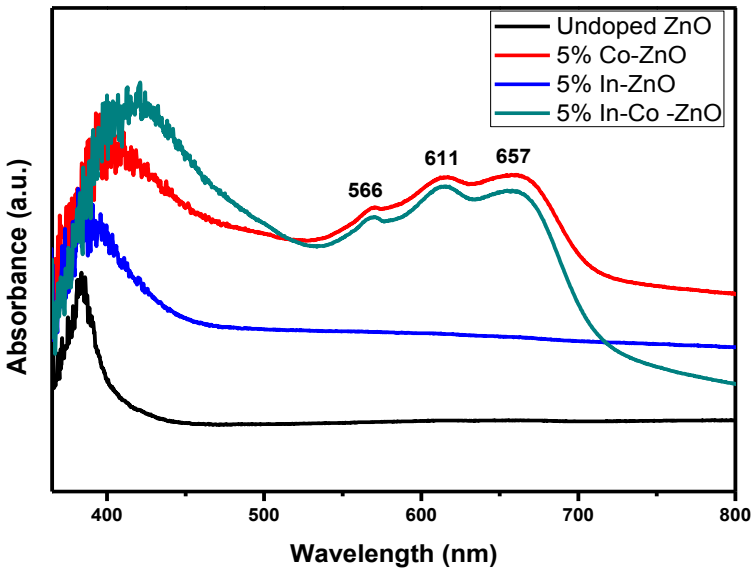


Figure 4. Optical Absorbance spectra of Undoped ZnO, 5% Co-ZnO, 5% In-ZnO and 5% In-Co-ZnO nanoparticles.

3.4. Raman spectroscopy

The vibrational properties of the Co and In single doped, Co and In co-doped ZnO nanoparticle and undoped ZnO nanoparticle samples were investigated with the excitation wavelength of 514 nm at room temperature. From Group theory it is predicted that the ZnO wurtzite hexagonal structure of a space group C_{6v}^4 will indicate the phonon modes near the centre of Brillouin zone given by the equation $\Gamma = A_1 + 2B_1 + E_1 + 2E_2$ [24]. Here A_1, E_1 and $2E_2$ are Raman active modes and $2B_1$ is the forbidden mode of ZnO. A_1 and E_1 are polar and are split into two, being transverse optical (TO) and longitudinal optical (LO) phonons were E_2 is nonpolar and is divided into E_2^{low} and E_2^{high} [24]. The Raman peaks at 329, 377, 407, 435, 536-558 and 581 cm^{-1} belong to the E_2^{high}- E_2^{low}, A_1 (TO), E_1 (TO), E_2^{high}, A_1 (LO), and E_1 (LO) modes, respectively. The Raman peaks at 435 cm^{-1} are thought to be confirming that the samples prepared are of the ZnO wurtzite hexagonal structure and similar results were obtained [25] [15], while the peaks around 581 cm^{-1} are thought to be from defects such as oxygen vacancies (VO), zinc interstitials (Zni) and/or free carriers [25] [26].

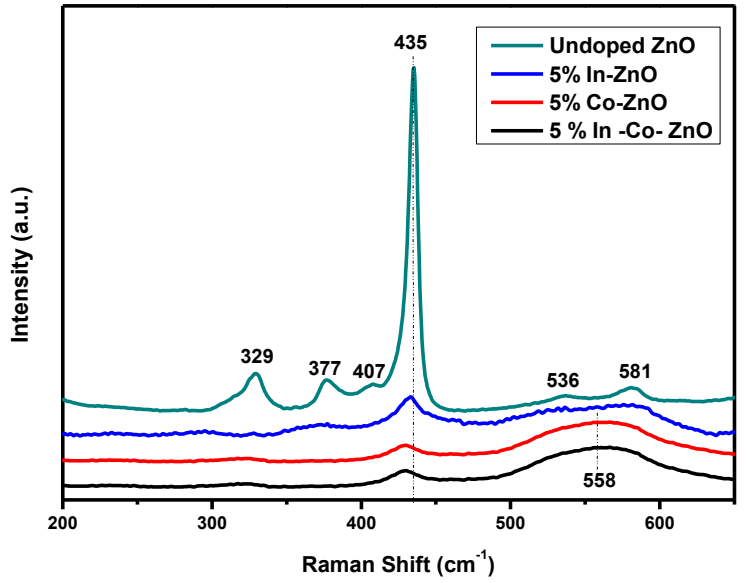


Figure 5. Raman spectroscopy of undoped doped ZnO, 5% of Co, 5% In and 5% of In/Co co-doped ZnO nanoparticles.

4. Conclusion

The In and Co single doped ZnO nanoparticle, In and Co co-doped ZnO nanoparticle and undoped ZnO nanoparticle samples were successfully synthesised using sol-gel method. The effect of doping In and Co on the structural and optical properties of ZnO were investigated. XRD results showed that all the prepared samples were of the ZnO wurtzite crystal structure and the results were corroborated by the Raman results. TEM images showed that the prepared samples were spherical in shaped. EDS confirmed that In and Co were indeed substitutionally replacing the Zn^{2+} ions in ZnO nanoparticles. The energy band gap of the prepared samples were determined and found to be smaller to the energy band gap of the bulk ZnO.

References

[1] C. Klingshirn, J. Fallert, H. Zhou, J. Sartor, C.Thiele, F. Maier-Flaig, D. Schneider, H. Kalt 2010 Phys Status Solidi B 247 1424

[2] A.N. Mallika, A. Ramachandra Reddy, K. Sowri Babu, Ch. Sujatha, K. Venugopal Reddy 2014 Opt. Mater 36 879

[3] S.J. Pearton, D.P. Norton, K. Ip, Y.W. Heo, T. Steiner 2004 J. Vac. Sci. Technol., B 22 932

[4] A.N. Gruzintsev, V.T. Volkov, E.E. Yakimov 2003 Semiconductors 37 259

[5] A. Kołodziejczak-Radzimska, T. Jesionowski 2014 Materials 7 2833

[6] H. Kim, J.S. Horwitz, W.H. Kim, A.J. Makinen, Z.H. Kafafi, D.B. Chrisey 2002 Thin Solid Films 420/421 539

[7] M. Chen, Z.L. Pei, C. Sun, J. Gong, R.F. Huang, L.S. We 2001 Mater. Sci. Eng., B 85 212

[8] J. Lee, H. Lee, S. Seo, J. Park 2001 Thin Solid Films 398/399 641

[9] D.S. Bohle,C.J.Spina 2010 J. Phys. Chem C 114 18139

[10] A. Al-Kahlout 2012 Thin Solid Films 520 1814

[11] S.Y. Kim, I.S. Lee, Y.S. Yeon, S.M. Park, J.K. Song 2008 Bull. Korean Chem. Soc. 29 1960

[12] B.K. Woo, W. Chen, A.G. Joly, R. Sammynaiken 2008 J. Phys. Chem C 1121 4292

[13] R.K. Dutta, P.K. Sharma, R. Bhargava, N. Kumar, A.C. Pandey 2010 J. Phys. Chem B 114 5594

[14] M.R.R. Madhukar Poloju 2013 in proceedings of the International Conference on Advanced Nanomaterials and Emerging Engineering Technologies, India

[15] Z. Pan, J. Luo, X. Tian, S. Wu, C. Chen, J. Deng, C. Xiao, G. Hu, Z. Wei 2014 J. Alloys Compd. 583 32

[16] D. Fang, P. Yao, H. Li 2014 Ceram. Int. 40 5873

[17] J. Nishino, T. Kawarada, S. Ohshio 1997 J. Mater. Sci. Lett. 16 629

[18] M. Sathyal, A. Claude 2010 advanced in applied science research 32591

[19] M. RezaeeRokn-Abadi, M. Behdani, H. Arabshahi and N. Hosseini 2009 IREPHY 12 103

[20] L. Xu, H. Zhang, K. Shen, M. Xu, Q. Xu 2012 Journal of Superconductivity and Novel Magnetism 25 1951

[21] H. weakliem 1962 J.Phys.chem 36 2117

[22] P. Li, S. Wang, J. Li, Y. Wei 2012 J. Lumin. 132 220

[23] K. Ando, H. Saito, Z. Jin, T. Fukumura, M. Kawasaki, Y. Matsumoto and H. Koinuma 2001, J. Appl. Phys 89 7284

[24] Ya-Ping. Du, Ya-Wen. Zhang, Ling-Dong. Sun, and Chun-Hua. Yan 2008 J. Phys. Chem. C 112 12234

[25] Sanjeev Kumar, C.L. Chen, C.L. Dong, Y.K. Ho, J.F. Lee, T.S. Chan, R. Thangavel, T.K. Chen, B.H. Mok, S.M. Rao, M.K. Wu 2013 J Mater Sci 48 2618

[26] K. Kawabata, Y. Nanai, S.Kimura, T. Okuno 2012 Appl Phys A 107 213

The modified interatomic potentials of FeS_2 in atomistic simulations

M A Mehlape[1], **P E Ngoepe**[1] **and S C Parker**[2]
[1] Materials Modelling Centre, School of Physical and Mineral Sciences, University of Limpopo, South Africa, Private Bag, X1106, SOVENGA, 0727
[2] Computational Solid State Chemistry, University of Bath, Claverton Down, Bath, Avon, UK, BA2 7AY

E-mail: mofuti.mehlape@ul.ac.za

Abstract. The modified interatomic potentials were used for both energy minimization and molecular dynamics to study the surfaces and the bulk structure of pyrite. With energy minimization we calculated the surface energies of the surfaces {100}, {110}, {111} and {210}. They revealed that {100} surface is the most stable surface. When we compared the surface energies calculated from the original potentials and the adjusted potentials, it is clear that the adjusted potentials improve the stability of the surfaces. It was also revealed that water stabilizes the surfaces, since the surface energies decreases when hydrated. Molecular dynamics (MD) was used to evaluate the effect of temperature on the bulk structure.

1. Introduction

Many properties of materials are found to originate from atomic level physical factors; hence it is necessary to understand the behaviour of these materials at the atomic level. As such, atomistic simulations are breaking through as a helpful tool for analysing and predicting various fundamental static and dynamic properties of materials [1]. Iron sulphide minerals exist in a variety of forms with varying stoichiometries that range from the sulfur deficient mackinawite (FeS_{1-x}) through iron deficient pyrrhotites ($Fe_{1-x}S$) to pyrite (FeS_2) and they are common in nature [2]. Among the iron-based sulphides, FeS_2 is the most abundant natural mineral, and it is available in two closely related polymorphic structures, namely, pyrite and marcasite [3]. The aim of this study is to determine the structure and dynamics of bulk and surfaces of pyrite FeS_2 using the modified interatomic potentials for atomistic simulations (i.e. energy minimization and molecular dynamics). The original potentials were first derived by de Leeuw and co-workers [4]. The modified interatomic potentials of pyrite have been able to improve some of our parameters, e.g. S-S bond length, from the original potentials. They have also improved the stability of the most stable surface, {100} of pyrite. The original potentials gave the surface energy of 1.23 J/m^2 while the modified potentials gave 1.04 J/m^2, which improves the agreement with density functional theory (DFT) calculations. Energy minimization was used to calculate the surface energies of the low index surfaces. MD was used to identify the melting temperature of the bulk structure. We also study the effect of water molecules on the mineral surfaces. Water on mineral surfaces plays an important role in nature, which include mineral weathering and corrosion. The properties of mineral surfaces, i.e. structure, surface energy and reactivity, change when in contact with water.

1.1. FeS2 (Pyrite) structure

Pyrite has a cubic structure and it belongs to the space group Pa3 [5]. The octahedrally coordinated Fe atoms are at the corners and face centers of the cube unit cell [6], and S atoms are arranged in 'dumb-bell' pairs centered at the mid-points of the cube edges and at the cube's body center, the structure of pyrite shown in Figure 1.

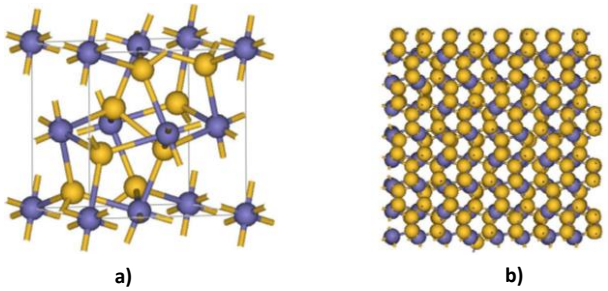


Figure 1. Snapshots of a) cubic bulk Pyrite and b) 4x4x4 supercell of bulk pyrite used in the MD calculations, purple atoms represent iron (Fe) and yellow atoms represent sulphur (S).

The four S-S joins are respectively parallel to four non-intersecting body diagonal directions. The cubic pyrite structure can be considered as a derivative of the face-centered rock salt structure, NaCl, where $(S_2)^{-2}$ occupies the Cl sites and Fe^{2+} occupies Na sites [7].

2. Methodology

This study is based on atomistic simulation of the iron sulphide. We use molecular dynamics (MD), which is coupled with the modified interatomic potentials. The potentials were modified using the GULP [8] code. In our MD calculations we used the MD code called DL_POLY [9]. The NPT and NVT ensembles were used and we considered cubic supercell consisting of 4x4x4 unit cells, which is equivalent to 768 atoms (shown in Figure 1b). The supercell used gives an approximately cubic simulation box of: 21.75 x 21.75 x 21.75 (in Å). The constant temperature and volume simulations were performed over the temperature range of 300 K to 1500 K with 200 K increments at zero pressure. The structure of the bulk and surfaces will be determined at different temperatures from the radial distribution functions (RDFs), diffusion coefficients. With these parameters we are able to determine the melting temperature of the bulk and the surfaces. The Born ionic model [10] was used and parameters were derived for short range interactions represented by the Buckingham potential, harmonic function and three body terms:

2.1. Buckingham Potential

In the Buckingham potential, the repulsive term is replaced by an exponential term and potential takes the form

$$U(r_{ij}) = A_{ij} * \exp^{-r_{ij}/\rho_{ij}} - \frac{C_{ij}}{r_{ij}^6} \qquad (1)$$

where A_{ij} and ρ_{ij} are parameters that represent the ion size and hardness, respectively, while C_{ij} describe the attractive interaction and r_{ij} is the distance between ion i and ion j. The first term is known as the Born-Mayer potential and the attraction term (second term) was later added to form the Buckingham potential. Very often, for the cation-anion interactions, the attractive term is ignored due to the very small contribution of this term to the short-range potential, or, alternatively, the interaction is subsumed into the A and ρ parameters.

2.2. Harmonic Potential

The interaction between the sulphur atoms of the S-S pair were described by a simple bond harmonic function:

$$U(r_{ij}) = \frac{1}{2} k_{ij} \left(r_{ij} - r_0\right)^2 \quad (2)$$

where k_{ij} is the bond force constant, r_{ij} the interionic separation and r_0 the separation at equilibrium.

2.3. Three-Body Potential

A further component of the interactions of covalent species is the bond-bending term, which is added to take into account the energy penalty for deviations from the equilibrium value. Hence, this potential describes the directionality of the bonds and has a simple harmonic form:

$$U(\theta_{ijk}) = \frac{1}{2} k_{ijk} \left(\theta_{ijk} - \theta_0\right)^2 \quad (3)$$

where k_{ijk} is the three-body force constant, θ_0 is equilibrium angle and θ_{ijk} is the angle between two interatomic vectors $i - j$ and $i - k$.

3. Results and Discussions

The interatomic potentials were modified, so that they can handle other parameters like elastic constants and most importantly the S-S intramolecular distance well. The properties, i.e. lattice parameters and elastic properties to validate the modified potential model of FeS_2 are discussed and also molecular dynamics simulation of bulk FeS_2 and the most stable surface at high temperatures are discussed.

3.1. Bulk FeS_2 (Pyrite)

To test our modified potential model we start by showing the calculated properties that were obtained using energy minimisation code GULP i.e., lattice parameters, bond lengths and elastic constants

Table 1. The calculated lattice constants, bond lengths and elastic constants in our work, compared with the calculated and experimental data in the given references.

Parameter	Modified Potential	Original Potential[a]	Calculated[b,c,d]	Experimental[e]
$a(\text{Å})$	5.385	5.357	5.370	5.418
$Vol(\text{Å}^3)$	156	154	155	159
Distances (Å)				
S – S	2.180	1.838	2.196	2.177
Fe – S	2.290	2.275	2.354	2.262
Fe – Fe	3.870	3.788	-	3.381
Elastic Properties (GPa)				
C_{11}	370	352	395	366
C_{12}	51	56	30	47
C_{44}	99	122	119	105
Bulk Modulus (GPa)	157	155	151	153

[a][11], [b][12], [c][5], [d][13], [e][14]

Table 1 shows the calculated lattice constants, bond lengths and elastic constants of modified potentials compared with those of original potentials, experimental results and theoretical results from

literature. The comparison of the calculated properties with those obtained experimentally shows that the modified potential is better than the original potentials and other theoretical results. Therefore, the modified potential gives more reliable results.

3.2. Molecular dynamics simulation of bulk FeS_2 (Pyrite)

The modified interatomic potentials were further tested by predicting the melting temperature of bulk structure using molecular dynamics technique. The melting point of the bulk structure is obtained from the radial distribution functions, structural changes and diffusion coefficients at various temperatures. Figure 2 shows the RDFs of bulk FeS_2 of the interactions Fe-S and S-S at different temperatures. From the radial distribution function plots we observe that at lower temperatures of 300 K, 700 K and 1100 K there are sharp peaks, implying a well-defined structure.

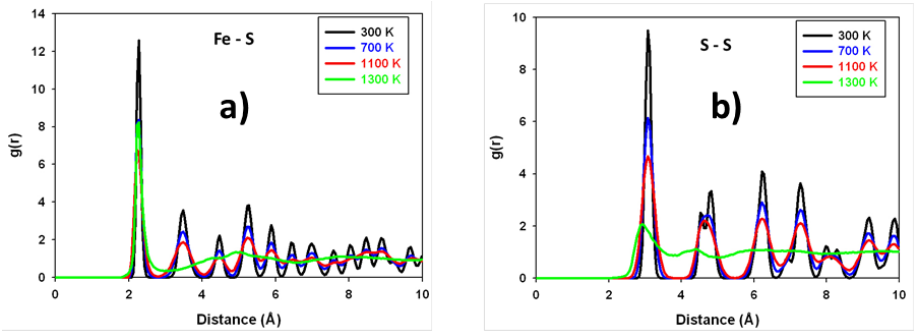


Figure 2. Radial distribution functions of pyrite FeS_2 for the a) Fe-S and b) S-S bonds at various temperatures.

However, as the temperature increased to 1300 K the peaks become broader and their number decreases, showing that the structure is experiencing the phase transition from a solid phase to a liquid phase. This is indicative of the melting temperature of bulk structure at 1300 K, which is not far from the experimental temperature of between 1450 K and 1461 K [15].

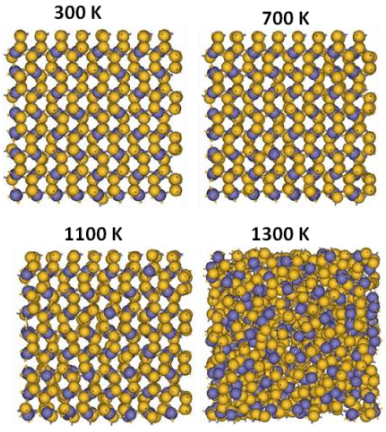


Figure 3. Snapshots of the 4x4x4 supercell of pyrite at different temperatures, as indicated.

In addition, figure 3 shows the snapshots for the 4x4x4 supercell structure of FeS_2 at different temperatures from MD simulations. From the pictures we observe that there is a phase transition from a lower temperature to a higher temperature. At 300 K, 700 K the structure is still intact and crystalline. However, as we increase the temperature to 1100 K, we observe that the structure loses its crystallinity, as the regular arrangement of atoms disappears. At 1300 K the arrangement of atoms has completely disappeared. This is in further support of the temperature proposition put forward by the

results of the RDFs in figure 2. The melting point of a simulated structure can also be located by increasing the temperature of a crystalline system until ion diffusion appears. In figure 4 we have plotted the diffusion coefficients of Fe and S ions as a function of temperature in the bulk structure. These ions start to diffuse at 1300 K, which is the melting point of the bulk structure deduced from the RDFs in figure 2.

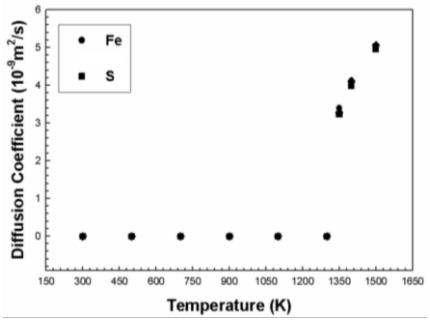


Figure 4. Diffusion coefficients of Fe and S ions as function of temperature for bulk structure of pyrite FeS_2.

3.3. Surface Energies of FeS_2 (Pyrite)

The modified potentials were also used for the calculations of surface energies of pyrite FeS_2. We calculated the surface energies from energy minimization. The surface energy is defined as the excess energy, per unit area of the surface compared to the bulk crystal. Thus any calculation of the surface energy also requires that we model a comparable amount of bulk crystal under identical condition.

Table 2. Calculated surface energies (J/m^2) compared with the original potentials, and previous DFT calculations.

Pure Surface	Modified Potentials	Original Potentials[a]	DFT[b]	DFT[c]
{100}	1.04	1.23	1.06	1.04
{110}	1.68	2.36	1.68	1.72
{111}	3.38	3.92	1.40	1.43
{210}	1.62	1.99	1.50	1.49

[a][4] [b][16] [c][17]

Table 2 gives the calculated surface energies using modified potential model compared with the original potentials, and previous DFT calculations. The modified potentials improve the stability of the surfaces, since it gives the lowest surface value. Amongst the high symmetry surfaces {100}, {111}, and {110}, {210} atomistic surface energy calculations proposed the {100} surface of FeS_2 as the most stable in agreement with those from DFT calculations.

Table 3. Calculated hydrated surface energies (in J/m^2) compared with those of calculated by original potentials.

Hydrated Surface	Modified Potentials	Original Potentials[a]
{100}	0.96	1.21
{110}	1.06	2.08
{111}	2.91	3.49
{210}	1.49	1.77

[a][4]

Table 3 shows the hydrated surface energies of low index surfaces of FeS_2. A comparison with the dry surfaces (Table 2) clearly depicts that hydration lowers surface energies; consequently the surfaces are more stable when hydrated. The {100} surface is still the most stable after hydration.

4. Conclusion

The modified interatomic potentials were used for atomistic simulations to study the surfaces and the bulk structure of pyrite. The simulated results of bulk FeS_2 compared well with experimental observations and other calculations reported in the literature. The surface energies of the miller indexes revealed that {100} is the most stable surface, with a lowest surface value for both pure and hydrated surfaces. When we compared the surface energies calculated from the original potentials and the modified potentials, it is clear that the modified potentials improve the stability of the surfaces by generating the lowest surface values. MD simulation predicted the melting temperature of the bulk structure at 1300 K, which is not far from the experimental. Therefore, the modified potential can be used reliably to investigate bulk properties, surface energies and even to investigate the effect of temperature.

Acknowledgements
The computations were performed at the Materials Modelling Centre (MMC), University of Limpopo. We also acknowledge the National Research Foundation (NRF) for funding.

References

[1] Lee B, Ko W, Kim H and Kim E 2010 *Calphad.* **34** 510
[2] Spagnoli D, Refson K, Wright K, and Gale J D 2010 *Phys. Rev. B* **81** 094106.
[3] Gudelli V K, Kanchana V, Appalakondaiah S, Vaitheeswaran G and Valsakumar M C 2013 *J. Phys. Chem. C* **117** 2120
[4] De Leeuw N H, Parker S C, Sithole H M and Ngoepe P E 2000 *J. Phys. Chem. B* **104** 7969
[5] Liu S, Li Y, Yang J, Tian H, Zhu B and Shi Y 2014 *Phys. Chem. Miner*. **41** 189
[6] Lennie A R and Vaughan D J 1992 *Am. Mineral.* **77** 1166
[7] Deer W A, Howie R A, and Zussman J 1992 *An introduction to the Rock-forming Minerals 2nd edition.* New York:Wiley
[8] Rohl A L and Gale J D 2003 *Mol. Simul.* **5** 291
[9] Smith W and Forester T R 1996 *J. Mol. Graphics* **14** 136
[10] Born M and Huang K 1954 *Dynamical Theory of Crystal Lattices 1st edition.* Oxford: University Press
[11] Sithole H M 2000 *PhD Thesis.* University of the North (now University of Limpopo)
[12] Reich M and Becker U 2006 *Chem. Geol.* **225** 278
[13] Muscat J, Hung A, Russo S and Yarovsky I 2002 *Phys. Rev. B* **65** 054107
[14] Benbattouche N, Saunders G A, Lambso E F and Honle W J 1989 *J. Phys. D: Appl. Phys* **22** 670
[15] *http://www.newworldencyclopedia.org/entry/Pyrite*. (n.d.). Retrieved May 2015
[16] Hung A, Muscat J, Yarovsky I and Russo S P 2002 *Surf. Sci.* **513** 511
[17] Sun R, Chan M K Y and Ceder G 2011 *Phys. Rev. B* **83** 235311

Néel temperature dependence of Cr + 1 at.% Al thin films on MgO(110), MgO(100) and fused silica

ZP Mudau[1], ARE Prinsloo[1], CJ Sheppard[1], AM Venter[2] and EE Fullerton[3]

[1]Department of Physics, University of Johannesburg, PO Box 524, Auckland Park, 2006, South Africa
[2]Research and Development Division, Necsa Limited SOC, P.O. Box 582, Pretoria 0001, South Africa
[3]Center 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. We have investigated the magnetic and transport properties $Cr_{99}Al_1$ films of thickness (t) from 28 to 450 nm deposited on fused silica, MgO(100) and MgO(110) by DC magnetron sputtering. The x-ray diffraction results were used to determine the crystallographic orientation of the thin films and the Néel transition temperature, T_N, was obtained using standard four-point probe resistance (R) measurements as function of temperature (T). Results show polycrystalline structure for sample prepared on fused silica while those on MgO(100) and MgO(110) shows good epitaxial growth. For samples deposited on fused silica no anomalies associated with the T_N were observed in $R(T)$. The $R(T)$ curves for films deposited on MgO substrates showed anomalies associated with T_N shaped as weak domes. The T_N values decreased with increase in t, levelling off at approximately 260 K for the 452 nm sample. This correlates with the T_N value expected from the magnetic phase diagram of bulk Cr-Al for this concentration.

1. Introduction

Chromium is an itinerant-electron antiferromagnetic (AFM) element which has an incommensurate spin density wave (ISDW) below the Néel transition temperature (T_N) of 311 K [1]. The unique characteristic of the spin density wave (SDW) is a result of the nesting between the electron and hole Fermi surfaces [1]. Cr alloys weakly diluted with metals retain their antiferromagnetism, but can have magnetic properties different to that of pure Cr. This is strongly dependant on the nature and concentration of the dilute [1]. Cr has an electron-to-atom (e/a) ratio of six. Alloying Cr with metals with different e/a values, affects the relative sizes of the electron and hole Fermi surfaces which in turn influences their nesting in the AFM phase. As the AFM nature of the Cr alloys are directly linked to the nesting of the Fermi surfaces, the effects are clearly seen in a variety of physical properties [1]. Alloying Cr with elements with an e/a less than six, such as V (with an e/a = 5), normally results in a decrease in T_N with the increase in diluent concentration, while alloying Cr with elements with e/a larger than six such as Mn

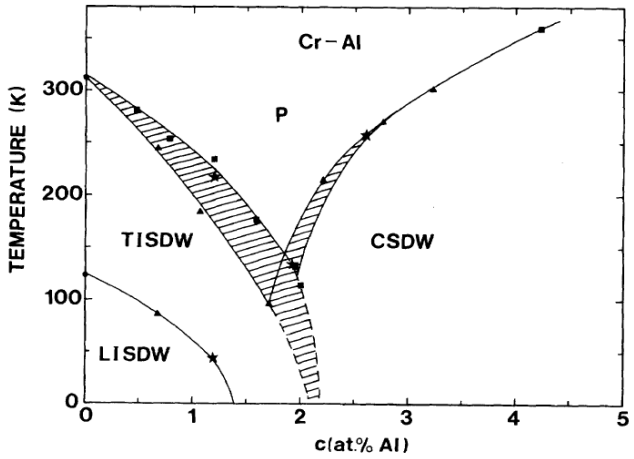


Figure 1: Concentration-Temperature (*c*-*T*) magnetic phase diagram of bulk Cr-Al alloy system that summarises the existence of three different spin density wave phases at a triple point around 2 at. % Al and 130 K, where the incommensurate spin density wave (ISDW), commensurate spin density wave (CSDW) and paramagnetic (P) phases coexist [3]. TISDW indicates the existence of a transverse polarised ISDW phase and LISDW a longitudinal polarised ISDW phase.

(with an $e/a = 7$) normally results in an increased T_N with increase in diluent concentration [1]. Cr-Al is an exception to this general rule. Aluminium (with $e/a = 3$) does not give a continuous decrease of T_N with increase in dilute concentration as expected, but instead the magnetic phase diagram shows a deep minimum around 2 at.% Al after which T_N increases sharply with further increase in concentration, as indicated in Figure 1. The origin of this unique behaviour in Cr-Al has been the subject of many studies [1, 2], but no conclusion has been reached in this regard.

At present there is no quantitative agreement in the literature concerning the position and the depth of the minimum [2, 3]. The *c*–*T* magnetic phase diagram also reveals the existence of three spin density wave phases converging to a triple point around 2 at.% Al and 130 K where all three phases coexist [2, 3]. The phase diagram shows the presence of an ISDW structure at 1 at.% Al, with a transition from longitudinal to transverse ISDW at a temperature of approximately 60 K [3]. In addition a recent study suggests that the triple point occurs at 2.2 at.% Al and 0 K; and most interestingly implies that this triple point may also be a quantum critical point [5].

Thin films of Cr alloys have shown unique properties not observed in bulk material [4]. Their magnetic properties are strongly influenced by the film thickness, surface and proximity magnetic effects. This study now extends such investigations to include $Cr_{99}Al_1$ thin films of different thicknesses deposited on different substrates.

2. Experimental

The Cr-Al thin films were prepared using DC magnetron sputtering at a substrate temperature of 973 K and argon pressure 3 mTorr, from elemental sources onto single-crystal MgO(110) and MgO(100), as well as fused silica substrates. The sample was prepared at a fixed concentration of Cr + 1 at.% Al with thicknesses from 28 to 452 nm. The film thickness was controlled through varying deposition rates and times and confirmed using x-ray diffraction (XRD) techniques. In order to investigate the growth orientation of the thin films symmetrical θ-2θ XRD was done using the Bragg-Brentano configuration set-up on Phillips X'Pert Pro Diffractometer. Electrical resistance measurements from 77 to 400 K using the standard DC four-probe method were employed to determine the Néel transition temperatures of these samples.

3. Results

Figures 2(a) to (c) summarise the representative XRD results for the $Cr_{99}Al_1$ thin films deposited on single crystalline MgO(110) with t = 452 nm, MgO(100) with t = 226 nm and polycrystalline fused silica with t = 226 nm, respectively. The XRD results indicate that all the samples prepared on MgO(100) and MgO(110) are epitaxial and exhibit a single crystallographic orientation, with the $Cr_{99}Al_1$ layers showing preferred growth direction of (100) and (211). Thin films prepared on fused silica substrates have a polycrystalline structure with no preferred orientation. The Néel temperature (T_N) of the films were determined from anomalies observed in the electrical resistance (R) measurements as a function of temperature (T). Representative R-T curves are shown in Figure 3 for samples deposited on (a) MgO(110) with t = 226 nm, (b) MgO(100) with t =226 nm and (c) fused silica with t = 226 nm.

Resistivity results of thin films prepared on MgO(100) and MgO(110) substrates show anomalies in the shape of weak domes that are ascribed to the formation of the ISDW on cooling down from the paramagnetic phase. T_N is usually taken at the inflection point of $R - T$ curves, but since the anomaly is relatively weak in the $Cr_{99}Al_1$ thin films, this point is determined from the dR/dT versus T curve which is calculated from the $R - T$ data. An example of such a curve is shown in Figure 3 (d) and T_N is taken at the temperature associated with the minimum. This is the standard method used to determine the T_N of Cr thin films [8]. However, it should be mentioned that the T_N values obtained maybe biased to some extent due to the extrinsic morphology contribution that dramatically changes the resistance [8]. No evidence is seen of the TISDW to LISDW phase transition in the R-T curves at low temperatures [1, 2], as was observed in bulk Cr-Al single crystals and reflected in Figure 1.

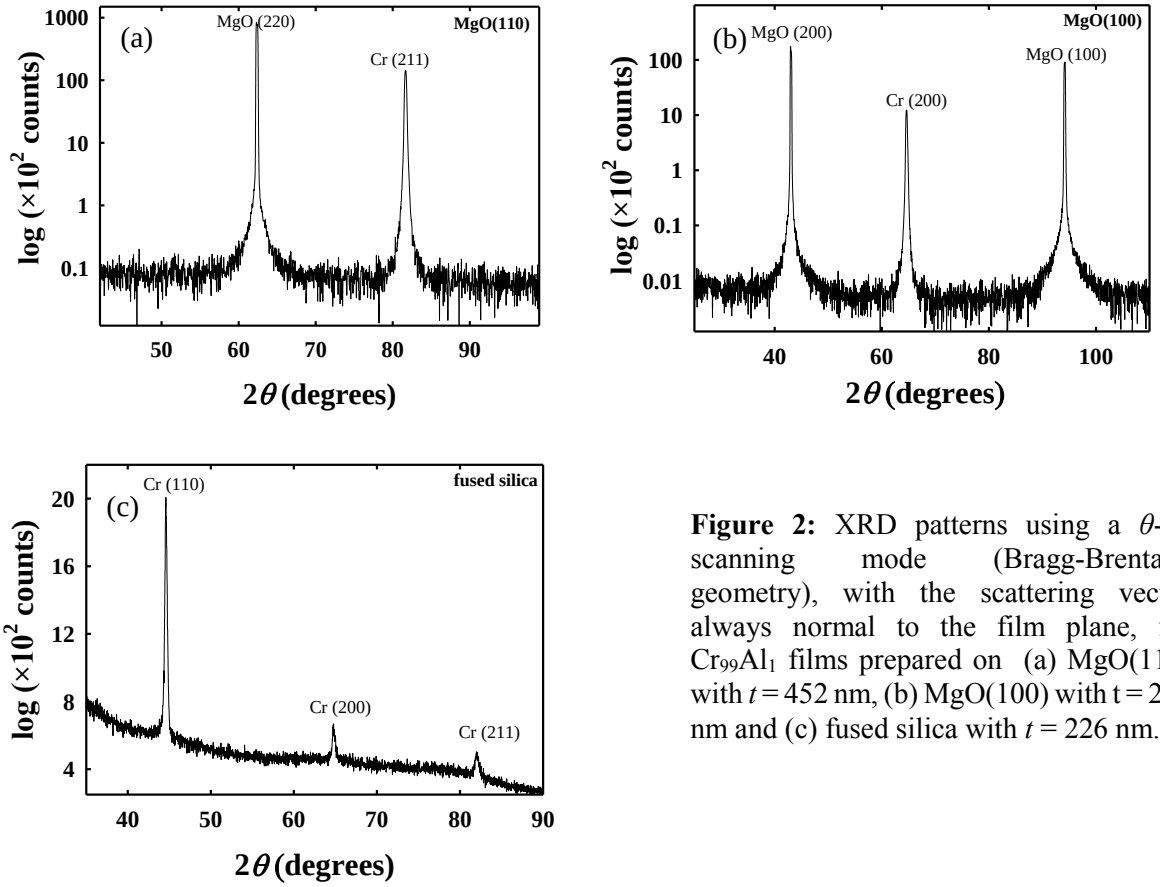


Figure 2: XRD patterns using a θ-2θ scanning mode (Bragg-Brentano geometry), with the scattering vector always normal to the film plane, for $Cr_{99}Al_1$ films prepared on (a) MgO(110) with t = 452 nm, (b) MgO(100) with t = 226 nm and (c) fused silica with t = 226 nm.

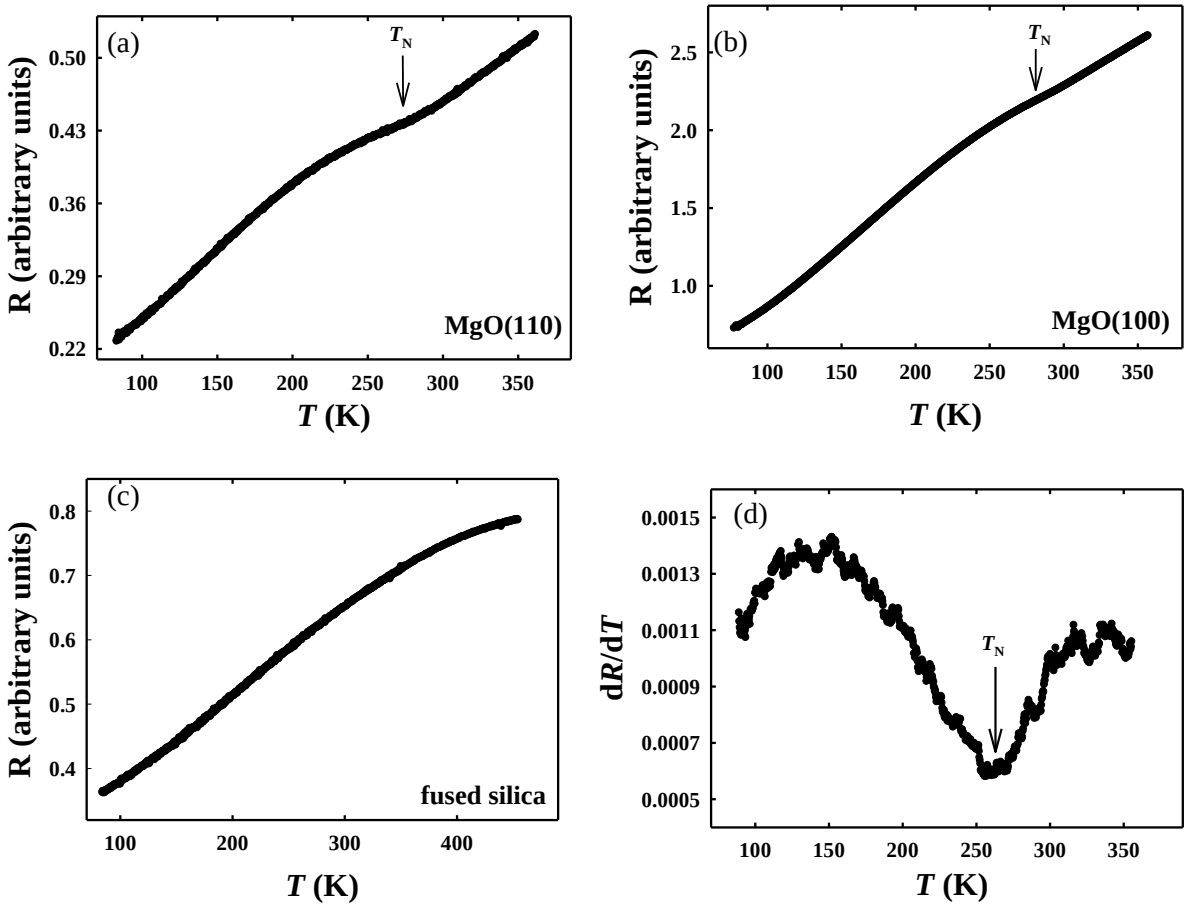


Figure 3: Resistance (R) versus temperature (T) graphs for $Cr_{99}Al_1$ thin films prepared on (a) MgO(110) with t = 226 nm, (b) MgO(100) with t = 46.5 nm, (c) fused silica with t = 226 nm and (d) show dR/dT versus T for the MgO(110) with t = 226 nm sample. The temperature associated with the minimum in the dR/dT versus T curves was taken as the T_N.

For $Cr_{99}Al_1$ thin films deposited on fused silica no anomalies were observed in the temperature range of 77 to 400 K. The T_N seemed to be higher than 400 K, from Figure 3 (c) it is clear there is a sudden increase of R, but the inflection point where it started to increase it is not observed in the measured temperature range. Measurements at higher temperatures were not considered as the experimental presently available in the laboratories did not alloy for measurements under high vacuum and this can result in a permanent physical damage of the thin films – in the form of sample peeling and oxidation.

Figure 4 shows the variation of T_N as a function of thickness (t) for the samples prepared on MgO(110) and MgO(100) substrates. For epitaxial thin films prepared on MgO(110), there is a slight increase in T_N with the increase in thickness, reaching a maximum around 280 K, followed by a decrease and then levelling off at approximately 260 K as shown in Figure 4 (a). For the $Cr_{99}Al_1$ epitaxial thin films prepared on MgO(100), the T_N is decreases with increase in t from around 290 K, giving rise to a minimum at t = 80 nm with the lowest T_N value of 245 K, finally levelling off at approximately 260 K for the thickest samples. The T_N values obtained for the $Cr_{99}Al_1$ thin films with t > 200 nm on MgO(110) and MgO(100) substrates is in agreement with what is expected from the c-T magnetic phase diagram

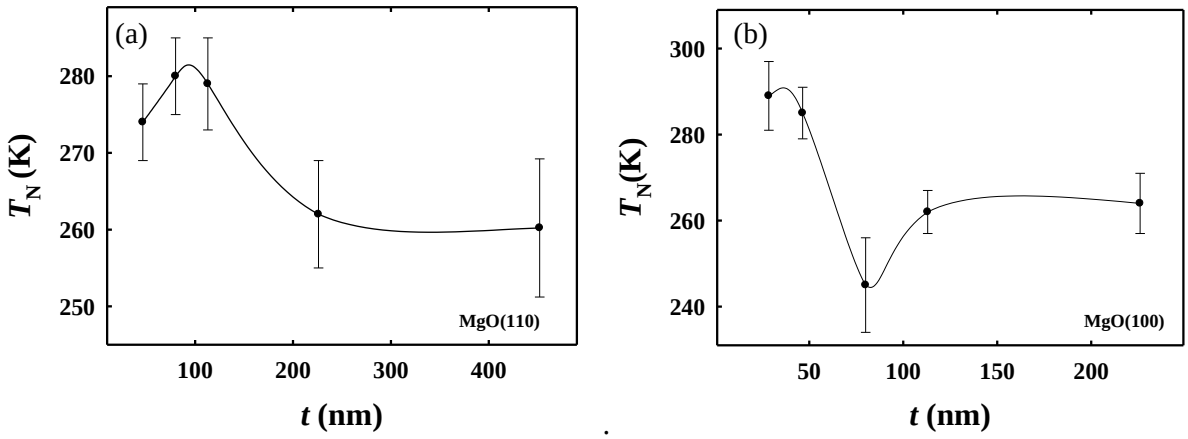


Figure 4: T_N versus $Cr_{99}Al_1$ layer thickness (t) as prepared on the substrates (a) MgO(110) and (b) MgO(100).

of bulk Cr-Al. The *c*-*T* magnetic phase diagram shows T_N to be around 240 to 260 K for bulk Cr + 1 at.% Al.

It is noted that the behaviour of $T_N - t$ for the $Cr_{99}Al_1$ sample series prepared on MgO (110) and MgO(100) differs slightly as the T_N values levels off at $t > 200$ nm for the MgO(110) series, but this levelling off already happens at $t > 100$ nm for the MgO(100) series. This might be attributed to the different crystallographic orientations obtained for the $Cr_{99}Al_1$ thin films prepared on the various substrates [4]. Previous studies also linked growth morphology to the behaviour of $T_N - t$ [6], this can be probed in further studies using AFM. Present results indicate that dimensionality strongly influence the T_N values in $Cr_{99}Al_1$. It is evident that the T_N values for samples prepared on MgO(110) is higher in the samples with $t < 200$ nm than the values reflected in phase diagram shown in Figure 1 for bulk Cr-Al. This is also the case for the samples with $t < 100$ nm prepared on MgO(100). This can be attributed to higher stress and strain in the thinner films [7]. For the thicker films the values of T_N is in agreement with those reflected in Figure 1, obtained for bulk $Cr_{99}Al_1$, as expected [4].

The proposal that the T_N values for the $Cr_{99}Al_1$ samples prepared on the fused silica substrates is above 400 K (and thus outside the measuring capabilities of the present experimental set-up), is in line with observations in previously reported [5, 6, 7, 9]. In studies [5, 6] performed on Cr and Cr-Ru thin films, the thin films deposited on fused silica also showed higher T_N values compared to samples deposited on single crystalline MgO(110) and MgO(100) substrates. These conclusions were confirmed in a later publication by Mudau *et al*. [7] where the residual stresses in Cr thin film samples used by Sheppard *et al.* [5] were investigated. The investigation revealed that the samples with enhanced T_N values were under tensile stresses of the order of 1 GPa [7]. Previous investigations [9] on polycrystalline Cr films, grown on Corning glass substrates, also show significant enhancement, up to 60 K above the bulk value, in T_N for $t \leq 30$ nm. This is ascribed to internal tensile stresses in the films that arise, in part, from the grain boundaries [10, 11]. It is therefore probable to assume that the higher T_N values observed in the $Cr_{99}Al_1$ thin films prepared on fused silica substates could also be atributed to higher internal stress/strains in the coatings and will be the subject of a future study.

The present study therefore indicate that the crystallographic orientation influence the directional transport properties, as well as the T_N values of $Cr_{99}Al_1$ thin films of various thicknesses. This is in line with results obtained in previous studies on thin magnetic films [4, 5, 6, 7, 8].

4. Conclusions

XRD investigations of $Cr_{99}Al_1$ thin films prepared on MgO(110), MgO(100) and fused silica substrates, with thickness 28.5 to 452 nm have been successful for structural characterization. The results indicated that the samples deposited on single crystalline substrates MgO(110) and MgO(100) are epitaxial and exhibit preferred growth direction, where-as those deposited on fused silica are polycrystalline. The T_N as function of t graphs for samples prepared on MgO(100) and MgO(110) show a decrease in the magnetic transition temperature with increase in thickness and the T_N values for the thickest samples being in good agreement with those found in bulk $Cr_{99}Al_1$. In the case of the fused silica samples, no T_N transitions were observed in the temperature range assessable with the equipment used. It is proposed that the higher T_N values in these samples are due to substantial tensile stresses in the coatings.

5. Acknowledgements

The authors wish to thank the National Research Funding of South Africa (Grant numbers 80928; 80631 and 93551) and the Faculty of Science of the University of Johannesburg for financial support. The use of the analytical facility at UJ, Spectrum, is acknowledged. Necsa Limited SOC is acknowledged for the use of the XRD facilities within their User Program.

References

[1] Fawcett E, Alberts HL, Galkin VY, Noakes DR and Yakhmi JV 1994 *Rev. Mod. Phys.* **66** 25
[2] Sheppard CJ, Prinsloo ARE, Alberts HL, Muchono B and Strydom AM 2014 *J. Alloys Compd.* **595** 164
[3] Baran A, Alberts HL, Strydom AM and P de V. du Plessis 1992 *Phys. Rev. B* **45** 10473
[4] Zabel H 1999 *J. Phys. Condens. Matter* **11** 9303
[5] Sheppard CJ, Prinsloo ARE, Fernando R R, Mudau ZP, Venter AM and Fullerton EE 2013 *SAIP Conf. Preceedings* 146
[6] Prinsloo ARE, Derrett HA, Hellwig O, Fullerton EE, Alberts HL and Van den Berg N 2010 *J. Magn. Magn. Mat.* **322** 1126
[7] Mudau ZP, Prinsloo ARE, Sheppard CJ, Venter AM, Ntsoane TP and Fullerton EE 2014 *SAIP Conf. Preceedings* 91
[8] Mattson JE, Fullerton EE, Sowers CH and Bader SD 1995 *J. Vac. Sci. Technol. A* **13**(2) 276
[9] Lourens JAJ, Arajs S, Helbig HF, Cherlet L and Mehanna EA 1988 *J. Appl. Phys.* **63** 4282
[10] Windischmann H 1992 *Crit. Rev. Solid State Mater. Sci.* **17** 547
[11] Boekelheide Z, Helgren E and Hellman F 2007 *Phys. Rev. B* **76** 224429.

Characterisation of traditional ceramic materials used in the Sotho culture (South Africa) for clay pot making

A F. Mulaba-Bafubiandi and P X D A Hlekane
Mineral Processing and Technology Research Centre, Department of Metallurgy, School of Mining, Metallurgy and Chemical Engineering, Faculty of Engineering and the Built Environment, University of Johannesburg

Email:amulaba@uj.ac.za

Abstract. Different types of clayey soils are readily available as a natural resource in South Africa. Some are used for brick making, some for eating in a geophagia (habit) while some others are used as cosmetic ingredients in cultural ceremonies or in traditional ceramic applications. Among the above mentioned numerous utilisations, clay pot making and trade are claimed to contribute to the improvement of the household income generation in the rural communities. Traditional clay pots are made in Africa by people from different cultures. This study focuses on the suitability of clayey soils used for clay pot production as traditionally made by the Sotho people of the Free State Province of South Africa. Clayey soil raw materials are collected from the local river banks and processed through a shaping, sun drying and firing set of subsequent processes. The knowledge of mineralogical composition, mineral phases formed during the processing of clays and the mastering of physio-chemical properties including plasticity helped to understand the thermal properties of the processed material generated, their forming and shaping, and the application of clay pots produced. Clayey soils, the main raw materials used, were randomly collected from the QwaQwa region of the Free State province and analysed using XRF, XRD, and FTIR. The clayey soil showed plastic limit of 34.36%.

1. Introduction

Ceramics have found large applications in prehistoric and in modern times. Clays are natural materials abundantly found and largely used in the manufacturing of ceramic products such as bricks, porcelain, sanitary ware, floor and roofing tiles and also various industrial applications [8]. With these vast numbers of usages, this paper aims to provide an insight into the characterisation of clayey soils used to make clay pots in QwaQwa, Free State province of South Africa. The sample was received from the Homelands of the Basotho people which lies in the heart of the Karoo sequence rocks, containing mudstones, shales, sandstones and the Drakensberg Basalts forming the youngest capping rocks in a province that is high lying with almost all the land being 1000m above the sea level [7]. In general clay can also be referred to as natural earthy fine grained materials which are powdery when dry, plastic when wet and stone like when baked [2]. In general, clay belongs to a wider group of minerals, however all clay minerals are classified as hydrous aluminium phyllosilicates group with a highly complex structures which are basically characterized by layers where each layer is composed of two types of structural sheets: an octahedral and a tetrahedral. The tetrahedral sheet is composed of silicon-oxygen tetrahedra linked to neighbouring tetrahedra by sharing three corners which is resulting in a hexagonal network on one side and on the other side, the remaining fourth corner of each tetrahedron forms a part to adjacent octahedral sheet. Usually the octahedral sheet is composed of aluminium or magnesium in six-fold coordination with oxygen from the tetrahedral sheet and with hydroxyl. The two sheets together form a layer, and several layers may be joined in a clay crystallite by interlayer cations, Vander Waals force, electrostatic force, or by hydrogen bonding. Based on the clay structure (Figure 1), classification

is according to the arrangement of tetrahedral and octahedral sheets. Therefore, it can be 1:1 when the clay mineral has one tetrahedral and one octahedral sheet per clay layer; 2:1 when the clay mineral contains two tetrahedral sheets and one octahedral sheet taken in sandwich between the two tetrahedral sheets; and 2:1:1 when the clay minerals are composed of an octahedral sheet adjacent to a 2:1 layer. Table 1 summarises the classification of clay mineral group in connection to their structures.

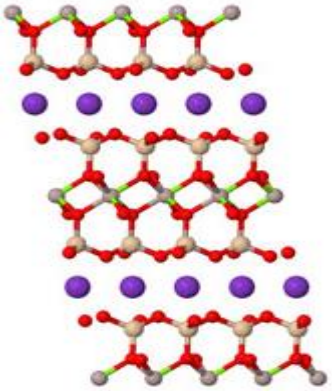


Figure 1. Structure of clay minerals showing tetrahedral and octahedral sheets.
This is a tri octahedral 2:1 [13].

Table 1. Sheet structure for the layers of clay minerals.

Clay mineral	Layer type
Kaolinite	1:1
Montmorillonite Or smectite	2:1
Illite	2:1
Vermiculite	2:1
Chlorite	2:1

Consequently, with respect to variety of the chemical species present in the clay and their arrangement in the clay structure, clay minerals may be divided into four major groups which are mainly in terms of the variation in the layered structure. These include the kaolinite group, the montmorillonite / smectite group, the illite group, with their related associated minerals and general formula as presented in Table 2. In the paper we interchangeably use bentonite, montmorillonite and smectite. Clays and their associated soils have a wide variety of physical characteristics such as plasticity, shrinkage under firing, under air drying, fitness of grain, colour after firing, hardness cohesion and capacity of the surface to take decoration. In general the plasticity is defined by means of its plastic limit and its liquid limit which are defined as the moisture content at which the soil passes from the friable to the plastic state and the moisture content at which the soil passes from the plastic to the liquid state respectively [4]. This study sets to investigate the physical, chemical and mineralogical properties of the clayey soils using x-ray diffraction (XRD), x-ray fluorescence (XRF) and Fourier-transform infrared spectroscopy (FTIR).

Table 2. Major groups of clay minerals [6].

Group name	Member minerals	General formula	Indications
kaolinite	kaolinite, dickite, nacrite	$Al_2Si_2O_5(OH)_4$	members are polymorphs (composed of the same formula and different structure)
montmorillonite or smectite	montmorillonite, pyrophyllite, talc, vermiculite, glauconite, saponite, nontronite	$(Ca,Na,H)(Al,Mg,Fe,Zn)_2(Si,Al)_4O_{10}(OH)_2\text{-}xH_2O$	X indicates varying level of water in mineral type
illite	illite	$(K,H)Al_2(Si,Al)_4O_{10}(OH)_2xH_2O$	X indicates varying level of water in mineral type

2. Materials and Methods

Clayey soils were received from the potter who collected them from the river banks of Phutaditshaba, QwaQwa (Figure 2). The sample was then oven dried at 105° for 24 hours. The dried sample was crushed and sieved using ASTM sieve. Materials > 45µm were utilised. Three spectroscopic techniques (XRF, XRD, and FTIR) were used. The chemical composition of clayey soils used in this work represented was studied using the X-ray fluorescence Rigaku Primus II. About 2g of the dry and pulverised sample used to determine the phase identification was studied using the XRD Rigaku Ultima IV. Vibration and stretching bands were determined using FTIR Thermo Scientific Nicolet iS10 with a spectral range of a wave number between 400 - 4000 cm^{-1}. The loss of ignition was calculated by placing 46.45g of the sample in a furnace at 1000°C for 2 hours. The weight of the dried sample was then measured as soon as the sample was removed from the furnace. This was done in order to determine the weight loss of the clayey soils. Plasticity of clayey soils was determined using the Atterberg limits. The ASTM procedure was used [1]. The plastic limit, liquid limit, moisture content and linear shrinkage were calculated from this experiment. The sample was placed in the liquid limit device, and 25 blows were counted in order to reach the point of contact for the clayey soils. The sample was collected at the junction point then put in the oven at 105° for 24 hours. This sample was then weighed, dried and the liquid limit was then measured

3. Results and Discussions

Figure 2. Map of South Africa showing the position of QwaQwa where clay pots are made

Figure 3. Photo showing a clay pot made in QwaQwa, Free State

QwaQwa clay pots generally have bright motives as seen on Figure 3.The dexterity of the potter is appreciated in order to produce such amenable and durable clay pots. Interviews conducted with the potters revealed that the exercise of clay pot making is an indigenous skill of the Sotho people and helps to create sustainable income for the potters from the rural lands of QwaQwa Free State. The knowledge and skill of pot making runs a risk of being lost because the current generation finds no interest in acquiring these pot making skills. The traditional uses of these pots include simple decoration in households, water storage and for everyday use of drinking water, grain storage, for drinking traditional beer at cultural events such as weddings, funerals and family gatherings.

The chemical composition of a soil type influences its technological properties. The results show major elements that include Al (24.33%) and Si (56.42%), leaving other minerals such as Fe (8.43%) and K (4.25%) in trace elements as per XRF results (Table 3). The large amount of silicate present in the clay is associated with the crystalline phase quartz which is combined to alumina in the alumina silicate structure [11]. The low Ca content (less than 5%) is an indication of non-calcareous clays therefore

limiting shrinkage. The presence of K, Ca, Mg and Ti help to melt silicates and bind particles together during firing. Aluminium oxide in kaolinite leads to increased plasticity [9]. The X-Ray Diffraction was used in order to explore the mineral assemblages of the clayey soil as depicted in Figure 4 with the presence of quartz, montmorillonite, illite and kaolinite. The semi-quantitative XRD denote 75% quartz, 4% illite, 5% Kaolinte.and 16% montmorillonite, these crystallite minerals are represented as peaks in figure 4. The high montmorillonite/kaolinite + quartz ratio of 3 makes this clayey soil usable in pot making as it has a higher plasticity index value of 64.85 which is in the range of a montmorillonite rich clay [1]. The higher amount of montmorillonite (16%) in the sample is then confirmed.The strength of the clay pot depends on the composition of the raw materials that are used. Quartz and montmorillonite are seen to be in abundance as compared to kaolinite and illite while other minerals are presented as trace minerals. The high presence of quartz further justified the high content of Si as seen in Table 3.

FTIR can determine many and very small mineral contents that are present within a sample. Vibration and stretching bands were identified as seen on Figure 5. Weak bands due to absorbed water are observed at wavenumbers 1633cm^{-1} and 1536cm^{-1}. Vibration bands were identified at different wavenumbers respectively, 993cm^{-1} corresponds to Si-O stretching band, 1416 cm^{-1} –OH-stretching band, 3693.63 cm^{1} –OH asymmetric band and weak band M-O at 688 cm^{-1} ,788 cm^{-1} and 910 cm^{-1} .C-H stretching vibrations which can be attributed to organic contribution are present at peaks 2848 cm^{-1} and 2919 cm^{-1}. All these peaks observed on Figure 5 were identified by comparing the observed wave numbers with available literature.

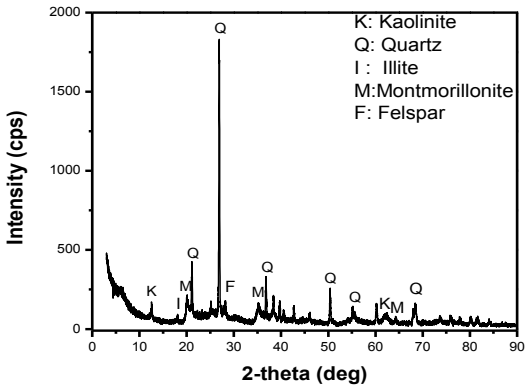


Figure 4. XRD spectrum of clayey soils showing their mineral contents .Smectite and kaolinite are observed.

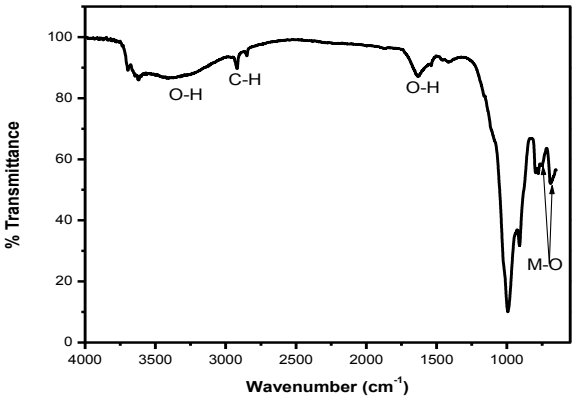


Figure 5. FTIR spectrum of pot clayey soils showing the vibration and stretching bands. O-H stretching band; C-H stretching band; Si-O stretching band.

Table 3. Main elements in the clayey soil used. XRF results

Element	Al	Ca	Fe	K	Mg	Mn	Na	P	Si	Ti	Zn
Mass%	24.33	2.58	8.43	4.25	1.38	0.04	0.60	0.06	56.42	1.38	0.02

In the present study, the plasticity was determined using the Atterberg limits, due to the useful information it provides on the plastic properties [1].The moisture content was calculated at 9.3 %, the high moisture content is associated with high plasticity, meanwhile a low moisture content will cause cracking during moulding. Because between 4-10% of linear shrinkage is within an acceptable range for clays, the QwaQwa pot clayey soils have a linear shrinkage calculated at 7.3% therefore classifying it as a good material to use that limits cracking and warping. Taking into consideration the heat factor during the firing process, heat is directly proportional to the shrinkage. The loss of ignition was calculated at 3.38%. The liquid limit is taken as the water content of the soil at which it will just begin to flow when jarred in a specific manner [1] .Liquid limit was calculated at 99.21% and plastic limit at

34.36%. The plastic index was therefore derived at 64.85%. When the plastic index is above 35, its high value is associated with low water permeability therefore making the QwaQwa pot clayey soils with 64.85% PI a good sealable product.. The property of plasticity is generally dependent on the mineralogical composition of the soil and smectite is generally sticky and plastic. The high plasticity of these clayey soils enables the potter to create different shapes in ceramic applications

4. Conclusion

The investigations performed on the clayey soils were done with an intention to understand the importance of the mineralogical and chemical composition of the clayey soils used to make clay pots. From this study it was obtained that clayey soils from QwaQwa are composed mainly of Si, Al and Fe and montmorillonite, kaolinite, illite and quartz as major minerals. The clayey soils were of excellent industrial quality for ceramic production making clay pots that were durable and suitable for diverse applications.

Acknowledgements

The authors wish to acknowledge the QwaQwa potters. The mineral processing research centre is acknowledged for the continuous support in the present research.

References

[1] Bain J.A. 1971 *Clay Minerals* **9** 1
[2] Bam S.A, G A. 2014. *Journal of Multidisciplinary Engineering and Technology* **5** 10
[3] Catuneau O, Wopfner H, Eriksson P.G, Cairncross B, Rubidge B.S, Smith R,M,H, & Hancox P.J. 2005 *Journal of African Earth Sciences* **211** 253
[4] Dumbleton M. & West G. 1966 *Clay Minerals* **6** 179
[5] Ekosse G.-I., de Jager L., & Ngole V. 2010 *Journal of Biotechnology*, **9** (47) 8058
[6] FAHEEM U. 2008 *The Minerals, Metals & Materials Society and ASM International,* 39A
[7] Johson M R, Van Vuuren C.J, Hegenberger W F, Key R, & Shoko U. 1996 *Journal of African Earth Sciences* **23** 3
[8] Mahanoharan C. S. 2012 *Journal Cerâmica* **58** 412
[9] Maria G M 2002 *Universitatis Babe¯***(**Bolyai,Geolgia) XLVII 1 77
[10] Murray H. 2007 *Applied Clay Mineralogy* (first edition ed. Netherlands)
[11] Nwajagu C O, Aneke L.E. 2001 *Journal of Nigerian Society of Chemical Engineering* ***5*** 1
[12] Sumner P D, Hall K J, van Rooy J L & Meik l. 2006. *Journal of African Earth Sciences* 236
[13] http://www.claysandminerals.com/minerals/clayminerals. [Accessed 3 June 2016]

Iron bearing minerals characterised with Mossbauer spectroscopy at the Mineral Processing and Technology Research Centre, University of Johannesburg, South Africa.

A F Mulaba – Bafubiandi
Mineral Processing and Technology research Centre, Metallurgy Department, Faculty of Engineering and The Built Environment, University of Johannesburg, Doornfontein campus, POBox 17011, Doornfontein, 2028, South Africa.

E-mail: amulaba@uj.ac.za

Abstract. Platinum group metals reserves have been mined in South Africa using the Merensky, UG2 and Plaatreef reefs. Recently the UG3 seams containing planar chromite platiniferous ore materials have been uncovered in the Uitkomst complex. The understanding of their mineralogy would inform on the processing and concentration routes to take leading to the recovery of associated base metals and to the production of platinum group elements therefrom. Base metals from such a feed are produced mainly through a dissolution process in acidic media. The knowledge of associated iron minerals informs on whether or not reductants or oxidants should be used before iron removal from the solution commonly by precipitation. ^{57}Fe Mossbauer spectroscopy has been extensively used at the mineral processing and technology research centre of the University of Johannesburg. Iron Mossbauer spectroscopy has been crucial in elucidating the presence of the two sites for Fe^{2+} and other two sites for Fe^{3+} in the flotation concentrate while only one site of Fe^{3+} and one site of Fe^{2+} were observed in the feeding ore. The existing gangue materials in the ore might have hindered the accessibility of the gamma by the probe. From the Mossbauer spectra obtained, one was able to determine the efficiency of the ferromanganese reduction process with coke and appreciate the effectiveness of the roasting of the nickel sulphide concentrate. An assessment of chromite types in the UG3 materials was achieved through then determination of the Fe^{3+}/Fe^{2+} ratios.

1. Introduction

Iron Mossbauer spectroscopy, a recoilless resonant emission and absorption of gamma rays where iron is used as a probe, has been used in earth science, materials development, materials characterisation and in monitoring of processes. Geology, chemistry, catalysis, materials engineering and science, even physics have been fields for its application. Up to the mid-eighties Gonser [1] extensively used the Mossbauer spectroscopy in physical metallurgy where phase transformations in steels and ferrous materials were investigated. The application of the Mossbauer spectroscopy in mineral processing or in metal extraction has been scarce. The Mineral Processing and Technology Research Centre from the University of Johannesburg was established in the late nineties for the training of postgraduate students and for service to the local mineral industry. In the past seventeen years, Mossbauer spectroscopy has been introduced to students at the metallurgy department where by-products generated during the drilling of hematite ore with tungsten carbide drill bits were simulated. Nanomaterials generated by mechanically alloyed minerals were also investigated. Since then the mineral processing and technology research centre has focussed its efforts on the mineral characterisation using Mossbauer spectroscopy in addition to the conventional x-ray diffraction, Fourier Transform Infrared spectroscopy, X-ray fluorescence and atomic absorption spectroscopy with the scanning electron microscopy.

2. Mossbauer spectroscopy in mineral processing and value recovery

As demonstrated in numerous technical confidential reports, published journal articles [2,3 and 4] and students dissertation/thesis, the effectiveness of the beneficiation and processing of minerals prior to the recovery or extraction of metals contained therein could be monitored with iron Mossbauer spectroscopy. Froth flotation of sulphide ores to produce concentrates or the conversion of sulphide minerals into their oxidised form through roasting could be followed by analysing the hyperfine interactions parameters of Mossbauer spectra of the crystalline as well as non-crystalline phases formed in the materials. Minerals enrichment through either magnetic separation or through gravity separation could be monitored. In collaboration with colleagues, it was interesting to see how Mossbauer spectroscopy could add value to the processing of locally found minerals in South Africa. Additionally, reported data on gold Mossbauer and on the application of Mossbauer spectroscopy to coal research are available [3, 6]. Platinum Mossbauer would be the future of Mossbauer spectroscopy in a mineral rich country such South Africa.

3. Materials and methods

Room temperature Mossbauer spectroscopy, in transmission mode with a ^{57}Co gamma rays source, was used. Locally found ores, concentrates or tailings were used while the efficiency of selected metallurgical processes was evaluated, and the quality of generated products was assessed using related hyperfine interactions parameters be it the isomeric shift, the quadrupole splitting or the magnetic field. Simulated plant treatment, processing and beneficiation processes of minerals were conducted at laboratory scale while some feeds were directly obtained from the production.

4. Minerals characterisation using Mossbauer spectroscopy at the Mineral Processing and Technology Research Centre.

Mossbauer spectroscopy has been used for the coal industry to identify and differentiate coal seam products as well as coal origins. The proportion of related coal iron bearing minerals (sulphate, pyrite or marcasite, clays) has been a decisive indicator [2] of a good coal in addition to the calorific contents, the ash content and the sulphur content. The coal abrasiveness index value should be related to the hyperfine interaction parameters. Mineral phase transformations occurring during the use of coal, gasification or pyrolysis have been investigated [2]. Mossbauer has also been used in the characterization of natural chromite samples from the Bushveld complex South Africa. The Fe^{3+} /Fe^{2+} ratios were determined. Two different types of chromites were identified and reported [3]. Type 1 chromites have low Fe^{3+}/Fe^{2+} ratios (between 0.2 and 0.4) and contains 3 cations for every 4 oxygen ions in the chemical formula while type 2 chromites have high Fe^{3+}/Fe^{2+} ratios (between 0.7 to 1.3) and contains less than 3 cations per oxygen ions in the formula). As noticed in this paper (Figure 1 and Table 1), Ferromanganese ores of different manganese contents (36 % and 44%) were studied with Mossbauer spectroscopy. From the lower manganese content ore only a super ferromagnetic doublet was observed while the ore with a higher manganese content showed the presence of an extra iron oxide which would possible be a hematite. The two manganese ores were reduced with metallurgical coke in a tube furnace at temperature between 1300^0C and 1400^0C. Mossbauer spectra (Figure 1) of the reduction products showed a Fe^{2+} component in both cases. Hematite contained in the 44% manganese ore (spectre ore 44) has been completely reduced into a simple oxide.

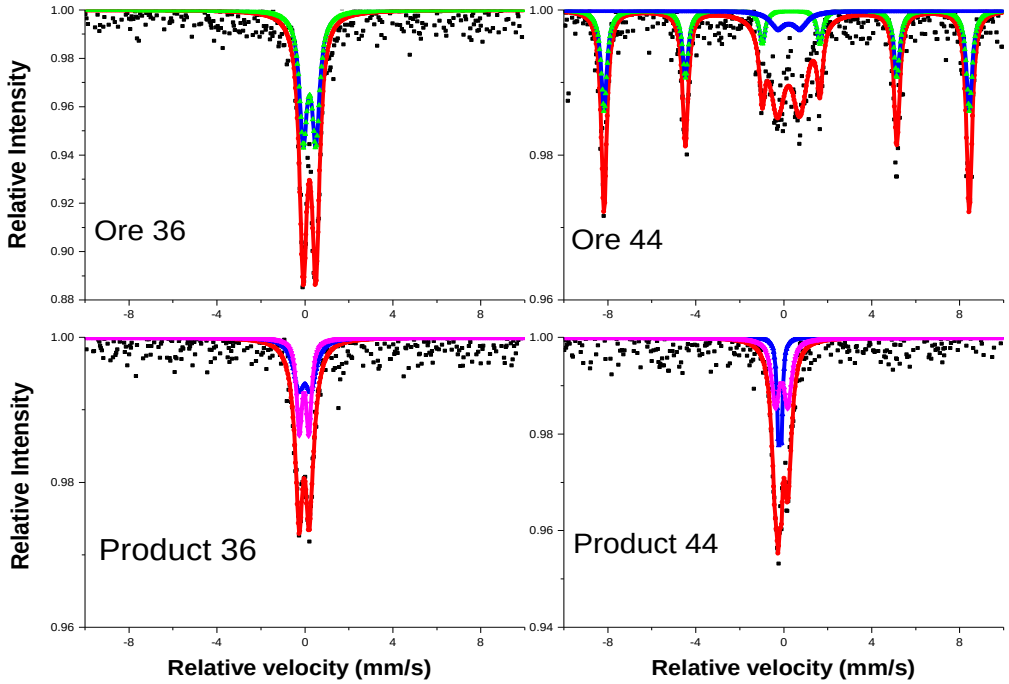


Figure 1. Mossbauer spectra of ferromanganese ore samples (Ore 36 and ore 44) and their reduction products with metallurgical coke in a tube furnace. The iron oxide contained in the ore 44 has been reduced as typified by only a superparamagnetic doublet in the spectrum of the product 44.

Table 1. Mossbauer parameters of ferromanganese ore samples and products from reduction.

Ferromanganese Ore	Isomeric shift δ_{Fe} **(mm/s)**	Quadrupole splitting **(Δ) (mm/s)**	BHF **Tesla (T)**	Fe^{3+}or Fe^{2+}	Abundance **(%)**
	0.36	-0.20	51.5	Hematite	62
Ore 44	0.35	1.03	----	Fe^{3+}	38
Reduction	-0.06	0.13.	----	Fe^{3+}	16
Product 44	0. 03	0.57	----	Fe^{2+}	84
Ore 33	0.33	0.57	----	Fe^{3+}	100
Reduction	0.12	0.56	----	Fe^{3+}	67
Product 33	0.09	0.44	----	Fe^{3+}	33

5. Mossbauer characterization of PGM containing chromite samples from the Bushveld complex

South Africa has the highest reserves of chromite which are approximately greater than 75% of the world's economic resources. All of these resources are mined from the Bushveld complex were several Chromite seams exist. The economically rich seams are the lower group seams, middle group 1 and 2 seams and the UG2 seam (upper group 2). The last of these is not of interest as a source of chromite alone but primarily as a source of platinum group metals (PGMs). A bulk of South African

chromite supply is from PGM producers since after the processing of the PGMs by froth flotation, the chromite is recovered from the tailings of the PGMs at a grade of approximately 45% Cr_2O_3 by means of gravity separation, making the tails easy and cheaper to extract chromite from them. In recent years, due to excessive exploitation PGM content of the UG2 Chromite seams is depreciating, PGM producers are slowly turning to the newly found UG3 seam for PGM production, characterization processes of the new seam are currently under way. Characterization of the newly discovered Bushveld complex UG3 seam ore and its froth flotation products was performed using Mossbauer spectroscopy technique and below are the Mossbauer spectra obtained from their analysis, together with the summary of their results showing, the isomer shift, quadrupole splitting, Fe oxidation states and % Fe abundance in the samples (Figure 2 and Table 2).

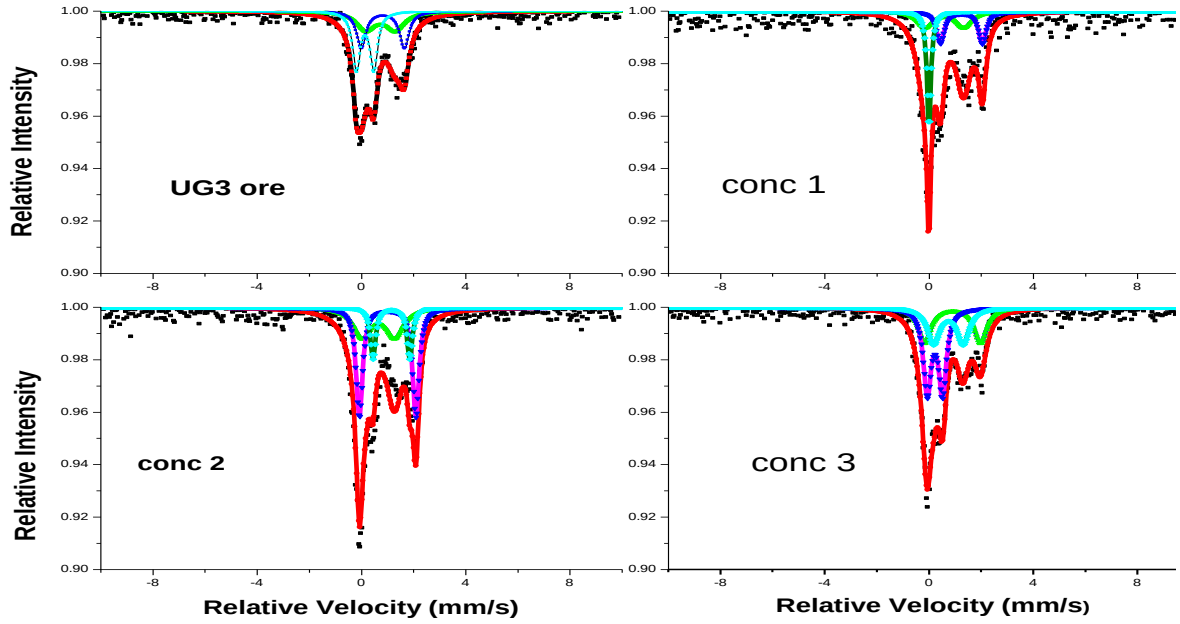


Figure 2. Mossbauer spectra of UG3 Chromite samples

MS spectra of UG3 ore samples and those of the flotation product looked similar but the difference is in the abundances of the Fe^{3+} and Fe^{2}+. UG3 ore has a ratio of 0.3 making it type 1 chromite. The concentrate 1 has a Fe^{3+}/Fe^{2+} ratio of 0.2, which indicates that it has high abundance of Fe^{2+} .The concentrate 2 has a Fe^{3+}/Fe^{2+} ratio of 1.4, making it a type 2 chromite and the concentrate 3 has a ratio of 0.5 making it a type 2 chromite.

From the UG3 chromite spectra (Figure 2), it was found that the ratio of the Fe^{3+} /Fe^{2+} in the ore is +-0.3, showing a high content of Fe^{2+} in the divalent site. Mossbauer has also been used in the characterization of natural chromite samples from the Bushveld complex South Africa. The Fe^{3+}/Fe^{2+} ratios were determined. Two different types of chromite were identified. Type 1 chromites have low Fe^{3+}/Fe^{2+} ratios (between 0.2 and 0.4) and contain 3 cat-ions for every 4 oxygen ions in the chemical formula, type 2 chromites have how Fe^{3+}/Fe^{2+} ratios between 0.7 to 1.3 and contains less than 3 cat ions per oxygen ions in the formula [3]. From the Mossbauer spectroscopy UG3 ore samples Fe^{3+}/Fe^{+2} ratio it can be said that the UG3 ore is the type 1 Chromite in the Bushveld complex.

Chromite from PGM ore processing and vanadium tailings from ferro-vanadium ore, copper converter slag from BCL and iron ore from Thabazimbi were characterized using Mossbauer spectroscopy at the Mineral processing and technology research centre. Iron ore was characterized by a magnetic sextet at room temperature. The copper smelter slag as well as the titanium slag showed iron in two sites as Fe^{2+} [4].

Table 2. Mossbauer parameters of UG3 Chromite samples

Ferromanganese Ore	Isomeric shift δ_{Fe} **(mm/s)**	Quadrupole splitting **(Δ) (mm/s)**	BHF **Tesla (T)**	Fe^{3+}or Fe^{2+}	Abundance **(%)**
	0.86	-1.13	----	Fe^{3+}	50
UG3Ore	0.94	1.65	----	Fe^{3+}	25
	0.26	0.67	---	Fe2+	25
UG3conc1	0.71	1.49	----	Fe^{2+}	57
	1.36	1.61	----	Fe^{2+}	28
	0.31	----	----	Fe^{3+}	15
	0.77	1.25	---		58
UG3conc2	1.12	2.17	----	Fe^{3+}	32
	1.27	1.41	-----		10
	1.03	2.14	-----	Fe^{2+}	34
UG3conc3	0.34	0.59	----	Fe^{3+}	32
	0.85	1.14	----	Fe^{2+}	34

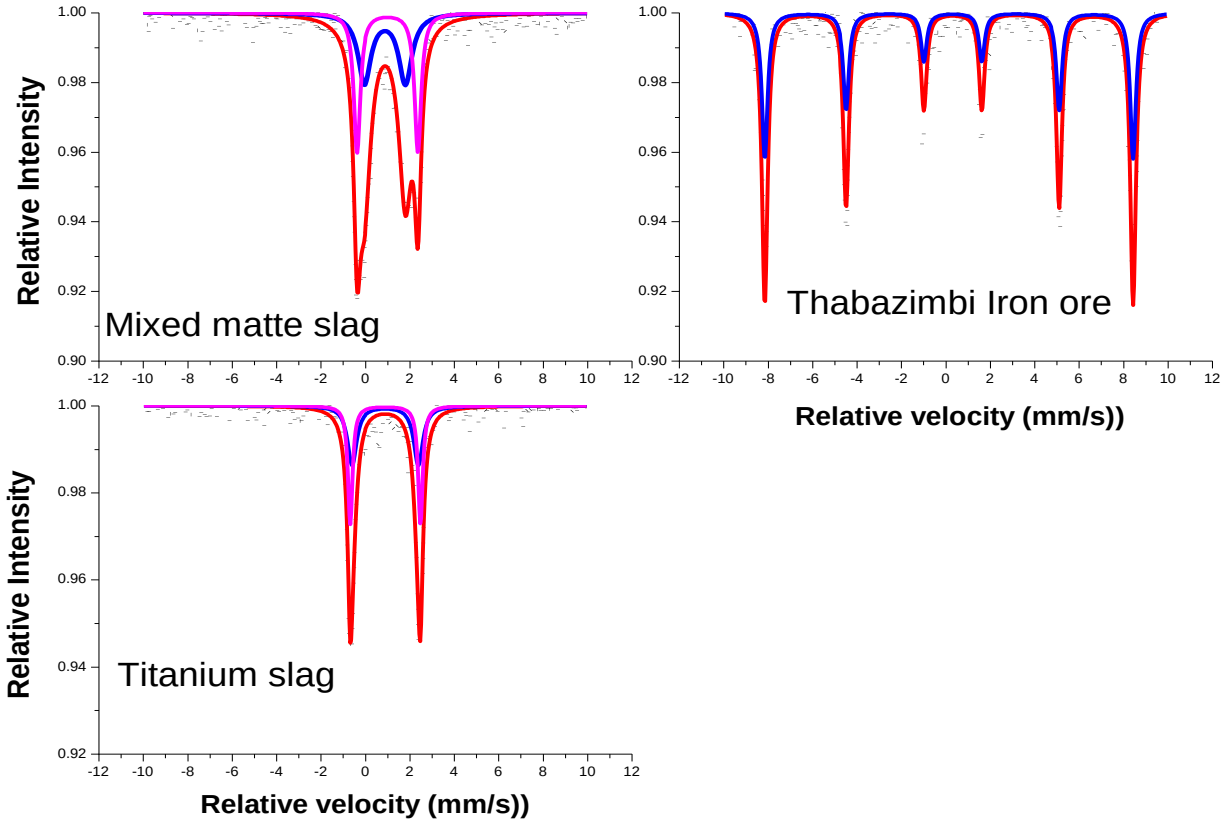


Figure 3. Various mineral processing materials

Mossbauer parameters of the Thabazimbi iron ore (Figure 3 and Table 3) reveal that the iron oxide in the studied ore is mainly hematite with an internal magnetic field of 51.1 Tesla.

Table 3. Mossbauer parameters of various mineral processing materials

	Isomeric Shift **(δ/Fe)+-0.01**	Quadrupole splitting **(Δ) +-0.01**	BhF	ion	% abundance
	0.81	0.96	----	Fe^{3+}	48
	0.26	0.65	---	Fe^{3+}	29
Chromite	0.92	1.56	----	Fe^{2+}	24
Iron ore Thabazimbi	0.35	-0.18	51.6	Fe^{3+}	hematite
Titanium	1.01	2.99	---	Fe^{3+}	66
Slag	1.02	3.17	---	Fe^{3+}	34
Matte-slag	1.00	1.85	---	Fe^{3+}	67
Mixture	1.11	2.76	---	Fe^{2+}	33

Nkomati mine in Mpumalanga produces nickel concentrate through froth flotation. Further beneficiation of base metal bearing minerals has being done using pyrometallurgical routes, these days hydrometallurgy is becoming an attractive tool because of the high thermal costs and pollution problems associated with Pyrometallurgy [2]. Mossbauer spectroscopy was used to characterize the products of roasting of the Nickel sulphide concentrate prior to leaching with sulphuric acid and hydrochloric acid [4].

6. Challenges of Mossbauer spectroscopy training at postgraduate level

Mossbauer spectroscopy is a delicate characterisation technique using nuclear transition leading to the emission of gamma rays which selected are resonantly absorbed by an absorber/sample conveniently prepared. The interpretation and discussion of the data obtained require a good understanding of the hyperfine interactions parameters (isomeric shift, quadrupole splitting, magnetic field etc…) and an insight on the material under the study. Often the type of candidate available to tackle the intrinsic and holistic comprehension challenge from this technology constitutes a problem. The understanding level of quantum and nuclear physics required from the candidate to comprehend the basics and fundamentals of the related interactions is often above the average of commonly encountered candidates. The technology and relevant nuclear spectrometer components also become very expensive.

7. Conclusion

Mossbauer spectroscopy is a characterization technique like any others. It can add value into the information or analysis obtained from other techniques. It supports and complements information from other techniques. From Mossbauer results South African chromites could be classified as type 1 or type 2. The reduction of the ferromanganese ores with coke could be easily monitored with the MS study. This paper shows data on the UG3 ore. It is for the first time that Mossbauer study of UG3 has been reported. From the research and analysis performed in this project, it was shown that one can monitor mineralogical phase transformations that are taking place in different mineral processing applications; it was also shown that Mossbauer spectroscopy can differentiate between types of ores and similarities between clays of different origins.

Acknowledgments

The author is thankful to Modikwa mine for providing the UG3 samples studied in this work. Nkomati mine provided the nickel bearing materials studied. The author is extensively collaborating with Professor F. Waanders from the North-West University (Potchefstroom campus). His crucial contributions in some measurements is sincerely acknowledged.

8. References

[1] Gonser U 1981 *Mossbauer spectroscopy II*, Germany Springer Verlag, 25. 1981, 1-301.
[2] Waanders F B, Mulaba-Bafubiandi A F and Lodya L 2013 *Hyp. Inter.* 1
[3] Nell J and Pollak H 1996 J.C Baltzer AG, Science Publishers. *Hyp Inter*. 427-31
[4] Waanders F B, Vinken E, Mans A and Mulaba-Bafubiandi A F 2003 *Hyp. Inter*. **148** 21-9

Electronic and Magnetic Structure of Cr doped Rutile and Anatase TiO_2; an *ab-initio* DFT+U study.

Winfred M. Mulwa[1]*, Cecil N.M.Ouma[2], Martin O. Onani[3], Francis B. Dejene[1]

[1] Department of Physics, University of the Free State, Qwaqwa Campus, South Africa

[2] Natural Ressources and Environment, CSIR, South Africa

[3]Chemistry Department, University of the Western Cape, South Africa

E-mail: mulwawinfred@gmail.com

Abstract. Semiconductor devices usually utilize the charge of electrons and holes, while magnetic materials used for recording information utilize electron spin. To make use of both charge and spin of electrons in semiconductors, a high concentration of magnetic elements can be introduced in nonmagnetic III-V semiconductors currently in use for devices. The usefulness of semiconductors resides in the ability to dope them with impurities to change their properties, usually to *p*- or *n*-type. This approach can be followed to introduce magnetic elements into nonmagnetic semiconductors to make them magnetic. This category of semiconductors, called diluted magnetic semiconductors, are alloys of nonmagnetic semiconductor and magnetic elements. In our case we investigate chromium-doped TiO_2 in the rutile and anatase phases. We realize that the transition from paramagnetism to ferromagnetism in both phases is at 6% doping. At this point, a large magnetic moment of 2.59 μB and 2.49μB per Cr atom in rutile and anatase respectively was realized.

1. Introduction

TiO_2 plays an important role in the fields of dilute magnetic semiconductors (DMS) and photocatalysis. For both applications TiO_2 requires tailor made electronic properties, which can be manipulated by selectively diluting the system by insertion of dopants. DMS have been investigated in the past in order to gain new insights in the functional principles and the manipulation of the magnetic and semiconducting properties. The search for magnetic semiconductors or for half-metals used for spin-injection led to the discovery of the new material class of d-electron magnets. The prospect of controlling the charge and the electron spin as information carriers would make it possible to combine information processing and storage at the same time [1-3]. The second major application is in the field of photocatalysis, where the favourable electronic gap range of 1.0 to 2.8 eV can be also achieved by doping**.** Impurity doping is an important technique for improving the response of TiO_2-based photoelectrolysis cells [2]. Chromium ions substituted for Ti^{4+} in the TiO_2, lattice have so far yielded optimal results. Substitutional Cr^{3+} in TiO_2 has a strong preference for ferromagnetic ordering. Stabilization of ferromagnetism is achieved through p-type doping. However, the description of the

transition from paramagnetic to ferromagnetic varies within different levels of dopant concentration. We examine the effects of varying Cr^{3+} concentration in the TiO_2 matrix.

In our study, we investigate TiO_2 anatase and rutile doped with chromium atoms, where the Cr is sited at the Ti site at various doping levels. Magnetic order is most common in metallic materials with narrow bands of d electrons therefore in this case, the carriers of the magnetic moments in doped semiconductors TiO_2 is the chromium atoms. The magnetic moment is produced by the d–electrons, this is called d-electron magnetism and has been widely studied in the recent past [4]. Regarding TiO_2 anatase from a photocatalytic aspect, we are mainly interested in the manipulation of the band gap by incorporation of impurities. TiO_2 anatase and rutile phases have band gaps of 3.2 eV [5] and 3.0 eV [6], respectively which corresponds to the ultraviolet region of the absorption spectrum [7]. The creation of electron hole pairs came up due to photon absorption. Photocatalytic and photoelectrochemical applications like the photoinduced decomposition of water on TiO_2 electrodes requires photon absorption as demonstrated by Fujishima and Honda [8]. In the photocatalytic utilization, creation of the electron hole pair makes valence band (VB) become the oxidative band and the conduction band becomes reductive band. Surrounding molecules can therefore be oxidized or reduced for example radicals which are harmful to organic compounds like bacteria and fungi. Therefore, TiO_2 can be used as a purifier removing gaseous or aqueous contaminants [9,10]. In addition to water and air purification, TiO_2 has other photocatalytic applications, like antifogging and self-cleaning surfaces [11]. Since UV light only makes up 5% of the sunlight spectrum we aim to shift the absorption spectrum of TiO_2 into the visible region so as to improve the absorption rate and therefore the photocatalytic efficiency. In this case we modify the band gap using chromium impurities.

2. Calculation details

Within DFT, the electron–electron interaction is expressed as the sum of the Hartree and exchange-correlation (XC) terms although, the XC term is usually approximated. Due to this approximation, the XC term cannot accurately account for electronic interactions in strongly correlated systems, hence the need for DFT+*U* technique where a Hubbard correction *U* is added to the standard DFT formalism 12]. Plane wave method based spin polarized DFT+U calculations were performed as implemented in the Quantum Espresso code [13]. The Perdew-Wang of the local density approximation with the Hubbard U correction (LDA+U) parameterization was utilized for the exchange correlation potentials. The norm-conserving pseudopotentials [14] were used for electron-ion interaction. Doping was performed by substituting one, two, three, four, five, six and seven Ti lattice atoms by Cr atoms corresponding to the Cr doping concentration of 1%, 3%, 4%, 5%, 6%, 8% and 10%, respectively. This was done in order to investigate the effect of doping concentration on the magnetic properties of Cr:TiO_2. A 72 atom supercell was modelled by $2\times 2\times 1$ repetition of the anatase unit cell. Plane waves with an energy up to 60 eV were included in the basis set, in order to avoid Pulay stress and other related problems. The Brillouin-Zone integration was performed using a $2\times 2\times 2$ Γ centered *k*-mesh with Gaussian smearing set to 0.05 eV. The total energy was converged at $1x10^{-6}$ eV. To compare and evaluate the structural changes induced by the impurities, calculations on pure TiO_2 were done initially to obtain structural and electronic data that can be used as a reference. Anatase is computationally stable than rutile and brookite phases therefore we investigated the properties of chromium doped anatase. To investigate the stability of the doped system we calculated the defect formation energies according to the following formula [15].

$$E^{f}[X] = E_{tot}[X] - E_{tot}[bulk] + n(\mu_0 - \mu_x) \quad (1)$$

where E_{tot} is the total energy of the doped supercell with one Ti atom replaced by the impurity *X*(*X*=Cr), E_{tot}[bulk] is the total energy for the pristine system. *n* indicates the number of Ti atoms that have be replaced in the supercell by the dopant atoms, μ_0 and μ_x are the corresponding chemical potentials. The chemical potentials were calculated with respect to the free atom.

In order to get more stable atomic configuration and lattice parameters, the geometry optimization was performed for pure and doped systems using Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-newton minimization scheme [16]. Supercells of optimised pure anatase TiO_2 and 6% Cr-doped TiO_2 are shown in figure1.

3. Results and discussions

3.1. Geometrically optimized structure properties

It is only rutile and anatase phases which play a key role in the applications of TiO_2 in industry because of their simple tetragonal structure compared to complicated orthorhombic brookite structure. Many of the technological applications utilize the anatase phase (figure 1) rather than the rutile one, as this appears to be more catalytically active and exhibits a higher electron mobility. Rutile and anatase are both tetragonal, containing 6 and 12 atoms per unit cell, respectively. In both structures each Ti atom is coordinated to six O atoms and each O atom is coordinated to three Ti atoms. In each case, the TiO_6 octahedron is slightly distorted, with two Ti-O bonds slightly greater than the other four, and with some of the O-Ti-O bond angles deviating from 90^o. The distortion is greater in anatase than in rutile. The structure of rutile and anatase crystals has been described frequently in terms of chains of TiO_6 octahedra having common edges [17-18]. Two and four edges are shared in rutile and anatase, respectively.
The calculated lattice parameters after geometry optimization are a = 4.5936Å, c = 2.9587Å and a = 3.784Å, c = 9.515Å for rutile and anatase, respectively from LDA+U calculations. The obtained lattice parameters are in good agreement with other theoretical calculations [19] which implies that our calculation methods are reasonable and the obtained results are trustworthy. The averaged bond lengths of the pure anatase TiO_2 and geometric optimized 0-10% Cr:TiO_2 are provided in table 1. From table 2 the LDA+U obtained lattice parameters confirms what was realized experimentally.
These results also agree very well with Deskins et al findings. [20].

Table 1: Average Mulliken bond lengths of pure and Cr:TiO_2 models after geometry optimization

System	Ti-O(Å)	O-O(Å)	Cr-O(Å)
Pure TiO_2	1.9759	2.719	
1%Cr:TiO_2	1.9445	2.68	1.9353
3%Cr:TiO_2	1.9483	2.678	1.9345
4%Cr:TiO_2	1.9485	2.7001	1.9342
6%Cr:TiO_2	1.9399	2.7101	1.9341
7%Cr:TiO_2	1.9392	2.7111	1.9340
8%Cr:TiO_2	1.9390	2.7111	1.9339
10%Cr:TiO_2	1.9391	2.7112	1.9339

Table 2: A comparison between calculated structural properties of rutile and anatase phases with experimental data.

Rutile	a (Å)	c (Å)
Experiment	4.587[21]	2.954[21]
This work	4.5936	2.9587
Anatase		
Experiment	3.782[21]	9.502[21]
This work	3.784	9.515

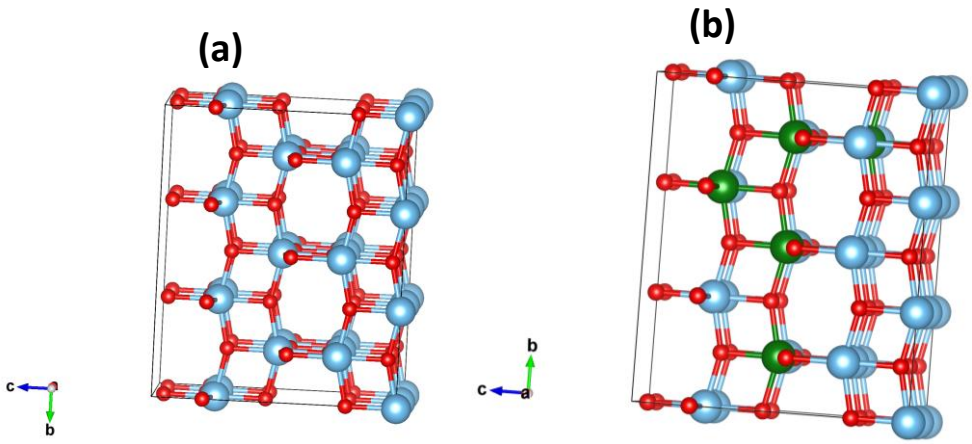
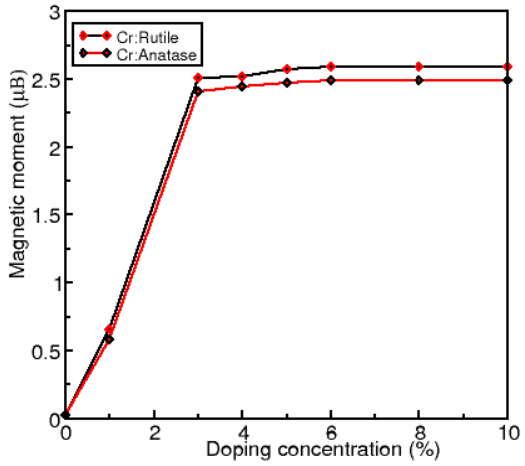


Figure 1: A supercell of a) Pristine TiO_2 b) Cr doped TiO_2 (Red, Blue and Green balls represents O,Ti and dopant atoms respectively).

The geometric optimized structures of pure anatase TiO_2 and 6% Cr-doped anatase were realized by the application of VESTA (Visualization for Electronic and Structural Analysis) [22] visualizing tool as shown in figure 1. In figure 1(a), the blue atoms represent titanium, while the red atoms represent oxygen. In figure 1 (b), the blue atoms are titanium, red atoms are oxygen and green atoms are chromium atoms sited at Ti site. Formation energy was calculated at various substitution sites and sites with the lowest formation energy were utilized.

3.2. Magnetic properties

From the projected density of states output file, at the Lowdin charges part, we were able to get the magnetic moment per Cr atom. A large magnetic moment of 2.59μB in rutile and 2.49μB in anatase per Cr atom is found for 6% Cr-doped rutile and anatase. After 6% doping of TiO_2 with chromium, both phase reached magnetic saturation as shown in figure 2.



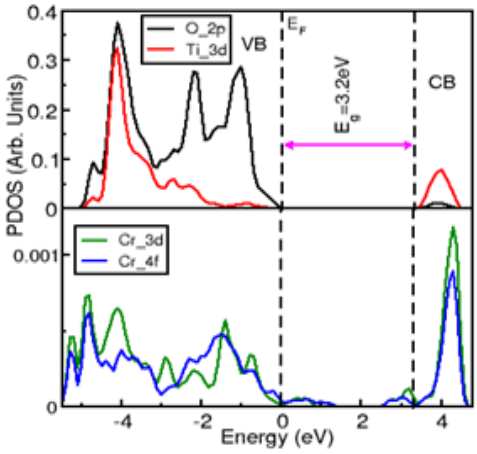


Figure 2: Magnetic saturation in rutile and anatase. **Figure 3:** PDOS of Cr doped TiO_2

The highest magnetic moment of 2.59μB in rutile and 2.49μB in anatase per Cr atom realized in our two phases is higher than expected. Chromium is magnetic, therefore there is a possibility that all the contribution to the magnetic moment is due to Cr ions but there are high chances that Ti moments contributed a certain amount of magnetic moment. When we look at the hybridization of Ti_3d and Cr_3d in figure 3, both Cr and Ti may have contributed to the realized high magnetic moment. The participation of Ti_3d electrons makes the understanding of magnetic properties in rutile and anatase more interesting.

3.3. Electronic properties

The top panel of figure 3 shows pure anatase with a band gap of 3.2 eV which is exactly same as experimental value. This confirms the accuracy of our results. Impurity states are realized just below the conduction band minimum (CBM) as well as above the valence band maximum (VBM), therefore the band gap has been reduced from 3.2 eV to 1 eV after 6% doping of anatase with chromium. Figure 4 (a) and (b) confirm what is realized in figure 3 from density of states point of view.
Figures 3 and 4 show an effective photocatalyst, that is, a band gap with no additional states inside. Additional states inside a band gap become electron-hole traps and recombination centers causing a reduction in the lifetime of the photoinduced charged carries, leading to a decrease of photocatalytic efficiency. Doping our system with chromium, we reduced the band gap in an efficient way, that is without introducing mid band gap states which make the recombination rate of the charge carriers to be maximum. Figure 4 (a) and (b) shows that after 6% doping (part f), the crystal structure undergoes deformation (part g and h). As seen in figure 4 (a) and (b), when the doping percentage is increased, more impurity states appear within the bad gap and the shape of the conduction band and valence band remains the same. After 6% doping (part f), any more increase of Cr in TiO_2 causes a change in the shape of the valence and conduction bands. This is attributed to the crystal structure undergoing deformation (part g and h).

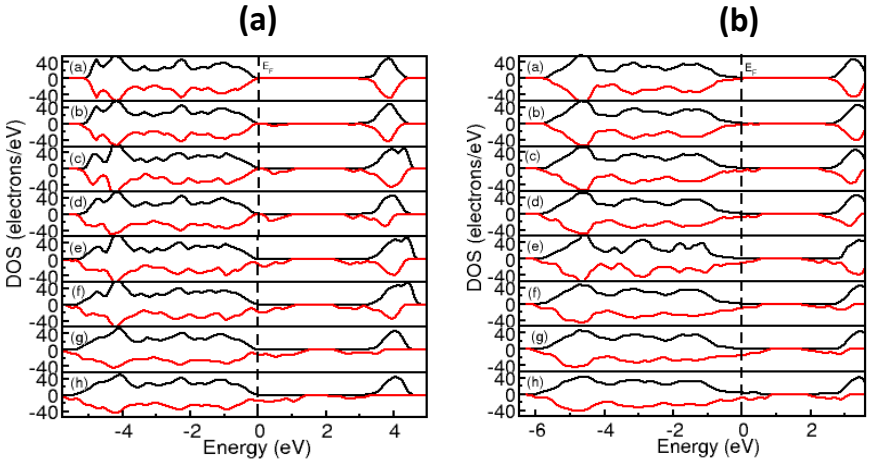


Figure 4: Calculated DOS of pristine and 1-10%Cr:TiO_2 as a function of energy.a) Rutile b) Anatase.

3.4 *Photocatalytic performance.*

The effect of Cr-doping on the photocatalytic effect (clean bandgap) is also shown in figure 4 (a) and (b). As the concentration of Cr is increased, more dopant states are seen to be encroaching into the bandgap both from valence band maximum and conduction band minimum. This represents a shift in the absorption band edge towards longer wavelength with the increasing Cr concentration (red shift).

The red shift is maximum at 4% doping of TiO_2 with Cr which is consistent with the experimental findings [23]. After which the size of the 'clean' band gap remains constant. Even after attaining optimal doping concentration at 4%, with further increase in doping, crystal deformation is achieved at 6% doping, which is also the magnetic saturation point as shown in figure 2.

4. Conclusion

Cr-doped rutile and anatase are magnetic semiconductors. They acquire magnetic moment saturation at 6% doping in each case with 2.59μB in rutile and 2.49μB in anatase. The transition point from paramagnetism to ferromagnetism is at 6% Cr doping in both phases**.** The large magnetic moments originates from two sources, that is, involvement of the 3d electrons of the Cr^{3+} ions and the formation of acceptor bound magnetic polarons, in which the spins of the holes and chromium are aligned through exchange interaction. From figures 3 and 4, we find enhanced photocatalysis because there are no mid band gap states. After 6% magnetic saturation is achieved and addition of more Cr^{3+} atoms does not improve magnetic properties but distorts the crystal lattice. This study clearly shows the importance of doping TiO_2 with Cr to enhance its magnetic and photocatalytic properties.

Acknowledgement

This research work is supported by the national research foundation of South Africa.

References

[1] Ohno H 1998 Science **281** 951.
[2] Matsukura F, Ohno H, Shen A and Sugawara Y 1998 Phys. Rev. B **57** 2037.
[3] DiVincenzo D P 1995 Science **270** 255.
[4] Y.Matsumoto, M.Murakami, T.Shono, T.Hasegawa, T.Fukumura, M.Kawasaki, P.Ahmet, T. Chikyow, S.Koshihara and Koinuma H 2001 Science **291**, 854.
[5] Tang H, Berger H, Schmid P E, Levy F, and Burri G 1993 Solid State Commun. **87**, 847
[6] Pascual J, Camassel J, and Mathieu H 1978 Phys. Rev. B **18**, 5606.
[7] Meng N, Micheal Leung* K H, Dennis Leung Y C, Sumathy K 2007 Renew. Sustainable Energy Rev. **11** 401-425.
[8] Fujishima A and Honda K 1972 Nature **238** 37.
[9] Anne M and Dulay T 1993 Chem. Rev. **93** 341.
[10] Hoffmann M R, Martin S T, Choi W and Bahnemann D W 1995 Chem. Rev. **95** 69.
[11] Lai Y, Tang Y, Gong J, Gong D, Chi L, Lin C and Chen Zc. 2012 J. Mater. Chem.**22** 7420-7426.
[12] Cococcioni M and De Gironcoli S 2005 Phys. Rev. B **71** 35105.
[13] Giannozzi P, Baroni S, Bonini N, Calandra M, Car R, Cavazzoni C, Ceresoli D, Chiarotti G L, M.Cococcioni, I.Dabo, A.Dalcorso, S.Fabris, G.Fratesi, S.deGironcoli F.Mauri, R.Mazzarello, Paolini S, Pasquarello A, Paulatto L, Sbraccia C, Scandolo S, Sclauzero G, Seitsonen A P, Smogunov A, Umari P and Wentzcovitch R M 2009 J.Phys.Condens. Matter **21** 395502.
[14] Hamann D R, Schluter M and Chiang C 1979 Phys. Rev.Lett, **43** 1494-1497.
[15] Van de Walle C G and Neugebauer J 2004 J. Appl. Phy. **95** 385.
[16] Pfrommer B G, Cote M, Loulle S G and Cohen M I 1997 J. Comput. Phys.**131** 233.
[17] Grant F A 1959 Rev. Mod. Phys. **31** 646.
[18] Kavan L, Grtzel M, Gilbert S E, Klemenz C and Scheel H J 1996 J.Am. Chem.Soc. **118** 6716
[19] Arroyo-de Dompablo M E, Morales-Garcia M and Taravillo 2011 J. Chem. Phys. **135** 54503.
[20] Aaron N, Deskins N and Michel D 2007 Phys. Rev. B **75** 195212.
[21] Burdett J K, Hughbanks T, Miller G J, Richardson Jr J W and Smith J V 1987 J.Am. Chem. Soc. **109** 3639.
[22] Momma K and F.Izumi F 2011 J. Appl. Crystallogr. **44** 1272.
[23] Yan-Hua P, Gui-Feng H* and Wei-Qing H* 2012 Adv. Powder Technol. **23** 8-12.

Exciton energies of chalcopyrites $AgAlX_2$(X=S,Se,Te) from GW and BSE calculations

GM Dongho Nguimdo and Daniel P Joubert

National Institute for Theoretical Physics, School of Physics and Mandelstam Institute for Theoretical Physics, University of the Witwatersrand, Private Bag 3, Johannesburg 2050, South Africa

E-mail: donghomoise@gmail.com

Abstract. Using state-of-the-art Density Functional and Many Body Perturbation Theories, we study electronic and optical properties of the chalcopyrites $AgAlX_2$(X=S,Se,Te). The Kohn-Sham Density Functional Theory (DFT) underestimates the fundamental and the optical gaps as a result of the particle number dependant discontinuity in the exchange-correlation potential [1]. Accurate estimates of fundamental gaps were obtained using post DFT Many Body Perturbation Theory at the GW level. Optical absorption spectra and optical gaps were determined from solutions of the Bethe-Selpeter Equation (BSE) in the Tamm-Dancoff approximation. Comparison of the BSE and the GW results were used to obtain exciton energies. The GW-level calculated bandgaps are in good agreement with experimental values. Exciton energies were estimated for the first time but we couldn't find any theoretical or experimental results for comparison.

1. Introduction

During the absorption of photons by the interband transitions, an electron and a hole are usually created in the conduction and valence bands respectively. The Coulomb interaction between the two particles leads to the formation of a new excitation in the crystal also known as exciton. It has been proved that these excitons can considerably impacted the optical properies of semiconductor materials and therefore are very important for opto-electronic applications [2,3]. The determination of the exciton energy has been a cumbersome task for the scientific community. Excitons are easily observed in pure samples. In real system, the impurities create unpaired electrons and holes which can screen the Couloumb interaction in the excitons and reduce the strength of the binding forces [2]. Hence, making the exciton very difficult to be observed experimentally. In theoretical point of view, accurate optical properties are not always directly accessible from first-principle DFT calculations. For example the eigenvalue energies $\epsilon_{v\mathbf{k}}$ and $\epsilon_{c\mathbf{k}}$ entering in the Fermi's golden rule expression [4]

$$\varepsilon_{2_M}(\omega) = \lim_{\mathbf{q}\to 0} \frac{8\pi^2}{\Omega \mathbf{q}^2} \sum_{vc\mathbf{k}} |\langle u_{c\mathbf{k}+\mathbf{q}} | e^{i\mathbf{q}\mathbf{r}} | u_{v\mathbf{k}} \rangle|^2 \delta(\omega - (\epsilon_{c\mathbf{k}+\mathbf{q}} - \epsilon_{v\mathbf{k}})) \quad (1)$$

for calculating the absorption spectra correspond to the Kohn-Sham (KS) highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) respectively. The energy difference between HOMO and LUMO yields the KS bandgap. However, it is now

well established that the KS bandgap is underestimated by more than 50% compared to the experimental measurement. As such, there is a need to go beyond the standard DFT in order to accurately predict the bandgap and the optical spectra. Reliable bandgap can be obtained by means of the GW approach by solving the *quasiparticle* (QP) equation [5]

$$[T + V_{ee} + V_{ext}]\,\varphi^{\mathrm{QP}}_{i,\mathbf{k}}(\mathbf{r}) + \int d\mathbf{r}'\Sigma(\mathbf{r},\mathbf{r}';\epsilon^{\mathrm{QP}}_{i,\mathbf{k}})\phi^{\mathrm{QP}}_{i,\mathbf{k}}(\mathbf{r}) = \epsilon^{\mathrm{QP}}_{i,\mathbf{k}}(\mathbf{r})\varphi^{\mathrm{QP}}_{i,\mathbf{k}}(\mathbf{r}) \tag{2}$$

where Σ is the self-energy which contains all the exchange and correlations effects among the electrons. Although this approach leads to a good approximation of the fundamental gap, it is

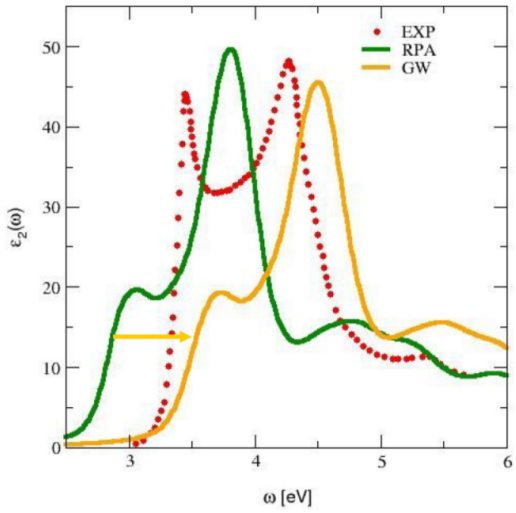


Figure 1. (Color online) Optical absorption of silicon from the experimental measurement (EXP), the GW and the DFT-LDA calculations within the random phase approximation (RPA) [6, 7].

still not sufficient for the correct description of the optical properties. For example in the case of silicon as described in Figure 1, GW gives a good onset of the optical absorption. However, the first peak occurs at a relatively high energy and its intensity is also underestimated as compared to that from experimental measurement. This is due to the fact that the GW only allows free electron-hole transition without accounting for the Coulomb interaction between the two particles. The electron-hole can be captured by solving the Bethe-Selpeter Equation (BSE) [8, 9].

In most chalcopyrite materials, such an interaction is yet to be carefully examined in contrast to the bandgap and the electronic properties where a huge effort has been made. Most of the previous works have been limited to the calculation of the dielectric response function in the independent particle picture within local field effect [10–12]. This is due to the difficulty of including the electron-hole interaction in the calculations. Moreover, these calculations are very computationally demanding because of the high number of empty bands which need to be included in order to reach the convergence. We are not aware of any work done in that regard. Our goal in this paper is to first estimate the fundamental gap and secondly investigate the importance of the excitons in the chalcopyrite materials $AgAlX_2$(X=S,Se,Te).

2. Methodology

All the relaxed structures were obtained from our previous work in Ref. [10]. In all the study, we use the plane wave projector augmented wave (PAW) [13] as implemented in the VASP package [14]. The self energy in Equation 2 is calculated using the many body perturbation

theory at the GW level ($\Sigma = iGW$) [15]. At the first order of the approximation, the Green's function G is built from the Kohn-Sham (KS) orbitals and eigenvalue energies as

$$G_0(\mathbf{r}', \mathbf{r}; \epsilon) = \sum_i \frac{\varphi_{KS}(\mathbf{r})\varphi^*_{KS}(\mathbf{r}')}{\epsilon - \epsilon_i + i\eta \ \mathrm{sgn}(\epsilon_i - \mu)} \tag{3}$$

while the dynamically screened Coulomb potential is given by

$$W_0(\mathbf{r}, \mathbf{r}', \epsilon) = \int \varepsilon^{-1}(\mathbf{r}, \mathbf{r}_1, \epsilon) v(\mathbf{r}, \mathbf{r}') d\mathbf{r}_1 \tag{4}$$

with $v(\mathbf{r}, \mathbf{r}')$ the bare Coulomb potential and ε^{-1} the dielectric function calculated within the RPA [16] at the DFT level. The QP energies are therefore calculated as follow.

$$\epsilon^{QP}_{i,\mathbf{k}} = \mathcal{R}e\left[\left\langle \varphi^{\mathrm{QP}}_{i,\mathbf{k}} \middle| H_{KS} + \Sigma(\epsilon^{KS}_{i,\mathbf{k}}) - v_{xc} \middle| \varphi^{\mathrm{QP}}_{i,\mathbf{k}} \right\rangle\right] \tag{5}$$

where H_{KS} is the KS Hamiltonian and v_{xc} the exchange-correlation potential. It is worth mentioning that the exchange-correlation potential v_{xc} has to be subtracted from the expression for the QP energy since it is included in the Kohn-Sham Hamiltonian H_{KS}. The generalised gradient approximation PBEsol [17] is used as exchange-correlation functional. The excitonic effect was included in the calculation of the dielectric function by solving the BSE. We used 696 bands and 300 eV cutoff energy for the response function for the calculations.

3. Results and discussions

We start our study by performing a single shot $\mathrm{G_0W_0}$ calculation. Our results (see Table 1) show an underestimation of the fundamental bandgap despite the large number of empty bands included in the calculations. Aguilera et al. [18] in their study on the bandstructure and optical properties of $\mathrm{CuGaS_2}$ also found such a failure. They claimed that the underestimation is due to an important contribution of the d electrons at the upper most valence band with a strong hybridisation of the p orbitals that the $\mathrm{G_0W_0}$ is not capable of reproducing. Our previous study on the density of states of $\mathrm{AgAlX_2}$ [10] showed that the d and p orbitals from the Ag and chalcogenide (X) atoms respectively were also hybridised at the top of the valence band. Hence, it should also be at the essence of the failure of $\mathrm{G_0W_0}$. We therefore proceed by doing a semi self-consistent GW calculation ($\mathrm{GW_0}$) where screened Coulomb interaction $\mathrm{W_0}$ remains at the RPA level and the Green's function updated following the Dyson's equation [19]

$$G\left(\mathbf{r}', \mathbf{r}; E\right) = G_0\left(\mathbf{r}', \mathbf{r}; E\right) + \iint G_0\left(\mathbf{r}, \mathbf{r}_1; E\right) \Sigma\left(\mathbf{r}_1, \mathbf{r}_2; E\right) G\left(\mathbf{r}_2, \mathbf{r}'; E\right) d\mathbf{r}_1 d\mathbf{r}_2 \tag{6}$$

where $\mathrm{G_0}$ refers to the Green's function of the Kohn-Sham system. The bandgaps obtained from this approach are also summarised in Table 1. One can note the discrepancy with the $\mathrm{G_0W_0}$ and an agreement within 4% with the experimental results. Point to note here is that similar results in chalcopyrite based materials were obtained using the self-consistent Coulomb Hole Screened Exchange followed by $\mathrm{G_0W_0}$ (sc-COHSEX+$\mathrm{G_0W_0}$) [20,21]. But this approach was not adopted in our study.

We present in Figure 2 the bandstructure from PBEsol and $\mathrm{GW_0}$ of $\mathrm{AgAlS_2}$. The two other compounds (not shown) have similar bandstructure. Both PBEsol and $\mathrm{GW_0}$ approximations show a direct bandgap at the Γ in accordance with experimental data [22]. The main difference between the bandstructure is the shift from 1.83 eV to 3.22 eV at the Γ. Note that, the bandgap does not only open because of the upshift of the conduction band as it seems to be the case in

Table 1. Bandgap from different methods, position of the exciton peak from BSE calculations and exciton binding energies of $AgAlX_2$. All the quantities are given in eV. * and * refer to Ref. [10] and Ref. [22] respectively.

	$AgAlS_2$	$AgAlSe_2$	$AgAlTe_2$
PBEsol*	1.83	1.11	1.03
G_0W_0	2.67	2.12	2.08
GW_0	3.22	2.46	2.22
EXP.**	3.13	2.55	2.27
First BSE peak	3.09	2.26	2.07
Binding energy	0.21	0.20	0.15

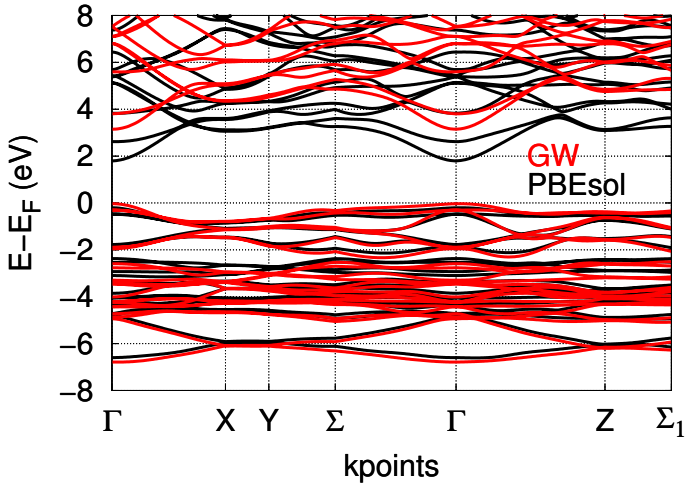


Figure 2. (Color online) Representative bandstrcture of $AgAlX_2$ from PBEsol and GW_0 calculation: Case of $AgAlS_2$.

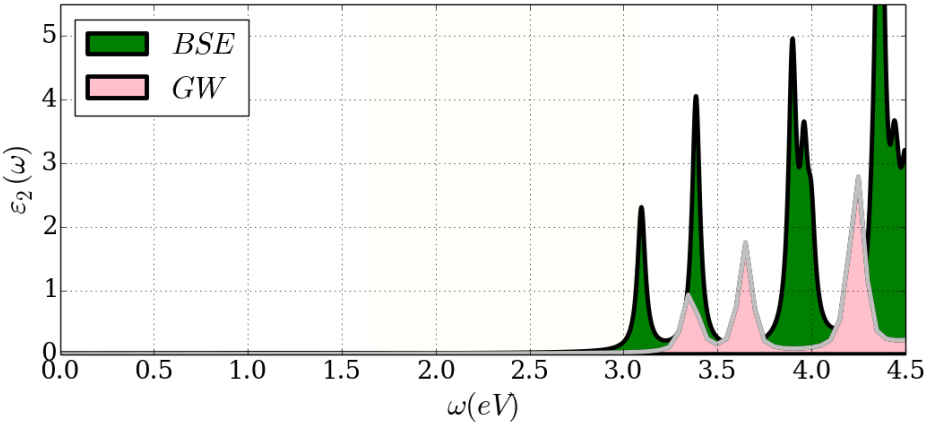


Figure 3. (Color online) $AlAgS_2$ imaginary part of the dielectric function (ε_2) from GW and BSE calculations.

Figure 2. Here the Fermi level is shifted to the upper valence band for the two approximations. Usually, the valence band also downshifts as a result of the cancellation of the delocalisation

error originated from the PBEsol wave function [10, 18]. We calculated the dielectric function in the frame work of many body pertubation theory at the GW and BSE level. We presented in Figure 3 the imaginary part of the dielectric ε_2 from the two approximations. Experimental data are not available for the sake of comparison. The onset of from GW lies at 3.19 eV, 2.44 eV, and 2.10 eV for $AgAlS_2$, $AgAlSe_2$ and $AgAlTe_2$ respectively. The values are in the range their respective bandgaps as seen in Table 1. However when the electron-hole interaction is incorporated in the calculation through the BSE, we observe a reduction of the onset with a peak appearing below the bandgap. Thus, attesting the presence of bound exciton in these materials. In direct bandgap materials, the binding energy of the exciton can easily be estimated. Here, electron and hole have the same group velocity ($v_{g_e}=v_{g_h}=\frac{1}{\hbar}\frac{\partial E}{\partial \mathbf{k}}$) since they appear at the same $\mathbf{k}$ vector [2]. In such condition the exciton energy can therefore be calculated as difference between the bandgap from GW and the position of the first excitonic peak from BSE. As summarised in Table 1, the excitonic binding energies range from 0.15 eV for $AgAlTe_2$ to 0.21 eV for $AgAlS_2$. Comparing the bandgaps and the excitonic binding energies, it turns out that the bigger the bandgap, the bigger the excitonic binding energy. Hence, the excitonic binding energy should be related to the nature of the chalcogenide atom in presence in the chalcopyrites.

4. Conclusion

We investigated the bandgap and optical spectra of the chalcopyrite materials $AgAlX_2$ by means of the many body perturbation theory at GW and BSE level. The single shot G_0W_0 method failed to predict correctly the bandgaps. Using the semi self-consistent GW_0, accurate bandgap within 4% agreement with the experimental data was obtained. The BSE calculation on top of the GW_0 allows us to predict the existence of exciton on these compounds. We also found that the excitonic energy was influenced by the size of the bandgap of the materials investigated.

Acknowledgement

This work was supported by the AIMS-DAAD Grant No: A/14/90078.

References

[1] Perdew J P and Levy M 1983 *Phys. Rev. Lett.* **51**(20) 1884–1887
[2] Fox M 2006 *Quantum Optics: An Introduction* (*Oxford Master Series in Physics* vol 15) (Oxford university press)
[3] Cudazzo P, Sottile F, Rubio A and Gatti M 2015 *Journal of Physics: Condensed Matter* **27** 113204
[4] Grosso G and Pastori-Parravicini G 2000 *Solid State Physics* (Academic, New York)
[5] Suleiman M S H 2014 Ph.D. thesis Faculty of Science, University of the Witwatersrand, Johannesburg
[6] Lautenschlager P, Garriga M, Vina L and Cardona M 1987 *Phys. Rev. B* **36** 4821–4830
[7] Gatti M 2015 Introduction to spectroscopy. lecture Note available at http://146.141.41.27/Lectures/Gatti-Monday-Intro-Spectrosscopy.pdf
[8] Salpeter E and Bethe H A 1951 *Phys. Rev.* **84** 1232
[9] Broido M M and Taylor J G 1969 *Journal of Mathematical Physics* **10** 184–209
[10] Dongho Nguimdo G and Joubert D P 2015 *Eur. Phys. J. B* **88**
[11] Laksari S, Chahed A, Abbouni N, Benhelal O and Abbar B 2006 *Comp. Mater. Sc.* **38** 223–230
[12] Nayebi P, Mirabbaszadeh K and Shamshirsaz M 2013 *Physica B* **416** 55–63
[13] Kresse G and Joubert D 1999 *Phys. Rev. B* **59** 1758
[14] Kresse G and Furthmüller J 1996 *Phys. Rev. B* **54** 169
[15] Hedin L 1965 *Physical Review* **139** A796
[16] Shishkin M and Kresse G 2006 *Phys. Rev. B* **74** 035101
[17] Perdew J P and Ruzsinszky A 2008 *Phys. Rev. Lett.* **100**(13) 136406
[18] Aguilera I, Vidal J, Wahnón P, Reining L and Botti S 2011 *Phys. Rev. B* **84** 085145
[19] Kohanoff J 2006 *Electronic structure calculations for solids and molecules* (Cambridge University Press)
[20] Bruneval F, Vast N and Reining L 2006 *Physical Review B* **74** 045102
[21] Körbel S, Kammerlander D and Sarmiento-Pérez 2015 *Physical Review B* **91** 075134
[22] Mishra S and Ganguli B 2011 *Solid State Commun* **151** 523–528

Collective Electronic Excitations in Ferromagnetic Metals

V Nolting

Vaal University of Technology, Vanderbijlpark

volkmarn@vut.ac.za

Abstract. Collective electronic excitations in the system of interacting conduction electrons of ferromagnetic metals (Fe, Co, Ni) are investigated. These conduction electrons stem from relatively narrow d-type bands and a suitable model to describe them is the Hubbard model. The tensor of the magnetic susceptibility is calculated within the Hubbard model. While the poles of the diagonal contribution represent the famous Stoner criterion for the stability of magnetic order the transverse susceptibility yields spin wave energies representing magnons. Magnetizations m as a function of temperature T and occupation number n, Curie temperatures T_C, and the temperature dependent exchange splitting ΔE_{ex} are numerically evaluated within reasonable agreement with experimental results.

1. Introduction

In ferromagnetic transition metals, e.g. Fe, Co, Ni both magnetism and the electric current are caused by the same electron group which stem from relatively narrow d-type conduction bands. Band ferromagnetism is a consequence of strong electron correlations where the correlation energy is the deviation of the exact ground state energy of the interacting electron system from the corresponding Hartree-Fock result and can only be approximately determined. Furthermore, correlation effects are not properly taken into account by standard band theory so that a full microscopic explanation of band ferromagnetism is still not available yet.

Even though density functional theory DFT is in principle an exact ground state theory its local density approximation LDA seems to underestimate electron correlation effects. The exchange splitting $\Delta E_{ex}\ (T = 0\ K)$ for example comes out far bigger than in the experiment. Also short range magnetic order in the paramagnetic phase is not reproduced in LDA in contrast to experimental results (e.g. Curie-Weiss law). The main shortcoming of LDA is its restriction to $T = 0\ K$. These difficulties could in principle be overcome by extending DFT to finite temperatures. An alternative strategy uses model Hamiltonians that are then approximately evaluated using many body techniques. A suitable model for band magnetism is the Hubbard model that is introduced in the following section. Collective electronic excitations in the interacting electron system are then investigated. One distinguishes between charge density waves (plasmons) and spin density waves (magnons). In this paper we concentrate on the latter which results in Section 3 into expressions for spin wave energies and the exchange splitting.

The exchange splitting is temperature dependent which is further investigated in Section 4 using a many body approach. From equations for the particle numbers $n_\uparrow$, $n_\downarrow$ results for the magnetization $m\ (T, n)$ and the Curie temperature $T_C\ (n)$ are derived. The results are compared with those of other methods.

2. The Model

Starting point for a theoretical description of ferromagnetic metals is the tight binding approximation which implies that both hopping integrals and Coulomb matrix elements are calculated from atomic wavefunctions. In the Hubbard approximation [1] only the dominant intraatomic matrix element is retained yielding the model Hamiltonian

$$H = \sum_{ij\sigma}\left(T_{ij} - \mu\ \delta_{ij}\right)\ a^{+}_{i\sigma}\ a_{j\sigma}\ + \frac{1}{2}\ U\ \sum_{i\sigma} n_{i\sigma}\ \ n_{i,-\sigma} \qquad (1)$$

Here $a^{+}_{i\sigma}$ denotes the creation operator for a σ-electron at lattice site R_i; $a_{i\sigma}$ is the corresponding annihilation operator. U is the intraatomic Coulomb matrix element mentioned above while the T_{ij} describe hopping integrals. Note that in the tight binding approximation the indices i, j are restricted to nearest neighbors only.

μ denotes the chemical potential that is both T and n-dependent. The temperature dependence follows from a statistical description of the thermodynamic properties of an ideal Fermi gas, i.e.

$$\mu\,(T)\ \cong\ \varepsilon_F\ \left(1 - \frac{\pi^2}{12}\ \left(\frac{k_B\ T}{\varepsilon_F}\right)^2\right) \qquad (2)$$

Note that for simple metals, e.g. Na, Cu, Au, etc at room temperature the ratio

$$\frac{k_B\,T}{\varepsilon_F}\ \sim\ 5\ \cdot\ 10^{-3}$$

So, the temperature dependence of μ becomes practically negligible. On the other hand, the n-dependence of μ is at low temperatures reasonably well described by the Stoner result

$$\mu\,(\,T = 0)\ =\ \varepsilon_F + \frac{1}{2}\ U\ n \qquad (3)$$

As a first approximation one may regard the d-band degeneracy as not particularly decisive and limit the calculation to narrow s-bands. The Coulomb interaction then obviously only acts if both electrons occupy the same Wigner-Seitz cell; these two electrons must then have opposite spins. The Hubbard model is thus the simplest model to describe both electronic and magnetic properties of the transition metals. It explains ferromagnetism on the basis of a spin-dependent band shift between the ↑ and ↓-density of states below the Curie temperature T_C [2]. This is further investigated in Fig 1 that schematically depicts the density of states for both ↑ and ↓-electrons according to the Stoner model.

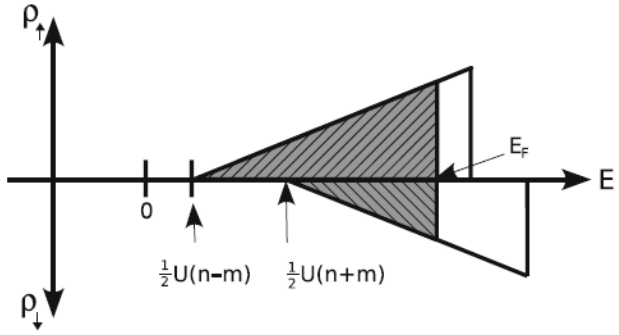


Fig 1: Schematic plot of the spin-dependent density of states $\rho_\uparrow$ and $\rho_\downarrow$ for a ferromagnetic metal according to the Stoner model. The figure is adopted from reference [3].

The bands are rigidly shifted against each other by an energy amount of $\Delta E = U\,m$; it will later be shown that ΔE defines the exchange splitting. The exchange splitting is temperature dependent and vanishes above T_C; this temperature dependence is further investigated in the next section. Furthermore, as both spin bands are filled up to the Fermi level ε_F it follows $n_\uparrow > n_\downarrow$ and thus a spontaneous magnetization $m = n_\uparrow - n_\downarrow > 0$ is observed. Despite its simplicity the Hubbard model remains a non-trivial many body problem and is in general not exactly solvable. Approximate solutions confirm a possible collective magnetic order under certain conditions for the external parameters, i.e. Coulomb coupling U/W where W denotes the band width, the lattice structure, and the band occupation n. M

Ulmke [4] numerically proved magnetic order within the model at finite temperatures in an extended parameter region using Quantum Monte Carlo calculations. W Nolting [5] points out that a correct inclusion of Fermi liquid properties is essential to obtain ferromagnetic solutions within the Hubbard model. The effect of electron correlations on the stability of magnetic order in the Hubbard model is also discussed in reference [6] where it is found that ferromagnetic order is generally suppressed by correlation effects. On the other hand, the Hartree-Fock approximation does not include the correlation energy and thus overestimates the strength of the ferromagnetic order.

Other interesting solutions discussed so far in the literature include
-The Mermin-Wagner theorem rules out ferromagnetic order at finite temperatures for dimensions $d \leq 2$ [7].
-For small band occupations n only a paramagnetic solution

$$n_{\uparrow} = n_{\downarrow} = \frac{1}{2} n$$

is observed. This agrees with the exact result of Kanamori [8].
-In the strong coupling limit $U \gg W$ a saturated ferromagnetic solution is expected for large band occupations n in agreement with the results of Nagaoka [9,10].
-In the zero bandwidth limit where $T_{ij} \rightarrow T_0 \ \delta_{ij}$ no ferromagnetic solutions are observed [11], the conventional explanation being that quasiparticle damping prevents any spontaneous magnetic order.
Current techniques used to improve the understanding of the microscopic origin of ferromagnetism in metals include the development of new mathematical methods, e.g. dynamic mean field theory or density matrix renormalizations [12,13] or alternatively new ab initio schemes [14]. Recent investigations deal with a possible linkage between magnetic order and structural phase transitions in Fe, Co, Ni [15]. The authors of ref [15] report a pressure induced suppression of magnetic order in Fe where the magnetic phase transition is accompanied by a structural phase transition from bcc ↔ hcp. The stability of the bcc-phase is due to the ferromagnetic order. Besides its application to band magnetism the Hubbard model is also successfully applied to describe metal insulator transitions and high temperature superconductivity [16, 17].

3. Spin Density Waves

Besides charge density waves another form of collective excitations in ferromagnetic metals is caused by the existence of the electron spin. The tensor of the magnetic susceptibility is determined from the Green functions involving the components of the spin operator S_i. The diagonal or longitudinal contribution $X_q^{zz}(E)$ yields information regarding the stability of the magnetic order. Within the Hubbard model it has the form

$$X_q^{zz}(E) = -\frac{\mu_0 \mu_B^2}{\hbar NV} \quad \frac{X_0(q,E)}{1-\frac{U}{2N\hbar} X_0(q,E)} \tag{4}$$

Here

$$X_0(q,E) = 2N\hbar \ \rho_0(\varepsilon_F) \tag{5}$$

denotes the temperature independent Pauli spin susceptibility; $\rho(\varepsilon_F)$ is the density of states at the Fermi level. Inserting Eq (5) into (4) yields

$$X_q^{zz}(E) = -\frac{\mu_0 \mu_B^2}{2V} \quad \frac{\rho_0(\varepsilon_F)}{1-U\,\rho_0(\varepsilon_F)} \tag{6}$$

The poles of the susceptibility represent the famous Stoner criterion

$$U \ \rho_0(\varepsilon_F) \geq 1 \tag{7}$$

regarding the stability of magnetic order. According to Eq (7) ferromagnetic order becomes energetically favourable if there is, firstly a large intraatomic Coulomb interaction U as this maximizes the gain in potential energy and secondly a large density of states at the Fermi level. In that case a lot of ↓-electrons can flip their spin and become ↑-electrons without increasing the kinetic energy of the system too much.

On the other hand, the transverse susceptibility $X_q^{+-}\ (E)$ is computed within the Stoner model as

$$X_q^{+-}(E) = \frac{\gamma}{N} \sum_k \frac{\langle n_{k+q,\downarrow}\rangle - \langle n_{k\uparrow}\rangle}{E - \Delta E_{\uparrow\downarrow}(k,q)} \quad (8)$$

The poles are identical to the spin flip excitation energies

$$\Delta E_{\uparrow\downarrow}(k,q) = \varepsilon(k+q) - \varepsilon(k) + U\,m \quad (9)$$

describing transitions between the two spin bands $E_\sigma(k)$. In the case of the Stoner model the quasiparticle energies $E_\sigma(k)$ are plotted as a function of wavevector k in Figure 2 below.

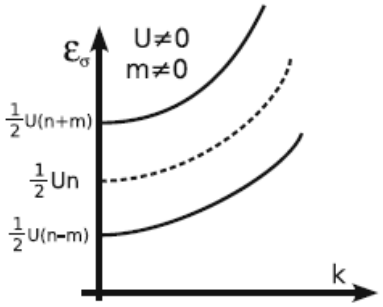


Figure 2: Quasiparticle energies $E_\sigma(k)$ as a function of wavevector k according to the Stoner model. Depicted is the most general case $U \neq 0, m \neq 0$. The figure is adopted from reference [3].

Particularly interesting is the case $U \neq 0, m \neq 0$ where the rigid energy difference between the two bands defines the exchange splitting

$$\Delta E_{ex} = E_\downarrow - E_\uparrow = U\,m \quad (10)$$

Note that the exchange splitting is temperature dependent. The $T = 0$-value can be used to estimate the Curie temperature T_C of the metal The parameter U is first fixed via the exchange splitting and the corresponding value is then inserted into the result from the strong coupling limit $U \gg W$, i.e.

$$k_B\ T_C \cong \frac{1}{4}\ U \quad (11)$$

The theoretically calculated values appear in the fifth column of Table 1 below and are compared with experimental values from the literature. Table 1 shows that our simple model calculation yields surprisingly accurate results for the Curie temperature T_C especially in the case of Ni and Co. Only for Fe the value for the Curie temperature is slightly overestimated due to the large exchange splitting at $T = 0\ K$. However, the proportionality $\Delta\ E_{ex} \sim m$ is not confirmed by experimental results.

metal	$m(T=0)$	$\Delta\ E_{ex}\ (T=0)$	T_C (exp. result)	T_C (theoret. result)
Fe	2.22 μ_B	2 eV	1043 K	1320 K
Ni	0.56 μ_B	0.35 eV	631 K	780K
Co	1.7 μ_B	1.5 eV	1388 K	1280K

Table 1: Experimental and theoretical values for magnetic properties of ferromagnetic metals

The Stoner model additionally suggests that the exchange splitting vanishes in the paramagnetic phase $T > T_C$. Especially in the case of Fe a persistent exchange splitting above T_C has been experimentally observed. The Stoner model also predicts a Pauli like susceptibility $X\ (T) = const$ in the paramagnetic phase with no indication of a Curie-Weiss behaviour again in contradiction to experimental evidence. One reason for the deficiencies of the Stoner model at finite temperatures is its suppression of spin waves. Spin waves in ferromagnetic metals are experimentally detected by inelastic neutron scattering. Expected is therefore at low temperatures $T \rightarrow 0K$ a deviation of the magnetization from saturation of the form

$$\frac{m\,(T)}{m_0} \sim 1 - C\,T^2 \qquad (12)$$

Eq (12) is, however, not reproduced by the Stoner model.

4. Results

Using the thermodynamic properties of an ideal Fermi gas the particle numbers $n_\uparrow$ and $n_\downarrow$ are calculated from the quasiparticle density of states $\rho_\sigma\,(E)$, i.e.

$$n_\sigma = \frac{N_\sigma}{N} = \int_{-\infty}^{+\infty} dE \;\; f_-\,(E,T) \;\; \rho_\sigma\,(E)$$

Here

$$f_-\,(E,T) = \frac{1}{e^{\beta(E-\mu)}+1}$$

denotes the Fermi function. As $N_\uparrow + N_\downarrow = N$ it follows $n_\uparrow + n_\downarrow = 1$. The magnetization is then defined as

$$m = n_\uparrow - n_\downarrow$$

and is plotted in Figure 3 below as a function of $k_B\,T$ with $n = const.$ Note that in the Stoner model Eq (11) applies. Figure 3 shows the typical Brillouin function behaviour for the magnetization $m\,(T)$ which shows that qualitatively the Stoner model describes the phase transition ferromagnetism ↔ paramagnetism reasonably well. The small kink in the magnetization curve close to T_C could be due to the fact that the particle dependence of the chemical potential according to Eq (3) is strictly only valid at small temperatures. Interesting is also the question in which region of n-values ferromagnetic solutions $m \neq 0$ can be expected in the Hubbard model. This is worked out in Figure 4.

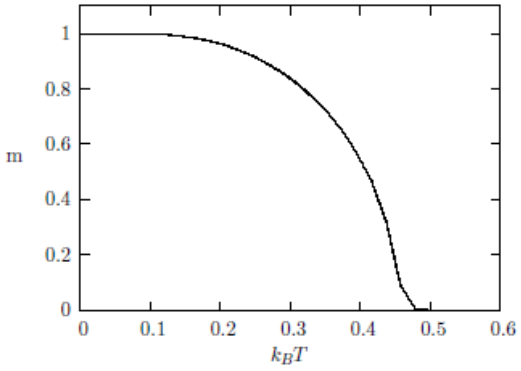


Figure 3: Magnetization m of a ferromagnetic metal as a function of $k_B\,T$ at fixed particle number n.

Note that for all band occupations n there is always a paramagnetic solution

$$n_\uparrow = n_\downarrow = \frac{1}{2}\,n$$

Additionally for particle numbers $n > 0.5$ (half- filled band) there is also a ferromagnetic solution with $m = n$ describing ferromagnetic saturation and similar findings are also reported for example in reference [18]. For small band occupations n a ferromagnetic solution is not possible as μ lies within the band and a solution is only obtained if $n_\uparrow = n_\downarrow$. This is consistent with the exact results of Kanamori [8].

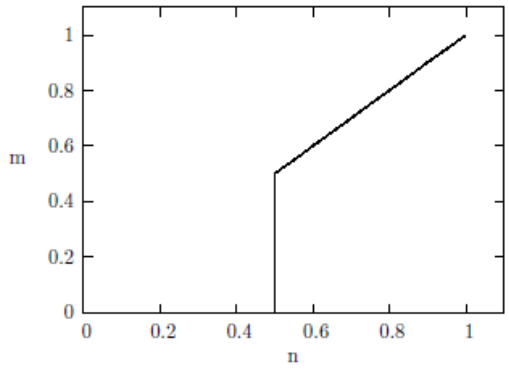


Figure 4: Magnetization m as a function of particle number n at fixed temperature $T = 0$.

The Curie temperature T_C depends on the strength of the ferromagnetic coupling which is reflected in the result

$$k_B\ T_C = \frac{W}{4\ \tanh^{-1}(W/U)} \tag{13}$$

In the strong coupling limit Eq (13) then reduces to Eq (11) meaning $k_B\ T_C \sim U$. A similar increase of T_C with U has also been obtained by other authors and qualitatively agrees with experimental results [19].

5. Conclusions

In this work collective electronic excitations in ferromagnetic metals (Fe, Co, Ni) are investigated. Using the Hubbard model as a theoretical description the exchange splitting between the ↑ and ↓ density of states is calculated within an RPA-approximation from the tensor of the magnetic susceptibility. A numerical evaluation of the magnetization $m\ (T, n)$ shows that our model calculation describes the phase transition ferromagnetism ↔ paramagnetism reasonably well. Ferromagnetic solutions are only obtained for band fillings $n > 0.5$ describing a half-filled band. The Curie temperature T_C comes out surprisingly accurate, and its dependence on the intraatomic Coulomb interaction U qualitatively agrees with experimental results.

References

[1] J. Hubbard, Proc Royal Society (London) A276, 238 (1963)
[2] W Nolting et al, arXiv: cond-mat/ 0107255 (2001)
[3] W Nolting and A Ramakanth, Quantum Theory of Magnetism, Springer (2009)
[4] M. Ulmke, Eur Phys J B1, 301 (1998)
[5] W Nolting et al, in Band Ferromagnetism, Springer (2001)
[6] P.A. Igoshev et al, arXiv: 1502.05839v1 (2015)
[7] D.K. Ghosh, Phys Rev Letters 27, 1584 (1971)
[8] J. Kanamori, Progr Theor Phys 30, 275 (1963)
[9] W. Nolting, Z Physik 255, 25 (1972)
[10] Y Nagaoka, Phys Rev 147, 392 (1966)
[11] H. Fukuyama and H. Ehrenreich, Phys Rev B7, 3266 (1973)
[12] W. Metzner and D. Vollhardt, Phys Rev Letters 62, 324 (1989)
[13] S. Daul and R. Noack, Phys Rev B58, 2635 (1998)
[14] V.I. Anisimov et al, Phys Cond Matt 9, 7359 (1997)
[15] V. Ista et al, Appl Phys Lett, 90, 42505 (2007)
[16] T. Yanagisawa, J Physics Conference Series 108, 12010 (2008)
[17] A. Isidori and M. Capone, Phys Rev B80, 115120 (2009)
[18] W Nolting and W Borgiel, Phys Rev B39, 6962 (1989)
[19] W Nolting, Theoretical Physics, Vol 7 Many Body Theory, Springer (2005)

Dependence of the photoionization cross-section of α-Al_2O_3:C on the measurement temperature

A. Nyirenda, M.L. Chithambo

Department of Physics and Electronics, Rhodes University, P.O Box 94, Grahamstown 6140, South Africa

E-mail: anyirenda@gmail.com

Abstract. We report the temperature dependence of the photoionization cross-section, that is, the effective area of interaction between incident photons and charge trapping states in α-Al_2O_3:C, a highly sensitive dosimetric material. Samples were exposed to 1.0 Gy of beta irradiation followed by measurement of linearly-modulated optically stimulated luminescence under 470 nm blue LED stimulation. The technique involves the linear increase of the power of stimulating signal from a minimum to some maximum value for which the luminescence appears as a spectrum of overlapping time-dependent peaks. The apparently single peak comprises of at least two components henceforth referred to as fast and slow components as determined by analytical deconvolution of the original spectrum. The position of each of the component peaks shifts to earlier times with increasing measurement temperature. In addition, the photoionization cross-section for both fast and slow components increases, due to increased electron-phonon coupling, from $(1.18 - 1.33) \times 10^{-18}$ cm^{-2} and $(4.05 - 5.22) \times 10^{-19}$ cm^{-2} respectively as measurement temperature increases from 30°C to 100°C.

1. Introduction

α-Al_2O_3:C is highly sensitive to ionizing radiation and is currently used for thermoluminescence and optically stimulated luminescence (TL/OSL) dosimetry specifically for personal and environmental monitoring. Thermoluminescence (TL) is a stimulated luminescence phenomenon which involves a non-isothermal scan of a pre-irradiated luminescent material to produce a spectrum of one or more temperature resolved peaks in luminescence intensity known as a 'glow-curve'. A TL peak is associated with the presence of certain species of point defects which may act as electron or hole traps during excitation of the material by an external energy source e.g. ionizing radiation. When the stimulating energy is optical in nature, the phenomenon is referred to as optically stimulated luminescence (OSL). OSL methods of dosimetry are preferred to TL methods since the former enable multiple read-outs of the OSL signal unlike the latter.

Linearly modulated optical stimulation of luminescence (LM-OSL) is an OSL method used for probing luminescence materials. In LM-OSL, the stimulating power from the light source is linearly increased from a minimum to some maximum value while monitoring the OSL throughout the whole stimulation period. The resulting OSL spectrum appears as a series of peaks with each one representing a component of the OSL signal with particular photoionisation cross-section [1]. The photoionisation cross-section is perhaps the most important parameter that governs trap ionisation and dictates the stability of a particular trap during optical stimulation [2]. It follows that at a particular stimulating wavelength, an electron trap with

a relatively large photoionisation cross-section will empty quickly and produce peaks at earlier times in the resultant curve. Thus, the LM-OSL signal allows more effective and accurate characterisation of each OSL component [1].

For first order kinetics i.e. negligible charge recapture at a trap following stimulation, the instantaneous LM-OSL intensity, $L(t)$, is given by the following equation:

$$L(t) = \gamma\ t\ \Sigma_i\ n0_i\ p_i \exp\left(-\frac{\gamma\ p_i\ t^2}{2}\right);\ \gamma = T_{stim}^{-1} \tag{1}$$

where γ is the stimulation power ramping rate, T_{stim} is the total stimulation time, $n0_i$ is the initial trapped charge population in the i^{th}trap is the detrapping probability, t is time in seconds. The summation is carried over all the component peaks [2]. The photoionization cross-section, σ_i, is obtained from the detrapping probability by using the relationship

$$p_i = \sigma_i\ \Phi \tag{2}$$

where Φ is the maximum stimulation photon flux. The maximum stimulation photon flux is calculated as

$$\Phi = \frac{maximum\ power\ density}{energy\ per\ photon} \tag{3}$$

where the energy per photon is hc/λ and h is the Planck's constant, c is the speed of light in vacuum. For blue LEDs ($\lambda = 470$ nm) with a maximum power density of 80 mWcm^{-2} at sample position, $\Phi = 1.89 \times 10^{17}$cm^{-2} s^{-1}.

The peak position, $t_{max,i}$, is given by

$$t_{max,i} = \sqrt{\frac{1}{p_i\ \gamma}} = \sqrt{\frac{T_{stim}}{\sigma_i\ \Phi}} \tag{4}$$

In cases where charge retrapping during optical stimulation is significant i.e. general order kinetics, the LM-OSL intensity is given by

$$L(t) = \gamma\ t\ \Sigma_i\ n0_i\ p_i \left[(b_i - 1)\frac{p_i\ \gamma\ t^2}{2} + 1\right]^{\frac{b_i}{1-b_i}} \tag{5}$$

where b_i is the kinetic order of the i^{th} component peak [2].

In this paper, we present the dependence of the photoionization cross-section of α-Al_2O_3:C on measurement temperature and the effect of this on other parameters that are related to the photoionization cross-section. Thus, this study represents an effort to characterize point defects involved in the OSL process.

2. Experimental details

All measurements were done using the Risø TL/OSL DA-20 Reader that measures both TL and OSL. The light detection unit consists of an EMI 9235QB photomultiplier tube and a 7-mm thick Hoya U-340 detection filter (transmission 270-380 nm FWHM). The irradiation unit is an in-built ^{90}Sr/^{90}Y β-source with a nominal dose rate of 0.1028 Gy/s. Samples used were aluminium disks measuring 5 mm × 1 mm (Rexon TLD Systems, Ohio, USA). The samples were exposed to 470 nm blue LED light whose stimulation power was linearly increased from 0 to 100% of the maximum stimulation power to produce an LM-OSL signal. The maximum power intensity at the sample position by blue LEDs was 80 mWcm^{-2}. A TL was then measured

at a heating rate of 1 °C/s to record the residual TL signal. All measurements were performed in a nitrogen atmosphere to inhibit spurious luminescence.

Samples were irradiated to 1.0 Gy of β-radiation followed by exposure to blue LEDs for 1000 s. A TL measurement was then taken at 1 °C/s to clear the residual dose i.e. trapped charge 'spared' during optical stimulation. The LM-OSL signal obtained was fitted with equation 1 for $i = 2$. The fitting was perfomed using a "NonlinearModelFit" function in Mathematica which uses the Levenburg-Marquadt algorithm [3] for nonlinear least squares curve-fitting. The goodness of fit was tested by the value of regression squared (R^2).

3. Results and Discussions

3.1. Deconvolution of the LM-OSL signal

Figure 1(a) shows the result of deconvoluting the LM-OSL signal (labelled 1) using equation 1 to produce two components of the LM-OSL peak namely, fast component (labelled 2) and slow component (labelled 3) and Figure 1(b) is a plot of the concomittant residuals. Figure 2(a) is the TL glow-curve before exposing the samples to LEDs and figure 2(b) is the glow-curve obtained due to residual dose after light exposure.

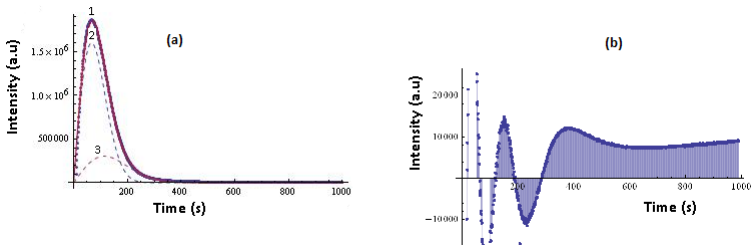


Figure 1. A deconvoluted LM-OSL signal (labelled 1) showing the fast component (labelled 2) and the slow component (labelled 3) (a) and the concomittant residuals (b), $R^2 = 0.9978$. The sample was irradiated to 1.0 Gy and the stimulating power was increased from 0 to 100% at a measurement temperature of 30°C.

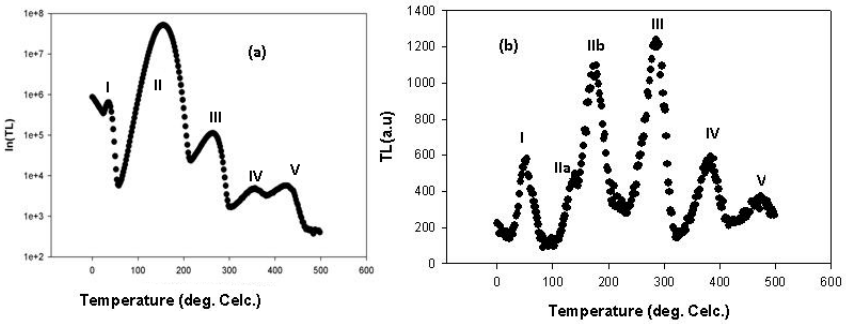


Figure 2. A TL glow curve before exposure to blue LEDs shown on a semi-log scale (a). Residual TL obtained immediately after exposing a sample to blue LEDs (b). The TL was, in each case, measured at 1°C/s after a 1.0 Gy dose.

Peak II in figure 2a is the main peak in α-Al_2O_3:C that is used for dosimetry applications. It has been reported that peak II is correlated to the LM-OSL peak [5, 6]. Consequently, we

associate the fast OSL component (labelled 2) and the slow OSL component (labelled 3) in figure 1 of the LM-OSL signal to peaks IIa and IIb respectively in figure 2b of the residual TL signal. Considering that the fast component occurs at earlier times than the slow component, it is expected that the photoionization cross-section of the fast component would be larger than that of the slow component. Other researchers [5, 7] fitted the LM-OSL signal with three first-order kinetics components. In our case, fitting with three first-order kinetics components did not yield meaningful parameters' values. The presence of only two first-order components in our investigations can be attributed to either differences in samples used for investigations or the maximum time of sample exposure to blue LEDs since the third component, according to Whitley and McKeever [5] and Dallas et al [7], has its maximum peak position beyond 2000 s whereas the maximum time of 1000 s was used in this study. However, looking at the residual plot, there is a low-intensity persistent signal that does not decay with time, but rather flattens out. This persistent signal could be the third component reported by Whitley and McKeever [5] and Dallas et al [7]. We attribute the fast and slow components obtained in this study to the double-component nature of the TL main peak as seen in figure 2(a). We associate the fast component with easy-to-bleach low temperature component (IIa in figure 2(a)) of the main peak. The slow component which we associate with high temperature component (IIb in figure 2(a)), may be a result of the substantially smaller photoionization cross-section rather than retrapping. We can speculate that the third component as observed by other researchers [5, 7] could be due to OSL from deep traps obtained from direct recombination of stimulated electrons at the luminescence centres or indirectly after retrapping at the main and shallow traps.

3.2. Dependence of photoionization cross-section on measurement temperature

The detrapping probabilities were obtained by fitting the experimental data with equation 1 and the corresponding values of photoionization cross-section were evaluated using equation 2.

Figure 3 shows plots, against measurement temperature, of the detrapping probability (in a) and the photoionization cross-section (in b) for the fast component. Figure 4 shows plots, against measurement temperature, of the detrapping probability (in a) and the photoionization cross-section (in b) for the slow component.

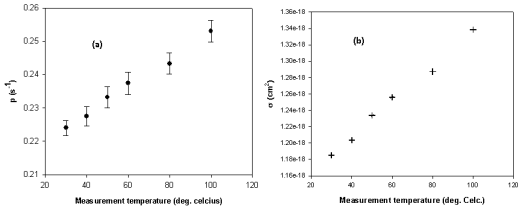


Figure 3. Detrapping probability against measurement temperature for the fast component (a). Photoionization cross-section against measurement temperature for the slow component (b).

Figures 3 and 4, representing the fast component and the slow component respectively, show that both the detrapping probability and the photoionization cross-section increase with measurement temperature. The photoionization cross-section for the fast component increases from 1.1852×10^{-18} cm^{-2} at 30°C to 1.338×10^{-18} cm^{-2} at 100°C representing a 13%increase. As shown in Figure 3, the photoionization increases almost linearly in the temperature range under investigation. On the other hand, the increase in photoionization cross-section for the slow component is nonlinear. The photoionization cross-section for the slow component increases from 4.0529×10^{-19} cm^{-2} at 30°C to 5.225×10^{-19} cm^{-2} at 100°C representing a

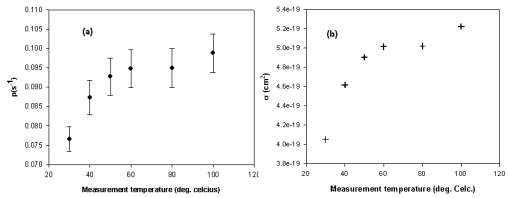


Figure 4. Detrapping probability against measurement temperature for the fast component (a). Photoionization cross-section against measurement temperature for the slow component (b).

29%increase. Whitley and McKeever [5] reported values of $(0.58-1.5)\times10^{-18}$ cm^{-2} for the first component (fast component herein) and $(0.9-3.7)\times10^{-19}$ cm^{-2} for the second component (slow component herein) at a measurement temperature of 70°C, which are consistent with the values reported here. However, Whitley and McKeever [5] used samples obtained from Landauer Inc. (Oklohama, USA) and Bicron NE (Ohio, USA). Furthermore, Whitley and McKeever [5] used a stimulation wavelength of 526nm (green) as compared to 470nm (blue) used in our investigations. The increase in the detrapping probability, which translates into the increase in photoionization cross-section, may be ascribed to a phonon assisted stimulation mechanism.As the measurement temperature increases, electron-phonon interactions become slightly stronger.This affects the detrapping probability and consequently, the photoionization cross-section. The temperature dependence of the electron-phonon interaction can be inferred from the following equation [8]:

$$S = S_0 Coth\left[\frac{\hbar\omega}{kT}\right], \tag{6}$$

where S is the temperature dependent electron-phonon interaction factor ($\gg 1$ for strong coupling), S_0 is the Huang-Rhys factor at absolute zero, k is the Boltzmann's constant, $\hbar$ is the Planck's constant, T is the absolute temperature, ω is angular frequency, and $\hbar\omega$ is equal to phonon energy. Chruścińska [9], using simulations, also showed that the photoionization cross-section increases with temperature, without considering any particular material.

3.3. Effect of temperature on LM-OSL peak position

Figure 5(a) is a plot of the peak position, t_{max}, against measurement temperature for the fast component whereas figure 5(b) is a similar plot for the slow component. As expected, the peak position moves to earlier times as the measurement temperature increases due to increase in photoionzation cross-section. A large photoionization cross-section ensures that traps are emptied quickly.

Figure 6(a) shows the plot of t_{max} against $1/\sqrt{\sigma}$ for the fast component and figure 6(b) is a similar plot for the slow component. From equation 4, t_{max}=$const/\sqrt{\sigma}$ where $const = \sqrt{T_{stim}/\Phi}$ is the proportinality constant. Thus, a plot of t_{max} against $1/\sqrt{\sigma}$ is expected to produce a straight line with slope *const*. For figure 6(a), $const = 7.22\times10^{-8}$ whereas $const = 7.30\times10^{-8}$ for figure 6(b) which are consistent. For the values used in this study i.e. $T_{stim} = 1000$s and $\Phi = 1.89\times10^{17}$ cm/s, $\sqrt{T_{stim}/\Phi} = 7.27\times10^{-8} \approx const$ as expected.

4. Conclusion

The dependence of the photoionization cross-section on measurement temperature in α-Al_2O_3:C has been investigated. The apparently single LM-OSL peak is a composition of at least two components i.e. the fast and a slow one. The photoionization cross-section of both

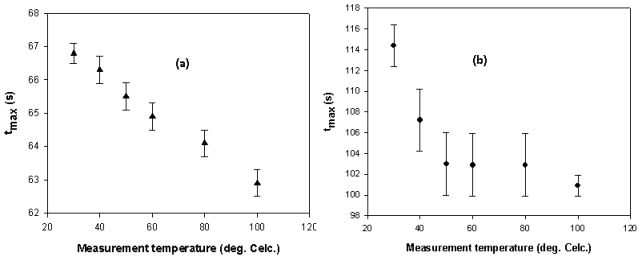


Figure 5. Peak position against measurement temperature for the fast component (a) and a similar plot for the slow component (b).

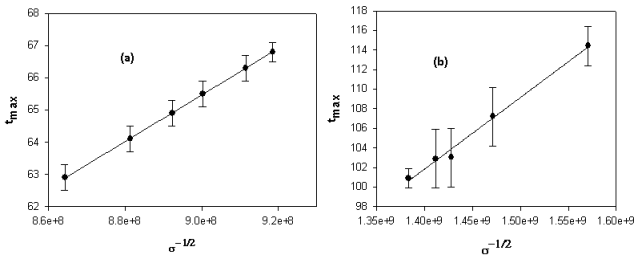


Figure 6. Peak position against the reciprocal of the square root of photoionization cross-section for the fast component (a) and a similar plot for the slow component (b).

components increases with measurement temperature, a behaviour that has been associated with increased electron-phonon coupling as temperature increases. As the measurement temperature is increased, the peak positions shift to earlier times due to increasing photoionization cross-section.

References

[1] Choi, J.H, Duller, G.A.T, Wintle, A.G, 2006 Analysis of Quartz LM-OSL signal Ancient TL **24** (9-20)
[2] Bøtter-Jensen, L, McKeever, S.W.S, Wintle, A.G, 2003 Optically stimulated luminescence dosimetry *Elsevier* (Amsterdam)
[3] Marquadt W.D., 1963 An algorithm for least-squares estimation of nonlinear parameters *J. SOC. INDUST. APPL., MATH* **11** 431-441
[4] Kitis, G, Pagonis, V, 2008 Computerized curve deconvolution analysis for LM-OSL. *Radiat. Meas.* **43** 737-41
[5] Whitley, V.H and McKeever, S.W.S, 2002 Linear modulation optically stimulated luminescence and thermoluminescence techniques in Al_2O_3:C. *Radiat. Prot. Dos.* **100** (61-6)
[6] Walker, F.D, Colyott, E, Agersnap Larsen, N, McKeever, S.W.S, 1996 The wavelength dependence of light-induced fading of TL from α-Al_2O_3:C *Radiat. Meas.* **26** (711-8)
[7] Dallas, G.I, Polymeris, G.S, Stefanaki, E.C, Afouxenidis, D, Tsirliganis, N.C, Kitis, G, 2008 Sample dependent correlation between TL and LM-OSL in Al_2O_3:C *Radiat. Meas.* **43** (335-40)
[8] Hayes, W., Stoneham, A.M., 1985, Defects and defect processes in nonmetallic solids *John Wiley and Sons, NewYork*
[9] Chruścińska A, 2010. On some fundamental features of optically stimulated luminescence measurements *Radiat. Meas.* **45** (991-9)

Simultaneous substitution of Ba, Mn and Co into Fe_3O_4 spinel structure: magnetic and electrochemical sensing properties of the synthesized nanoparticles

Nadir S E Osman*[1], Neeta Thapliyal[2], Thomas Moyo[1] and Rajshekhar Karpoormath[2]

[1]School of Chemistry and Physics, Westville campus, University of KwaZulu-Natal, Durban 4000, South Africa
[2]Department of Pharmaceutical Chemistry, Discipline of Pharmaceutical Sciences, College of Health Sciences, University of KwaZulu-Natal, Durban 4000, South Africa

E-mail: osmann@ukzn.ac.za

Abstract. Single phase structured $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ nanoferrite was synthesized. The phase formation was confirmed by X-ray diffraction. The morphology and particle distribution of the synthesized nanoferrite particles were observed using transmission electron microscopy. The magnetic hyperfine parameters were obtained from room temperature ^{57}Fe Mössbauer spectroscopy. Vibrating sample magnetometer measurements were performed to investigate the magnetic properties of the sample. The electrochemical properties of the sample were also investigated. Cyclic voltammograms of ferricyanide oxidation showed that the modified electrode by $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ nanoparticles exhibited improved electrochemical activity as compared to the bare electrode. These high-performance electrodes are expected to lead to the development of a novel group of electrochemical sensors.

1. Introduction

Spinel ferrites have several properties [1] which make them useful in many applications such as in photoelectric devices [2], microwave devices [3, 4], catalysis [5] and sensors [6]. Electrical and magnetic properties of spinel ferrites depend on their chemical composition, cation distribution and synthesis method. In this respect numerous efforts have been undertaken in order to improve the basic properties of spinel ferrites [7]. It is known that by substituting different ions in spinel ferrite structure, their magnetic and electrical properties can be changed [8-9]. In the present work, we have synthesized a single phase structure of $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ ferrite nanoparticles and investigate the associated structural, magnetic and electrochemical properties.

2. Experimental details

The $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ ferrite nanoparticle sample was produced by glycol-thermal technique using a Watlow series model PARR 4843 stirred pressure reactor. The starting materials were high purity ($BaCl_26H_2O$: 99 %, $MnCl_24H_2O$: 99 %, $CoCl_24H_2O$: 98 % and $FeCl_2.6H_2O$: 99 %) all purchased from Sigma-Aldrich. A well-mixed solution of the starting materials was precipitated with NH_4OH solution until a PH of 9.0 was reached. Details of the synthesis method have been reported elsewhere [10]. The phase and structural characterizations of the sample was obtained by a Phillips X-ray diffractometer type Model: PANalytical, EMPYREAN using CoKα radiation. The morphology and micro-structure of

the nanoparticles were investigated by high-resolution transmission electron microscopy (HRTEM) (type: Jeol_JEM-1010). ^{57}Fe Mössbauer spectra were obtained by a conventional spectrometer using a ^{57}Co source sealed in Rh matrix and vibrated at constant acceleration. Magnetic measurements were performed at room temperature using a LakeShore model 735 vibrating sample magnetometer (VSM). Potassium ferricyanide and potassium chloride were supplied by Merck. The voltammetric experiments were conducted on an electrochemical analyzer (800B Series, CH Instruments, Inc.) in a three-electrode electrochemical cell, consisting of a Ag/AgCl reference electrode, a platinum counter electrode $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$/GCE and a bare/modified glassy carbon working electrode. The modified electrode was prepared by drop-casting a 0.5 mg/mL solution of $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ nanoparticles in dimethylformamide (DMF) solvent onto the electrode surface.

3. Results and discussion

Figure 1(A) shows XRD patterns of the synthesized nanoparticles. All the peaks are well indexed based on spinel structure (JCPDS file No. 22-1086) [11]. No impurity peaks were observed on the XRD pattern. The highest peak intensity in Figure 1 was identified as 311 and was used to calculate the lattice parameter a using Bragg's equation $a = d(h^2 + k^2 + l^2)^{1/2}$ where d is the inter-planar spacing and hkl are the Miller indices [12]. The average crystalline size D was calculated using Scherrer's formula $D = 0.9\lambda / \beta \cos\theta$, where β is the full-width at half-maximum of the (311) XRD peak and θ is the Bragg's angle [13]. The microstrain ζ was determined using the Williamson-Hall plot [14]. The calculated values of the lattice parameter, crystallite size and microstrain for the as-prepared $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ were found to be 0.839±0.003 nm, 8.37±0.06 nm and 0.0015±0.00012, respectively. The morphology of the synthesized nanoparticles was investigated by high resolution transmission microscopy (HRTEM). The synthesized nanoparticles show a degree of shape uniformity. Furthermore, some agglomeration is observed as it seen from Figure 1(B). The average particles size of 9 ± 2 nm was estimated from the HRTEM image.

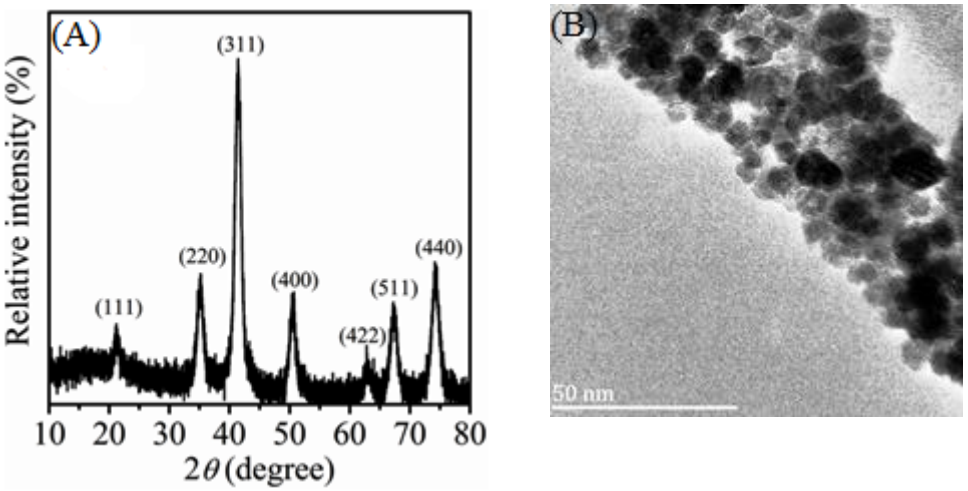


Figure 1. (A) X-ray diffraction patterns and (B) HRTEM images for as-prepared $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ ferrite nanoparticles.

Figure 2 shows the magnetic properties of the as-prepared $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ investigated at room temperature by Mössbauer spectroscopy and magnetization measurement. In Figure 2(A), the

Mössbauer spectrum is fitted by two Zeeman sextets which are related to the iron ions in tetrahedral A and octahedral B sites [15]. The s-shape magnetization is consistent with superparamagnetic behavior.

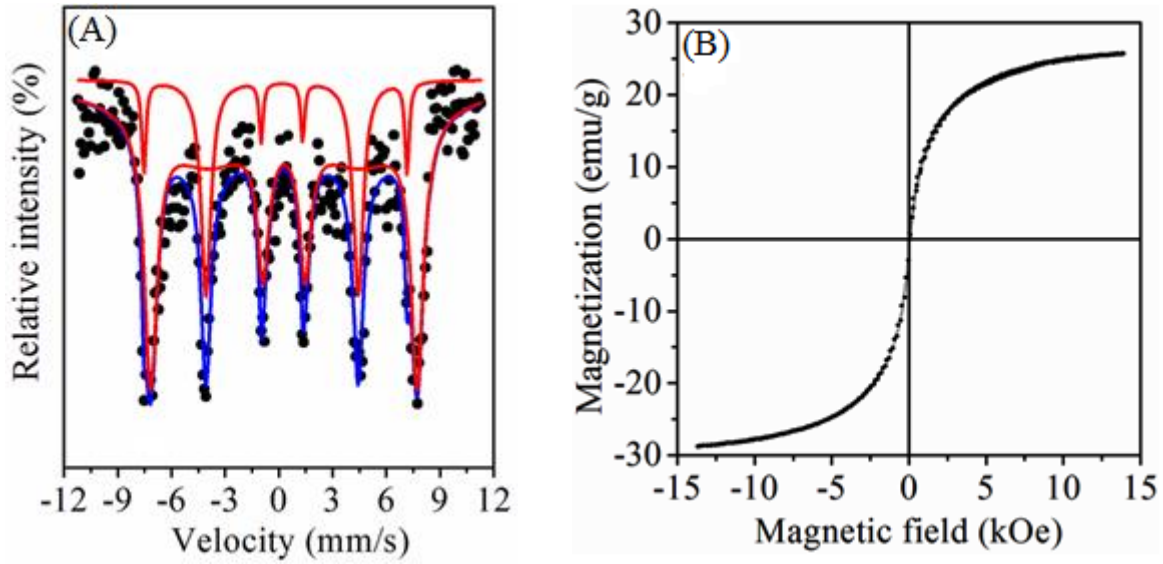


Figure 2. Room temperature Mössbauer measurement (A) and typical magnetic hysteresis loop measure at room temperature (B).

The magnetic hyperfine fields (H) were calculated to be 456±3 kOe and 462±3 kOe at A- and B-sites, respectively. The values of H are expected to be high due to super-exchange interactions between the atomic moments of the sample [16]. The isomer shift value at B-site is greater than at A-site. This is because of the bond separation between O^{2-} and Fe^{3+} is larger at B-site [17]. Hence the overlapping of Fe^{3+} ion orbitals is smaller at B-site. The magnetic hyperfine parameters are presented in Table 1.

Table 1. Isomer shifts (δ), hyperfine magnetic fields (H), line widths (Γ) and Fe^{2+} fraction population (f) on A-site and Fe^{3+} on B- site for $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ nanoparticles.

Site	δ (mm/s)	H (kOe)	Γ (mm/s)	f (%)
A	0.09±0.0	456±3	0.13±0.04	18.4
B	0.39±0.1	462±3	0.38±0.02	81.6

The magnetic hysteresis loop of the as-prepared $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ was carried out in an external applied magnetic field of about 14 kOe at 300 K. The *M-H* (where *M* is the magnetization) curve (Figure 2 (B)) has a coercivity value of 43±1 Oe. The obtained values of maximum magnetization, remanence and hysteresis loop squareness are 27±2 emu/g, 1.5±0.1 emu/g and 0.037, respectively.

The electrochemical activity of the synthesized $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ nanoparticles was investigated using potassium ferricyanide $K_3[Fe(CN)_6]$ as the redox probe. Figure 3 A shows the cyclic voltammetric response observed at bare glassy carbon electrode (GCE) and $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$/GCE in the presence of 1 mM $K_3[Fe(CN)_6]$ in 0.1 M KCl solution. Well-defined redox peaks were observed at both electrodes. However, the $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$/GCE was found to exhibit improved electrochemical behavior, exhibiting an increased peak current response and a negative shift in peak potential, as

compared to the bare electrode. This suggests that the $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ nanoparticles effectively electrocatalyze the redox process of $K_3[Fe(CN)_6]$. This is attributed to an enhanced rate of electron transfer owing to excellent electrical conductivity and high surface area of the synthesized $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ nanoparticles.

To establish that $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ nanoparticles improve the surface area of the electrode, the effective surface area (A) of bare GCE and $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$/GCE were determined by recording cyclic voltammograms in 1.0 mM $K_3[Fe(CN)_6]$ solution at different sweep rates (Figure 3 B). Using the Randles-Sevcik equation: $i_p = 2.69\times10^5 n^{3/2} A D^{1/2} C v^{1/2}$ where i_p is the peak current (in Amperes), n is the number of electrons transferred in the electrochemical process, D (in $cm^2\ s^{-1}$) is the diffusion coefficient (D = 7.6 × 10-6 $cm^2\ s^{-1}$ for $[Fe(CN)_6]^{3-}$), C (in mol cm^{-3}) is the concentration of the electroactive species and v is the scan rate (in V s^{-1}) and employing the slope of i_{pa} versus $v^{1/2}$ plot (inset of Figure 3 B), Surface area was found to be 0.070 cm^2 and 0.218 cm^2 for bare GCE and $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$/GCE, respectively. The surface area of the modified electrode was at least 3 times larger than that of the bare GCE electrode, indicating superior conductivity of the $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ nanoparticles. Thus, it is expected that the novel nanoferrites can lead to the development of a new group of electrochemical sensors.

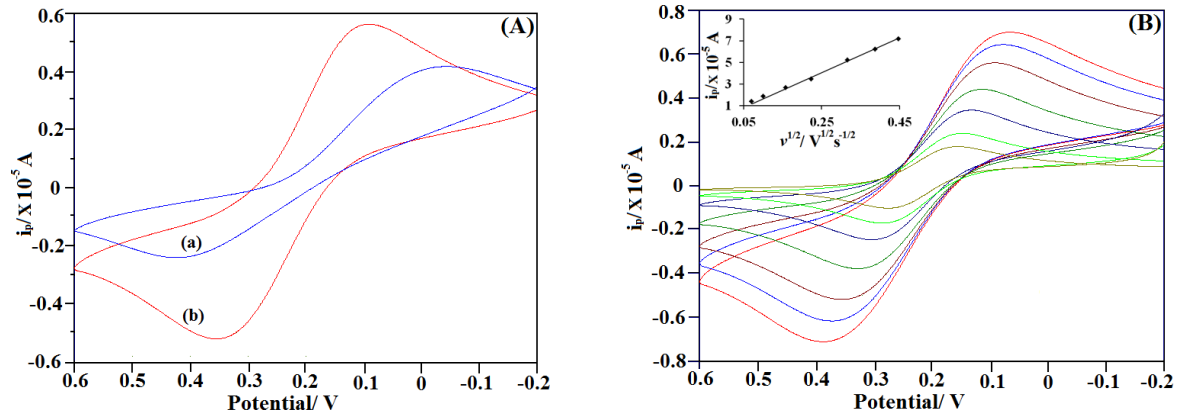


Figure 3 (A) Cyclic voltammograms of 1.0 mM $K_3[Fe(CN)_6]$ in 0.1 M KCl solution at (a) bare GCE and (b) $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$/GCE at 100 mV/s. (B) Cyclic voltammograms of 1.0 mM $K_3[Fe(CN)_6]$ in 0.1 M KCl solution at $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$/GCE at sweep rates 5–200 mV s^{-1} (Inset displays plot of peak current versus square root of sweep rate).

4. Conclusions

$Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ nanoparticles ferrite was successfully produced by glycol thermal method. The obtained values of lattice parameter, crystallite size and microstrain are found to be 0.839±0.003 nm, 8.37±0.06 nm and 0.0015±0.0001 respectively. We attribute the microstrain as due to the synthesis method and conditions. HRTEM images show that the synthesized nanoparticles have average particle sizes of 9±2 nm. This is close to the crystallite size estimated from the XRD measurements. The Mössbauer spectrum was fitted only with two sextets associated with the Zeeman splitting with Fe^{2+} and Fe^{3+} on A- and B-sites respectively. The synthesized sample shows small value of coercive field of 43±1 Oe as deduced from the room temperature hysteresis loop. $Ba_{1/3}Mn_{1/3}Co_{1/3}Fe_2O_4$ nanoferrite modified electrode showed good electrocatalytic activity. These high-performance electrodes are expected to lead to the development of a novel group of electrochemical sensors.

Acknowledgments
The authors would like to thank the National Research Foundation (NRF) of South Africa for equipment grants, Department of Pharmaceutical Chemistry (College of Health Sciences, University of KwaZulu-Natal, South Africa) for financial support (NT) and Sudan University of Science and Technology for the study leave (NSEO).

References

[1] Xu Q, Wei Y, Liu Y, Ji X, Yang L and. Gu M 2009 *Solid State Sci*. **11(2)** 472-478

[2] Hu J, Li L, Yang W, Manna L, Wang L and Alivisatos A P 2007 *J. Magn. Magn. Mater*. **308** 198

[3] Baykal A, Kasapoglun, Durmus Z, Kavas H, Toprak M S and Koseoglu Y 2009 *Turk J Chem*. **33** 33-45

[4] Gunjakar J L, More A M, Gurav K Vand Lokhande C D 2008 *Appl Surf Sci* **254(18)** 5844-5848

[5] Sloczynski J, Janas J, Machej T, Rynkowski J and Stoch J 2000 *Appl Catal* B **24(1)** 45-60

[6] Pena M A and Fierro J L G 2001 *Chem Rev* **101(7)** 1981-2018

[7] Smit J and Wijn H P J 1959, *Ferrites* (Eindhoven: Netherlands) p 278–282

[8] Goldman A 1999, *Handbook of Modern Ferromagnetic Materials* (New York: Springer)

[9] Willard M A, Kurihara L K, Carpenter E E, Calvin S and Harris V G 2004 *Int. Mater*. *Rev* **49** 125-170

[10] Osman N S E, Moyo T and Abdallah H M I 2014 *Proceedings of SAIP* 128-133 Available online at http://events.saip.org.za

[11] Osman N S E, Thapliyal N, Alwan W S, Karpoormath R and Moyo T.2015 *J Mater Sci: Mater Electron* **26** 5097-5105

[12] Lemine O M, Bououdina M, Sajieddine M, Al-Saie A.M, Shafi M, Khatab A, Al-hilali M and Henini M 2011 *Physica* B **406** 1989–1994

[13] Thapliyal N, Osman N S E, Patel H, Karpoormath R, Goyal R N, Moyo T and Patel R 2015 *RSC Advances* DOI: 10.1039/C5RA05286F

[14] Salunkhe A B, Khot V M, Phadatare M R, Thorat N D, Joshi R S, Yadav H M and Pawar S H 2014 *J. Magn. Magn. Mater* **352** 91-98

[15] Parvatheeswara Rao B, Subba Rao P S V, Murthy G V S and Rao K H 2004 *J. Magn. Magn. Mater* **268** 315–320

[16] Ghasemi A, Jr A P and Machado C F C 2012 *J. Magn. Magn. Mater* **324** 2193–2198

[17] Lakshman A, Subba Rao P S V, Rao KH 2006 *Mater Lett* **60** 7–10

An ab-initio study of the metastability of the boron-vacancy complex in silicon

Cecil N M Ouma and Walter E Meyer
Department of Physics, University of Pretoria, Private Bag X20, Pretoria, 0002, South Africa

E-mail: walter.meyer@up.ac.za

Abstract. Using the experimentally proposed model of the C centre in silicon (Si), we have calculated the electronic properties of the C centre defect in Si using Perdew-Burke-Ernzerhof (PBE) with a Γ-point (PBE- Γ) as well as a 2×2×2 k-mesh (PBE-2×2×2) sampled Brillouin-zone (BZ). The choice of BZ sampling was found to significantly influence the calculated defect properties. PBE-2×2×2 did not predict any metastability of the C centre, PBE-Γ predicted charge-state controlled metastability of the C centre under both bias on and bias off conditions. The C1 and C2 configurations, in which the boron-vacancy pairs are in the nearest- and next-nearest-neighbour positions, have been predicted as the metastable configurations of the C centre in Si.

1. Introduction

The semiconductor electronics industry has evolved from macro to micro and recently to nano electronics. As the circuits become more miniaturized and the packing density becomes larger, there is need for more understanding of the fundamental properties of the semiconductors in use. Other semiconductors including GaAs, GaN, AlGaN, graphene, boron nitrine and ZnO are also currently intensively being studied both experimentally and theoretically as possible alternatives to silicon (Si). However, interfacing these semiconductors with the Si drive electronics is still costly and problematic.[1]

Defects can be beneficial or detrimental in the functionality of semiconductor devices, and need to be taken into account in device design. A fundamental understanding of these defects and their properties is essential in defect engineering. Most defects in semiconductors are known to be stable in one specific configuration is all charge states. However, the same defects may occur in more than one configuration, at either of the charge states. These types of defects are referred to as metastable defects.[2–5] In silicon, several metastable defects have been experimentally observed, some of these include; the FeAl pair[3], the CsCi defect pair[4,5] and the boron-vacancy (B-V) pair, also known as the C centre,[2] among others.

The C centre has been observed using deep-level transient spectroscopy (DLTS) by Chantre[2] in ultra-fast-quenched boron-doped Si to be metastable. In the study, Chantre observed that this defect centre introduced two deep donor-hole traps at 0.50 eV and 0.36 eV above the valence band maximum (VBM). These two donor-hole traps were associated to the next-nearest- and nearest-neighbour B-V pairs in silicon respectively.

Theoretical studies have also been done on both the isolated B_{Si} and V_{Si} as well as the B-V pair (complex) in Si.[6–8] However a survey of the available literature on the theoretical studies on B-V

complex (C centre) in Si reveals that these studies have only been on impurity- or vacancy-mediated diffusion in Si.[6,7] None of these studies have investigated the metastability of the C centre defect in Si.

Although density functional theory (DFT) has been used to study defects in semiconductors,[9–11] it has been shown that DFT inaccurately predicts defect thermodynamic transition levels among other properties[12–15]. These failures of DFT stem mainly from the approximations made to the exchange-correlation (XC) functional. The local density approximation (LDA)[16,17] and the Perdew-Burke-Ernzerhof (PBE) form of generalized gradient approximation (GGA),[18] are the most commonly used approximations of the XC functionals. Although several attempts have been made to remedy some of the failures of DFT, [12,14,19–22] standard DFT (using LDA or GGA) can still be used to investigate some defect properties in semiconductors but with additional corrections.[9,23]

2. Computational details

We have used PBE form of GGA within standard DFT We have used the projector-augmented wave (PAW) method,[24,25] as implemented in the VASP code.[26,27] The unit cell was optimized using an energy cutoff of 600 eV and a Γ centred 8×8×8 Monkhorst-Pack grid of *k*−points using PBE. We obtained a lattice constant of 5.45 Å. From this, supercells consisting of 64 silicon atoms were constructed. PBE defect calculations were done using the 64 silicon atoms supercell with a kinetic energy cutoff of 600 eV, Methfessel-Paxton (MP) smearing of 0.02 eV, and a Γ-point (hereto referred as PBE-Γ) as well as a Γ cantered 2×2×2 Monkhorst-Pack grid of *k*−points (hereto referred as PBE-2×2×2) to sample the Brillouin zone (BZ). Band gaps of 0.65 eV (PBE-2×2×2) and 0.50 eV (PBE-Γ) were obtained. The experimental band gap of silicon is 1.12 eV.[28] While the Γ-point sampling of the BZ is often regarded as the simplest scheme of sampling the BZ, Puska *et al.*[29] using LDA XC functionals showed that in some cases (e.g. vacancy in silicon), this sampling scheme gives a correct qualitative description of the defect compared to a 2×2×2 *k*−points sampled BZ.

We considered two configurations of C centre namely; C1, where the silicon vacancy (V_{Si}) is at the nearest-neighbour position w.r.t. the substitutional boron (B_{Si}), and C2 in which V_{Si} is in the next-nearest-neighbour position w.r.t. the B_{Si} as suggested by Chantre[2] (see Figure 1). Formation energies of the defects were calculated using the Zhang and Northrup formalism[30] i.e., the formation energy of the C centre in silicon at charge state q is

$$E^{f,q}(BV) = E_{D,q} - E_p + 2\mu_{Si} - \mu_B + q(E_F + E_V + \Delta V) \tag{1}$$

Where $E_{D,q}$ is the total energy of the defect supercell at charge state q, E_p is the total energy of the pristine 64 atoms silicon supercell, μ_{Si} and μ_B are the chemical potentials for Si and B respectively, E_F is the Fermi level, E_V is the valence band maximum (VBM) and ΔV is the potential alignment term. For the charged supercells, a jellium charge compensating background of opposite sign was used, as implemented in the VASP code. Monopole correction[23] was done on the total energies and potential alignment between the pristine and the defect supercell was done using the technique of ref [9]. The thermodynamic transition level, ε, for q' and q charge sates was obtained as

$$\varepsilon(q/q') = \frac{E^{f,q'}(BV) - E^{f,q}(BV)}{q - q'}, \tag{2}$$

giving the Fermi energy at which the formation energies of the defect in the two charge states are equal.

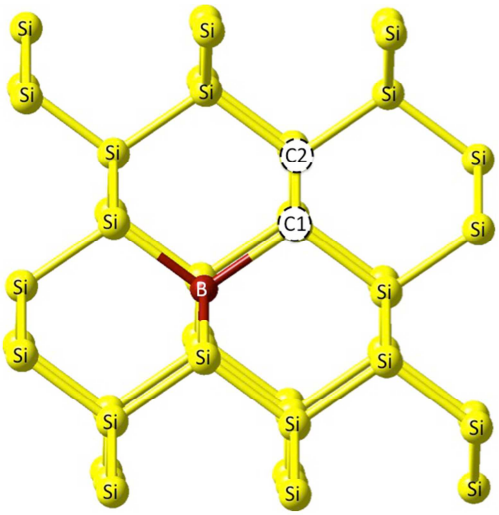


Figure 1. The V_{Si} position for the C1 and C2 configurations of the C centre in silicon.

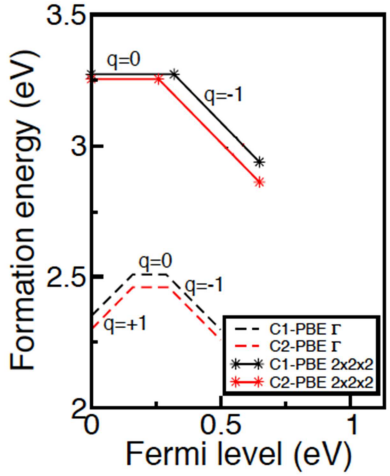


Figure 2. Figure showing the thermodynamic transition levels for the C1 and C2 configurations of the C centre in silicon.

2.1. V_{Si} defect in silicon

The isolated vacancy in silicon, V_{Si} has been extensively studied via *ab initio* LDA within DFT by Puska *et al.*[29] we have calculated both the formation energies at various charge states using equation 3.2 (Table 1) and the thermodynamic charge transition levels using equation 3.3. Our calculated formation energies, for both PBE-Γ and PBE-2×2×2, are in good agreement with those of Puska *et al.*[29] Although the calculated thermodynamic transition levels are in good agreement with those calculated by Puska *et al*.,[29] they are not consistent with the experimentally reported values[31], which report no defect levels associated with donor states only and with acceptor states.

2.2. B_{Si} defect in silicon

Most of the theoretical studies in literature have investigated interstitial boron in silicon (B_i). Substitutional boron in silicon (B_{Si}), is a shallow acceptor in silicon. Similar to case of V_{Si}, we calculated the formation energies of B_{Si} defect in various charge states (see Table 1) from which its thermodynamic transition levels were determine (see Figure 2). Formation energy calculated

Table 1. Calculated formation energies in eV at various charge states (figures in bold indicate minimum energy configurations).

Defect	Functional	-1	0	1
V-Si	PBE-2×2×2	3.01	2.80	2.83
	Ref [29]-2×2×2	3.68	3.42	3.72
	PBE-Γ	2.85	2.43	2.1q
	Ref [29]-Γ	3.56	2.86	2.51
B-Si	PBE-2×2×2	1.33	0.75	0.68
	PBE-Γ	0.16	0.22	0.33
C1	PBE-2×2×2	3.59	3.27	3.41
	PBE-Γ	3.29	2.69	**2.29**
C2	PBE-2×2×2	**3.51**	**3.26**	**3.32**
	PBE-Γ	**2.94**	**2.64**	2.48

Table 2. Calculated thermodynamic transition levels in eV for the defects referenced to the valence band edge.

Defect	Functional	0/–	+/0
V-Si	PBE-2×2×2	0.20	--
	Ref [29]-2×2×2	0.26	--
	PBE-Γ	0.42	0.32
	Ref [29]-Γ	0.70	0.35
	Experiment[31]	--	0.05
B-SI	PBE-2×2×2	0.57	0.07
	PBE-Γ	--	--
	Experiment[32]	0.05	--
C1	PBE-2×2×2	0.32	--
	PBE-Γ	0.28	0.23
	Experiment[2]	--	0.36
C2	PBE-2×2×2	0.26	--
	PBE-Γ	0.30	0.21
	Experiment[2]	--	0.50

according to PBE-Γ was the lowest, while PBE-2×2×2 yielded the highest formation energy. Only PBE-2×2×2 predicted thermodynamic transition levels at $E_V + 0.07$eV and $E_V + 0.57$eV respectively. It is evident that PBE-Γ gives the incorrect description of B_{Si}. However, its important to mention that experimentally, B_{Si} has only an acceptor level at 0.045 eV.[32] Since B_{Si} is a shallow acceptor in silicon, no alignment term for the defect levels was used in calculating the formation energies for its charged states. This is because the alignment of shallow defect levels is still the subject of on-going debate.[23,33,34] Furthermore, due to the large hydrogenic orbitals of the shallow donor, it is not clear that the super cell chosen in this study is large enough.

2.3. C centre in silicon

B_{Si} is a shallow accepter in silicon, while V_{Si} is known to be highly mobile in silicon. When these two defects interact they form the C centre defect in silicon. We have calculated the formation energies (Table 1) at various charge states as well as the thermodynamic transition levels (Table 2 and Figure 2) for two different configurations of the C centre. The main aim of this study is to investigate the metastability of the C centre and if the metastability is charge-state controlled. From our study we found that both the formation energy and the thermodynamic transition levels are dependent on the position of vacancy in the silicon lattice (configuration dependent). This is in agreement with what was observed experimentally.[2] In that study, Chantre [2] experimentally observed two donor-hole traps that were associated to next-nearest- and nearest-neighbour configurations of the C centre. C2 was found to be the most stable configuration for all the charge states using PBE-2×2×2. For the case of PBE-Γ, C2 was the most stable configuration for the -1 and neutral charge state while C1 was the most stable configuration for the +1 charge state.

2.3.1. Metastability of the C centre and comparison to experiment. Chantre[2] using DLTS experiments posited that the C centre in Si was configurationally bistable. From that experiment, it was observed that the C centre had donor levels at $E_V + 0.54$ eV and $E_V + 0.36$ eV that were associated with two DLTS peaks. Chantre[2] labeled these peaks H_1 and H_2 respectively. The same experiment identified the H_1 and H_2 peaks as associated with the next-nearest- and nearest-neighbour B-V pairs, respectively. In our study, as mentioned in section 2. we refer to the nearest- and next-nearest-neighbour configurations as C1 and C2 respectively.

In *p*-type material, under zero bias (referred to as *bias off* by Chantre[2]), the Fermi level is close to the valence band edge (E_V), i.e. $E_F \approx E_V$. This leads to defects emitting electrons (i.e. filled with holes). Under reverse bias (referred to as *bias on* by Chantre[2]), the Fermi level in the depletion region of a Schottky diode is close to the conduction band edge (E_C), i.e. $E_F \approx E_C$. This leads to defects capturing electrons (i.e. emitting holes). Since it was experimentally observed that the emission rates of the H_1 and H_2 peaks did not depend on the electric field, it was concluded that these two peaks correspond to donor levels,[2] e.g. $+/0$.

The annealing by Chantre[2] was performed at approximately 300 K, which leads to a value of kT of 0.026 eV. This implies that, for the occupation of two defect configurations to differ significantly, the difference in their formation energies should be at least approximately kT. A difference in formation energy of approximately $3kT$ ($\approx$ 0.1 eV) or greater would lead to the ratio of occupation of the two configurations to be 1:10 or greater (assuming no degeneracy). This means that, experimentally, one level will dominate.

As seen in Figure 2, PBE-2×2×2 does not predict a $+/0$ thermodynamic transition level. Ignoring this, we still went ahead to analyze whether there was any sign of metastability predicted by PBE-2×2×2, using the $q = 0$ and $q = -1$ charge states. For the $0/-$ thermodynamic transition, the C1 transition level is deeper than the C2 transition level w.r.t. E_V. This implies that, in a DLTS experiment, C1 will emit at a lower rate than the C2, and therefore the emission by the C1 configuration will be observed at a higher temperature than that of the C2.

Under *bias off*, $q = 0$, the C2 configuration was the most stable configuration of the C centre. The formation energy difference between the two configurations was 0.018 eV, which is less than 3*kT* at 300 K. This implies that both the C1 and C2 defect levels will be observed since they will have roughly similar occupation probabilities. This is not consistent with what was reported by Chantre[2], who observed only one dominant peak. Furthermore, the PBE-2×2×2 prediction here implies that the dominant peak, the C2 in this case, is likely to be observed at lower temperatures, which is also not consistent with the reported experimental results.[2]

Under *bias on*, $q = -1$, the C2 configuration is still the most stable configuration. The formation energy difference between C1 and C2 configurations is 0.075 eV. This implies that the C2 configuration, which gives rise to the lower temperature DLTS peak, is predicted to be more dominant. This is also not consistent with what was reported by Chantre[2], who observed both peaks at approximately equal heights under these conditions.

PBE-2×2×2 thus predicts variation in peak heights of the H_1 and H_2 peaks after annealing under different bias conditions. However, the relative (temperature) positions of the peaks and their relative heights do not agree with experiment.[2]

PBE-Γ predicted both the $+/0$ and $0/-$ thermodynamic transition levels (see Figure 2). In both cases, the transition level for the C2 configuration is closer to E_V compared to that of the C1 configuration. This implies that there will be faster (hole) emission by C2 and hence, C2 is likely to be observed at lower temperature compared to the C1. This differs from the experimental observation by Chantre[2], who associated the higher temperature peak (H_1) with the C2 configuration.

Under *bias off*, $q = +1$ and PBE-Γ predicted the C1 as the most stable configuration of the C centre. The formation energy difference between the C1 and C2 configurations was 0.190 eV, which is greater than 3*k*T at 300 K. This implies that, under these conditions, only the C1 configuration will be occupied and hence experimentally observable. As discussed earlier, the C1 configuration leads to a DLTS peak at higher temperature. This agrees with the experimental observations of Chantre[2], who observed the higher temperature DLTS peak under *zero bias*.

Under *bias on*, $q = 0$ and C2 is the most stable configuration. The formation energy difference between the two configurations is only 0.049 eV, which is about 2*kT* at 300 K implying that both the C1 and C2 configurations will be experimentally observable, with the occupation of the C2 configuration (at lower temperature) being significantly higher than that of the C1 (ignoring the effects of degeneracy). If degeneracy is taken into account, the occupation of C2 would be increased by a factor 2.5. Chantre[2] experimentally observed both the H_1 and H_2 peaks, with the H_1 (at higher temperature) having a much lower peak height, as a shoulder of the *H*(0.44).

PBE-Γ predicts some metastability, which agrees with the experimental observations[2], but not with the identification of the defect configurations.

3. Conclusions

We used DFT with GGA functionals to investigate the metastability and other properties of the C centre (B-V complex) in silicon and compared our results to the experimentally observed values. In addition, we have also investigated the properties of isolated defects that form the B-V defect (V_{Si} and B_{Si}). We have compared how the choice of *k*-points used in sampling the BZ influenced the predicted defect properties of $V_{Si,}$ B_{Si} and the C centre in silicon.

Our predicted properties of V_{Si} defect in silicon were consistent with those predicted by Puska *et al.*[29] The choice of BZ sampling significantly influenced the predicted defect properties. We found that the PBE functional using Γ-point sampling (PBE- Γ) gave a better qualititative description of the defect compared to the PBE functional using a 2×2×2 *k*-mesh (PBE-2×2×2). For B_{Si}, only PBE-2×2×2 predicted thermodynamic transition levels in the band gap.

Although PBE-Γ predicted some form of charge-state controlled metastability of the C centre that was consistent with experimental observations, the identity of the defect configurations was not

consistent with other experimental observations. PBE-2×2×2 only predicted variations in peak heights under different bias conditions that were not consistent with the experimental observations.

Acknowledgement
This work is based on the research supported in part by the National Research Foundation of South Africa (Grant specific unique reference number (UID) 76938) Opinions, findings and conclusions or recommendations are that of the authors, and that the NRF accepts no liability whatsoever in this regard. The authors would also like to acknowledge Prof. Nithaya Chetty for helpful discussions.

References

[1] Kenyon A J 2005 *Semicond. Sci. Technol.* **20** R65–84
[2] Chantre A 1985 *Phys. Rev.* B **32** 3687–94
[3] Chantre A and Bois D 1985 *Phys. Rev.* B **31**
[4] Song W, Benson W and Watkins G D 1990 *Phys. Rev.* B **42**
[5] Song L W, Benson B W and Watkins G D 1987 *Appl. Phys. Lett.* **51** 1155
[6] Nichols C, Van de Walle C and Pantelides S 1989 *Phys. Rev.* B **40** 5484–96
[7] Windl W and Stumpf R 2008 *Mater. Sci. Eng.* B **155** 198–201
[8] B., Cowren N. E., F., K. T. F. Janssen, G. F. A. van de Walle D J G 1990 *Phys. Rev. Lett.* **65** 2434–7
[9] Van de Walle C G 2004 *J. Appl. Phys.* **95** 3851
[10] Van de Walle C G and Janotti A 2011 *Phys. Status Solidi* **248** 19–27
[11] Stampfl C, Van de Walle C, Vogel D, Krüger P and Pollmann J 2000 *Phys. Rev.* B **61** R7846–9
[12] Komsa H-P and Pasquarello A 2011 *Phys. Rev.* B **84** 075207
[13] Komsa H-P, Broqvist P and Pasquarello A 2010 *Phys. Rev.* B **81** 205118
[14] Parker W D, Wilkins J W and Hennig R G 2011 *Phys. Status Solidi* **248** 267–74
[15] Lany S and Zunger A 2010 *Phys. Rev.* B **81** 113201
[16] Ceperley D M 1980 *Phys. Rev. Lett.* **45** 566–9
[17] Wang Y and Perdew J 1991 *Phys. Rev.* B **44** 13298–307
[18] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865–8
[19] Oba F, Togo A and Tanaka I 2008 *Phys. Rev.* B **77** 245202
[20] Alkauskas A, Broqvist P and Pasquarello A 2011 *Phys. Status Solidi* **248** 775–89
[21] Shang H, Li Z and Yang J 2011 *J. Chem. Phys.* **135** 034110
[22] Ouma C N M, Mapelu M Z, Makau N W, Amolo G O and Maezono R 2012 *Phys. Rev. B* **86** 104115
[23] Lany S and Zunger A 2008 *Phys. Rev.* B **78** 235104
[24] Blöchl P E 1994 *Phys. Rev. B* **50** 17953–79
[25] Kresse G and Joubert D 1999 *Phys. Rev.* B **59** 1758–75
[26] Kresse G and Hafner J 1993 *Phys. Rev.* B **48** 13115–8
[27] Kresse G and Furthmüller J 1996 *Comp. Mat. Sci.* **6** 15–50
[28] Streetman B G and Banerjee S 2000 *Solid state electronic devices* vol 4 (Prentice Hall New Jersey)
[29] Puska M, Pöykkö S, Pesola M and Nieminen R 1998 *Phys. Rev.* B **58** 1318–25
[30] Zhang S and Northrup J 1991 *Phys. Rev. Lett.* **67** 2339–42
[31] Watkins G D and Pantelides S T 1986 *Gordon Breach, New York* 147
[32] Morin F and Maita J 1954 *Phys. Rev.* **96** 28–35
[33] Freysoldt C, Neugebauer J and Van de Walle C G 2011 *Phys. Status Solidi* **248** 1067–76
[34] Freysoldt C, Neugebauer J and Van de Walle C 2009 *Phys. Rev. Lett.* **102** 016402

Physical properties of $Cr_{78}Al_{22}$ thin films

ARE Prinsloo[1], CJ Sheppard[1], HA Derrett[1], N van den Berg[2] and EE Fullerton[3]

[1]Department of Physics, University of Johannesburg, PO Box 524, Auckland Park, 2006, South Africa
[2]Department of Physics, University of Pretoria, Private bag X20, Hatfield, 0028, South Africa
[3]Center for Magnetic Recording Research, University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093-0401, USA

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

Abstract. The $Cr_{100-x}Al_x$ alloy system shows astonishing behaviour at higher Al concentrations [1]. Very high Néel temperatures (T_N > 800 K) are observed in samples with x > 20. The SDW amplitude for these alloys has a value of about $1\mu_B$ – larger than in other Cr alloy systems. In addition both the Hall coefficient and the resistivity for samples in the concentration range $15 \leq x \leq 25$ are large. In this concentration range the resistivity has a negative temperature dependence and is in form characteristic of narrow-bandgap semiconductors. Combining these unique bulk characteristics with exceptional thin film properties seen for Cr and its alloys, appears to be a way forward in order to improve modern technologies, such as used in data storage. For this reason the present study focus on $Cr_{78}Al_{22}$ thin films in a thickness (t) range 12 to 400 nm, prepared on MgO(100), MgO(110), a-plane sapphire and quartz substrates, deposited using DC magnetron sputtering. AFM results on the fused silica samples indicate interesting growth patterns with cubic structures forming in the thicker samples. This is supported by XRD results indicating that for the samples prepared on fused silica substrates preferred Cr(110) growth occurs for $t \geq 100$ nm. XRD results also show good epitaxial growth of the films prepared on the MgO and a-plane sapphire substrates. Resistance (R) as function of temperature (T) investigations were done using the standard four-point probe method in a temperature range 77 K < T < 400 K and show negative temperature dependence. Interestingly, the behaviour of $R(T)$ differs for those samples prepared on MgO(100), as the film with t = 400 nm shows characteristics not associated with narrow-bandgap semiconductors.

1. Introduction

The richness of magnetic phenomena of Cr and Cr alloys has attracted considerable interest for many years [1]. Bulk Cr is an itinerant electron antiferromagnetic material that forms an incommensurate (I) spin density wave (SDW) phase below its Néel temperature T_N = 311 K [1]. The ISDW period varies from 78 Å at T_N to 60 Å at 10 K [2, 3]. Thin films and hetero-structures of Cr and Cr alloys show fascinating properties, not observed in the bulk [3, 4]. These properties include dimensionality, surface and proximity effects, as well as the mediating role of Cr thin films in exchange coupled super-lattices and in giant magneto-resistive (GMR) materials [2].

The $Cr_{100-x}Al_x$ alloy system has shown interesting behaviour recently explored further [5] around the triple point where the commensurate SDW (CSDW), ISDW and paramagnetic phases co-exit [1]. The

study on $Cr_{100-x}Al_x$ single crystals [5] is considered a detailed investigation of the magnetic behaviour of Cr alloy systems in the proximity of the triple point concentration on the concentration-temperature magnetic phase diagram. The data indicates that the triple point minimum on this diagram seen at $x \approx 2.3$ might be deeper and much narrower than previously thought, resembling quantum critical behaviour at this point. However, at higher Al concentrations new and astonishing behaviour begin to emerge [1]. Very high Néel temperatures ($T_N > 800$ K) are observed in samples with $x > 20$. The SDW amplitude for these alloys has a value of about $1\mu_B$ – larger than in other Cr alloy systems. In addition both the Hall coefficient and the resistivity for samples in the concentration range $15 \leq x \leq 25$ is large [1]. In this concentration range the resistivity of bulk samples have a negative temperature dependence and has a similar behaviour as that of narrow-gap semiconductors [1].

Combining these unique bulk characteristics of Cr-Al with exceptional thin film properties seen for Cr and its alloys [2], appears to be a way forward in order to extend the understanding of the behaviour of the SDW in this system and also provides basic information that can possibly improve modern technologies in future, such as those applicable to data storage. This paper focuses on the structural, surface and electrical properties of epitaxial and polycrystalline $Cr_{78}Al_{22}$ thin films in a thickness (t) range 12.5 to 400 nm.

2. Experimental

The Cr-Al thin films were deposited using Direct-Current (DC) magnetron sputtering at a substrate temperature of 973 K and argon working pressure of 3 mTorr. The Cr and Al were co-sputtered from elemental sources onto single-crystal MgO(110), MgO(100), a-plane sapphire, and amorphous quartz substrates. The Al-concentration in the Cr-Al films was obtained by adjusting the Al target deposition power. The samples were prepared at a fixed concentration of Cr + 22 at.% Al with thicknesses varying from 12 to 400 nm. Atomic force microscopy (AFM) imaging was used to study the topographic properties of the $Cr_{78}Al_{22}$ thin films and the film structure was characterized using X-ray diffraction (XRD) techniques. Electrical resistance measurements using standard DC four-probe method from 77 to 400 K were employed to investigate the electrical transport properties of these samples.

3. Results

The XRD measurements indicate that all the films prepared on MgO(100), MgO(110) and a-plane sapphire, are epitaxial and exhibit a single crystallographic orientation, with the Cr-Al showing preferred growth directions of (100), (211) and (110) for the respective substrates. The samples prepared on quartz were polycrystalline for thicknesses 12.5 nm < t < 100 nm and t = 400 nm, however for thickness t = 100 and 200 nm preferred Cr(110) growth orientation is shown. Examples of XRD measurements for the samples are shown in Figs. 1 (a), (b), (c) and (d) for the various substrates.

The full-width at half-maximum (FWHM) of the Bragg peaks was used to determine the length scale (in growth direction) over which the film is structurally coherent [6]. The Debye-Scherrer formula was used to calculate the coherence lengths after removing the instrument resolution. For $Cr_{78}Al_{22}$ thin films ($12 \leq t \leq 400$ nm) grown on MgO(100) substrates the crystal coherence length in growth directions (100) of the films increased with increasing t reaching a maximum at t = 100 nm thereafter it decreases and levels off at approximately 13.5 nm for $t \geq 200$ nm. For the MgO(110) substrates the crystal coherence length in growth directions (110) generally increased with increasing t. Samples prepared on a-plane sapphire show an increase in coherence length with increase in sample thickness reaching a maximum for a 100 nm sample and then the coherence length decreased again. Results obtained for the $Cr_{78}Al_{22}$ thickness series are summarized in Table 1.

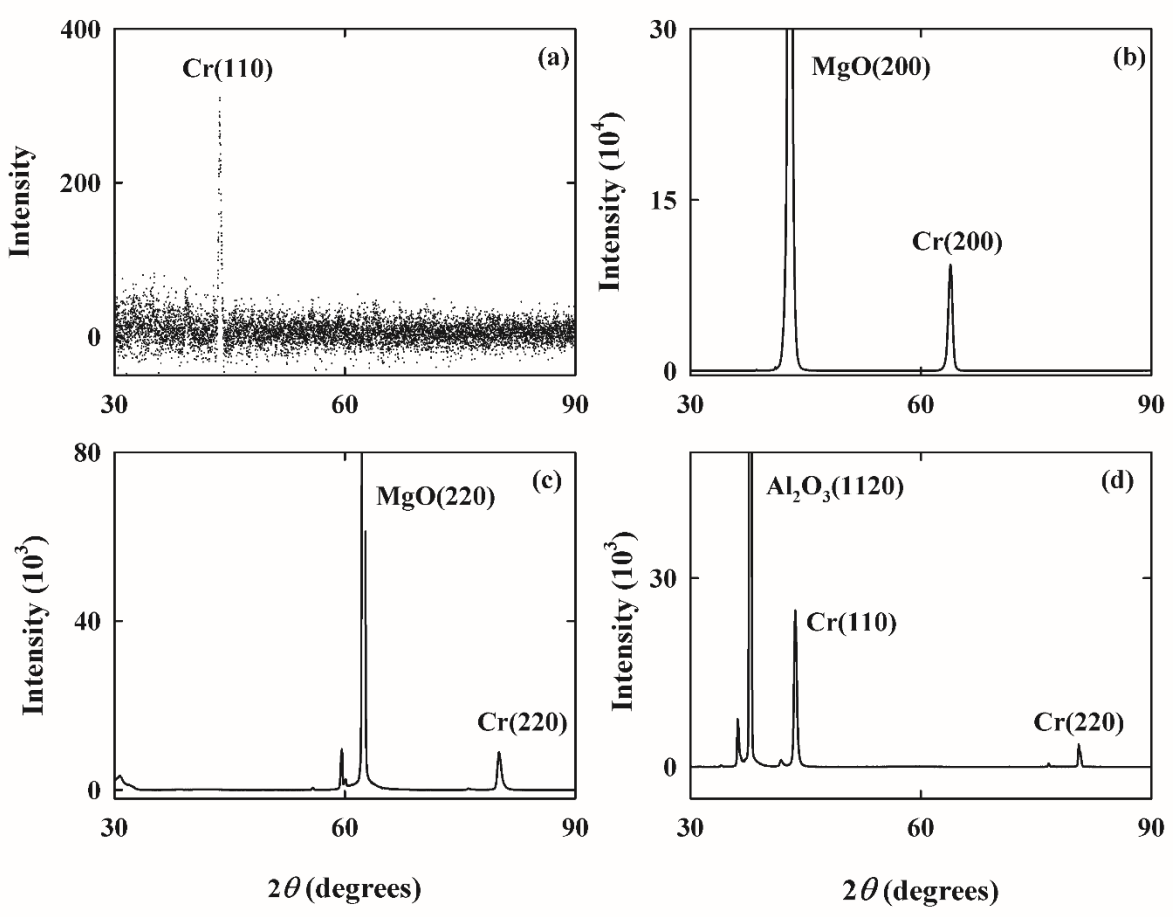


Figure 1. θ-2θ XRD scans, with the scattering vector normal to the plane of the film, for the $Cr_{78}Al_{22}$ thin films prepared on (a) quartz (t = 200 nm); (b) MgO(100) (t = 100 nm), (c) MgO(110) (t = 200 nm) and (c) a-plane sapphire (t = 50 nm) substrates.

	MgO(100)			MgO(110)			a-plane sapphire		
t (nm)	FWHM (200)	T_{coh} (nm)	Mosaic (200)	FWHM (211)	T_{coh} (nm)	Mosaic (211)	FWHM (110)	T_{coh} (nm)	Mosaic (110)
12.5	1.98°	8.3	1.18°	1.15°	9.5	0.89°	0.66°	13.5	0.28°
25	1.64°	11.5	0.85°	0.84°	12.9	1.47°	0.45°	19.1	0.56°
50	1.32°	15.9	0.61°	0.73°	14.9	1.18°	0.40°	22.4	0.46°
100	0.97°	16.0	0.61°	0.75°	14.6	1.06°	0.37°	24.2	0.42°
200	1.16°	13.6	1.16°	0.57°	19.0	0.50°	0.40°	22.2	0.39°
400	0.72°	13.6	0.72°	0.60°	18.0	0.70°	0.43°	21.0	0.46°

Table 1. The measured out-of-plane full-width at half-maximum (FWHM) XRD parameters of the selected Bragg peaks and the mosaic spread of this peak, for epitaxial $Cr_{78}Al_{22}$ films of thickness t prepared at 800°C on MgO(100), MgO(110) and a-plane sapphire substrates, respectively. The coherence length, T_{coh}, was calculated using the Debye-Scherrer equation after removing the resolution of the instrument.

The mosaic spread was determined from the FWHM of the rocking curve [6]. For $Cr_{78}Al_{22}$ thin films prepared on MgO(100) substrates (Table 1) the general tendency is a decrease in the mosaicity with increase in *t*, except for the film of thickness 200 nm that does not fit into this pattern, with a mosaicity of 1.16°. This indicates a general trend of improved crystal quality with films thickness, but since the values exceeds 0.5° there is not a high degree of crystallographic alignment in these samples [6, 7] which is typical for Cr and Cr alloys epitaxially grown onto MgO. No clear tendency is seen in the mosaicity of the thickness series prepared on MgO(110). The films prepared on the a-plane sapphire however, had a high degree of crystallographic alignment with the mosiacities less than 0.5°, except for the 35 nm sample with a mosiacity of 0.56°.

AFM studies on the $Cr_{78}Al_{22}$ monolayers thickness series indicate that polycrystalline film with t = 12.5 nm has a relatively smooth surface morphology, depicted in figure 2 (a). As t increases, grains become more visible in the AFM images accompanied by an increase in surface roughness. For the sample with t = 200 nm clear cubic crystals are visible (depicted in figure 2 (b)). Interestingly this sample also showed preferred Cr(110) growth. On increasing the thickness to t > 200 nm columnar growth observed, with a clear increase in surface roughness. Figure 2 (c) show the AFM image for the epitaxial film with t = 100 nm prepared on MgO(110) and columnar structures consisting of merged cubical units are observed. However, for the samples prepared on a-plane sapphire the surface appears relatively smooth with no columnar structures (as an example consider figure 2 (d) for the sample with

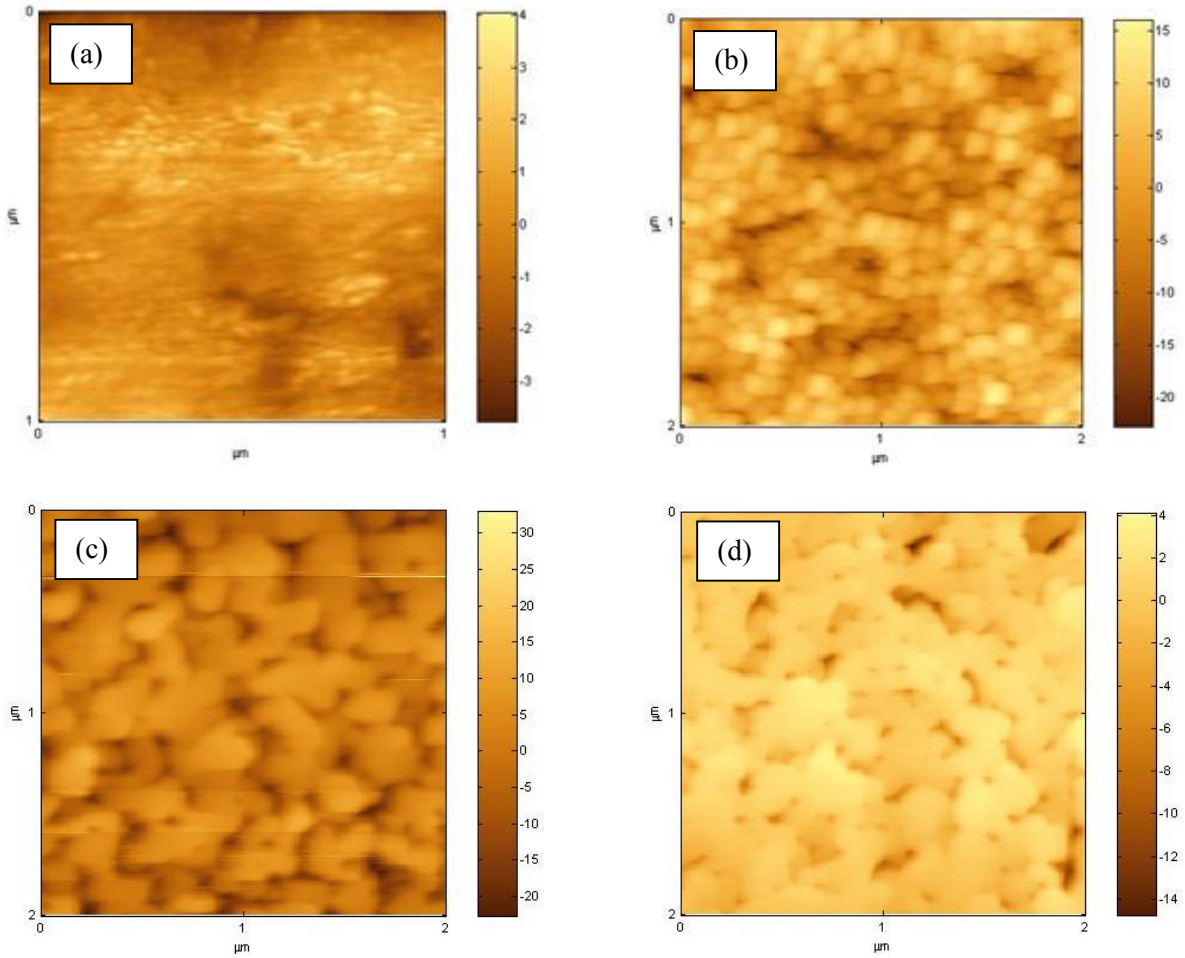


Figure 2. AFM images for $Cr_{78}Al_{22}$ thin films prepared on (a) quartz (t = 12.5 nm); (b) quartz (t = 200 nm); (c) MgO(110) (t = 100 nm) and (d) a-plane sapphire (t = 400 nm). The height range (in nm) are indicated to the right of each figure.

thickness 400 nm prepared on a-plane sapphire). This corresponds with the low mosaicity and high crystallinity of the samples prepared on a-plane sapphire reported from the XRD results for this thickness series.

The electrical resistance (R) as a function of temperature (T) measurements were used to electrically characterise the samples. T_N could not be determined for the samples investigated as the transition from antiferromagnetic to paramagnetic phase it is expected to occur at $T > 800$ K for this Al concentration. The $R - T$ curves of the samples prepared on quartz and MgO(110) show negative gradients in the temperature regions investigated, indicative of narrow-band semiconductor behaviour – as was also observed for bulk $Cr_{78}Al_{22}$ [1]. Representative curves are shown in figure 3 for the samples prepared on (a) quartz ($t = 25$ nm) and (b) MgO(110) ($t = 50$ nm). Although only representative data are shown figure 3, all the samples in each series were characterized and results indicate that resistance decreases with increase in layer thickness as can be expected. The $R - T$ behaviour of the thickness series prepared on MgO(100) however show a changing behaviour as the thickness is varied. Figures 3 (c) and (d) shows the behaviour for the $Cr_{78}Al_{22}$ thin film samples prepared on MgO(100) with thicknesses 50 nm and 400 nm, respectively. Over this series the negative gradient observed for the thinner samples gradually makes way for a clear change in gradient in the $R - T$ curve from $dR/dT > 0$ to $dR/dT < 0$ at a temperature of approximately 230 K. Interestingly, this implies that the structural and dimensionality effects in the thin film samples prepared on MgO(100) resulted in new characteristics different from the semiconductor behaviour observed in bulk samples of this concentration.

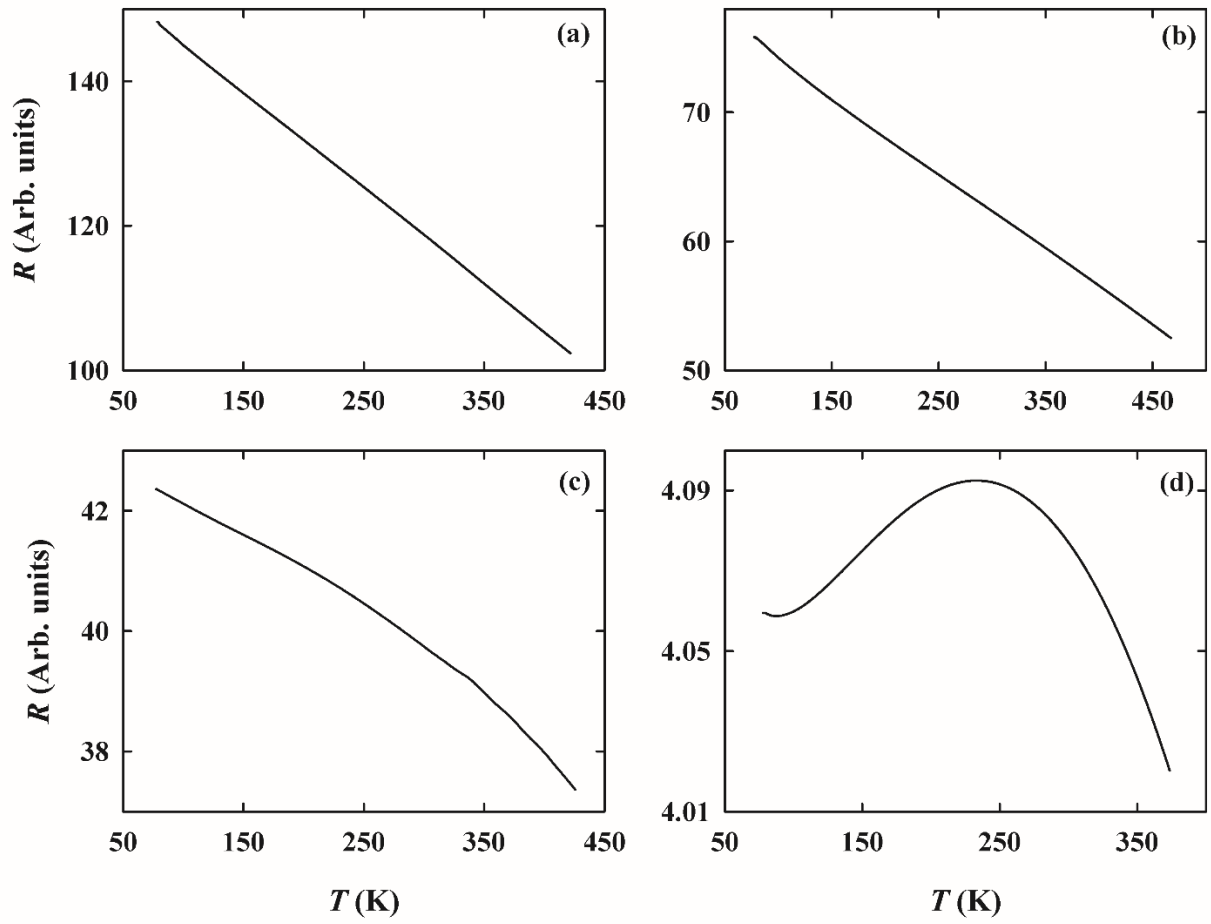


Figure 3. Resistance (R) versus temperature (T) graphs for a $Cr_{78}Al_{22}$ thin films prepared on (a) quartz ($t = 25$ nm); (b) MgO(110) ($t = 50$ nm); (c) MgO(100) ($t = 50$ nm) and (d) MgO(100) ($t = 400$ nm).

4. Conclusions

The results show clear difference between films depending on substrate and film thickness, both in the structure and transport properties. For the $Cr_{78}Al_{22}$ thickness series prepared on MgO(110) XRD results indicated a general decrease in mosaicity and an increase in coherence length with increase in the layer thickness. Narrow bandgap semiconductor behaviour was inferred to from the $R - T$ measurements for the samples prepared on MgO(110) and quartz. The series prepared on a-plane sapphire showed high crystallinity with optimal coherence length at t = 100 nm. Preliminary AFM studies show that the films prepared on the a-plane sapphire was more smooth than those prepared on MgO(110) that show more columnar growth.

All the samples in the $Cr_{78}Al_{22}$ thickness series on quartz also showed narrow bandgap semiconductor behaviour in the temperature range studied. AFM studies showed that the smooth surface of the thin samples is gradually replaced by bigger grain structures and an increase in the surface roughness. Cubic crystallites are observed at t = 200 nm, followed by columnar growth seen in the sample with t = 400 nm. The XRD and AFM results for this series is complementary.

For the $Cr_{78}Al_{22}$ thickness series prepared on MgO(100) a general decrease in mosaicity with increase in layer thickness is observed, with the coherence length of the samples optimal at a thickness of 100 nm. The electrical characteristics of this sample series gradually changed with layer thickness, and the $R - T$ semiconductor behaviour disappeared in the thickest film for temperatures below approximately 230 K.

More work is required to fully understand the complex dimensionality effects behaviour seen in the $Cr_{78}Al_{22}$ thickness series prepared on MgO(100). In order to do so it is suggested that the sample series be extended to include more samples with different sample thicknesses so that strong correlations can be drawn between the morphological, structural and electrical properties of these thin films.

5. Acknowledgement

The authors wish to thank the National Research Funding of South Africa for financial support towards this study (Grant number 80928; 80631 and 93551) and the financial contributions from Faculty of Science from the University of Johannesburg is also acknowledged.

References

[1] Fawcett E, Alberts HL, Galkin VY, Noakes DR, Yakhmi JV 1994 *Rev. Mod. Phys.* **66** 25
[2] Zabel HJ, J. Phys.: Condens. Matter 11 (1999) 9303
[3] Fullerton EE, Robertson JL, Prinsloo AER, Alberts HL, Bader SD 2003 *Phys. Rev. Lett.* **91** 237201
[4] Kummamuru RK, Soh YA 2008 *Nature* **452** 859
[5] Sheppard CJ, Prinsloo ARE, Alberts HL, Muchono B and Strydom AM 2014 *J. Alloys Compd.* **595** 164
[6] Mattson JE, Fullerton EE, Sowers CH, Bader SD 1995 *J. Vac. Sci. Technol.* **A13** (2) 276
[7] Mattson JE, Brurnitt B, Brodsky MB, Ketterson JB 1990 *J. Appl. Phys.* **67** (9) 4889

Electrical characterization of undoped and niobium-doped n-type silicon diodes

M J Thebe[1], S J Moloi[1 (a)] and M Msimanga[2]
[1] University of South Africa, Department of Physics, P.O. Box 392, Pretoria 0003, South Africa.
[2] Tshwane University of Technology, Department of Physics, Private Bag X 680, Pretoria 0001, South Africa.

Email: mohapit@vut.ac.za

Abstract. This work presents the effects of niobium impurities on electrical characteristics of silicon diodes. The work is carried out with the ultimate aim of improving radiation-hardness of silicon. Schottky diodes were fabricated on niobium- doped *n*-silicon substrates. The diodes were also fabricated on pure *n*-silicon for control purpose. The fabricated diodes were characterised by current-voltage and capacitance-voltage measurements. The diodes fabricated on niobium-doped *n*-silicon show Ohmic behaviour and low-voltage peak. These two features indicate that silicon has become relaxation material after doping with niobium. Relaxation material is radiation-hard since the effects of radiation on device based material are suppressed.

1. Introduction

The dominance of silicon in a wide variety of applications such as in high energy physics detectors, microelectronics, solar cells and photodiodes cannot be argued against [1]. This dominance is ascribed to well established methods that are available for manufacturing silicon into single crystals which make it cost effective when compared to other semiconductor materials. This study focuses on silicon diodes that are used as radiation detectors. When they are used continuously under harsh radiation environments, these detectors fail to operate efficiently [2]. The incident radiation damages the detectors and this result in a change in electrical characteristics of the devices [3]. The detector exhibits degradation in charge collection efficiency and increase in leakage current and in depletion voltage [4] after the damage. These changes are due to the defect levels that are created in the band gap of silicon material by incident particles [5]. It is then required that the characteristics of detectors are stable throughout their operational lifetime.

Defect engineering has been found to be an effective strategy to mitigate the failure experienced by detectors operating in harsh radiation conditions [6]. Defects in silicon, in this case, are intentionally introduced by metal doping or by irradiation before the fabrication of detector. Metals that have been used for doping are gold and platinum [7-10]. Erbium and niobium in silicon have been found to have similar effects as gold and platinum [11]. These metals create 'midgap defects' in the band gap of silicon and have been found to be responsible for making silicon radiation tolerant [7, 10]. Pre-irradiation, on the other hand, involves exposure of silicon material to heavy radiation prior to fabrication or processing of the devices for radiation detector [12]. Doping with metals and pre-irradiation generate defect levels in the energy gap that are responsible for relaxation behaviour of the

material [10]. Relaxation material is radiation-hard and differs from lifetime material in terms of the magnitude of carrier lifetime. In relaxation material $\tau_D >> \tau_0$ while in lifetime material $\tau_D << \tau_0$ [13]. In these relations τ_D is dielectric relaxation time and τ_0 is the minority carrier lifetime. The work presented here aims to consolidate and build on studies presented earlier [9-11].

2. Experimental Details

A 7.62 cm diameter *n*-type silicon (111) oriented wafer was acquired from Semiconductor Wafer Inc. The resistivity of the wafer was quoted by the manufacturer as 1 - 20 Ω.cm and the thickness as 275± 25.0 µm. The wafer was diced into 0.5 cm x 0.5 cm pieces using a laser cutter. The standard procedure of cleaning silicon samples was followed [14]. After cleaning process the two pieces were mounted in the chamber for niobium implantation. The implantation was carried out using an ion implanter set up at iThemba LABS (Gauteng). Niobium was implanted onto the front (or polished) side of silicon pieces at an energy of 160 keV. The implanted dose was measured by Rutherford Backscattering Spectrometry (*RBS*) technique and found to be 0.93 x 10^{15} cm^{-2}[15]. A detailed material process has been outlined elsewhere [15] and will not be repeated here.

Devices in this study are silicon diodes that were fabricated on undoped and niobium-doped *n*-type crystalline silicon. Prior to diode fabrication, silicon pieces were cleaned using the standard procedure [14]. The pieces were then loaded into an evaporation system for formation of Schottky contacts. The contacts were achieved by evaporation and deposition of 100 nm palladium through a mask with 0.6 mm diameter holes. The deposition was carried at 10^{-6} mbar at the rate 1Å/s. The Ohmic contact was realised by rubbing Indium Galinide (InGa) onto the back (unpolished) surface of the pieces. The finished devices each consists of 16 diodes on a piece and with one common Ohmic contact.

3. Results and discussion

The fabricated diodes have been characterized by current-voltage (*I-V)* and capacitance (*C-V)* techniques in the dark and at room temperature. The current through the diode of series resistance, *R* is given [16] as

$$I = I_S \left[\exp\left(\frac{q(V-IR)}{\eta k_B T}\right) - 1\right] \quad (1)$$

where I_S is the saturation current, q is the electronic charge, V is the applied voltage, η is the ideality factor, k_B is the Boltzmann constant, and T is the absolute temperature. The term IR in equation 1 is the voltage drop across R. The above equation indicates that the current varies exponentially with applied voltage. It can also be noticed from the equation that the magnitude of the current obtained in reverse bias is lower than the one obtained in forward bias at the same voltage.

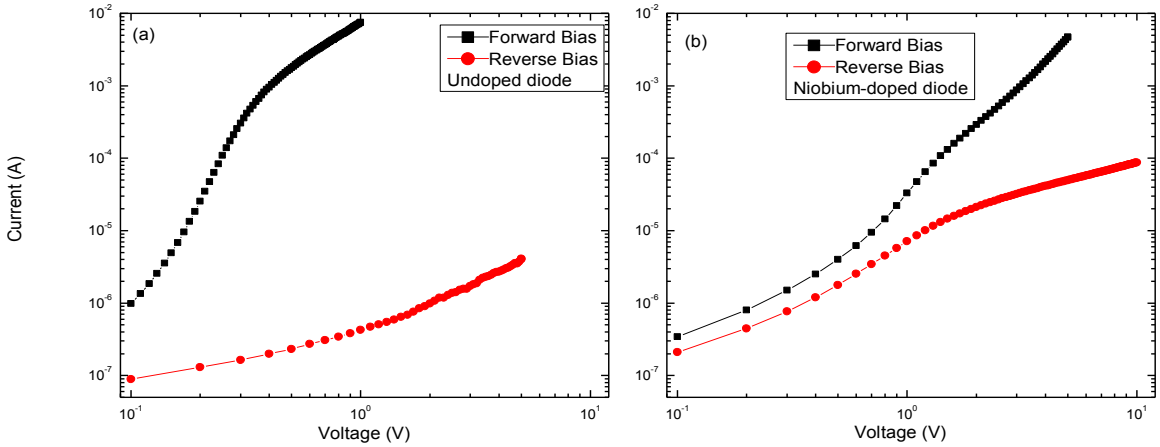


Figure 1: *The I-V characteristics of diodes fabricated on undoped (a) and niobium-doped (b) n-type silicon at 300 K.*

Figure 1 shows the *I-V* characteristics of the diodes fabricated on undoped and niobium-doped *n*-type silicon in a logarithmic scale. Figure 1(a) shows typical characteristics of silicon diodes with a big gap between reverse and forward bias trends. This gap indicates that the current measured in reverse bias is much lower than that measured in forward bias. These results indicate that the diode is well fabricated since they exhibit typical behaviour of silicon diodes [16-17]. This typical behaviour is justified by the evaluated ideality factor (of 1.19 ± 0.04) close to the unity. Thus, the diodes are fabricated on lifetime material [16-17].

Figure 1(b) shows *I-V* characteristics of the diodes fabricated on niobium-doped *n*-type silicon in logarithmic scale. The plots show a notable difference from those presented in figure 1(a) for undoped *n*-type silicon. The trends in figure 1(b) are close to each other especially, at lower voltages, ranging from 0.1 to 0.6 V. This shows that the diodes have changed from typical exponential to Ohmic behaviour after doping with niobium. This Ohmic behaviour has been explained in terms of "midgap defect" [9-11]. This defect is found at the centre of the energy gap (0.56 eV) and is responsible for change in silicon material from lifetime to relaxation behaviour [9-11]. Similar results were obtained on the diodes that were fabricated on *p*-type silicon material [11]. Thus, niobium doping induces similar effects in both types of the material.

Capacitance for Schottky diodes is represented in terms of the reverse voltage (V_R) [20] as

$$C = A\sqrt{\frac{q\varepsilon_s\varepsilon_o N_D}{2(V_{bi} + V_R)}} \qquad (2)$$

where A is the active area of the diode, ε_s (= 11.8) is the dielectric constant of silicon, ε_o is the dielectric constant of free space, N_D is the doping density and V_{bi} is the built in voltage of the diode. Equation (2) shows that $C \propto \sqrt{V_R^{-1}}$ for a constant N_D.

Figure 2 shows *C-V* characteristics of the diodes fabricated on undoped and niobium-doped *n*-type silicon in linear scale. Figure 2(a) shows typical characteristic of silicon diode with rapid decrease of capacitance at low voltages. A tendency of the capacitance towards saturation observed at 1 V indicates a gradual approach of the device towards full depletion. These *C-V* results complement

those of *I-V* and indicate that the diode is well fabricated since it exhibits typical behaviour of silicon diode [16-17]

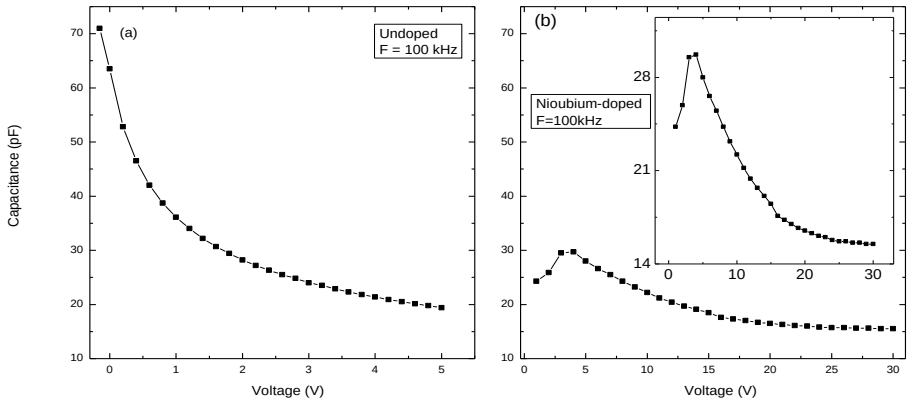


Figure 2: *The C-V characteristics of diodes fabricated on undoped (a) and niobium-doped (b) n-type silicon measured at 300 K and 100 kHz.*

Figure 2(b) shows *C-V* characteristic of the diode fabricated on niobium-doped *n*-type silicon. The figure shows a notable difference from the one presented in figure 2(a) for undoped *n*-type silicon. The trend in figure 2(b) exhibit a capacitance peak at lower voltages and thereafter a gentle decrease in capacitance is observed up to 15 V. At voltages higher than 16 V the capacitance saturates showing the diode has become fully depleted. This gentle decrease in capacitance indicates that the doping density has increased. The existence of the peak, on the other hand, has been a subject of discussion for many years and has not been fully understood. According to Ho *et al*. [21], Sahin *et al*. [22] and Ouennoughi and Sellai [23] the existence of the peak is due to interface states while the work presented by Chattopadhyay and Sanyal [24] explains the peak in terms of series resistance. We argue here that the peak is due to defects that are induced by niobium in the band gap of silicon. If the cause was due to surface effects or interface states, the peak would have also been observed on the diodes that were fabricated on undoped *n*-type silicon. This peak was also observed on the diodes that were fabricated on gold-doped *n*-type silicon and were interpreted in terms of relaxation theory [9]. Similar to gold in silicon, niobium generates defect levels that are responsible for change in material behaviour from lifetime to relaxation behaviour.

4. Conclusion

Schottky diodes were well fabricated on undoped and niobium *n*-type silicon using palladium as Schottky contacts. Current and capacitance measurements have been used to show that in silicon niobium is responsible for change in electrical properties of the diodes from exponential to Ohmic *I-V* behaviour. The diodes fabricated on niobium-doped *n*-type silicon also show a low voltage capacitance peak. These two features have also been observed in the work presented previously involving gold-doping [9, 11] and were explained in terms of midgap defect that is generated by gold in the energy gap of silicon. This defect level is situated very close to the centre of the energy gap (~ 0.56 eV) and is responsible for relaxation behaviour of the material [18]. Relaxation material is radiation-hard, since properties of the material based diode are not affected by incident radiation [7, 10, 19]. The results presented here show that doping with niobium has similar effects as doping silicon with gold. It is, therefore, expected that the material based diodes would be resistant to radiation environment.

It is intended to perform deep level transient spectroscopy (DLTS) on the diodes to investigate the exact properties of defect levels generated by niobium in silicon. The fabricated diodes will also be irradiated to determine any suppression of radiation effects as it has been done before with gold-doped silicon diodes [10]. This work would then assist in the fabrication of the ultimate radiation detectors.

Acknowledgments
This work has been made possible by financial assistance from Unisa Postgraduate Bursary Department. We would also like to thank iThemba LABS (Gauteng) for assistance with sample preparations.

References

[1] Martini M and McMath T A 1970 *Nucl. Instr. Meth.* **79** 259.
[2] Van Lint V A J 1987 *Nucl. Instr. Meth. A* **253** 453.
[3] Edmonds T 1990 *Appl. Phys. Lett.* **57** 487.
[4] Li Z, Chen W and Kraner H W 1991 *Nucl. Instr. Meth. A* **308** 585.
[5] Moll M *et al*. 1997 *Nucl. Instr. Meth. A* **388** 335.
[6] Moll M, Fretwurst E and Lindstrom G 1999 *Nucl. Instr. Meth. A* **426** 87.
[7] Dixon R L and Ekstrand K E 1986 *Radiation Protection Dosimetry* **17** 527.
[8] Kwon Y K, Ishikawa T and Kuwano H 1987 *J. Appl. Phys.* **61** 1055.
[9] Msimanga M, McPherson M and Theron C 2004 *Radiation Physics and Chemistry* **71** 733.
[10] McPherson M, Sloan T and Jones B K 1997 *J. Phys. D: Appl. Phys*. **30** 3028.
[11] Moloi S J and McPherson M 2009 *Physica B* **404** 3922.
[12] Litovchenko P G *et al.* 2006 *Nucl. Instr. Meth. A* **568** 78.
[13] Haegel N M 1991 *Appl. Phys. A* **53** 1.
[14] Cimilli F E, Saglam M and Turut A 2007 *Semicond. Sci. Technol.* **22** 851.
[15] Thebe M J, Moloi S J and Msimanga M 2015 *Depth profiling of niobium into n- silicon for radiation-hard detectors*, in preparation.
[16] Sze S M 1981 *Physics of Semiconductor Devices*, 2nd ed. (Wiley: New York).
[17] Streetman B G 1990 *Solid State Electronic Devices*, 3rd ed. (Prentice Hall: London).
[18] Jones B K, Santana J and McPherson M 1997 *Nucl. Instr. Meth. A* **395** 81.
[19] Brudnyi V N, Grinyaev S N and Stepanov V E 1995 *Physica B* **212** 429.
[20] Grove A S, 1967 *Physics and Technology of Semiconductor Devices*, Wiley New York.
[21] Ho P S, Yang E S, Evans H L, and Wu Xu 1986 *Phys. Rev. Lett.* **56** 177.
[22] Sahin B, Cetin H and Ayyildiz E 2005 *Solid-State Commun.* **139** 490.
[23] Ouennoughi Z and Sellai A 2010 *Inter. Journ. Electr.* 5 **83** 571.
[24] Chattopadhyay P and Sanyal S 1995 *Appl. Surf. Sci*. **89** 205.

Electrical characterisation of 5.4 MeV alpha-particle irradiated; low doped, n-type gallium arsenide

S M Tunhuma ,F D Auret ,M J Legodi and M Diale[1]
Physics Department, University of Pretoria, Private bag X20, Hatfield 0028, South Africa.
E-mail:malven.tunhuma@up.ac.za

Abstract. We have investigated the effects of alpha particle irradiation on the electrical characteristics of Au/n-GaAs Schottky diodes The diodes were irradiated with an Am-241 alpha-particle source up to a fluence of 2.56×10^{10} cm^{-2} at room temperature. The induced defects were studied using deep level transient spectroscopy (DLTS) in the 15-300 K range and revealed the defects $E_{0.04}$, $E_{0.14}$, $E_{0.17}$ and $E_{0.38}$. The current-voltage *(I-V)* characteristics remain largely unchanged after irradiation, whilst capacitance-voltage (*C-V)* characteristics showed a decrease in net doping density**.**

1. Introduction

Defect engineering has enabled the development of optimized semiconductor material structures [1]. Understanding the physical properties and occurrence of defects will potentially lead to optimal device design. Particle irradiation induced defects modify the electronic properties of semiconductors [2, 3]. These modifications lead to applications such as carrier lifetime control and device isolation [1, 4].
GaAs based devices have been used in space applications, where they are exposed to irradiation from highly energetic particles. In the past, studies have been carried out in order to gain insight on the effects of particle irradiation on GaAs devices [4-8]. In this study we have investigated the defects introduced in Au-n/GaAs Schottky barrier diodes by alpha particle irradiation using DLTS. This study was limited to only low doped GaAs. DLTS is a convenient and powerful technique for studying defects in semiconductor materials [8].

2. Experimental details

(MOVPE) grown, silicon doped *n*-GaAs with an average free carrier density of 1.0×10^{15} cm^{-3} and <100> orientation, as specified by Epi Materials limited the supplier was used. The wafers were degreased and etched chemically using trichloroethylene and isopropanol. Au-Ge (88%:12%) ohmic contacts were resistively deposited on the n^{+}- backsides of the samples [1]. Thereafter 1000 Å thick Au contacts, 0.6 mm in diameter, were resistively deposited on the epitaxial layer. Contact quality was evaluated using *I-V* and *C-V* measurements.
The samples were irradiated with 5.4 MeV He ions (α-particles) from an Am-241 radionuclide for 20 min to a fluence of 2.56×10^{10} cm^{-2}. DLTS spectra were recorded at a scan rate of 3 K/min in the 15–300 K temperature range. The quiescent reverse bias was –2 V, filling pulse amplitude 0 V and filling pulse width 1 ms.

[1] To whom any correspondence should be addressed.

The DLTS defect signature (energy level in the band gap, E_t and apparent capture cross section, σ) were determined from the slope and y-intercept, respectively, of log (e_n/T^2) versus (1000/T) Arrhenius plots according to the equation

$$e_n = \sigma_n \langle v_{th} \rangle \frac{g_o}{g_1} N_c \exp\left(-\frac{E_c - E_t}{k_B T} \right)$$

where e_n is the emission rate, k_B is the Boltzmann constant, T the temperature, $\langle v_{th} \rangle$ is the thermal velocity of electrons, N_c is the density of conduction band states, g_0 and g_1 are the degeneracy terms for the states before and after electron emission.

3. Results and discussion

The parameters deduced from the *I-V* and the *C-V* measurements of the Schottky diodes before and after irradiation are shown in Table 1. The *I-V* characteristics do not show any changes in ideality factor and barrier height after particle irradiation. However, the *C-V* characteristics show a decrease in the net doping density. The as-deposited value of carrier concentration is different to the one specified by the suppliers something which can be attributed to several factors chief amongst them a smaller active diode area. The decrease after irradiation has been explained as due to the electron traps introduced by the irradiation, trapping carriers [4].

Table 1. Schottky diode *I-V* and *C-V* parameters obtained from the Schottky diode before and after α-particle irradiation.

	Ideality factor	Barrier height (eV)	Nd ($\times 10^{14}$ cm^{-3})
As deposited	1.02	0.89	7.64
5.4 MeV irradiated	1.02	0.89	6.68

Figure 1 depicts the DLTS scans obtained at a frequency of 200 Hz. Curve (a) is the control spectrum obtained from the as-deposited samples before irradiation. It shows no defects present in measurable quantities in the 15-300 K temperature range. Curve (b) reveals emission peaks from defects induced by α-particle irradiation superimposed on a "skewed baseline'. The peaks are labelled $E_{0.04}$, $E_{0.14}$, $E_{0.17}$ and $E_{0.38}$ based on their energy levels.

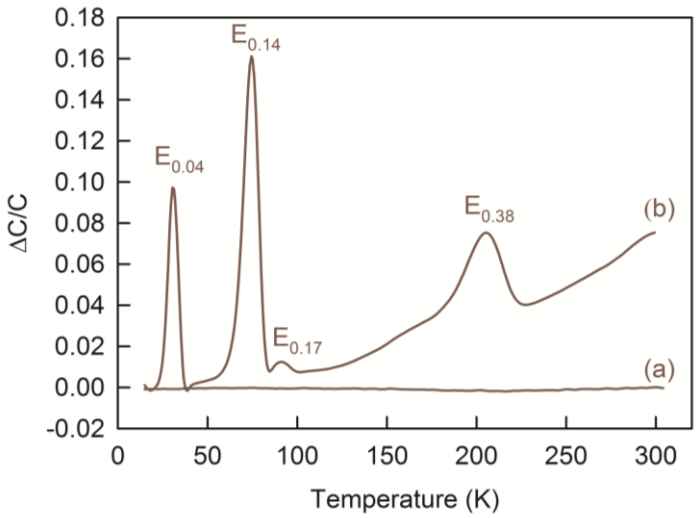


Figure 1. (a) DLTS spectra of un-irradiated (b) 5.4MeV alpha particle irradiated samples recorded at a quiescent reverse bias of 2 V at a rate window of 200 Hz and filling pulse 0 V and pulse width 1 ms.

A metastable defect that has been observed by other researchers was not observed in our scans [9]. This is because its presence and magnitude is highly dependent on the bias conditions, temperature and incident particle type [1, 5]. The "skewed baseline' in curve (b) has been observed only in α-particle irradiated gallium arsenide but not in other semiconductor materials exposed to the same type of radiation. It can therefore not be explained in terms of ion-solid interactions. A similar baseline was observed by Janse van Rensburg *et al* [10] in europium and xenon implanted GaN thin films without conclusion on its cause. The forward bias filling pulse in all the measurements done for the purpose of this study was -0.5 V in order to avoid the capacitance signal from surface states. This also shows that the baseline is not a consequence of surface states [11]. Surface states are sometimes confused with, or prevent, the detection of the presence of deep levels in the bulk of the material [12].
Irradiation induced defects are a result of collision of energetic particles. The Frenkel pair, is created when a particle imparts energy to a lattice atom to displace it forming a vacancy-interstitial pair [13]. However considering the energy transferred to atoms of the crystal by alpha-particles not only point defects, but defects consisting of clusters of atoms displaced from their normal lattice sites can be formed [7]. The Arrhenius plots for the defect characteristics are shown in Figure 2, from them defect 'signatures' shown in Table 2 were deduced.
The defects $E_{\mathbf{0.04}}$, $E_{\mathbf{0.14}}$ have been confirmed to be caused by primary defects such as the single and double charge states of the vacancy in the As sub-lattice. It has been proposed that both are different charge states (o/+) and (+/o) of the isolated vacancy V_{As} in GaAs. The $E_{\mathbf{0.38}}$ is speculated to be related to close As vacancy interstitials whilst the $E_{0.17}$ is a metastable defect with transformational kinetics that displays charge state dependent first order behavior [12, 14] .

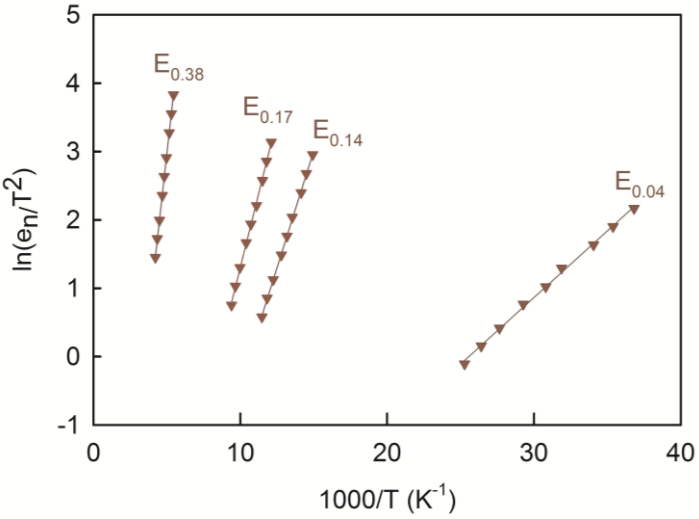


Figure 2. Arrhenius plots of defects observed after alpha particle irradiation.

$E_{0.04}$, $E_{0.14}$ and $E_{0.38}$ have the same electronic structure and are point defect in nature as they are observable with the same signatures after being induced by different radiation types [11]. The charge state of the $E_{0.38}$ is highly field dependent [4]. Further, the defects are dependent on the growth technique and carrier density implying that they emanate from the defects and impurities on the as grown GaAs [11].Table 2 summarizes the defect signatures. The defect characteristics shown in the table corresponds to those that have been observed by other researchers [7].

Table 2 . Characteristics of alpha-particle irradiation induced defects detected by DLTS in OMVPE-grown *n*-GaAs.

Defect label	Activation Enthalpy (meV)	Capture cross section (cm^{-2})
$E_{0.04}$	38.7	1.6×10^{-17}
$E_{0.14}$	135	3.0×10^{-15}
$E_{0.17}$	171	3.4×10^{-13}
$E_{0.38}$	382	7.4×10^{-16}

4. Conclusions

We have used DLTS to study the defects in the 15 -300 K range in α-particle irradiated GaAs. Four defect peaks were observed namely the $E_{0.04}$, $E_{0.14}$, $E_{0.17}$ and $E_{0.38}$ defects which are all point defect in nature. They have been associated with solid-ion interactions. The defects have been previously shown to be related to isolated vacancies and vacancy-interstitial pairs in the As sub-lattice. The $E_{0.38}$ defect has been observed to be field dependent. The *(I-V)* characteristics remained significantly constant confirming the radiation hardness of GaAs. A decrease in carrier density was observed from the (*C-V*) characteristics after irradiation. We therefore can conclude that alpha particle irradiation affects the electrical characteristics of *n*-GaAs based devices.

Acknowledgements

The authors gratefully acknowledge financial assistance from the University of Pretoria, the South African National Research Foundation and Johan Janse van Rensburg for technical assistance.

References

1. Tunhuma, S.M., et al., *The fine structure of electron irradiation induced EL2-like defects in n-GaAs.* Journal of Applied Physics, 2016. **119**(14): p. 145705.
2. Paradzah, A.T., et al., *Electrical characterization of 5.4MeV alpha-particle irradiated 4H-SiC with low doping density.* Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 2015. **358**(0): p. 112-116.
3. Omotoso, E., et al., *The influence of high energy electron irradiation on the Schottky barrier height and the Richardson constant of Ni/4H-SiC Schottky diodes.* Materials Science in Semiconductor Processing, 2015. **39**(0): p. 112-118.
4. Meyer, Walter E., F.D. Auret, and Stewart A. Goodman, *Electric Field Enhanced Emission from Two Alpha-Particle Irradiation Induced Traps in n-GaAs.* Japanese Journal of Applied Physics, 1996. **35**(1A): p. L1.
5. Goodman, S.A., et al., *Electrical and defect characterization of n-Type GaAs irradiated with α-particles using a van de graaff accelerator and an Am-241 radio-nuclide source.* physica status solidi (a), 1993. **140**(2): p. 381-390.
6. Goodman, Stewart A., F.D. Auret, and Walter E. Meyer, *Electric Field Effect on the Emission of Electron-Irradiation-Induced Defects in n-GaAs.* Japanese Journal of Applied Physics, 1994. **33**(4R): p. 1949.
7. Goodman, S.A., F.D. Auret, and W.E. Meyer, *The effect of alpha-particle and proton irradiation on the electrical and defect properties of n-GaAs.* Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 1994. **90**(1–4): p. 349-353.
8. Legodi, M.J., F.D. Auret, and S.A. Goodman, *Dopant-related metastable defects in particle irradiated n-GaAs.* Physica B: Condensed Matter, 1999. **273–274**(0): p. 762-765.
9. Auret, F.D., S.A. Goodman, and W.E. Meyer, *New electron irradiation induced electron trap in epitaxially grown Si-doped n-GaAs.* Applied Physics Letters, 1995. **67**(22): p. 3277-3279.
10. Janse van Rensburg, P.J., et al., *Electrical characterization of rare-earth implanted GaN.* Physica B: Condensed Matter, 2009. **404**(22): p. 4411-4414.
11. Auret, F.D., et al., *Electrical characterization of defects introduced in n-GaAs by alpha and beta irradiation from radionuclides.* Applied Physics A, 1993. **56**(6): p. 547-553.
12. Auret, F., D, et al., *Electrical characterization of particle induced damage in n-GaAs*. STEDCON '92: International Conference on Science and Technology of Electron Devices : PapersVol. 16. 1993: South African Journal of Physics.
13. Coelho, S.M.M., et al., *Electrical characterization of defects introduced in n-Ge during electron beam deposition or exposure.* Journal of Applied Physics, 2013. **114**(17): p. 173708.
14. Auret, F.D., M.E. Rudolph , and A.G. Stewart *A Metastable Alpha-Particle Irradiation Induced Defect in n-GaAs.* Japanese Journal of Applied Physics, 1994. **33**(4A): p. L491.

A quantitative evaluation of the depth resolution of AES depth profiling of Cu/Ni multilayer thin films using the MRI model.

X.L. Yan [1], Y. Liu [2], H.C. Swart [1], J.Y. Wang [2], J.J. Terblans [1]*
[1] Department of Physics, University of the Free State, PO Box 339, Bloemfontein, 9300, South Africa
[2] Department of Physics, Shantou University, 243 Daxue Road, Shantou, 515063 Guangdong, China

E-mail: terblansjj@ufs.ac.za

Abstract: Depth profiles of as-deposited Cu/Ni multilayer thin films were investigated using Auger electron spectroscopy (AES) in combination with Ar^+ ion sputtering. The Cu/Ni multilayer structures were deposited on a SiO_2 substrate by means of electron beam evaporation in a high vacuum. The measured AES depth profiles of the as-deposited Cu/Ni multilayer was quantitatively fitted by the MRI model assuming that the roughness parameter has linearly increased with the sputtered depth. The roughness values extracted from the depth profiling data fits, agreed well with those measured by atomic force microscopy (AFM). The depth-dependent depth resolution upon depth profiling of the Cu/Ni polycrystalline multilayer was quantitatively evaluated accordingly.

1. Introduction

Auger electron spectroscopy (AES) in combination with ion beam sputtering is widely used for the determination of the composition-depth profiles of thin films. The quality of depth profiling can be characterized by the so-called depth resolution Δz, which defines the depth range to which a certain composition has to be assigned (see section 3) [1]. The measured depth profile is always influenced by factors such as atomic mixing, sputter-induced surface roughness, information depth, preferential sputtering, segregation, etc. [2]. Often the measured depth profile as compared to the true concentration-depth profile is described by a so-called depth resolution function (DRF), which can be experimentally determined or theoretically estimated. The Mixing-Roughness-Information depth (MRI) model is a theoretical description of the DRF taking into account atom mixing, sputter-induced surface roughness and information depth. The model developed by Hofmann was first given in reference [3]. Since then, numerous extensions were developed to enable its applicability for preferential sputtering [4], non-stationary mixing, information depth as well as for non-Gaussian roughness [5,6], and the analytical depth resolution function for thin delta layer [7]. Profile reconstruction from measured depth profile data can be obtained by the MRI model, where the original in-depth distribution can be obtained by a least square approach optimizing the fit of the calculated profile to the measured data.

* To whom any correspondence should be addressed.

The aim of this study is to test the MRI model on a Cu/Ni polycrystalline multilayer structure from measured AES depth profiling data with assuming a linear roughness with sputter depth to fit. The fitting roughness values are compared with those measured by atomic force microscopy (AFM). The depth-dependent depth resolution upon depth profiling of the Cu/Ni polycrystalline multilayer are quantitatively evaluated accordingly.

2. Experimental

The Ni/Cu multilayer structures composed of four pairs of Ni and Cu layers were prepared by electron beam physical vapour deposition onto passivated silicon (100) substrates (SiO_2) at a base pressure of $<7\times10^{-6}$ Torr, resulting in polycrystalline multilayer thin films. The layer thickness was monitored by an Inficom XTC thin film monitor as Cu(8 nm)/Ni(8 nm)/Cu(11 nm)/Ni(11 nm)/Cu(15 nm)/Ni(12 nm)/Cu(13 nm)/Ni(14 nm)/SiO_2. The topographies measured by AFM with scanning area 3×3 μm^2 are shown in Fig.1(a) for the surface of the as-deposited sample and in Fig.1 (b) for the crater centre for sample sputtered at the depth of 100 nm (substrate SiO_2 surface), respectively. The corresponding root mean square (RMS) roughness values were determined as 0.91 nm and 5.54 nm, respectively.

The AES depth profiles were measured using a PHI 600 SAM without sample rotation at a base pressure < 10^{-9} Torr. A static primary electron beam of 10 keV and beam current of 200 nA with beam size diameter of 9.2 μm was used. Ion sputtered was performed with 2 keV Ar^+ ions at an incidence angle of 60° to the normal of the sample surface, the beam current density was 0.127 A/m^2 with a raster area 2×2 mm^2. The Auger peak-to-peak heights (APPHs) were recorded as a function of sputtering time for Ni(680-740 eV), Cu(880-989 eV), C(240-295 eV), O(480-510 eV) and Si(70-105 eV). In order to apply the MRI model, the APPH-sputter time profile was converted into normalised APPH depth profiles. The Cu APPH was normalized to the maximum of the Cu 922 eV APPH in the as-prepared sample. The time scale was converted into the depth scale using the average sputtering rate of 0.05nm/s determined by depth profiling data with the known thicknesses of the Cu/Ni layers.

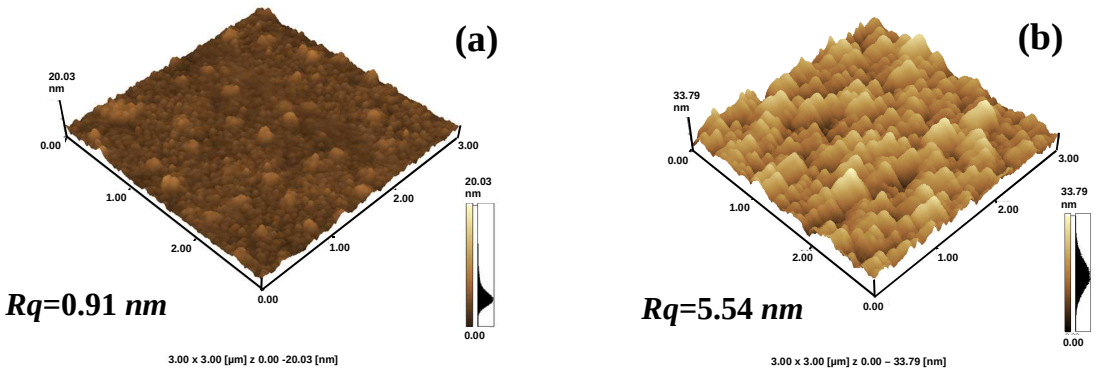


Figure 1. AFM images of (a) the surface of the as-deposited sample and (b) the crater centre for sample sputtered at the depth of 100 nm (substrate SiO_2 surface). The corresponding root mean square (RMS) *Rq* roughness values were determined as 0.91 and 5.54 nm, respectively.

3. Depth resolution and MRI model

The depth resolution Δz is defined conventionally as the depth range over which a 16-84% (or 84-16%) change in signal by a specified amount when profiling an ideally sharp interface between two media (see Fig.2) [8]. This definition has a physical meaning only for a Gaussian shape of the depth resolution function in Fig. 2. However, if the concentration saturation levels of 100% and 0% for the signal

analyzed no longer occur, for example, in case of sputter depth profiling of a thin multi-layered film, the determination of Δz defined above is no longer possible. Using the MRI model for the calculation of sputter depth profiles, it is shown that, for the same Δz, different resolution functions can be obtained with different full width at half-maximum (FWHM) values [9].

The depth resolution function in the MRI model takes into account three effects: atom mixing, escapes depth of Auger electrons, surface roughness and is described as [3]:

Mixing length (w): $g_w = \frac{1}{w}\exp\left[-\left(z - z_0 + w\right)/w\right]$ (1);

Roughness (σ): $g_\sigma = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-(z - z_0)^2/2\sigma^2\right]$ (2);

Information depth (λ): $g_\lambda = \frac{1}{\lambda}\exp\left[-\left(z_0 - z\right)/\lambda\right]$ (3).

where w is atomic mixing length, z is sputter depth, z_0 is the running depth parameter, σ is the surface roughness and λ is the information depth parameter.

Based on the above-discussed refinements of the DRF in terms of symmetric (Gaussian functions) and asymmetric (non-Gaussian functions), it is necessary to clarify the contribution to the depth resolution Δz 16-84%. According to this MRI model, a symmetric contribution to the depth resolution function originates from the intrinsic roughness and the surface roughening by ion sputtering, which both are described by a Gaussian smearing function (see Eq.(2)), characterized by its standard deviation: the surface roughness parameter σ ($\Delta z_\sigma^2 = 2\sigma^2 = 2\sigma_i^2 + 2\sigma_s^2$, where σ_i is the intrinsic roughness, σ_s is the sputter induced roughness.). For the asymmetric broadening functions, the atomic mixing is described by an exponential function (see Eq.(1)) and this exponential function is characterized by the atomic mixing parameter w ($\Delta z_w = 1.668w$); the information depth of the Auger electrons (for AES) is also described by an exponential function (see Eq.(3)) and this exponential function is characterized by the information depth parameter λ ($\Delta z_\lambda = 1.668\lambda$). Thus, three parameters (σ, w and λ) suffice to characterize the total smearing, Δz (see Ref.[1]). Values for these three parameters may be obtained experimentally and/or calculated theoretically (see the section 4).

According to the above-discussion in the MRI model, the total "symmetric" and "asymmetric" contribution to the depth resolution, Δz, can approximately be written as [10]:

$$\Delta z = \left(\left(2\sigma\right)^2 + \left(1.668w\right)^2 + \left(1.668\lambda\right)^2\right)^{1/2} \quad (4).$$

Fitting the MRI model to experimental sputtering-depth profiles in principle leads to values for σ, w and λ in Eqs. (1-3). Then Δz can be calculated using Eq.(4).

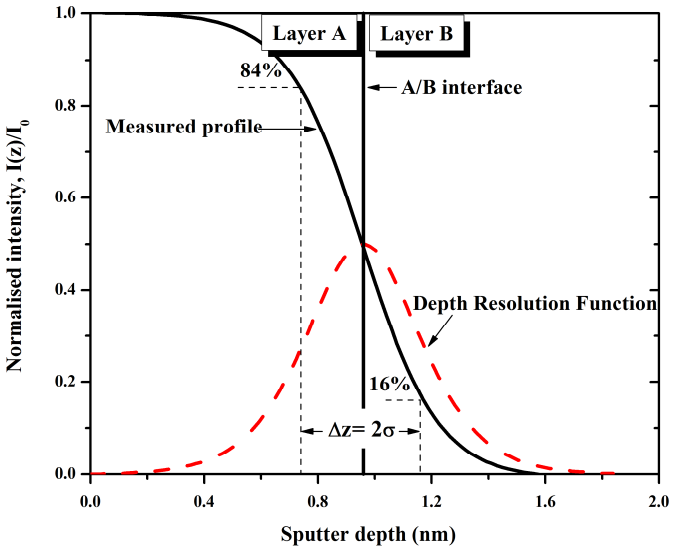


Figure 2. Schematic definition of the depth resolution, Δ*z* (84-16%), at a sharp interface in the broadening profile and a Gaussian resolution function (dashed line).

4. Results and discussion

The measured Cu depth profiling data of the as-deposited Cu/Ni multilayer is shown in Fig. 3 as open circles. A close examination of Fig. 3 revealed that the measured maximum/minimum value of each sublayer decreased/increased with the sputtered depth. This effect is related to the development of the ion bombardment induced roughness for polycrystalline material upon stationary depth profiling because of the dependence of ion sputtering yield on the orientation of the crystal. The measured Cu depth profiling data for the as-deposited sample was fitted and shown as a solid line in Fig. 3 by the MRI model assuming simply that the roughness parameter increased linearly with the sputtered depth and taking the other two MRI parameters into account. The atomic mixing length *w* was estimated as 1.0 nm by the TRIM code [11] and the Auger electron escape depth λ (effective attenuation length times cos(θ), where θ is the angle of emission of the detected electrons) for Cu (922 eV) was calculated as 0.8 nm [12]. The fitted roughness parameter was obtained as σ = (0.9+0.049×*z*) nm, which agrees very well with the measured RMS *Rq* values of 0.91 nm and 5.54 nm by AFM at the depth of *z* = 0 nm and *z* =100 nm as indicated in Fig.1(a) and (b) , respectively.

Using the aforementioned fitted parameters in the MRI model, the depth resolution Δz, as given by Eq. (4), for the as-deposited sample have been calculated as a function of the sputter depth in Fig.4 and is shown as a solid line. The fitted roughness σ for the as-deposited sample agrees very well with the *Rq* roughness values measured by AFM at the depth of *z* = 0 nm and *z* =100 nm as closed squares denoted in Fig.4 and its corresponding twice value 2σ (contribution in depth resolution, see the first term of right side in Eq. (4)) is presented in Fig.4 as dashed lines.

As is shown in Fig.4, when the sputter depth increased, the values of depth resolution Δ*z* was closer to the values of roughness 2σ.This result implies that the roughness contribution is the dominant factor degradation the depth resolution at the deeper sputtered depth. We should also note that the MRI fitting roughness parameter σ=5.0 nm at the depth *z*=100 nm is larger than the other two parameters (*w*=1.0 nm and λ=0.8 nm).

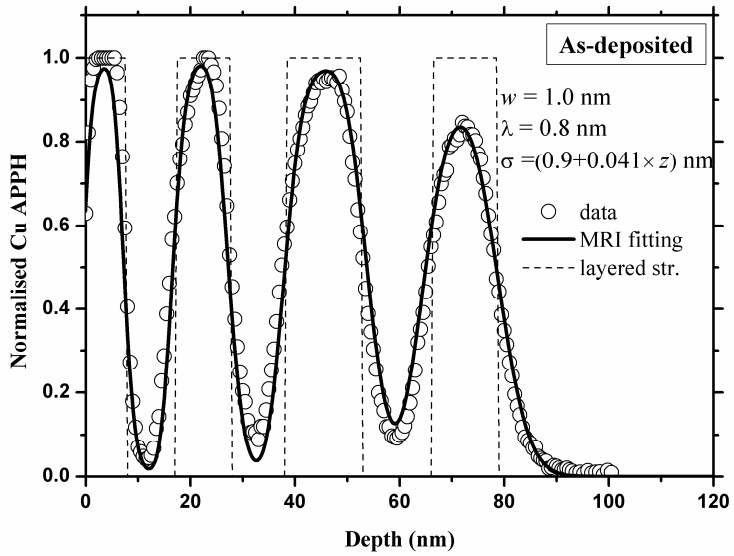


Figure 3. The normalised Cu depth profiling data and the best fit to the measured data by the MRI model for the as-deposited sample. The dashed line drawn represents the as-deposited Cu sublayer structure.

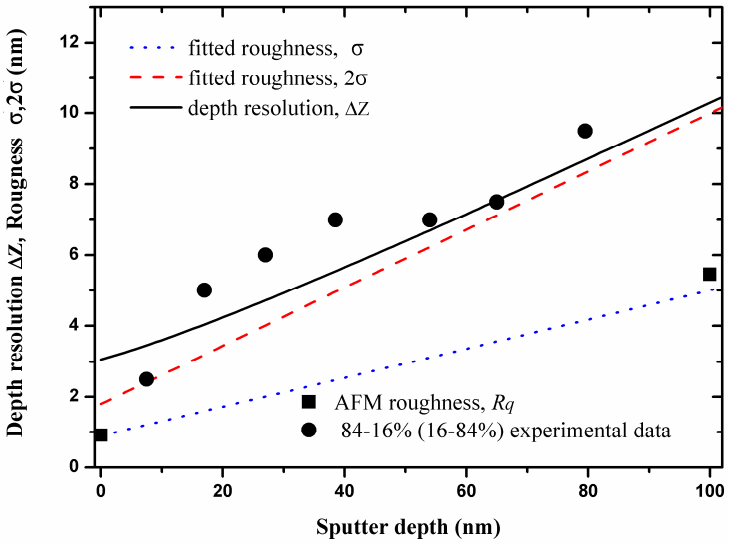


Figure 4. The depth resolution Δz (solid line) and the fitted roughness parameter σ and its double value 2σ (dotted and dashed lines, respectively) as a function of the sputter depth for the as-deposited sample. The closed circles and squares represent the values of the depth resolution by 84-16%(16-84%) method and the *Rq* roughness values determined by AFM (see Fig.1), respectively.

In additional, the values of the depth resolution in the MRI calculation are in accordance satisfactorily with the ones in the experimental data by the 84-16% (16-84%) method denoted as closed circles in Fig.4. This confirms further the capability of the MRI model with the linear roughness in the Cu/Ni polycrystalline multilayer structure. The calculated Δz/2 value of 4.1 nm at the last Cu/Ni

interface (corresponding to a depth of 78 nm) for the as-deposited sample is less than the thickness of the last sublayer Cu/Ni under the present measurement conditions.

5. Conclusions

The application and capability of the MRI model to evaluate different contributions to the depth resolution were demonstrated for the Cu/Ni polycrystalline multilayer structures by fitting the measured AES depth profiling data. The roughness parameter was assumed linearly increased with the sputtered depth. And the roughness values extracted from the MRI model fitting agreed well with those measured by AFM. The depth-dependent depth resolution upon depth profiling was quantitatively evaluated accordingly.

Acknowledgements

This work is based on the research supported by the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation of South Africa. We would like to thank the National Research Foundation and the cluster of the University of the Free State for financial assistance.

References

[1] Hofmann S 1999 *Surf. Interface Anal.* **27** 825
[2] Kovac J, Zalar A and Pracek B 2003 *Appl. Surf. Sci.* **207** 128
[3] Hofmann S 1994 *Surf. Interface Anal.* **21** 673
[4] Hofmann S (by edited), Chap. 7: *Quantitative Compositional Depth Profiling, in Auger-and X-ray Photoelectron Spectroscopy in Materials Science*, Springer Verlag Heidelberg, New York, Dordrecht, London, 2013
[5] Wang JY, Liu Y, Hofmann S and Kovac J 2012 *Surf. Interface Anal.* **44** 569
[6] Liu Y, Jian W, Wang JY, Hofmann S and Kovac J 2013 *Appl. Surf. Sci.***276** 447
[7] Liu Y, Hofmann S and Wang JY 2013 *Surf. Interface Anal.***45** 1659
[8] Hofmann S 1998 *Rep. Prog. Phys.* **61** 827
[9] Wang JY, Starke U and Mittemeijer EJ 2009 *Thin Solid Films* **517** 3402–3407
[10] Wang JY, Hofmann S, Zalar A and Mittemeijer EJ 2003 *Thin Solid Films* **444** 120
[11] Ziegler JF, TRIM Code, IBM Corporation, Yorktown Heights, NY, (http:// www.srim.org)
[12] NIST Electron EAL, 2001 *NIST Database* 82

Study of the interdiffusion in Cu/Ni multilayer thin films by Auger electron spectroscopy depth profiling

X.L. Yan [1], Y. Liu [2], H.C. Swart [1], J.Y. Wang [2], J.J. Terblans [1]*

[1] Department of Physics, University of the Free State, PO Box 339, Bloemfontein, 9300, South Africa

[2] Department of Physics, Shantou University, 243 Daxue Road, Shantou, 515063 Guangdong, China

E-mail: terblansjj@ufs.ac.za

Abstract: The interdiffusion upon annealing of Cu/Ni multilayer structures at 325 ºC, 350 ºC and 375 ºC for 30 min was studied by Auger electron spectroscopy (AES) depth profiling. The measured AES depth profiles of the unannealed and annealed samples were quantitatively fitted by the MRI model and the interdiffusion parameters, pre-exponential factor D_o and activation energy E_a, were extracted. The depth-dependence of the interdiffusion coefficient in the Cu/Ni multilayer structures was characterized with relating to the reduction of the activation energy at deeper depth during sputtering.

1. Introduction

Multilayer thin films have been increasingly used in both basic research and applications because of their specific properties that differ from those of bulk materials and single-layer thin films [1]. The interdiffusion at interfaces has a strong impact on the properties of a multilayer thin-film system [2]. In addition, diffusion in polycrystalline thin films is much faster than in bulk materials due to a high density of defects such as dislocations, vacancies, and grain boundaries, which act as a short-cut for diffusion. So diffusion in polycrystalline thin films cannot be described by the extrapolation of data obtained for bulk materials at higher temperatures.

The Auger electron spectroscopy (AES) depth profiling technique is one of the most commonly used methods for the study of diffusion in multilayer thin films due to its high surface sensitivity. In the past decades, several methods have been proposed to extract the interdiffusion coefficient from AES depth profiling data. For example: the plateau-rise method [3], the center-gradient method [4], and the interface-width method [5] for a bilayer sample and a Fourier series method for a multilayer sample [6]. In these four methods, only parts of the measured depth profile are used for extracting the interdiffusion coefficient values. Recently, based on the Mixing-Roughness-Information depth (MRI) model [7], a new method for fitting the entire measured AES depth profile has been proposed [8] and widely been used to extract the interdiffusion coefficient in layered structures [9,10]. The aim of this study was to determine the interdiffusion coefficients in the Cu/Ni multilayer structures from measured AES depth profiling data by the MRI model.

* To whom any correspondence should be addressed.

2. Experimental

2.1. Sample preparation and annealing

The Ni/Cu multilayer structures, composed of four pairs of Ni and Cu layers, were prepared by electron beam physical vapour deposition onto passivated silicon (100) substrates (SiO_2) at a base pressure of $<9.3\times10^{-4}$ Pa, resulting in polycrystalline multilayer thin films. The passivated silicon (100) substrates (SiO_2) was the barrier layer for preventing diffusion into the substrates. The Ni layer films were deposited at a rate of ~0.2 nm/s, while the Cu layer films were deposited at a rate of ~0.3 nm/s. The layer thickness was monitored by an Inficom XTC thin film monitor as being Cu(8 nm)/Ni(8 nm)/Cu(11 nm)/Ni(11 nm)/Cu(15 nm)/Ni(12 nm)/Cu(13 nm)/Ni(14 nm)/SiO_2. Then the samples were annealed at 325 ^{o}C, 350 ^{o}C and 375 ^{o}C for 30 min, respectively in a high vacuum system ($< 5\times10^{-5}$ Pa).

2.2. AES depth profiling measurement

The depth profiles of the unannealed and annealed samples were measured using a PHI 600 Scanning Auger Microscope (SAM) without sample rotation at a base pressure $< 1.3\times10^{-7}$ Pa. A static primary electron beam of 10 keV and a beam current of 200 nA was used. The samples were sputtered employing 2 keV Ar^+ ions, incidence at an angle of 60^{o}, beam current density of 0.127 A/m^2 with a 2×2 mm^2 raster. The Auger peak-to-peak heights (APPH) were recorded as a function of sputtering time for Ni(680-740 eV), Cu(880-989 eV), C(240-295 eV), O(480-510 eV) and Si(70-105 eV). The average sputter rate of the Cu/Ni multilayers was 0.05 nm/s, as determined by depth profiling data of the unannealed sample with known thicknesses of the Cu/Ni layers.

2.3. Conversion of the intensity-sputter time profile into a normalized APPH depth profile

In order to apply the MRI model, the APPH-sputter time profiles were converted into normalised APPH depth profiles. According to the standard Auger spectra of pure Cu and Ni [11], the whole range of Ni and Cu Auger peaks overlapped except for the Cu 922 eV peak. The Cu APPH was normalized to the maximum of the Cu 922 eV APPH in the unannealed sample. The time scale was converted into the depth scale using the average sputtering rate of 0.05 nm/s (see previous paragraph). Conversions of the normalised APPH-sputter time profiles are shown in Fig.1 for the unannealed and annealed samples.

3. Results and discussion

For the unannealed sample, the normalised APPH of Cu depth profiling data and the best fit to the measured data by the MRI model are shown in Fig. 1(a) as open circles and a solid line, respectively. The atomic mixing length w is estimated as 1.0 nm by the TRIM code [12] and the Auger electron escape depth λ (effective attenuation length times cos(θ), where θ is the angle of emission of the detected electrons) for Cu (922 eV) is calculated as 0.8 nm [13]. The fitted roughness parameter, assumed as a linear increase with depth in the Cu/Ni polycrystalline multilayer structure, is obtained as $\sigma_0 = (0.9+0.049\times z)$ nm, which agrees very well with the AFM measured data in Ref.[14]. In addition, Fig. 1(a) shows that the measured Cu concentration in the near substrate region is lower than that in the near surface region due to ion beam sputter-induced roughness, same as that in depth profiling of Cu/Pd, Ag/Pd polycrystalline multilayer films [15, 16].

The unannealed Cu layered structure is also shown as a dashed line in Fig. 1(a). For the annealed samples, the normalised APPH Cu depth profiling data and the best fit to the measured data are shown in Fig. 1(b-d). The fitting parameters w and λ were kept the same as for the unannealed sample and only the roughness parameter was changed. The determined linear depth dependence of the roughness parameter σ, at the different temperatures, has been indicated in Fig. 1(b-d), respectively.

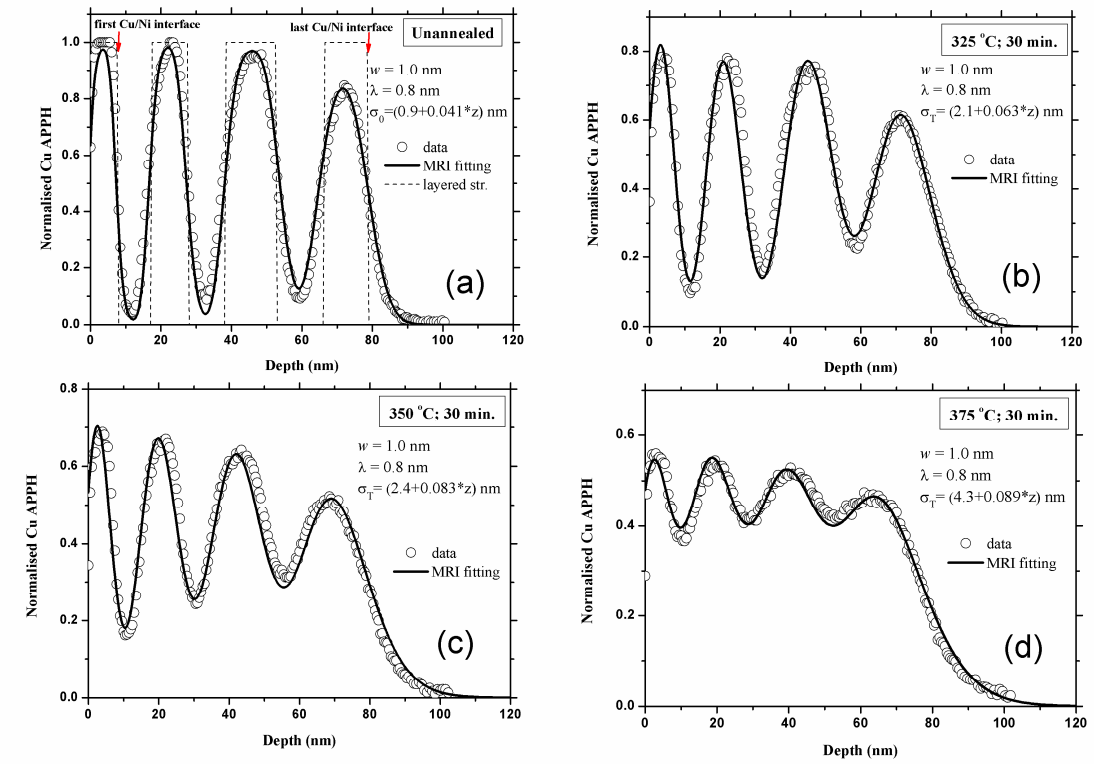


Figure 1. The normalised Cu depth profiling data and the best fit to the measured data by the MRI model for (a) the unannealed sample, and the annealed samples for 30 min at (b) 325 ºC (c) 350 ºC and (d) 375 ºC. The dashed line drawn in Fig. 1(a) represents the unannealed Cu sublayer structure.

According to the MRI model, the interface roughness broadening is described by a Gaussian function ~exp[-$z^2/2\sigma^2$] [7]. In an initial stage of diffusion, the diffusion-induced concentration profile can be written as a Gaussian function ~exp[-z^2/4Dt], where z represents the diffusion depth and the diffusion length is $(4Dt)^{1/2}$. Therefore, the square of the diffusion length can be expressed in terms of the interface roughness as [17]:

$$2Dt = \sigma_T^2 - \sigma_o^2 \qquad (1)$$

where t is the annealing time, and σ_T and σ_o are the values of the roughness parameter after and before annealing at temperature T, respectively. Using Eq. (1), the interdiffusion coefficient D can be determined from the fitted interface roughness parameter by the MRI model. Using the roughness parameter values presented in Figs. 1(a-d), the calculated values of the interdiffusion coefficient according to Eq. (1) at the two locations, namely at the first and the last Cu/Ni interface as indicated in Fig.1(a), are presented in Fig. 2 by open and solid circles, respectively. The corresponding values of the interdiffusion parameters, the pre-exponential factor D_o and the activation energy E_a determined from the Arrhenius plot are indicated in Fig. 2 and are compared with the literature values listed in Table1.

The interdiffusion in the Cu/Ni multilayer at the first Cu/Ni interface was determined as 6.2×10^{-13}exp(-101.4 *kJ/RT*) m^2/s, and at the last Cu/Ni interface as 6.3×10^{-14}exp(-79.0 *kJ/RT*) m^2/s. The dash-dotted line shown in Fig. 2 was also obtained from measured AES depth profiling data for Cu/Ni multilayer (m) films deposited on glass substrates [19]. The dashed line shown in Fig. 2 was obtained for a Cu(~100 nm)/Ni(~500 nm) bilayer (b) films sample [20].

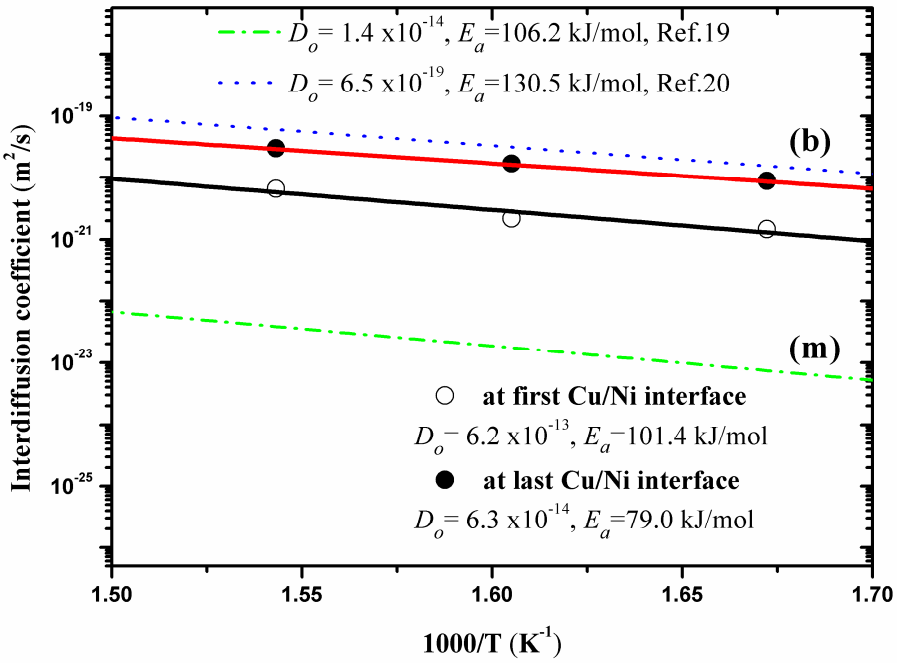


Figure 2. Arrhenius plots for the interdiffusion coefficients calculated from the fitted roughness parameters (see Fig. 1) by applying Eq. (1). The open and solid circles represent the values of the interdiffusion coefficient at the locations of the first and last Cu/Ni interface, respectively. The corresponding interdiffusion parameters (pre-exponential factor D_o and activation energy E_a) are indicated in bilayer (b) and multilayer (m) thin films.

Table 1. Diffusion parameters obtained in the present work in comparison with the results of other workers.

Source	D_o (m^2/s)	E_a (kJ/mol)
Present work (first Cu/Ni interface)	**6.2×10^{-13}**	**101.4**
Present work (last Cu/Ni interface)	**6.3×10^{-14}**	**79.0**
Ref.18	2×10^{-5}	228.7
Ref.19	1.4×10^{-14}	106.2
Ref.20	6.48×10^{-9}	130.5
Ref.21	6×10^{-12}	94.1
Ref.22	6.93×10^{-7}	90.4

The very low activation energy E_a values of 101.4 kJ/mol (79.0 kJ/mol) obtained in the present work in comparison with an E_a value of 228.7 kJ/mol extracted from a single crystalline copper [18], suggests that grain boundary and high defect density play a dominant role in the diffusion process of the

investigated multilayer thin films. The obtained diffusion parameters, measured at the low temperature range of 325-375 ºC, agree reasonably well with the ones listed in Table 1 for the polycrystalline bi-/multilayer films samples [19-22], which support the grain boundary diffusion mechanism at low temperatures for polycrystalline Cu/Ni .

A close examination of Fig. 2 reveals that the interdiffusion in a Cu/Ni multilayer was faster across the last Cu/Ni interface (near the substrate) than across the first Cu/Ni interface (near the surface) with a difference in activation energy value of 22.4 kJ/mol. In order to quantitatively evaluate the depth-dependent interdiffusion coefficient, Fig.3 shows the interdiffusion coefficient as a function of the sputtered depth for the same annealing time 30 min at the three different investigated annealing temperatures, on the basis of Eq. (1) with a linear dependence on depth adopted for σ_T (see above). The open circle, square, triangle and solid circles represent the values of the interdiffusion coefficient at the locations of the first, second, third and last Cu/Ni interfaces, respectively. It is found that the interdiffusion coefficient increase with the sputtered depth at the three different annealed temperature. This result is related to the reduction of the activation energy with sputtered depth for interdiffusion at the first, second, third and last Cu/Ni interfaces as 101.4 kJ/mol, 87.7 kJ/mol, 81.9 kJ/mol and 79.0 kJ/mol, respectively. The use of ion sputtering may involve changes in both the creation of vacancies, and energy of removed vacancies in the grain boundary and/or near grain boundary. During the sputtering process the larger concentration of vacancies at the deeper sputter depth might be created. This reduction of the activation energy with sputtered depth was also observed in Cu/Ag, Au/Ag, Pd/Au and Pd/Cu multilayer structure with AES measured by Bukaluk [15, 16]. In addition, enhanced grain boundary diffusion could be involved in the multilayer closer to the substrate; this however needs to be investigated further.

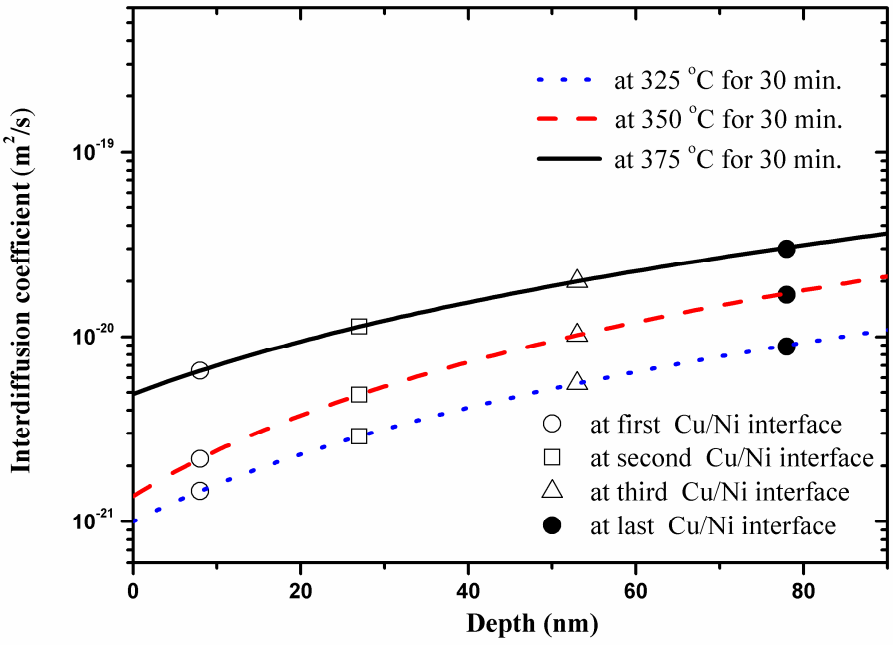


Figure 3. The interdiffusion coefficient as a function of the sputter depth for the annealed sample for 30 min at 325 ºC, 350 ºC and 375 ºC, respectively. The open circles, squares, triangles and solid circles represent the values of the interdiffusion coefficient at the locations of the first, second, third and last Cu/Ni interfaces, respectively.

4. Conclusions

The MRI model was applied for extracting the interdiffusion coefficients in Cu/Ni multilayers from measured AES depth profiling data. The interdiffusion coefficient of Cu/Ni multilayer was determined by $6.2\times10^{-13}\exp(-101.4\ kJ/RT)$ m^2/s at the first Cu/Ni interfaces and by $6.3\times10^{-14}\exp(-79.0\ kJ/RT)$ m^2/s m^2/s at the last Cu/Ni interfaces.

Acknowledgements

This work is based on the research supported by the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation of South Africa. We would like to thank the National Research Foundation and the cluster of the University of the Free State for financial assistance.

References

[1] Jeon IJ, Hong JH and Lee YP 1994 *J. Appl. Phys.* **75** 7825
[2] Pretorius R, Marais TK and Theron CC 1993 *Mater. Sci. Eng. Rep.* **10** 1
[3] Bukaluk A 1983 *Surf. Interface Anal.* **5** 20
[4] Hall PM and Morabito JM 1976 *Surf. Sci.* **54** 79
[5] Highmore RJ, Evetts JE, Greer AL and Somekh RE 1987 *Appl. Phys. Lett.* **50** 566
[6] Pamler W 1987 *Appl. Phys.* A **42** 219
[7] Hofmann S 1994 *Surf. Interface Anal.* **21** 673
[8] Kesler V and Hofmann S 2002 *J. Surf. Anal.* **9** 428
[9] Wang JY, Zalar A, Zhao YH and Mittemeijer EJ 2003 *Thin Solid Films* **433** 92
[10] Wang JY, Zalar A and Mittemeijer EJ 2004 *Appl. Surf. Sci.* **222** 171
[11] Hedberg CL (by edited), *Handbook of Auger Electron spectroscopy, Third Edition,* Physical Electronics Industries, Eden Prairie, MN, USA,1976
[12] Ziegler JF, TRIM Code, IBM Corporation, Yorktown Heights, NY, (http:// www.srim.org)
[13] NIST Electron EAL, 2001 *NIST Database* 82
[14] Yan XL, Liu Y, Swart HC, Wang JY and Terblans JJ, Proceedings of SAIP2015. This volume.
[15] Bukaluk A 2001 *Vacuum* **63** 119
[16] Bukaluk A 2000 *Surf. Interface Anal.* **30** 597
[17] Wang JY, Hofmann S, Zalar A and Mittemeijer EJ 2003 *Thin Solid Films* **444** 120
[18] Van Dijk T and Mittemeijer EJ 1977 *Thin Solid Films* **41** 173
[19] Venos R, Pamler W and Hoffmann H, 1988 *Thin Solid Films* **162** 155
[20] Joubert HD, Terblans JJ and Swart HC 2010 *Surf. Interface Anal.* **42** 1281
[21] Abdul-Lettif AM 2007 *Phys.B* **388** 107
[22] Divinski S, Ribbe J, Schmitz G and Herzig C 2007 *Acta Mater.* **55** 3337

Division B – Nuclear, Particle and Radiation Physics

High-momentum particle production at hadron colliders

D M Adamiak[1,2],W A Horowitz[1]

[1]Department of Physics, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa

E-mail: [2] daniel.m.adamiak@gmail.com

Abstract. We compute the distributions of charged particles at large transverse momenta in pp, $p\bar{p}$, and pA collisions at Fermilab. Our calculations are performed using leading order perturbative quantum chromodynamics (pQCD), with both the usual parton distribution functions (PDFs) and nuclear PDFs, which encapsulate the modifications of the usual PDFs by the presence of multiple nucleons in a nucleus. We find that our results consistently describe the data at Fermilab, across multiple orders of magnitude in centre of mass energy $\sqrt{s}$, and over many orders of magnitude in transverse momentum. We then examine the transverse momentum dependence of the partonic contributions to these cross sections, which provides the critical input spectra for theoretical predictions for the suppression of charged particle spectra in heavy ion collisions.

1. Introduction

One of the current bleeding edges of physics concerns the question what arises from the emergent, many-body physics of quantum chromodynamics (QCD) at high-energies[1], such as that of the quark-gluon plasma. Analytically, one of the tools we use to make predictions for these phenomena is perturbative QCD (pQCD)[2]. Experimentally, one of the only available measurable quantities is the cross-section[3], in this a case a measurement of the spectrum of hadrons produced from an experiment. At colliders, such as those located at Fermi Lab[4] and CERN[5], we are able to to measure the charged hadron production of high energy collisions between protons, anti-protons and heavy nuclei. It is our area of interest to investigate how the partons, the constituent particles of the protons, behave in these collisions.

In high energy collisions it is impossible to directly observe the processes that the constituent partons undergo during the collision itself [6]. We are, however, able to control the input into the collision (particles used in the collision and their energies) as well as measure the output (hadron production and energy spectrum). Using pQCD we make predictions about how the partons behave during these collisions, with this behaviour producing its own prediction spectrum that we are able to compare with the data.

We use the result derived by Eskola and Honkanen [7] using Leading Order pQCD to calculate the differential cross-section of charged hadron production in pp collisions. The calculation is performed numerically, with the results being compared to Eskola and Honkanen' computation. Since they have already demonstrated that the theory agrees strongly with data, we compare only with their calculation to confirm whether we have performed our computation correctly. Agreement with their calculation within machine precision implies agreement with data.

After finding strong agreement with the experimental data, we break down the cross-section into the contributions due to each parton and their dependence on transverse momentum. The long-term goal is to use these results as an input for theoretical predictions for the suppression of charged particle spectra in heavy ion collisions [8].

Both calculations can also be done for the pA collision case, a collision between a proton and either a lead or a gold nucleus. This calculation for pp collisions is modified by introducing a nuclear PDF (nPDF) into our calculations.

2. Leading Order pQCD Calculations

We calculate the inclusive cross-section for production of a parton of a flavour f and a rapidity y_f in Mathematica. The expression for this cross-section, derived using pQCD, is given by [7]

$$\begin{aligned}\frac{d\sigma^{AB\to h+X}}{dq_T^2 dy_f} &= K(\sqrt{s})J(m_T,y)\int\frac{dz}{z^2}\int dy_2 \sum_{<ij><kl>}\frac{1}{1+\delta_{kl}}\frac{1}{1+\delta_{ij}}\times \\ &\times\left\{x_1 f_{i/A}(x_1,Q^2)x_2 f_{j/B}(x_2,Q^2)\left[\frac{d\hat{\sigma}}{d\hat{t}}^{ij\to kl}(\hat{t},\hat{u})\underset{k\to h}{D}(z,\mu_F^2)+\frac{\hat{\sigma}}{d\hat{t}}^{ij\to kl}(\hat{u},\hat{t})\underset{l\to h}{D}(z,\mu_F^2)\right]\right. \\ &+ \left. x_1 f_{j/A}(x_1,Q^2)x_2 f_{i/B}(x_2,Q^2)\left[\frac{d\hat{\sigma}}{d\hat{t}}^{ij\to kl}(\hat{u},\hat{t})\underset{k\to h}{D}(z,\mu_F^2)+\frac{\hat{\sigma}}{d\hat{t}}^{ij\to kl}(\hat{t},\hat{u})\underset{l\to h}{D}(z,\mu_F^2)\right]\right\}\end{aligned} \tag{1}$$

where y_2 and y_f are the rapidities of the outgoing partons. $x_1 = \frac{p_T}{\sqrt{s}}(e^{y_f}+e^{y_2})$, $x_2 = \frac{p_T}{\sqrt{s}}(e^{-y_f}+e^{-y_2})$. z is a parameter equal to the ratio of the energy of the quark, f, to hadron h. $\sqrt{s}$ is the root center of mass energy, the energy in the rest frame of the interaction. Present are the differential cross-sections of parton interactions, $\frac{d\hat{\sigma}}{d\hat{t}}$ and the Mandelstam variables, $\hat{t},\hat{u}$. Also present are the PDFs, f and fragmentation functions, D. J is a function defined by

$$J(m_T,y) = (1-\frac{m^2}{m_T^2\cosh^2 y})^{-1/2} \tag{2}$$

where m is the mass of the outgoing hadron and m_T is the transverse mass of the outgoing hadron. The sum over i,j,k,l is a sum over every possible interaction, where i,j,k,l stand for labels of the the different partons(quarks, anti-quarks and gluons). The constant, K, is the ratio of the experimental cross-section to the leading order cross-section. It was found by Eskola and Honkanen in their paper [7] that K was a function only of $\sqrt{s}$ and encapsulates the next-to-leading order contributions.

What is being described by this integral, roughly, is that the probability of h being formed at momentum q_T, is equal to the sum over every possible interaction that could form h with momentum q_T.

2.1. Parton Distribution Functions

When a proton is moving with some momentum, each of it's constituent partons carries a fraction of that momentum. These constituent partons include the uud quarks, the gluons binding them and all the virtual quarks inside the proton as well. This way it is possible for, say, the strange quarks to carry some momentum fraction of the proton as well. The ratio of momentum each constituent parton carries to the total momentum is known as the Parton Distribution Function (PDF)[9]

These are not analytic functions, but rather a polynomial interpolation of recorded PDF values. Originally written in FORTRAN, the code was adapted to Mathematica by us, available

on request. The PDF is a function f that takes two input parameters, fractional momentum x and energy scale Q. While, in general, the PDF does not have to be of a proton, that is what we are considering here. In our calculation, the parton distribution function of [10] is used.

For pA collisions, we modify the PDF by a nuclear parton distribution function (nPDF). It is another polynomial interpolation originally written by in FORTRAN [11], adapted to Mathematica by us. The way the nPDFs were constructed is that, given an input energy, transverse momentum and nucleus size, they produce a modification factor which we multiply into the PDF of the parton that we choose to come from the heavier nucleus. This will give us the differential cross-section of a pA collision.

2.2. Fragmentation Functions

Fragmentation functions, D, describe transition of a parton, f, with some fractional momentum, z, at some energy scale, μ_F, into a hadron, h.[12]. Again, we have no analytic function at hand. We use the fragmentation functions found in [13]. We use the same functions adapted for Mathematica

3. Modification to extract partonic contributions

We now modify our calculation so that we can calculate the individual contribution to the total cross-section from each parton. The modification is quite simple. Examining equation(1), we eliminate the fragmentation functions, since these are used to tell us about the formation of hadrons. This leaves us with a calculation about the production of partons.

The second modification is that, rather than summing over all the partons, we record the contribution due to each parton separately. Our calculation then looks like

$$\begin{aligned}\frac{d\sigma^{AB\to k+X}}{dq_T^2 dy_f} = \int dy_2 \sum_{<ij><kl>} \frac{1}{1+\delta_{kl}} \frac{1}{1+\delta_{ij}} \times \\ \times x_1 f_{i/A}(x_1, Q^2) x_2 f_{j/B}(x_2, Q^2) \left[\frac{d\hat{\sigma}^{ij\to kl}}{d\hat{t}}(\hat{t}, \hat{u}) \right]\end{aligned} \tag{3}$$

where, once again, i, j, k, l represent each of the participating partons where we now save the result for the k outgoing partons individually.

4. Results

In figure 1, we compare our calculated differential cross-section with the calculation from figure 5 of [7], which was, in turn, used to make a prediction for [14]

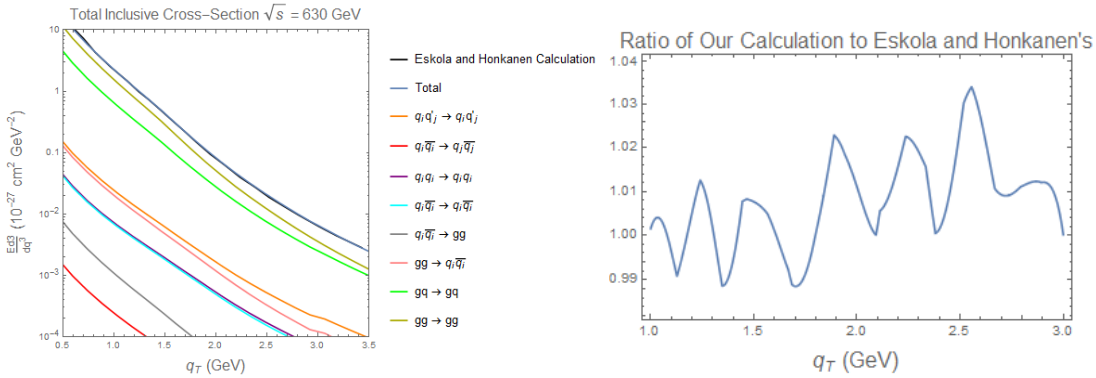


Figure 1: (Left) Calculated differential cross-section of $p\bar{p}$ production at $\sqrt{s} = 630$ GeV(blue), compared with the calculation from Eskola and Honkanen (black), as well as the contributions to the total cross-section from each of the possible partonic interactions. (Right) The ratio of our calculation to the calculation from Eskola and Honkanen.

The first plot of figure 1 is a plot of our calculated differential cross-section of charged hadron production of $p\bar{p}$ collisions at $\sqrt{s} = 630$ GeV as a function of q_T compared with the calculation performed in [7]. Also included are the contributions from the various processes that make up the total cross-section when summed over. The plot on the left of figure 1 is a plot of the ratio of our calculated cross-section to the calculation in 1. Agreement within four percent implies that we have implemented our calculation correctly. The four percent discrepancy is attributed to machine precision error and is small compared to other sources of uncertainty, for example in figure 5 of [7], their K value has an uncertainty of 18%

Next we give a result of our parton contribution calculation. We perform the calculation at many energies to match the energies used in experiment, but, in figure 2, we just show the result for $\sqrt{s} = 200$ GeV.

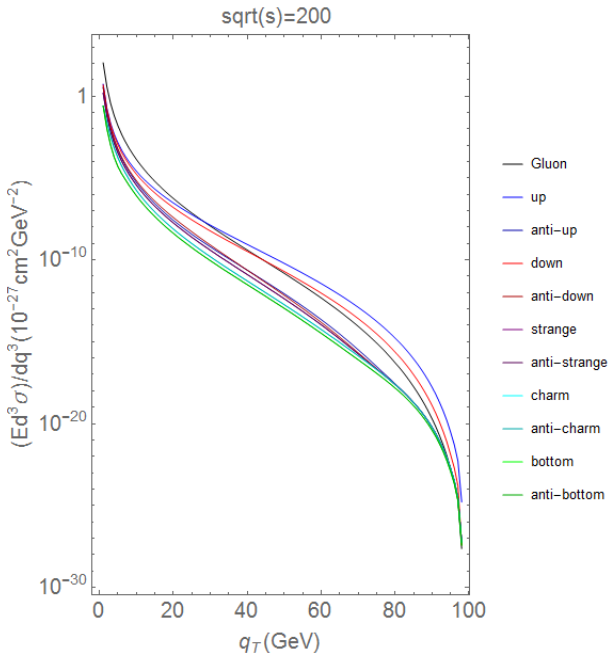


Figure 2: Parton conrtibutions to the total differential cross-section

We find that at low energies, the gluons are the leading source to the cross-section. At higher energies we find that the valence quarks dominate the spectrum. This is inferred from the spectrum containing a leading contribution from the up and down quarks in a 2-1 ratio, the make-up of a proton.

5. Conclusion

Using the methods of pQCD, we have made a prediction for the behaviour of partons within a $p\bar{p}$ collision. These partonic contributions were used as an input to compute the total differential cross-section. When compared with the calculation in [7], strong agreement was found. This implies that our calculation agrees with the data strongly. From this we conclude that either our partonic contributions were calculated correctly, or that any unaccounted for effects perfectly cancelled each other out when we computed the cross section. Assuming the former, we can use this partonic contributions as a critical input for energy loss calculations.

The next step would be to modify the calculations for pA collisions to see if they match the measured data and predicted results. The partonic contributions from a pA, when used as an input for energy loss calculations, can be used to explore the effect of the medium, if any, present in pA, but not in $p\bar{p}$ collisions.

References

[1] The Frontiers of Nuclear Science, A Long Range Plan. 2008.

[2] K. Kajantie and Hannu Kurki-Suonio. Bubble growth and droplet decay in the quark-hadron phase transition in the early universe. *Phys. Rev. D*, 34:1719–1738, Sep 1986.

[3] L.P. Remsberg T. Abbott, L. Kowalski. Rapidity and invariant cross sections.

[4] James Green. A Measurement of High Transverse Momentum Charged Hadron Production Using a Pion Beam and Hydrogen Target. 1981.

[5] The ATLAS collaboration. Charged hadron production in p+Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV measured at high transverse momentum by the ATLAS experiment. 2014.

[6] B. Povh, M. Lavelle, and M. Rosina. *Scattering and Structures: Essentials and Analogies in Quantum Physics*. Number p. 66713 in Scattering and Structures: Essentials and Analogies in Quantum Physics. Springer Berlin Heidelberg, 2005.

[7] K. J. Eskola and H. Honkanen. A Perturbative QCD analysis of charged particle distributions in hadronic and nuclear collisions. *Nucl. Phys.*, A713:167–187, 2003.

[8] K. Aamodt et al. Suppression of Charged Particle Production at Large Transverse Momentum in Central Pb-Pb Collisions at $\sqrt{s_{NN}} = 2.76$ TeV. *Phys. Lett.*, B696:30–39, 2011.

[9] M Dittmar, S Forte, A Glazov, S Moch, S Alekhin, Guido Altarelli, J Andersen, R D Ball, J Blmlein, Helmut B Bttcher, T Carli, Marcello Ciafaloni, D Colferai, A Cooper-Sarkar, Gennaro Corcella, L Del Debbio, G Dissertori, J Feltesse, A Guffanti, C Gwenlan, J Huston, G Ingelman, M Klein, J I Latorre, T Lastoviicka, G Lastoviicka-Medin, L Magnea, A Piccione, J Pumplin, V Ravindran, B Reisert, J Rojo, Agustin Sabio Vera, Gavin P Salam, F Siegert, A M Stasto, H Stenzel, C Targett-Adams, R S Thorne, A Tricoli, J A M Vermaseren, and A Vogt. Introduction to parton distribution functions. 2005.

[10] H. L. Lai, J. Huston, S. Kuhlmann, J. Morfin, Fredrick I. Olness, J. F. Owens, J. Pumplin, and W. K. Tung. Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions. *Eur. Phys. J.*, C12:375–392, 2000.

[11] K. J. Eskola, H. Paukkunen, and C. A. Salgado. EPS09: A New Generation of NLO and LO Nuclear Parton Distribution Functions. *JHEP*, 04:065, 2009.

[12] D de Florian and D Milstead. 17. fragmentation functions in e e-, ep and pp collisions.

[13] Bernd A. Kniehl, G. Kramer, and B. Potter. Fragmentation functions for pions, kaons, and protons at next-to-leading order. *Nucl. Phys.*, B582:514–536, 2000.

[14] F. Abe et al. [CDF Collaboration]. *Phys. Rev. Lett. 61 1819*, 1988.

Online energy reconstruction on ARM for the ATLAS TileCal sROD co-processing unit

Mitchell A. Cox and Bruce Mellado

School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africa

E-mail: mitchell.cox@students.wits.ac.za

Abstract. Modern Big Science 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 Hadronic Tile Calorimeter (TileCal) will increase by 200 times to over 40 Tb/s. This increase requires more advanced processing on the raw data in order to harness a larger quantity of good quality physics data. An algorithm called Optimal Filtering (OF) is currently used in the TileCal front-end for online energy reconstruction of the digitised photo-multiplier tube signals and is currently implemented on Digital Signal Processors (DSPs) and Field Programmable Gate Arrays (FPGAs) which are difficult to program and are expensive. It is proposed that a cost-effective, high data throughput and general purpose Processing Unit (PU) can be developed by using several commodity ARM processors while maintaining minimal software design difficulty for the end-user. This PU could be used for a variety of high-level algorithms other than OF on the high data throughput raw data to combat the issue of out of time pile-up and for online data quality testing. OF and histogram algorithms have been implemented in C++ and several ARM platforms have been tested and shown to have good CPU to external I/O balance.

1. Introduction

Projects such as the Large Hadron Collider (LHC) generate enormous amounts of raw data which presents a serious computing challenge. After planned Phase-II upgrades in 2023, the raw data output from the ATLAS Hadronic Tile Calorimeter (TileCal) will increase by 200 times to over 40 Tb/s (Terabits/s) [1, 2]. It is infeasible to store this data for offline computation and so an online triggering system is used to reduce the quantity of data before storage.

The LHC Run 0, as of 2012 had a peak luminosity of 7.7×10^{33} cm^{-2}s^{-1} with a center of mass energy of $\sqrt{s} = 8$ TeV. The Phase-II upgrades in 2023 assume a maximum instantaneous luminosity of 7×10^{34} cm^{-2}s^{-1} [3].

The average number of interactions per bunch crossing also known as "in time pileup", $\langle\mu\rangle$, can be calculated. As of 2012, $\langle\mu\rangle = 20.7$. This is assumed to increase to $\langle\mu\rangle = 200$ for Phase-II [3]. By increasing the amount and also the precision of the data read out from TileCal, more sophisticated algorithms may be used for energy reconstruction in a high pile-up environment.

The sROD (Super Read out Driver) is a FPGA-based device which is located in the back-end. It is the interface to the upgraded TileCal front-end read-out electronics and is responsible for its management and control. The sROD will become the new Level-0 trigger, performing energy reconstruction and then reducing the data rate from the 40 MHz bunch crossing rate to 500 kHz

[2, 3]. The upgraded Level-1 trigger will then reduce the data to 200 kHz as opposed to the current 100 kHz. A diagram of the upgraded TileCal read out architecture is shown in Figure 1.

Photomultiplier tubes (PMTs) are linked to layers of scintillator in TileCal by wavelength shifting optical fibres. The electrical signal that is produced by a PMT when a particle interacts with the scintillator is stretched, amplified and sampled with an analog to digital converter operating at 40 MSa/s (Million Samples per Second). Seven samples are taken of each pulse. Figure 2 shows an ideal pulse shape with arbitrary amplitude and phase shift with seven samples. From the figure it is clear that out of time pileup can dramatically influence the shape of the pulse. It is the job of the filtering algorithm to negate this. Optimal Filtering (OF) is currently used to perform energy reconstruction on the raw data [4].

The OF algorithm simply finds two dot products of these seven samples and two pre-calculated "weights" vectors. Each vector is highly specific and the result is the peak amplitude of the pulse as well as the phase shift of the pulse from the expected time.

Under higher pile up conditions OF loses accuracy and so a third calculation is done to estimate the quality factor of the results. If this quality factor is low then the raw data is stored for later offline processing with a more accurate algorithm.

Several new energy reconstruction algorithms are in development and still more are expected in years to come, each with advantages and disadvantages. A Matched Filter based technique shows similar performance to OF, and there is also a signal response deconvolution based technique shows promise [5, 6]. The deconvolution based technique cannot be implemented on the existing read out electronics due to lack of processing power.

An ARM System on Chip (SoC) based co-processing unit (PU) for the TileCal sROD is being developed [7]. ARM is an alternative CPU architecture to Intel x86 and is commonly used in mobile devices due to its energy efficiency and low cost. This PU can be used to monitor the raw data that passes through the sROD and to verify the sROD energy reconstruction data quality. Since the PU is based on general purpose processors, it is easy to implement new algorithms and functionality using a programming language such as C or even Java. A more ambitious use for the PU could be to update the energy reconstruction parameters used by the sROD in real-time or to take over energy reconstruction in its entirety.

The results of a C++ OF algorithm on ARM are presented in Section 2 and the implementation of a simple PCI-Express CPU interconnect and preliminary performance is given in Section 3. Section 4 concludes with a brief discussion of future work.

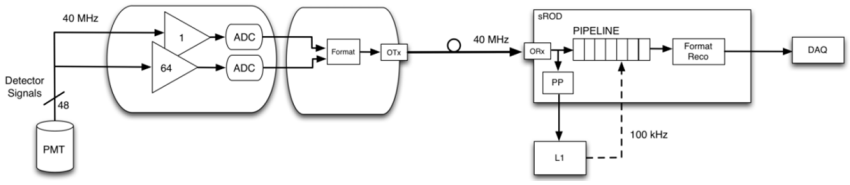


Figure 1: ATLAS TileCal upgraded read out architecture showing the sROD [2].

2. Optimal Filtering on ARM

Optimal Filtering (OF) is a simple algorithm that solves Equation 1 and 2 to find the peak amplitude, A, and the time shift, τ, given two weights vectors, a_i and b_i and the samples vector S_i [8]. n is equal to 7 in the existing TileCal implementation. The pedestal or noise floor, p, can be estimated as a constant or calculated with its own set of weights.

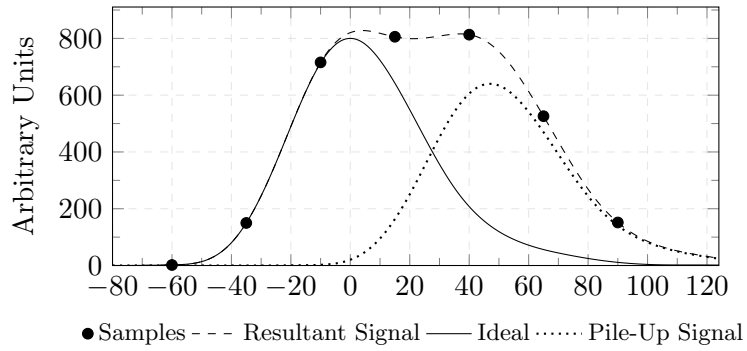


Figure 2: Ideal filtered PMT pulse showing seven samples.

$$A = \sum_{i=1}^{n} a_i(S_i - p) \qquad (1)$$

$$\tau = \frac{1}{A}\sum_{i=1}^{n} b_i(S_i - p) \qquad (2)$$

This algorithm was implemented using the EIGEN C++ library which is one of the highest performing linear algebra libraries that also supports ARM [9]. Vector sizes of eight were used to correspond to the current TileCal read-out padded with an additional zero, and two dot products were performed with random weights and data to represent the amplitude and phase filters. A CPU is able to perform the dot product more efficiently with a vector size of eight because it fits into its Single Instruction Multiple Data (SIMD) registers better.

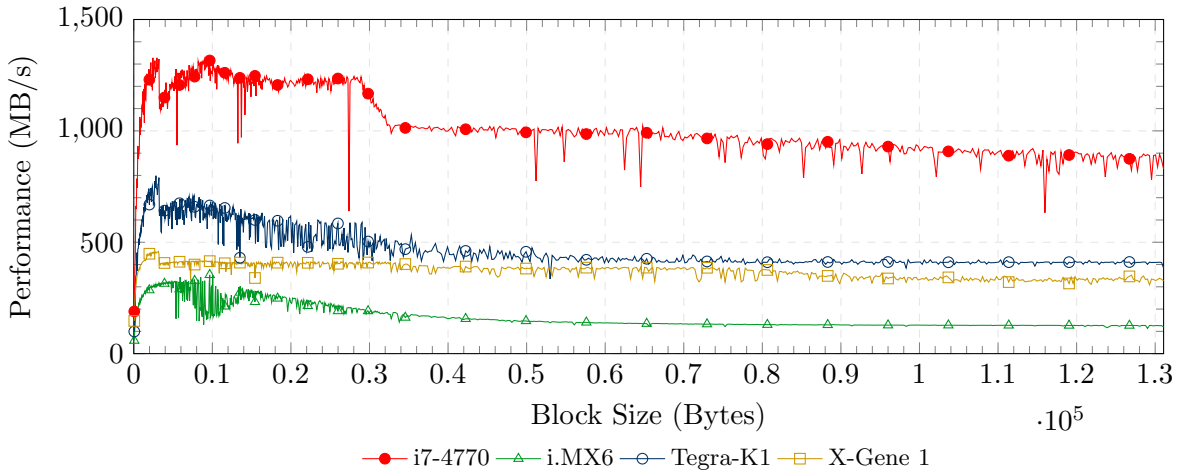


Figure 3: Single thread performance of an OF implementation with seven samples and two outputs with varying block size. An Intel i7-4770 CPU and three ARM platforms are shown. The results are normalised to a 1 GHz CPU clock.

Figure 3 shows the performance which has been converted to MegaBytes per second (MB/s) of filtered data where one megabyte is equal to 1024×1024 bytes. A high-end Intel i7-4770 CPU has been compared to an ARM Cortex-A9 (Freescale i.MX6), an ARM Cortex-A15 (NVIDIA Tegra-K1) and an ARMv8 (APM X-Gene 1) platform. All results have been normalised to 1 GHz. The i7 normally runs at 3.5 GHz "turbo boost", the i.MX6 at 1 GHz, the Tegra-K1 at 2.3 GHz and the X-Gene at 2.4 GHz.

Batches of sample and weight vectors were combined into blocks for more efficient processing. Block sizes from 32 B (one set of samples) to 128 kB were tested to find the maximum performance of the OF algorithm on the different platforms.

It is clear that the Intel performance is better than the ARM platforms. It is interesting to note that the performance of the Tegra-K1 is better than the X-Gene. This should not be the case but it was found that the compiler is not yet mature for the new ARMv8 architecture. Using GCC (GNU C Compiler) 4.9 instead of GCC 4.8 makes approximately 10% performance difference but for uniformity GCC 4.8.2 was used on all platforms.

The Wandboard platform (i.MX6 SoC - similar to the SoC found in a Samsung Galaxy SIII smart phone, for example) is capable of a peak of 350 MB/s and the other platforms all achieve over 1 GB/s of OF throughput. This is typically more than the external I/O available on the SoCs. It is important to note that the results are for a single thread and so if multiple cores are used then the results will scale accordingly because the algorithm can be run in parallel on separate data.

3. PCI-Express External I/O

The external I/O interface of a computing system should be well balanced with the available processing power. Because this is highly dependent on the algorithms used, benchmarking is important. Based on the results in Section 2, even the lowest performing system, the Wandboard, requires up to 350 MB/s external I/O for good system balance. Gigabit Ethernet - which is almost always available on a SoC - only allows up to 125 MB/s. This is clearly insufficient for optimal system balance in this application.

PCI-Express is a high-bandwidth external I/O interface that is energy efficient and simple to implement for the system designer as it only requires several PCB traces and very little special and potentially expensive hardware. For the connection to the sROD, PCIe can be used directly between the sROD and ARM SoCs.

PCI-Express throughput tests for the lower end i.MX6 SoC have been performed and are described in detail in another paper [7]. A summary of the results is shown in Table 1. The theoretical maximum throughput available on the SoC is 500 MB/s but based on the test results only slightly more than half of that is attainable on the i.MX6 because the Direct Memory Access (DMA) is not usable for generic PCIe data transfers.

Table 1: PCI-Express throughput results of an i.MX6 pair.

	Memory Copy	**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 X-Gene and Tegra-K1 system on chips have more advanced PCI-Express controllers which do support DMA. A test system for these platforms has not been built but in theory there should be no significant issues. The theoretical data throughputs for these platforms is approximately 15 GB/s and 2 GB/s respectively.

4. Discussion, Conclusions and Future Work

High data throughput 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 operation of the LHC.

An OF algorithm was implemented and tested on ARM Cortex-A9, A15 and X-Gene (similar to Cortex-A57) SoCs. The slowest SoC, the Cortex-A9, achieved 350 MiB/s throughput on the algorithm. The Tegra-K1 peaked at 1.8 GB/s at the maximum clock speed with the X-Gene being slower because of the lack of compiler optimisation.

A PCI-Express interface will be used for the raw data transfer between the sROD and the PU. 2.8 GB/s of data throughput is required to sustain the raw data from the sROD prototype [2]. Initial throughput measurements presented for a pair of Freescale i.MX6 quad-core Cortex-A9 SoCs is 283 MB/s for the available x1 link. The theoretical data throughput of the Tegra-K1 and X-Gene SoCs are 2 GB/s and 15 GB/s respectively.

With multi threaded operation, the Tegra-K1 could sustain 1.8×4 GB/s but is limited by its 2 GB/s maximum PCIe throughput. A more sophisticated filtering algorithm may be effective given the excess of CPU processing power for the OF algorithm. The X-Gene could sustain eight times the measured throughput and so there would theoretically be no I/O bottleneck.

Testing still needs to be done on more complex filtering algorithms as well as algorithms for other sROD co-processor uses such as histograms.

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] The ATLAS Collaboration, *Readiness of the ATLAS Tile Calorimeter for LHC collisions*, *The European Physical Journal C* **70** (Dec., 2010) 1193–1236.

[2] F. Carrió et al., *The sROD module for the ATLAS Tile Calorimeter Phase-II Upgrade Demonstrator*, *Journal of Instrumentation* **9** (Feb., 2014) C02019–C02019.

[3] The ATLAS Collaboration, *Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment*, Tech. Rep. CERN-LHCC-2012-022. LHCC-I-023, CERN, Geneva, Dec., 2012.

[4] E. Fullana et al., *Optimal Filtering in the ATLAS Hadronic Tile Calorimeter*, 2005.

[5] B. Peralva, *The TileCal energy reconstruction for collision data using the matched filter*, in *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC)*, pp. 1–6, IEEE, oct, 2013.

[6] L. M. A. Filho, D. O. Damazio, and A. S. Cerqueira, *Calorimeter Signal Response Deconvolution for Online Energy Estimation in Presence of Pile-up*, *ATLAS Internal Note* (Sept., 2012).

[7] M. A. Cox, R. Reed, and B. Mellado, *The development of a general purpose ARM-based processing unit for the ATLAS TileCal sROD*, *Journal of Instrumentation* **10** (Jan., 2015) C01007–C01007.

[8] W. Cleland and E. Stern, *Signal processing considerations for liquid ionization calorimeters in a high rate environment*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **338** (Jan., 1994) 467–497.

[9] B. Jacob and G. Guennebaud, *Eigen C++ Library*, 2014. http://eigen.tuxfamily.org.

Measuring the orientation of the Ge Crystals of the iThemba LABS Segmented Clover Detector

T. S. Dinoko[1], T. D. Bucher[1], E. A. Lawrie[1], J. Easton[1], N. Erasmus[1], S. P. Noncolela[1], O. Shirinda[1]

[1]iThemba LABS, National Research Foundation, P. O. Box 722, Somerset West 7129, South Africa

tdinoko@tlabs.ac.za

Abstract. The iThemba LABS segmented detector includes four Ge crystals in a clover configuration inside a single cryostat [1]. The orientation of the Ge crystals inside the cryostat needs to be measured. It plays a crucial role in the simulation of the charge collection in the detector and generating realistic pulse shapes. The rise times of the core signals corresponding to two arbitrary interaction points were measured. They show a distinct difference, indicating that the rise times of the segmented clover detector are sensitive to interaction position and can be used to determine the crystal orientation of the Ge crystals.

1. Introduction

One of the primary parameters to describe a Ge detector for pulse shape analysis is the orientation of the crystallographic axes of the cubic centered Ge crystal. The drift mobility for the electron-hole pairs in Ge depends on the orientation of the electrical field with respect to these axes and causes deviation in the collection times of up to 30% [2]. The intention of a team at iThemba LABS is to characterize the segmented iThemba LABS Ge detector, see Figure 1, and to develop a pulse shape analysis technique able to determine the position of each energy deposition caused by the interaction of a gamma-ray in the segmented Ge crystal [3]. This necessitates a measurement of the lattice orientation of the four crystals of the detector.

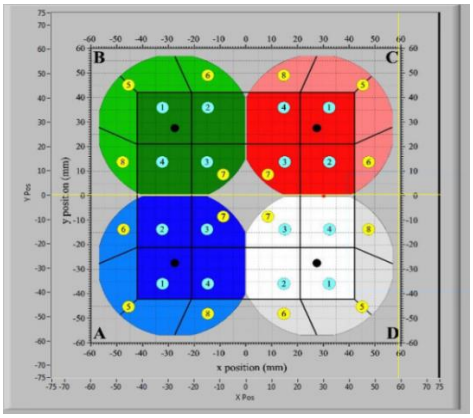


Figure 1. Schematic diagram of the HPGe crystals of the iThemba LABS segmented clover detector. The four crystals are labelled as A, B, C, D.

2. Method

To determine the orientation of a Ge crystal one has to scan the detector measuring its time response e.g. T10, T30, T60 and T90, which are the times for the pulse to rise to 10%, 30%, 60% and 90% of its maximum amplitude, respectively, on a grid with a small step. Examples for such measurements, performed at TRIUMF for a similar detector are shown [4]. Based on these measurements it was concluded that the lattice of crystal D of their detector is rotated by 5° clockwise in the (x,y) plane [4]. To prepare for such measurements new sorting code was developed at iThemba LABS. To test the detector and the sorting code, initial measurements were done for two arbitrary positions with a collimated source placed on the scanning table. A ^{137}Cs source was installed underneath a collimator with an inner hole of 3 mm in diameter and 100 mm in length. Position (0,0) corresponds to the centre of the detector, see Figure 1. Two arbitrary locations, (-5,27) and (-23,27) in crystal B were measured with at least 10^3 counts in the 662 keV gamma-ray peak of ^{137}Cs. The measurements typically took 30 minutes per location. Both measurements were in front of segment 2 of crystal B, see Figure 1.

3. Results and Discussion

The events with energy of E = 662 keV registered on the core were selected and the corresponding traces are shown in Figures 2 and 3 for positions 1 and 2 respectively.

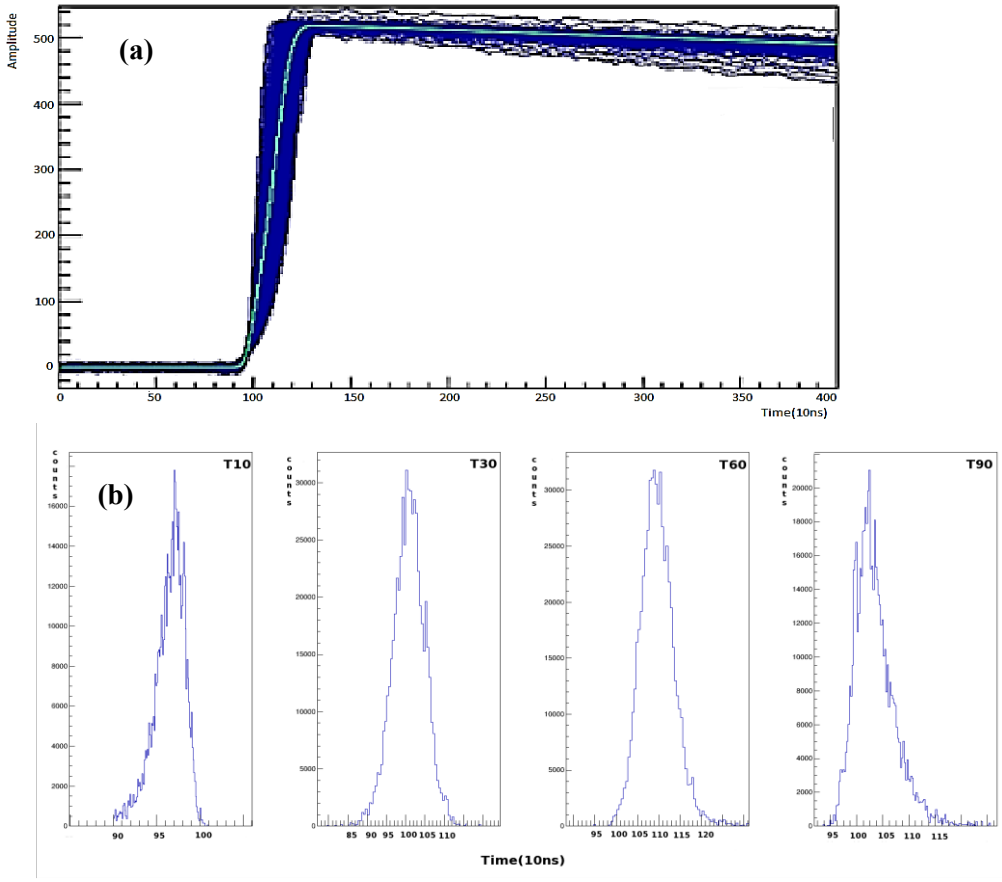


Figure 2. The average (green line) and the individual (blue lines) traces for position (-5, 27), which is near the surface of crystal B is shown in (a). The measured individual rise times are shown in (b).

The rise times T10, T30, T60 and T90 were determined on the average traces for these positions and are listed in table 1. In addition, the distributions of the time spectra T10, T30, T60 and T90 (10% to 90% of the amplitude) rise times were measured, by evaluating the rise times from each trace. They are shown in Figures 2 and 3. Note, that using a collimated source the x- and y-coordinates of the interaction sites are determined, while the depth (the z-coordinate) remains unrestricted.

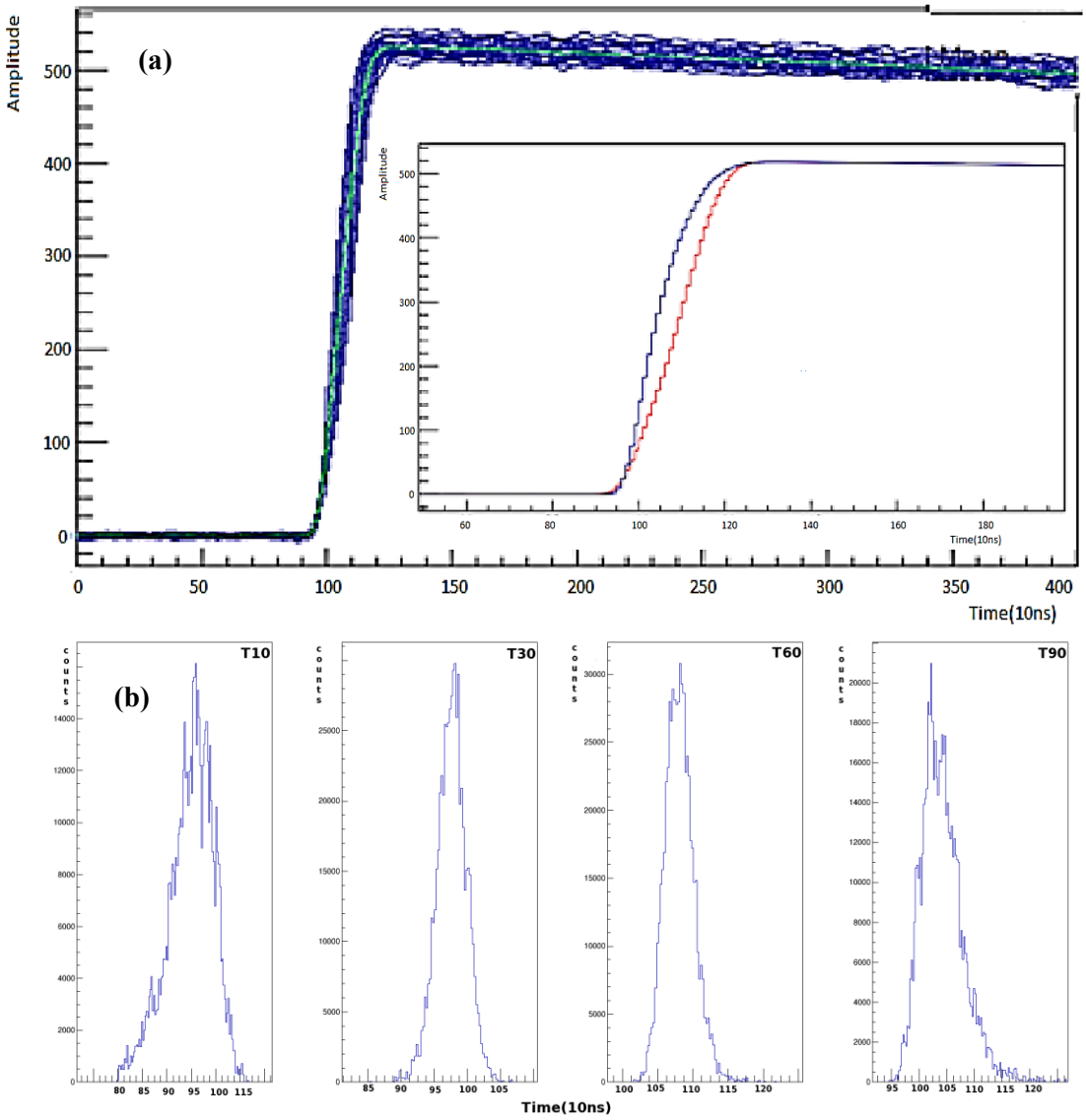


Figure 3. The average (green line) and the individual (blue lines) traces for position (-23, 27), which is near the core of crystal B are shown in (a). The insert shows the average traces for positions 1 (blue) and 2 (red) . The measured individual rise times for position 2 are shown in (b).

Table 1. Measured rise times in units of time clocks (10ns) for the average core signal registered for the two positions.

	T10	T30	T60	T90
Position 1	97	100	105	114
Position 2	97	103	110	117

The insert in Figure 3 shows the average traces for the two arbitrary positions 1 and 2. It is clear that the rise times are distinctly different. Thus the sorting software is completed. We were able to determine average traces and to extract the T30, T60 and T90 rise times. These rise times show a distinct difference for the two arbitrary interactions. As a next step a scan of the detector will be carried out in order to determine the rise times as a function of (x,y) coordinates of the interaction position. The results will be used to determine the lattice crystal orientation.

References

[1] B. Bruyneel, et al., Nucl. Instr. And Meth., A569, 774 (2006)
[2] B. Bruyneel, et al., Nucl. Instr. And Meth., A569, 764 (2006)
[3] T. D. Bucher, et al., Progress Report (2014)
[4] H. C. Scraggs, et al., Nucl. Instr. And Meth., A543, 431-440 (2005)

Analysis and quantification of modelling errors introduced in the deterministic calculational path applied to a 2-node problem

T.P. Gina[1], S.A. Groenewald[2] , W.R. Joubert[2] , S.H. Connell[1]

[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: thembelanipg@gmail.com

Abstract. The objective of this work is to analyse and quantify modelling errors introduced by simplifications made in the deterministic calculational path as applied to nuclear reactor simulation. These simplifications are introduced in order to make a model practically solvable with a diffusion code, and they are classified as follows: spatial homogenization, energy (spectral) condensation, diffusion approximation and environmental dependency. In this work, a two-node model consisting of a SAFARI-1 reactor fuel assembly next to a water node is modelled because it is a typically encountered configuration and fairly sensitive to spatial and spectral approximations. The analysis and quantification of modelling errors introduced in the calculational path were performed. Errors introduced by the four approximations in the calculational path are quantified by investigating the effective multiplication factor (k-eff) as well as calculational time as an integral measure of difference between two models. All calculations were performed with the neutron transport codes NEWT, the Monte Carlo Serpent and the diffusion code MGRAC. The results indicate that for the fuel-water model, environmental dependency and the diffusion approximation are the largest contributors to the total calculational error.

Keywords: condensation, homogenization, diffusion, environmental error

1. Introduction

For the development of computer codes that accurately predict the flux distribution in nuclear reactors, modelling error analysis is necessary. The majority of publications on the modelling error analysis [3][4][5] focused on power reactors. Not much work has been published on modelling error analysis for research reactor (such as material testing reactors (MTRs)) simulation. This work forms part of a bigger study focused on improving the errors made in modelling MTRs. As a first step in this larger study, modelling error analysis is done on a 2-node problem. In future work, the approach defined will be applied to a full-core MTR model. This paper is arranged as follows: Section 2 discusses the theoretical background and calculational path. Section 3 provides a layout of the research methodology and the codes used. Section 4 discusses the results and the final section discusses concluding remarks for this work.

2. Theoretical background and calculational path

In order to predict neutron flux distribution in the reactor core, neutronics modelling of the reactor is performed by solving the neutron transport equation. The neutron transport equation describes the motion and interaction of neutrons with nuclei in a nuclear reactor [2] and is given by:

$$\frac{1}{\nu}\frac{\partial\phi}{\partial t}(\bar{r},\hat{\Omega},E,t)+\hat{\Omega}\cdot\bar{\nabla}\phi(\bar{r},\hat{\Omega},E,t)+\Sigma_t\phi(\bar{r},\hat{\Omega},E,t) = \int_{4\pi}d\hat{\Omega}'\int_0^{\infty}dE'\Sigma_s(\hat{\Omega}'\rightarrow\hat{\Omega},E'\rightarrow E)\phi(\bar{r},\hat{\Omega}',E',t)+s(\bar{r},\hat{\Omega},E,t). \quad (1)$$

Equation (1) has 7 independent variables, position $\bar{r}(x,y,z)$, angular $\Omega(\theta,\phi)$, energy (E) and time (t) variables. The dependent variable is the angular neutron flux $\phi(\bar{r},\hat{\Omega},E,t)$. All terms on the left hand side of the equal sign are neutron loss terms and terms on the right hand side are neutron gain terms. Σ_t and Σ_s are total and scattering macroscopic cross-sections of the system being modelled, and they are measured in cm^{-1}. The source term $s(\bar{r},\hat{\Omega},E,t)$ includes external neutron sources and neutrons from fission, and is thus dependent on the neutron flux. Typically, for day-to-day reactor calculations, a deterministic approach is used to solve this equation because of its computational efficiency compared to other approaches such as the stochastic approach. The deterministic approach involves discretizing the variables of the neutron transport equation to a set of discrete equations that are numerically solvable [2] and is applied to reactor analysis calculations via a two step process.

A reactor core is discretized into nodes (assemblies). In the 2-step deterministic approach; the first step, a 2D steady state transport calculation is performed for each detailed assembly (node) type. The transport solution is used to produce energy collapsed, and spatially homogenized assembly parameters. During the homogenization and energy group condensation process the node-averaged parameters are conserved. The energy collapsed and homogenized parameters will be used in the second step by the full core diffusion solver. The homogenized group-collapsed parameters are generated in an approximated environment instead of the true physical conditions of the assembly in the reactor or full-core calculation. The four simplifications introduced in the deterministic calculational path are discussed in more detail in the following three subsections.

2.1. Energy group condensation and spatial homogenization

The simplifications made in the energy and geometric representation in the node involve performing a fine-group (100s) heterogeneous transport calculation. The transport solution (group heterogeneous neutron flux) is used to collapse the number of energy groups to few-group (less than 10) and homogenize the geometry over the node so that, each node has a constant set of few-group homogenized parameters that preserve the transport solution in an average sense. For few-group homogenized node to be equivalent to the fine-group heterogeneous node, the node-averaged reaction rates and surface transport leakages must be preserved. Therefore, for the homogeneous model to preserve the same average parameters calculated by heterogeneous model, the homogeneous cross-sections ($\tilde{\Sigma}^i_{\alpha,g}$) per group (where α is the type of nuclear reaction) must satisfy the following condition.

$$\int_{V_i}\int_{E_g}^{E_{g-1}}\tilde{\Sigma}^i_{\alpha,g}(\bar{r},E)\tilde{\phi}(\bar{r},E)d\bar{r}dE=\int_{V_i}\int_{E_g}^{E_{g-1}}\Sigma^i_{\alpha,g}(\bar{r},E)\phi(\bar{r},E)d\bar{r}dE. \quad (2)$$

However, because we do not have the few-group homogeneous flux ($\tilde{\phi}$) to accurately preserve heterogeneous solution, the few-group homogeneous cross-sections are calculated with approximated neutron flux before being used in the diffusion full core calculation, and this contributes to the spectral and spatial homogenization errors.

2.2. The diffusion approximation

The diffusion equation is a simplified form of the neutron transport equation. The main simplification of the diffusion theory is that, it eliminates the need to solve for angular flux as in most transport approximations. The derivation of the diffusion equation uses various approximations, one of them being Fick's law (Equation 3 which relates neutron current ($\bar{J}$) to neutron flux (ϕ) [2].

$$\bar{J}(\bar{r},t) = -D(\bar{r})\bar{\nabla}\phi(\bar{r},t). \tag{3}$$

$$\frac{1}{\nu}\frac{\partial\phi}{\partial t} - \bar{\nabla}\cdot D(\bar{r})\bar{\nabla}\phi + \Sigma_a(\bar{r})\phi(\bar{r},t) = S(\bar{r},t). \tag{4}$$

Equation (4) is the single group diffusion equation, and is valid when: there is an isotropic neutron source and scattering, angular flux distribution is linearly anisotropic and current density varies slowly with time. Here, D is called diffusion coefficient, and S is the source term. In both Equations (3 and 4) the energy variable has been suppressed for clarity.

2.3. Environmental dependency

Macroscopic cross-sections are generated in an environment that is not an exact match to the environment where the assembly will be used in the core calculation. For example, fuel cross-sections are typically generated in an infinite fuel environment (i.e. a fuel element with reflective boundary conditions), but in a reactor core, fuel elements are surrounded by fuel of, potentially different burn-up history and composition or even non-fuel elements (such as control rods or reflectors). By using cross-sections from an infinite environment for the fuel elements in a different core environment (i.e using approximate boundary conditions), an environmental error is introduced in the model. Not only fuel cross-sections suffer from environmental errors but also other cross-sections (control rods, reflectors, moderator, etc.) may suffer the same environmental errors.

The spectral, homogenization, and diffusion approximation errors are typically addressed by using equivalence theory (ET). ET is able to mitigate these errors by reproducing the node-integrated properties (reaction rates and leakage rates) of the known heterogeneous solution by introducing discontinuity factors in the homogenization procedure [3]. However, ET cannot address environmental error.

3. Methodology

The four simplifications introduced in the deterministic calculational path are investigated for a 2-node model consisting of fuel assembly next to a water node as shown in Figure 1. Reflective boundary conditions were applied to all outer sides.

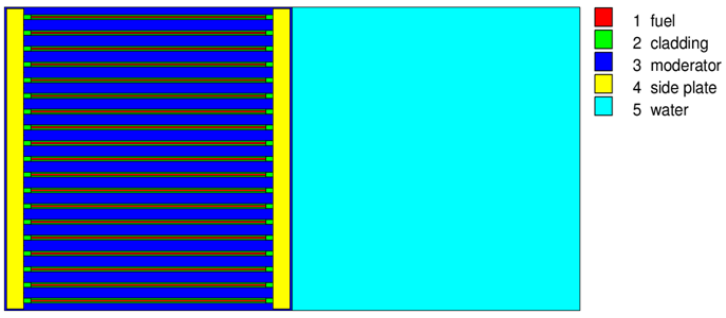


Figure 1. Heterogeneous 2-node model.

Three codes were used in this work, namely NEWT, MGRAC and Serpent. NEWT is part of the SCALE code system, MGRAC is part of the OSCAR-4 code system, and Serpent is a Monte Carlo code system. NEWT (NEW Transport code) is a 2D discrete ordinate transport code [8]. MGRAC (Multi-group Reactor Analysis Code) is a 3D nodal diffusion code [1][6] and Serpent uses Monte Carlo stochastic approach to reactor modelling [9].

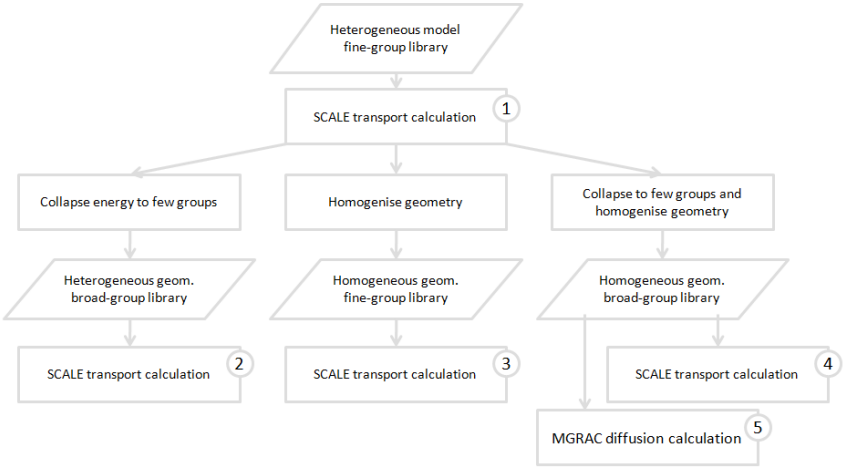


Figure 2. Flow diagram of breakdown of errors in the calculational path.

The first three simplifications can be quantified in a step-wise procedure as described in Figure 2. The NEWT and MGRAC codes are used in this part of the study. In Figure 2, calculation (1) is the reference transport calculation with a fine-group (238 groups) cross-section library and heterogeneous geometry. Calculation (5) is the few-group homogenized diffusion calculation after all simplifications have been made to the model. The errors induced by condensation, homogenization, and the diffusion approximation are determined by comparing the results of the five calculations marked in Figure 2. Table 1 is a summary of the breakdown of errors introduced in the calculational path. Only NEWT and MGRAC were used for this part of the study.

Table 1. Breakdown of error in calculational path

Compare calculations	Quantity to investigate
1 and 2	Spectral error
1 and 3	Homogenization error
2 and 4	Homogenization error in few-group
4 and 5	Diffusion approximation error in few-group
1 and 5	Total simplification error in calculational path

To investigate environmental error, the Serpent code was used to generate fuel and water cross-sections in the correct environment (fuel next to water) and continuous in energy (from which the reference k-eff is taken) and also to generate fuel cross-sections in the approximated environment (infinite fuel environment). The cross-sections generated by Serpent were used in MGRAC diffusion solver to calculate the k-eff in few-group (6 groups). Two MGRAC calculations were

set up; one with no environmental error (fuel and water cross-sections from 2-node Serpent, exact same environment) and one with environmental error where the cross sections were taken from an infinite fuel calculation in Serpent and placed in the 2-node model in the MGRAC. The comparison was made between k-eff calculated in the correct environment and k-eff calculated in the approximate environment for 6 groups. Serpent was used here because the functionality exists to generate nodal equivalence parameters for MGRAC (including discontinuity factors) from the Serpent calculation, but not from SCALE calculation. The generated equivalence parameters were used to resolve the first three errors and only the environmental error need to be studied further and corrected.

4. Results and discussion

The k-eff is a measure of the reactivity of a model and serves as an integral parameter to describe a model. The error will be described as the difference in k-eff (Δk) is measured in pcm (per cent mille) and calculated as: $\Delta k = \frac{k_1 - k_2}{k_1 \times k_2} \times 10^5$ where $\Delta k > 500$pcm is considered large.
The heterogenous transport calculation in 238 energy groups yields the reference reactivity k-eff = 1.17073 in 4118 seconds.

Table 2. Breakdown of errors introduced by simplifications in the 2-node problem

	2-groups	4-groups	6-groups	238-groups
Spectral error (pcm)	3267	819	63	N.A
Homogen. error (pcm)	-214	-141	-81	-22
Diffusion error (pcm)	-6525	-5084	-4094	N.A
Total error (pcm)	-3472	-4406	-4113	N.A

Table 3. Calculational time

	2-groups	4-groups	6-groups	238-groups
Spectral (second)	37	50	60	4118
Homogen.(second)	8	9	17	1313

Table 3 shows the time each calculation took to complete. In Table 2, the spectral error introduced by energy group condensation increases as the number of groups decreases. The reactivity increases with energy group collapsing because physical processes (nuclear reactions) are under represented in the group structures.
Homogenization induces somewhat smaller errors compared to the larger spectral errors observed. Unlike energy condensation, homogenization reduces reactivity, and this error increases as the number of groups decreases. It can be observed that, the 2-group case gives the largest error in reactivity of 214 pcm.
A large error in reactivity is introduced when using a diffusion solver instead of a transport solver. The 6-groups case, yields a reactivity of 4000 pcm lower than the transport calculation. The error increases to just above 6000 pcm as the energy groups are further collapsed. All three simplifictions reduce the calculational time significantly when applied to a 2-node problem.
In Table 4, environmental error was investigated in 6-groups only. Notice that the continuous energy, heterogeneous Serpent calculation and the 6-group homogenized diffusion calculation are equivalent (within some statistical margin) because ET was used in the cross-section generation.

Table 4. Environmental dependency effect in 6 energy groups

Codes	k_{eff}correct env.	k_{eff}Approx. env.	Δk (pcm)
Serpent	1.17050	-	N.A
MGRAC	1.17047	1.16051	733

This shows that ET can mitigate the error of the first 3 simplifications if the cross-sections are used in the exact same environment as that in which they were generated. We see that even though an error of about 4000 pcm has been negated by the use of ET, we are left with an environmental error of 733 pcm.

5. Conclusion

All four simplifications introduced in the deterministic calculational path were investigated for a two node fuel-water problem. From the results presented above, it can be seen that the diffusion approximation and environmental dependency are major contributors to the total error in the deterministic calculational path for a fuel-water model. However, the first 3 simplifications are mitigated through the use of equivalence theory by introducing equivalent nodal parameters to preserve not only node-averaged reaction rates but also the node-leakages. The equivalence theory only addresses the first 3 errors and not the environmental, therefore, the environmental error mitigation is still an area of active research. The findings can be used to develop models that can address environmental and diffusion approximation errors. Future work will look at modelling error analysis and quantification for other configurations (fuel-absorber, fuel-reflector, etc.) found in SAFARI-1 and go to the full core calculations.

References

[1] Joubert W.R, Reitsma F, Tomašević D.I 2002 *PHYSOR 2002, Comparative Study Of Nodal Cross Section Models Applied to MTR Core Analysis* (Seoul, Korea)

[2] Duderstadt J.J, Hamilton L.J 1976 *Nuclear Reactor Analysis*, John Wiley & Sons New York, N.Y

[3] Smith K.S 1986 *Progress in Nuclear Energy*, **17** 303

[4] Müller E.Z, Jourbet W.R, 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

[5] Sargeni A, Santamarina A 2005 *Annals of Nuclear Energy* **32** 89-99

[6] Prinsloo R.H 2012 *OSCAR-4 system overview.* RRT-OSCAR-REP-12002, Necsa internal report.

[7] Stander G, Prinsloo R.H, Müller E.Z, Tomašević 2008 *International Conference on Reactor Physics, Nuclear Power: A sustainable Resource* (Interlaken, Switzerland)

[8] Parks C.V 2009 *A modular code system for performing standardized computer analysis for licensing.* ORNLUNUREG/CSD-21R6, UNR Commission, Radiation Safety Information computational Centre.

[9] Leppänen J 2012 *Psg2/Serpent-a continuous-energy Monte Carlo Reactor Physics Burn-up Calculation Code* (VTT Technical Research centre of Finland)

Jet substructure: a discovery tool

Deepak Kar

School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africa

E-mail: deepak.kar@cern.ch

Abstract. Jets are the collimated bunches of hadrons measured in our detectors, created at high energy particle collisions. As we go to higher energies at the Large Hadron Collider (LHC), Higgs bosons, or yet undiscovered heavy particles are produced with very high energy and the decay products from these "boosted" particles tend to be contained in large radius jets. The internal structure of these jets is exploited to identify the original objects. In this talk, I will motivate the use of substructure techniques for probing new physics at the LHC. I will then discuss the recent experimental results on substructure measurements, including a very new and promising method called *shower deconstruction*.

1. Introduction: jets and jet substructure

Most of the interesting collisions at a hadron collider, such as at the Large hadron Collider (LHC) at CERN, involve production of quarks and gluons. When quarks and gluons are produced in a collision, they radiate other quarks or gluons, which in turn radiate again. As they cannot remain as free particles, they form colour-neutral hadrons by combining with other quarks and gluons, which are often created spontaneously from the vacuum. This is observed in the detector as a collimated spray of hadrons, which are termed as jets. There are different algorithms to combine the input objects (i.e. clusters of energy deposits in the calorimeter or simulated particles) in to jets [1]. Along with the specific jet reconstruction algorithm, a jet is also defined by a *size* measure, often termed as radius.

The main aim of the LHC is to discover new particles, which are heavier than already observed Standard Model (SM) particles. Typically, the heavy particles we are interested are not observed directly, but they decay into other particles that are captured by our detectors. Now if this heavy particle is produced with a very high energy, it will be *Lorentz boosted* in the detector center-of-mass frame, resulting in collimated decay products. If the boosted particles decay into a quark anti-quark pair for example, this means the resultant two jets will be largely overlapping, resulting in a single large radius jet in the detector. This discussion is not just relevant for new unseen heavy particles, but top quarks, W, Z and Higgs bosons produced with a high enough energy and decaying hadronically can yield such large radius jets.

The idea is then to exploit the internal (sub)structure of these large radius jets to determine their origin, and discriminate the signal against the background. For example, a large radius jet from a hadronically decaying top-quark should have three smaller radius (sub)jets inside, while a large radius jet created from a light quark and gluon should not have such a distinctive characteristic. As the upcoming LHC Run 2 will have more center-of-mass energy compared to Run-1, boosted objects will be more prevalent, and the chances of a discovery might depend on reconstructing them successfully.

Jet grooming and jet tagging are two classes of substructure techniques that have been developed to achieve this, which will be discussed in the subsequent sections.

2. Jet grooming

The large-radius jets not only capture the decay products from the heavy boosted object, but also include uncorrelated soft energy deposits, mostly from multiple proton-proton collisions in the same bunch-crossing (termed as *pile-up*), from extra activity associate with the hard scatter (termed *underlying event*) or from initial state radiation. The former is the biggest concern, since the rate of pile-up is going to increase in Run 2.

Three commonly used procedures (collectively called jet grooming) to remove these extra energy deposits, mass-drop filtering [2], trimming [3], and pruning [4] are schematically depicted in Fig. 1, following [5]. They all involve undoing the jet-forming, and using the initial jet constituents.

The mass-drop filtering procedure seeks to isolate concentrations of energy within a jet by identifying relatively symmetric subjets, each with a significantly smaller mass than that of the original jet. The trimming algorithm discards softer constituents with below a certain p_{T} fraction compared to the jet. The pruning ignores soft or wide-angle radiations.

In Fig. 2, an example result of the jet grooming, as well as the impact of pile-up is shown, following [5]. The left plot shows a comparison of jet mass distribution from simulated signal and background samples before and after grooming. The signal is a simulated Z' particle decaying to quark anti-quark pair produced with HERWIG [6]+JIMMY [7] generator setup, while the background is POWHEG [8, 9, 10]+PYTHIA6 [11] dijet events. It can be seen that after trimming, the Z' mass peak is clearly visible over the background (solid lines), whereas that was not the case before trimming (dotted lines). In the right plot, it can be seen that the ungroomed jet mass increases with the number of pile-up vertices, while after mass-drop filtering (with three different settings), the jet mass has essentially no dependence on pile-up.

3. Jet tagging: shower deconstruction

The next step is to come up with observables, distributions of which would be different for large-radius jets coming from signal and background events. Many such observables have been studied with varied amount of success [12]. These observables, and some other dedicated algorithms are used as *taggers*, where a probability of being a signal jet is assigned to each large-large radius jet. A relatively new tagger, called *Shower Deconstruction (SD) [13]*, has yielded very promising results for tagging boosted top quarks [14], and this will be discussed in this section.

The SD algorithm constructs a discriminant, χ_{SD}, optimised to distinguish jets produced in decays of signal particles (S) from jets produced by background processes (B). In the following discussion, the signal process used will be a hadronic top quark decay, and the background process is a jet originating from a single gluon. The SD algorithm is based on the assumption that a final state configuration containing N subjets with four-momenta $\{p\}_N = \{p_1, p_2, ..., p_N\}$ can be generated in many different ways in parton shower approach, and each of these constitutes a possible *shower history*. Then SD calculates the probability that a given shower history was realised in a given event. The shower histories are used to construct the likelihood ratio χ_{SD} by

$$\chi_{SD}(\{p\}_N) = \frac{P(\{p\}_N|S)}{P(\{p\}_N|B)} \tag{1}$$

where $P(\{p\}_N|S)$ is the probability of obtaining of $\{p\}_N$ given the signal hypothesis, and $P(\{p\}_N|B)$ is then the probability of obtaining $\{p\}_N$ from background jets arising from background processes. $P(\{p\}_N|S)$ and $P(\{p\}_N|B)$ are calculated as the sum of the probabilities for each shower history.

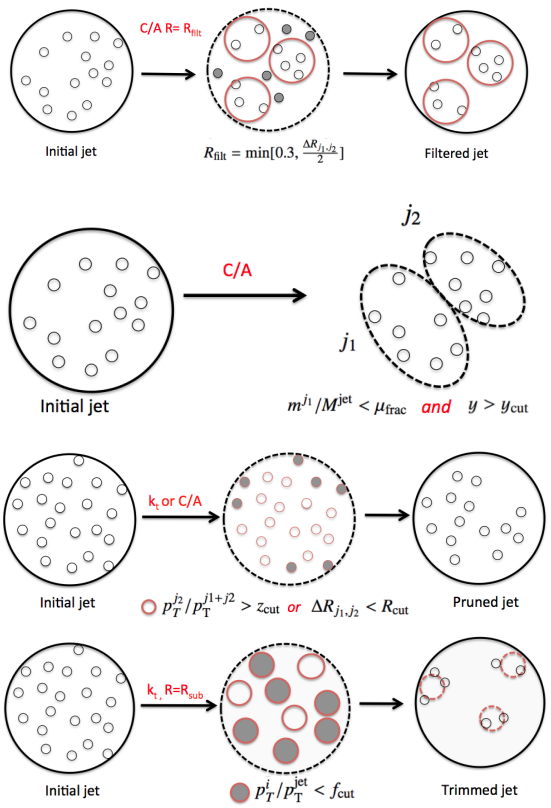


Figure 1. Diagrams depicting the the two stages of the mass-drop filtering, jet trimming and jet pruning procedures, from top to bottom.

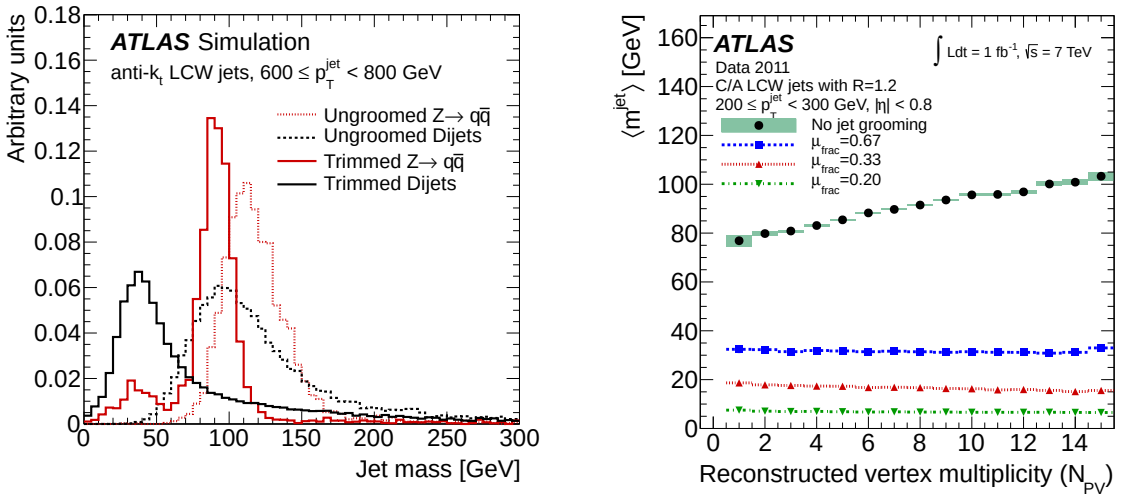


Figure 2. Comparison of ungroomed and groomed leading jet mass (left) and stability of groomed jet mass with number of reconstructed vertices (right), using simulated events.

Experimentally the SD is implemented by decomposing the large-radius jet into 0.2 radius subjets, which are used as inputs to the algorithm. The stability of χ_{SD} against pile-up is an important test of the algorithm. In Fig. 3, top left, it can be seen that indeed $\log \chi_{SD}$ is fairly flat as a function of number of reconstructed vertices. In the top right plot, the mass of the leading composite jet (i.e the jet formed by adding the subjets) is shown, and the peak corresponds to top quark mass peak. These comparisons are done with ATLAS data at a centre-of-mass energy of 8 TeV, corresponding to an integrated luminosity of 14.2 fb^{-1} and signal and background MC samples used in [15].

For the next part, a sample of simulated high-p_{T} top quarks is used to determine the tagging efficiency. These are generated through a sample of Z' with a mass of 1.75 TeV decaying exclusively to $t\bar{t}$ in the semi-leptonic channel, modelled using PYTHIA8 [16] generator.

The logarithm of χ_{SD} for the signal and background sample is shown in the bottom left figure, where a clear distinction can be observed. Finally, in the bottom right plot, the Receiver Operating Characteristic (ROC) curve for SD is compared to other commonly used top tagging algorithm [17]. This is obtained by varying the cut on $\log \chi_{SD}$, where background rejection is defined as the reciprocal of the efficiency. The SD algorithm yields the best background rejection over a wide range of signal efficiencies.

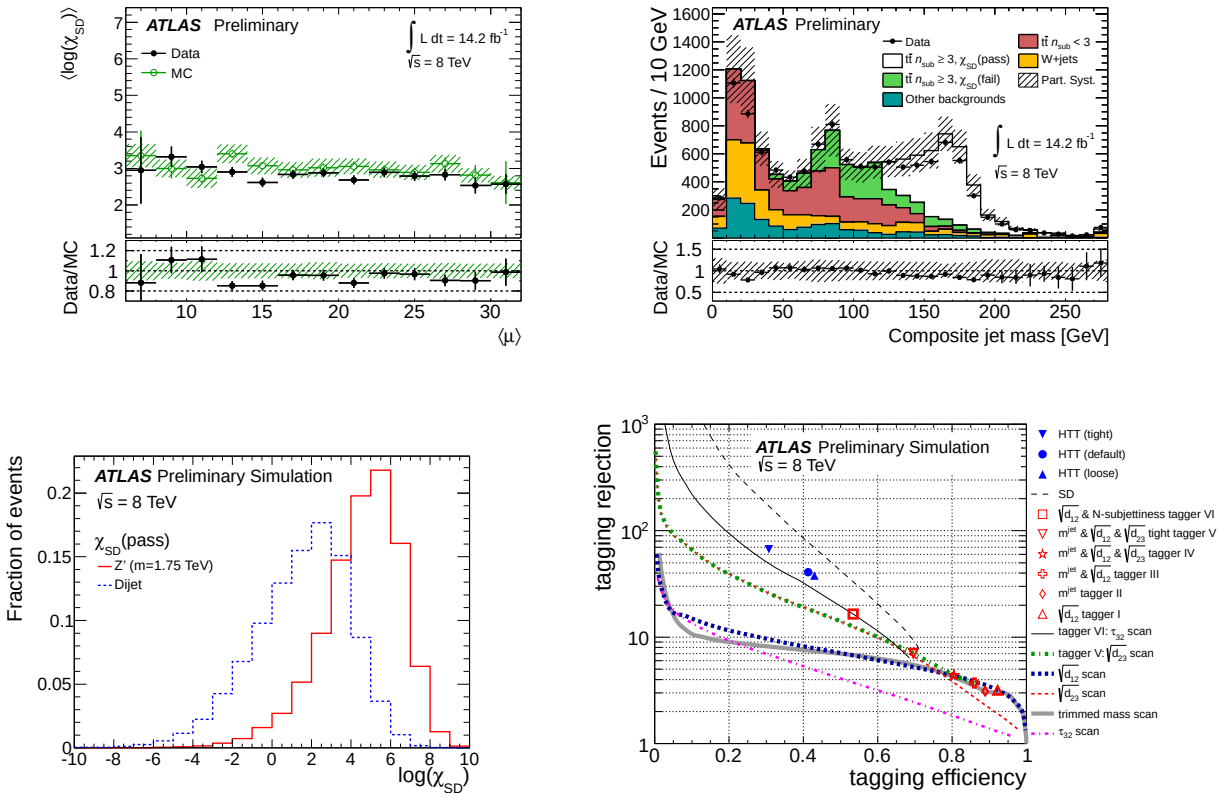


Figure 3. Different aspects of SD performance: stability of the $\log \chi_{SD}$ discriminant against pile-up (top left), reconstruction of top quark mass by subjets used in SD (top left), signal and background discrimination using $\log \chi_{SD}$ variable(bottom left) and comparison of SD ROC curve with other taggers (bottom right).

4. Summary

Substructure algorithms will be critical in searches for new physics in Run-2. The initial target will be validating the algorithms with Run-2 data, derive experimental uncertainties, and then use them in searches. The Shower Deconstruction algorithm, using the a maximum information approach is one of the promising algorithms, and expected to be used widely.

The performance of this algorithm has been examined in detail for data and MC samples of events predominantly arising from top-quark pair production observed in the lepton plus jets final state. The data were compared to simulation for observables, such as the composite jet mass defined by the mass of sum all of the subjet four-vectors considered by the SD algorithm, and the χ_{SD} observable. Satisfactory agreement was found between data and simulation as well as stable performance as a function of the pile-up conditions. The expected performance of the SD algorithm and of other top-tagging and substructure techniques has been estimated using samples of simulated high-p_{T} top quarks from Z' with a mass of 1.75 TeV as the signal and dijets as the background. For this scenario, the SD algorithm shows the best light quark and gluon jet background rejection over a wide range of top-jet signal efficiencies, when systematic uncertainties are not considered.

Shower deconstruction can easily be extended to tag boosted Z, W or Higgs bosons. ATLAS collaboration recently reported an excess in diboson mass spectra in hadronic decay channel [18]. SD and other optimised substructure algorithms will be used to probe this excess in Run 2, and confirm a discovery, if that is the case.

References

[1] Salam G P 2010 *Eur.Phys.J.C* **67** 637–686 (*Preprint* 0906.1833)
[2] Butterworth J M, Davison A R, Rubin M and Salam G P 2008 *Phys.Rev.Lett.* **100** 242001 (*Preprint* 0802.2470)
[3] Krohn D, Thaler J and Wang L T 2010 *JHEP* **1002** 084 (*Preprint* 0912.1342)
[4] Ellis S D, Vermilion C K and Walsh J R 2010 *Phys.Rev.D* **81** 094023 (*Preprint* 0912.0033)
[5] ATLAS Collaboration 2013 *JHEP* **1309** 076 (*Preprint* 1306.4945)
[6] Corcella G *et al.* 2012 Herwig 6.5 release note (*Preprint* hep-ph/0210213)
[7] Butterworth J M, Forshaw J R and Seymour M H 1996 *Z. Phys.* **C72** 637–646 (*Preprint* hep-ph/9601371)
[8] Alioli S, Nason P, Oleari C and Re E 2010 *JHEP* **06** 043 (*Preprint* 1002.2581)
[9] Frixione S, Nason P and Oleari C 2007 *JHEP* **11** 070 (*Preprint* 0709.2092)
[10] Nason P 2004 *JHEP* **11** 040 (*Preprint* hep-ph/0409146)
[11] Sjostrand T, Mrenna S and Skands P 2006 *JHEP* **05** 026 (*Preprint* hep-ph/0603175)
[12] Altheimer A, Arce A, Asquith L, Backus Mayes J, Bergeaas Kuutmann E *et al.* 2014 *Eur.Phys.J.C* **74** 2792 (*Preprint* 1311.2708)
[13] Soper D E and Spannowsky M 2013 *Phys.Rev.D* **87** 054012 (*Preprint* 1211.3140)
[14] ATLAS Collaboration 2014 Performance of shower deconstruction in atlas (*Preprint* ATLAS-CONF-2014-003)
[15] ATLAS Collaboration 2013 A search for $t\bar{t}$ resonances in the lepton plus jets final state with atlas using 14 fb^{-1} of pp collisions at $\sqrt{s} = 8$ tev (*Preprint* ATLAS-CONF-2013-052)
[16] Sjostrand T, Mrenna S and Skands P 2008 *Comput. Phys. Commun.* **178** 852–867 (*Preprint* 0710.3820)
[17] ATLAS Collaboration 2013 Performance of boosted top quark identification in 2012 atlas data (*Preprint* ATLAS-CONF-2013-084)
[18] ATLAS Collaboration 2015 (*Preprint* 1506.00962)

Ambient gamma dose rate measurements at Manyoni uranium deposits, Singida, Tanzania

K F Kasoga[1], D A Mwalongo[2], S F Sawe[2], M M Nyaruba[2] and U Dammalapati[1]

[1]Department of Physics, The University of Dodoma, P O Box 338, Dodoma, Tanzania.
[2]Directorate of Nuclear Technology, Tanzania Atomic Energy Commission (TAEC), P O Box 743, Arusha, Tanzania.

E-mail: kasoga_kasoga@yahoo.com.sg

Abstract. High levels of natural radiation have been measured at the proposed Manyoni uranium mines in Singida, Tanzania. Pre-mining baseline data of the ambient gamma radiation dose rates have been carried out and serve as reference information during and after the uranium mining. The absorbed gamma dose rates on contact i.e. at the surface and in air (one meter above the surface) were measured with a survey meter. The sampling coordinates were located by a GPS meter. The dose rates due to the naturally occurring radionuclides varied from 131 to 1678 nGy h^{-1} (98 to 1657 nGy h^{-1}) with the mean value of 458 nGy h^{-1} (436 nGy h^{-1}). The maximum value measured is about thirty times the world average of 59 nGy h^{-1}. The annual effective dose rates for the region range from 0.16 to 2.06 mSv y^{-1}. High gamma radiation background is detected at Mwanzi and Kinangali villages and there is a need for conducting studies on health risk assessment for the general public near the proposed uranium mining site.

1. Introduction

Studies on radioactivity and radiation in the environment play an important role in our day to day life and are essential for better protection measures to be taken against the health hazards due to this radiation. Contribution to the background radiation in the environment is mainly from the naturally occurring radioactive materials associated with the formation of the earth's crust and atmosphere (cosmic rays). These materials release radiation everywhere in our surroundings to which humans are exposed. In addition, the exposure to radiation by general public is due to testing of nuclear weapons, disposal of radioactive waste, leakages from nuclear reactors and manufacture of radioactive isotopes [1]. Investigations of background radiation present in natural environmental samples such as vegetation, water, air and soil provide vital information as they help to monitor the radiation levels present in the surroundings and also give an indication about changes in the radiation levels due to human activities [2]. These studies help also to identify and evaluate the distribution of radionuclides in the area [1].

Uranium is a naturally occurring radioactive, heaviest element available in nature, is found in almost all types of soils, rocks, sands and water. Taking into account that uranium (^{238}U), thorium (^{232}Th) and potassium (^{40}K) are always present in soil, their gamma radiation causes external exposures with the consequent absorbed doses to humans. Excessive natural radiation

exposure from uranium mining can also contribute to a variety of health hazards such as skin cancer and genetic effects [3].

In the recent years, extensive uranium exploration and feasibility studies in Tanzania have found several sites with economically viable uranium deposits. In 2009, deposits of uranium were discovered at Mkuju, Namtumbo district, southern Tanzania [4]. This discovery was followed by Manyoni uranium deposits (Singida region) and Bahi uranium deposits (Dodoma region) both in Central Tanzania [5]. Furthermore, several studies of natural radioactivity have been conducted in different parts of Tanzania to assess environmental radioactivity levels in areas with proven uranium deposits and phosphate mining sites. For example, background radioactivity levels and elemental composition at Mkuju uranium deposit was reported by Mwalongo [6]. Baseline data of external ionizing radiation dose at proposed Mkuju uranium mining site and its neighbouring residential areas was reported by Lolila [7]. Natural uranium levels in Bahi district was reported by Orazaliyev [8]. Elisadiki [9] measured the terrestrial background radiation dose rate in the vicinity of proposed Manyoni uranium project using CaF_2 thermoluminescent dosimeters.

Uranium exploration in Manyoni district brought concerns to public residing around the exploration sites and nearby villages. Concerns are mainly on the radiological health hazard which is associated with uranium exploration. During uranium mining and milling operations, large quantities of waste containing naturally occurring radioactive materials (NORMs) are generated and may contaminate the environment and therefore increase ambient radiation level. For this reason, we have carried out measurements of ambient gamma radiation doses in the proposed uranium mining area and surrounding villages. This pre-mining baseline data on the ambient gamma radiation doses serves as a reference information to assess any changes in the ambient background radioactivity levels during and after uranium mining. We present results of our measurements and discuss them comparing background radiation level in regions of Tanzania and other countries and assess the health risks to the workers and general public.

2. Materials and methods

2.1. Description of the study area

Manyoni District is located in the central part of Tanzania. Its geographical coordinates are 5 45' 0" South, 34 50' 0" East. It has an area of 28,620 km^2 with population of 296,763 people [10]. The Manyoni project area (Fig. 1) is situated in the northern section of the Bahi regional uranium province near the town of Manyoni, which is 120 km north-west of Dodoma, the capital of Tanzania. The region incorporates an extensive closed draining system developed over weathered uranium rich granites. This drainage captures dissolved uranium leached from underlying rocks and transports it to suitable precipitation trap sites. The uranium targets in the area are described as calcrete-hosted uranium mineralization near to the surface and sandstone-hosted deposits within buried fluvial channel systems [5].

2.2. Measurement of gamma dose rates

Dose rate measurements were taken over a large area on the sites stretching from the expected mining zones to the the neighbouring residential areas (town and villages). The dose rates on contact (on the ground surface) as well as in air (at 1 meter above the ground) in the study areas were measured using a gamma radiation survey meter (Model: RAM DA-3-2000, Serial no. 8312-035) provided by the Tanzania Atomic Energy Commission (TAEC), Arusha. It was calibrated by the national calibration laboratory for ionizing radiation which is traceable to the International Measurement System through the International Atomic Energy Agency (IAEA) with an overall uncertainty of 10%. The energy response of the detector ranged from 60 keV to

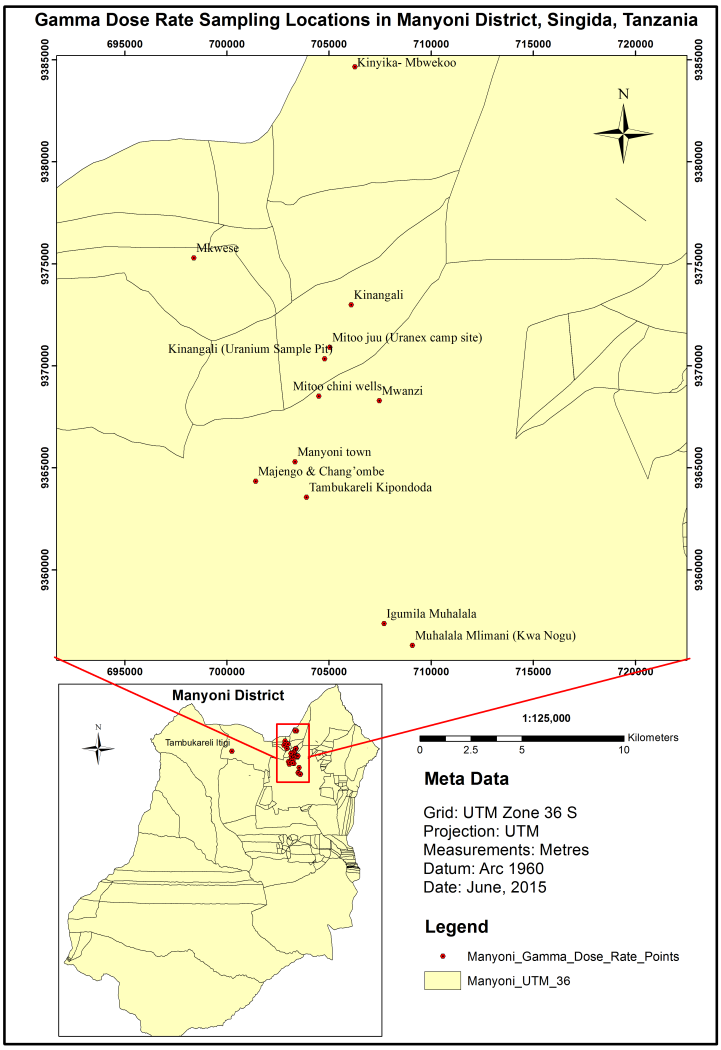


Figure 1. Map of Manyoni district, Singida, Tanzania with sample locations where ambient gamma dose rates are measured. Tambukareli Itigi is shown in the Manyoni district map and the remaining locations are shown in the expanded view.

1.3 MeV. The measurements were taken by positioning the survey meter at each point, switched on and allowed to stabilize for one minute and the readings were recorded. At each location, fifteen data points (n = 15) were recorded. Coordinates of all measurement points were located and recorded by using a Global Positioning System (GPS) meter. For Masigati (row 12 of Table 1) coordinates are not recorded.

2.3. Calculation of annual effective absorbed dose rate

The measured gamma dose rates were converted into annual effective dose rates using the equation below [3]:

$$EDR = D \times f \times v \times T \times 10^{-6} \quad (1)$$

where, EDR = Annual effective absorbed dose rate in mSv y^{-1}, D = measured absorbed dose rate in outdoor (nGy h^{-1}), f = conversion coefficient (0.7 Sv Gy^{-1}), v = outdoor occupancy

factor (0.2), T = Time in hours (8760 hours per year) and the factor 10^{-6} is to convert nano Sievert to milli Sievert.

3. Results and Discussion

The data taken at different locations are shown in Table 1. Presented in the table are the minimum and maximum of the data at each location. The absorbed dose data taken at the surface (column 2), one meter above the surface (column 3), the average dose of all the data taken at each location (column 4), annual effective dose rates on surface and at 1 m (column 5 and 6, respectively) are presented in this table. The uncertainty in the measured absorbed dose rates lies between 10 - 15%.

Table 1. Absorbed and average radiation dose rates and annual effective dose at different locations in Manyoni area. The number of data points at each location are fifteen (n = 15) unless otherwise indicated.

Location	Absorbed dose (nGy h^{-1})		Average dose	Effective dose (mSv y^{-1})	
	surface	at 1 meter	(nGy h^{-1})	surface	at 1 meter
Manyoni town	185 - 338	164 -316	234 - 254	0.23 - 0.41	0.20 - 0.39
Mitoo Chini wells	131 - 283	98 - 262	181 - 204	0.16 - 0.35	0.12 - 0.32
Mwanzi (n = 21)	927 - 1678	883 - 1657	1317 - 1343	1.14 - 2.06	1.08 - 2.03
Kinangali	785 - 1439	741 - 1428	1098 - 1132	0.96 - 1.76	0.91 - 1.75
Tambukareli Kipondoda	262 - 382	251 - 360	305 - 323	0.32 - 0.47	0.31 - 0.44
Changombe & Majengo	174 - 273	153 - 262	205 - 226	0.21 - 0.33	0.19 - 0.32
Kinyika-Mbwekoo	382 - 829	360 - 807	533 - 562	0.47 - 1.02	0.44 - 0.99
Igumila-Muhalala	174 - 305	164 - 273	216 - 235	0.21 - 0.37	0.20 - 0.33
Muhalala Mlimani	273 - 458	262 - 436	334 - 355	0.33 - 0.56	0.32 - 0.53
Mitoo juu	164 - 240	131 - 229	179 - 201	0.20 - 0.29	0.16 - 0.28
Tambukareli Itigi	196 - 305	164 - 283	222 - 242	0.24 - 0.37	0.20 - 0.35
Masigati	371 - 501	338 - 491	406 - 424	0.45 - 0.62	0.41 - 0.60

From Table 1, Mwanzi and Kinangali (row 3 and 4) showed the highest levels of gamma radiation dose. The effective dose rates at Mwanzi and Kinangali are twice the recommended value of the annual effective dose rate for the general public. The total absorbed dose estimated on the surface varies from 131 to 1678 nGy h^{-1} with an average value of 458 nGy h^{-1}. The maximum value measured is about 30 times higher than the world average of 59 nGy h^{-1} while the obtained average is 8 times higher than the world average.

The corresponding outdoor annual effective dose range from 0.16 to 2.06 mSv y^{-1} with an average value of 1.14 mSv y^{-1}; while the world wide average annual effective dose is approximately 0.5 mSv y^{-1} and the results for individual countries being generally within the 0.3 to 0.6 mSv y^{-1} range for outdoors [9]. The total absorbed dose estimated at 1 meter above the ground varies from 98 to 1657 nGy h^{-1} with an average value of 436 nGy h^{-1}. The maximum value measured is about 30 times higher than the world average of 59 nGy h^{-1} while the obtained average is 7 times higher than the world average.

In Table 2 absorbed doses reported at other locations in Tanzania and other countries are given. The values at Mwanzi and Kinangali are about the same as the average doses reported at Minjingu phosphate mine. The minimum absorbed dose reported at Mkuju uranium mining site is more than four times higher than the maximum measured at Mwanzi. However, the absorbed

Table 2. Comparison of gamma radiation absorbed and average dose rates at different locations in Tanzania and from other countries. The dose rates given are minimum and maximum for the location.

Country	Location	Absorbed dose rate (nGy h^{-1})	Average (nGy h^{-1})	Reference
Tanzania	Manyoni region	131 - 1678	458	This work
	Bahi region	121 - 594		[8]
	Minjingu phosphate mine	1375-1475		[11]
	Mkuju river	24.8 - 260	99.8	[7]
	Likuyu village	24 - 148	55.5	[7]
	Mkuju uranium mining site	7804.8 - 23360		[6]
Other	Ramsar, Iran	69.63 - 29630.37	1141.55	[12]
	Kerala, India	20.55 - 1100.46	55.94	[3]
	Norway	20.55 - 1200.91	73.06	[3]
	Italy	3.42 - 228.31	74.20	[3]
	Hongkong	51.37 - 119.86	86.76	[3]
	World average	18.26 - 93.61	59	[3]

dose reported from Bahi area, Mkuju river and Likuyu village are well below our values. The comparative results of absorbed dose rates for high background radiation areas around the world are shown in Table 2. All data reported in Table 2 for Tanzania are higher than the world average with the exception of Likuyu village.

4. Conclusions

In conclusion, we have measured the ambient gamma dose rates in Manyoni town and eleven locations near the proposed uranium mines at Manyoni, Tanzania. At two locations: Mwanzi and Kinangali, the measured values are about thirty times higher than the world average and the remaining locations also have higher background radiation levels. The annual effective dose rate at two places is twice the limit to the general public. These places need a full analysis of all exposure pathways (radon, diet, dust etc.) that may lead to health hazard to the members of public residing in these areas.

Acknowledgments

The authors are grateful to Frank E S Mahuve, Department of Geology, The University of Dodoma for his help with drawing of maps. One of the authors U D acknowledges kind help from Arun K Thazathveetil, Department of Chemistry, Northwestern University, Evanston, USA and Peane Maleka, iThemba LABS, South Africa for careful reading of the manuscript.

References

[1] IAEA 2010 *Radioelement mapping* nuclear energy series No. NF-T-1.3 (Vienna: IAEA)
[2] Hu Q -H, Weng J -Q and Wang J -S 2010 *J. Environ. Radioactiv.* **101** 426-437
[3] UNSCEAR 2000 *Sources and effects of ionizing radiation, United Nations Scientific Committee on the Effect of Atomic Radiation* (New York: United Nations)
[4] Mantra EIS 2010, *Mantra Tanzania Limited Environmental Impact Statement for the Proposed Uranium Mining Project at Mkuju River Project*, Namtumbo, Vol. 1, Final Report

[5] Uranex 2010 *New Uranium Mineralization discovered at Manyoni* (website www.infomine.com/index/pr/Pa872980.PDF)
[6] Mwalongo D A 2011 *Determination of background radioactivity levels and elemental composition at Mkuju uranium deposit in Tanzania* MSc Thesis, University of Dar es Salaam, Tanzania
[7] Lolila F 2011 *Establishment of baseline data of external ionizing radiation dose at proposed uranium mining sites and their neighbouring residential areas in Tanzania: The case study Mkuju river* MSc Thesis, University of Dar es Salaam, Tanzania
[8] Orazaliyev Y 2013 *Report on natural uranium level in Bahi district, Tanzania* University of Munich, Germany
[9] Elisadiki J 2013 *Terrestrial background radiation dose rate in the vicinity of the proposed Manyoni uranium project* MSc Thesis, University of Dar es Salaam, Tanzania
[10] United Republic of Tanzania 2012 *United Republic of Tanzania Population and Housing Census (PHC)*
[11] Banzi F P, Kifanga L D and Bundala F M 2000 *J. Radiolog. Prot.* **20** 41-51
[12] Ghiassi-ejnad M,Mortazavi S M J, Cameron J R, Niroomand-rad A and Karam P A 2002 *Health Physics* **82** 87-93

An electronics test-bench for the certification of the Tile Calorimeter of the ATLAS detector

C O Kureba, X Ruan, M Spoor, M Govender, I Hofsajer, B Mellado and C Sandrock

School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africa

E-mail: Chamunorwa.Kureba@wits.ac.za

Abstract. The instantaneous luminosity of the LHC at CERN is envisaged to be increased by up to 5-7 times after its upgrade in the year 2022. Termed the upgrade Phase-II, the High Luminosity LHC will bring with it a mandatory complete re-design of the read-out electronics in the Tile Calorimeter of the ATLAS detector. Here, the new read-out architecture is expected to have the front-end electronics transmit fully digitized information of the detector to the back-end electronics system. This will allow more sophisticated reconstruction algorithms which will contribute to the required improved trigger efficiencies at high pile-up. Currently, the MobiDICK test-bench is being used to test the front-end electronics of the TileCal. The MobiDICK will be replaced by a new test-bench for the new Phase-II electronics. The new test-bench is being designed, developed and assembled by the University of the Witwatersrand and installed at CERN to gain experience with prototype electronics. The new PROMETEO is a portable, high-throughput electronic system for full certification of the new front-end electronics of the ATLAS TileCal. A description is given in this article, of the overall design of the new PROMETEO, and its crucial role in TileCal electronics upgrade.

1. Introduction

The European Organization for Nuclear Research, well known as CERN, is a research organization that operates the largest particle physics accelerator in the world, the Large Hadron Collider (LHC). The LHC accelerates and collides protons, and also heavy ions. ATLAS [1] is one of two general-purpose detectors at the LHC. It investigates a wide range of physics, from the search for the Higgs boson to extra dimensions and particles that could make up dark matter. The Tile Calorimeter (TileCal) [2] is the central hadronic calorimeter of the ATLAS detector. Alternating steel plates and plastic scintillator tiles make up each cell of TileCal. The energy of the detected particles is sampled by the scintillators and signals are collected from the photomultipliers (PMTs) by the front-end electronics located in super-drawers, which is the outermost part of the detector.

The readiness of the TileCal for data collection is determined by testing each of its modules, hence, the need for an electronics test-bench. The High Luminosity LHC will lead to an inevitable complete re-design of the read-out electronics in the TileCal. Figure 1 shows a comparison of the current and future front-end electronics set-up for the TileCal. In the current front-end, pipeline memories are used to store digitized data samples before they can be trigger-selected. The read-out electronics in each super-drawer are daisy-chained, resulting in the sharing of a data connection to the Read-Out Driver (ROD). In the future electronics

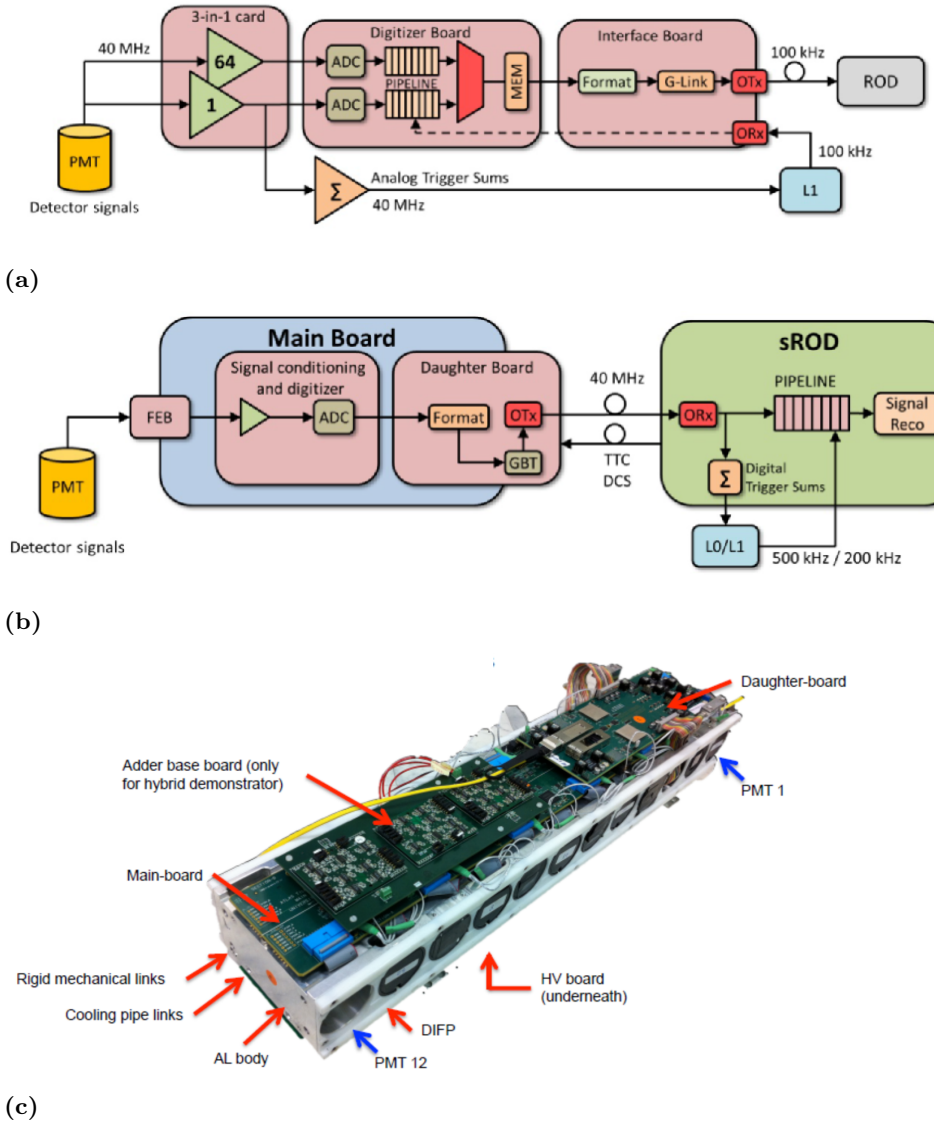


Figure 1: **(a)** Current front-end electronics; **(b)** Equivalent electronics for Phase-II upgrade; **(c)** A single mini-drawer, recently built at CERN for the upgrade of the TileCal electronics.

design, each super-drawer is divided into 4 mini-drawers. One mini-drawer hosts 12 PMTs and 1 daughter board. Each daughter board has one link to the Tile Pre-Processor (TilePPr), thereby reducing potential failure rate by 25% with respect to ATLAS. There are up to 48 PMTs in one super-drawer, in groups of 6 in a digitizer board. There are 16 digitizers for one super-drawer, interfaced by 1 interface card. It is envisaged that the future electronics will have:

- a Tile Pre-Processor (TilePPr), 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 the LHC Phase II upgrade, the current Mobile Drawer Integrity ChecKing (MobiDICK) system [4, 5] test-bench will be replaced by the next generation test-bench for the TileCal super-drawers, the new PROMETEO (A Portable Read-Out ModulE for Tilecal ElectrOnics). The MobiDICK system faces challenges against ageing and new technologies [6]. The PROMETEO system is designed to certificate the TileCal front-end electronics by performing multiple tests. The PROMETEO's prototype is being assembled and designed by the University of the Witwatersrand and installed at CERN to gain system experience.

2. PROMETEO (new test-bench): Hardware design

PROMETEO is inspired by the currently used MobiDICK test-bench whose main board is a Xilinx ML507 evaluation board with a Virtex-5 FPGA. The MobiDICK test-bench has a server running on Power PC, which connects to the client via ethernet. It records digital data at a maximum rate of 100 kHz. It uses "slow" canbus to control the current integrator circuit and the HV applied to every PMT and a custom designed ADC board to sample the analogue signals.

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 mini-drawer be self-contained and portable for maintenance inside the detector, and be low-cost and scalable for network usage. In the current design being deployed for the hybrid demonstrator that will also have analogue output, PROMETEO will make use of an ADC board. The current version is using a Xilinx VC707 board. Figure 2 shows a block diagram of the PROMETEO.

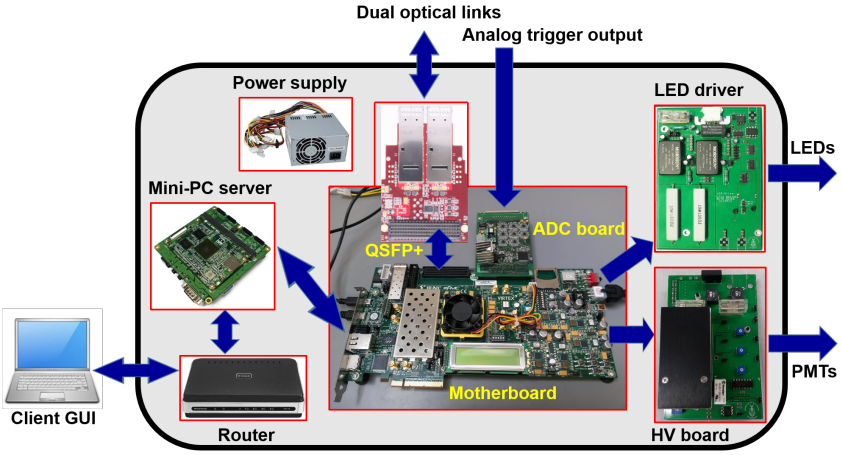


Figure 2: Block diagram of the PROMETEO system.

The PROMETEO system 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 mini-drawers,
- Analogue to Digital Converter (ADC) FMC card: for the digitization of the analog trigger output from the the hybrid demonstrator adder cards,
- 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,
- Mini-PC server: to provide direct connection to the test results when a user logs onto the webpage. It is a low-cost Cortex-A9 ARM processor board based on the Freescale i.MX6 family of System on Chips (SoCs). The mini-PC server uses a Linux, ROOT and Apache environment to host the PROMETEO webpage.
- Ethernet router,
- Aluminum enclosure: whose dimensions are 50 cm × 35 cm × 20 cm in terms of length, width and depth, thickness of 3 mm and weight of 8 kg.

3. PROMETEO: Software design

The software design of PROMETEO is based on direct hardware access technology for FPGAs with direct ethernet connection. Figure 3 shows a block diagram of the PROMETEO system software structure. An IPbus [7] ip-core module on the FPGA that enables direct communication

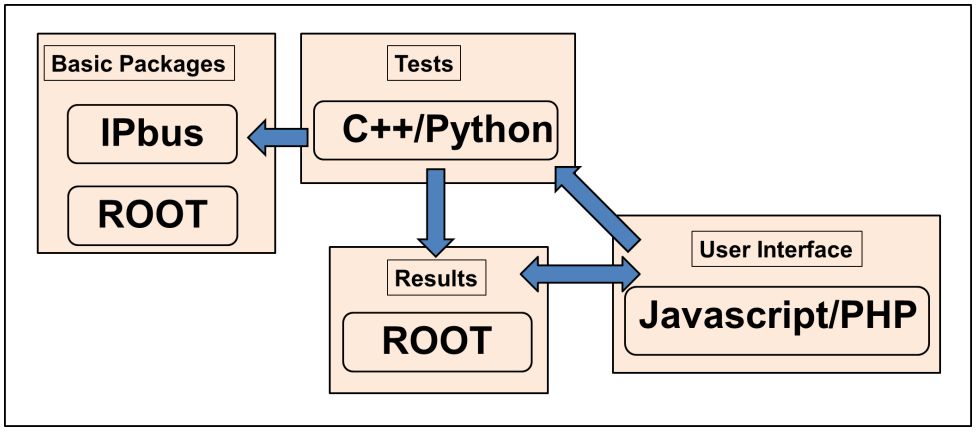
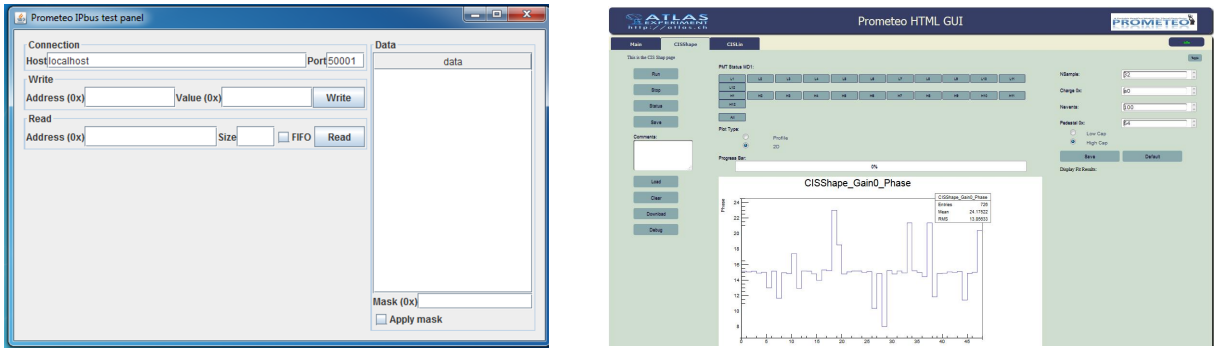


Figure 3: Block diagram of the PROMETEO system software structure.

of the registers is employed. A custom built software version of the IPbus client that is implemented in C++, Java and Python is used to allow multiple developments and tests to perform in parallel. At this stage there are two implementations of the software being developed, namely, one in pure Java and another one in C++/Python. The Java software is inspired by the one in the MobiDICK GUI where a core window loads different tests as plugin tabs, and each one executes and displays a different test. The C++ one implements the tests directly as command line applications. A front-end HTML/Python interface is used to execute the tests and assess the results. Figure 4 shows snapshots of the Java and HTML GUIs used by the client to execute the PROMETEO system tests.

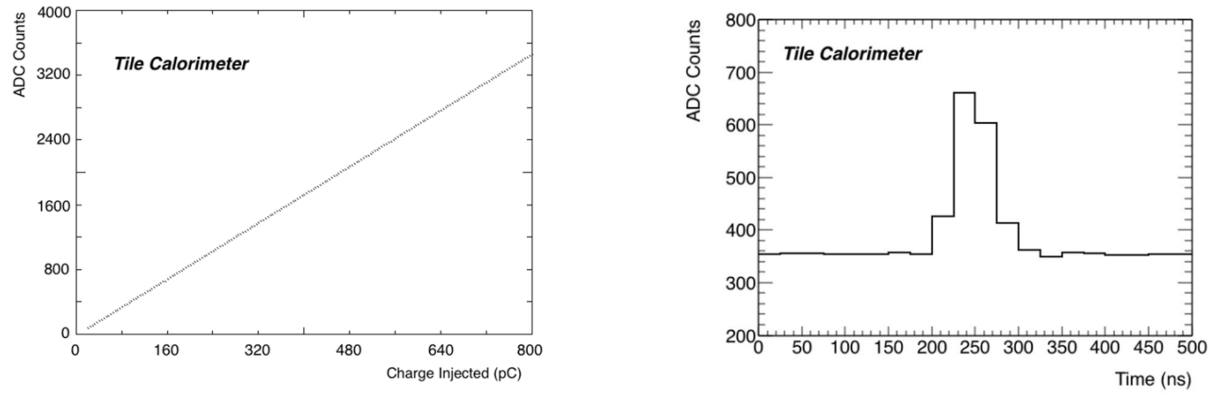
Calibration methods are generally employed in order to assess the performance of the TileCal Demonstrator electronics using the PROMETEO, namely:

- **Charge Injection System (CIS)**: Here, the electronics is tested directly through the injection of a known charge into the front-end boards, thereby simulating a PMT signal. Figure 5a shows a typical result of such a measurement. One can see a good linearity in the Demonstrator's High- and Low-Gain fast readout channels.
- **LASER calibration system**: LASER pulses of known intensity are sent to the PMTs and the PMT response is used to calibrate their gain as well as settings of the trigger timing. A typical response of the PMTs in the new Demonstrator system is shown in Fig. 5b.



(a) **(b)**

Figure 4: Java **(a)** and HTML **(b)** GUIs used by the client to execute the PROMETEO system tests



(a) **(b)**

Figure 5: Results for calibration of the Demonstrator using two methods: **(a)** electronics response to a known charge; **(b)** a triggered LASER pulse. [9]

These results are used for diagnosis of faulty PMTs, 3 in 1 cards, mainboards, daughter boards, optical links and analogue cards. Most of the tests are implemented in the HTML interface, thereby leaving Java as an excellent tool for debugging at expert level.

4. Conclusions

PROMETEO, the next generation test-bench for the certification of the new TileCal electronics has been designed and manufactured in South Africa, in preparation for the LHC phase-II upgrade. PROMETEO is inspired by the current MobiDICK test-bench. It is capable of performing multiple tests on the electronics modules of the upgraded superdrawers. PROMETEO is currently in a prototyping phase and currently serves a hybrid demonstrator of the future Phase-II electronics. All components of PROMETEO have been manufactured and are undergoing tests as they are added to the system. Further tests are being developed and tested in parallel with the development of the demonstrator. PROMETEO software is well advanced and used everyday to assess the stability of the first demonstrator.

References

[1] ATLAS Collaboration 2008 The ATLAS Experiment at CERN Large Hadron Collider *JINST* **3** S08003
[2] ATLAS Collaboration 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 Giangiobbe V 2008 Performance of the TileCal superdrawers from a global analysis of the MobiDICK tests *ATL-TILECAL-PUB-2008-007*
[5] Carrio F, Kim H Y, Moreno P, Reed R, Sandrock C, Schettino V, Shalyugin A, Solans C, Souza J, Usai G and Valero A 2014 Design of an FPGA-based embedded system for the ATLAS Tile Calorimeter front-end electronics test-bench *ATL-TILECAL-PROC-2013-017*
[6] Solans C, Carrio F, Kim H Y, Moreno P, Reed R, Sandrock C, Ruan X, Shalyugin A, Schettino V, Souza J, Usai G, Valero A on behalf of the ATLAS Tile calorimeter system 2014 Computing challenges in the certification of ATLAS Tile Calorimeter front-end electronics during maintenance periods *J. Phys.: Conf. Ser.* **513** 012035
[7] Frazier R *et al* 2013 The IPbus Protocol: An IP-based control protocol for ATCA/μTCA *Draft 6*
[8] Moreno P *et al* 2013 A new portable test-bench for the ATLAS Tile Calorimeter front-end electronics *JINST* **8** C02046
[9] *https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsTile*

The search for Dark Matter in association with the Higgs boson with the di-photon decay

C. O. Kureba, X. Ruan and B. Mellado

University of Witwatersrand,
1 Jan Smuts Avenue Braamfontein 2000,
Johannesburg, South Africa

E-mail: cokureba@gmail.com

Abstract. The ATLAS and CMS experiments at the Large Hadron Collider discovered a Higgs boson like particle in 2012. The differential and fiducial cross sections of the Higgs boson are measured using $20.3\,\mathrm{fb}^{-1}$ 2012 data taken at $\sqrt{s} = 8\,\mathrm{TeV}$ after the discovery by ATLAS. The measurement is focusing on the Higgs boson kinematics and jet activity, including the Higgs boson transverse momentum, rapidity and the Higgs boson+jet production mode. The Higgs boson candidates are extracted by fitting the two-photon invariant mass spectrum. The observed kinematic distribution of the Higgs boson is translated to particle level to reduce the detector efficiency and resolution, using bin-by-bin unfolding method. A distortion of the Higgs boson transverse momentum is found in comparison with the state-of-the-art predictions. One of the explanations is the Higgs boson production in association with invisible particles, such as the dark matter. The observation indicates that the missing particle has intermediate energy and same order of production cross-section as the Standard Model gluon-gluon fusion to the Higgs boson process. The search for dark matter in association with the Higgs boson will be performed in the 2015 data taking in the Higgs boson decaying into two-photon channel. The study will focus on the Higgs boson production associated with intermediate missing transverse energy. The knowledge and understanding of the missing transverse energy reconstruction is critical.

Keywords: ATLAS; Higgs boson; Diphoton; Differential; Fiducial; Dark matter

1. Introduction

The Standard Model (SM) Higgs boson-like particle was observed in 2012 by ATLAS and CMS [1] [2]. After the discovery, the Higgs boson differential cross section and fiducial cross section were measured [3].

The ATLAS detector [4] is one of the general purpose detectors at the LHC. Its four sub-detectors: the inner detector, the electromagnetic calorimeter, the hadronic calorimeter and the muon spectrometer provide the precise track and energy measurements. Electrons, muons, photons and jets are reconstructed in the detector. The measurements of the Higgs boson production cross section and mass were performed by combining the $20.3\,\mathrm{fb}^{-1}$ $\sqrt{s} = 8\,\mathrm{TeV}$ data taken in 2012 and $4.5\,\mathrm{fb}^{-1}$ $\sqrt{s} = 7\,\mathrm{TeV}$ data taken from 2011. The differential and fiducial cross section measurements were derived only using 2012 data to avoid large statistical fluctuation. The photon energy was re-calibrated in 2013 including implementing the multivariate analysis (MVA) method to improve the identification performance. The unfolding technique was

implemented to extract the Higgs boson distribution from data to the particle level. The results are compared with several Monte Carlo (MC) predictions using higher order QCD calculations.

In the measured result, a small excess in the Higgs boson transverse momentum (p_T) spectrum were observed. This excess provided evidence for the Higgs boson production associated with missing particle. A dark matter particle, which is not observable via electromagnetic spectrum, is one of the possible candidate. Figure 1 shows one of the possible explanation of the distortion of Higgs boson p_T spectrum. The search based on the 2011 and 2012 run I result will be performed in 2015 $\sqrt{s} = 13\,\text{TeV}$ run II new data.

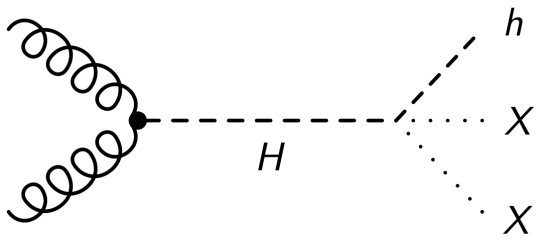


Figure 1. The Feynman diagram of a kind of the Higgs boson production in association with the dark matter particles. The SM Higgs boson is h and the dark matter candidates are Xs. The h and Xs are coupled with a heavy mediator H, which is produced by gluon-gluon fusion process. This model provides moderate missing transverse energy.

2. The Event Selection

In the Higgs boson decaying into the diphoton channel, events are required to have at least two photons. The events must pass the trigger criteria that the transverse energy (E_T) should be larger than 35 GeV and 25 GeV for the leading (highest E_T) and sub-leading photons, respectively. The photons are restricted to be within the fiducial calorimeter region of $|\eta| < 2.37$ [1] and to exclude the transition region between the barrel and the endcap calorimeters, $1.37 < |\eta| < 1.56$. The relative p_T: $p_T/m_{\gamma\gamma}$ of the leading and subleading photon should be larger than 0.35 and 0.25, respectively, in which the $m_{\gamma\gamma}$ is the invariant mass of the two photons. The photons are well identified and isolated. The jets are required to have $E_T > 25\,\text{GeV}$ for $|\eta| < 2.4$ and $E_T > 30\,\text{GeV}$ for $2.4 < |\eta| < 4.5$.

3. The Extraction of the Higgs Boson

To extract the Higgs boson in the continuous background, the signal plus background fit was performed on the diphoton invariant mass spectrum. The signal shape was modelled by fitting the MC signal shape using Crystal-ball+Gaussian function. The signal MC consists of gluon-gluon fusion (ggF), vector boson fusion (VBF), the Higgs boson produced in association with W boson (WH), Z boson (ZH) and top quarks ($t\bar{t}$H). The ggF and VBF are produced using Powheg [5] generator and showered by Pythia [6], the WH, ZH and $t\bar{t}$H are produced and showered using Pythia. The signal shape and predicted event yields are parameterised as a function of the Higgs boson mass.

The background shape are fitted using exponential, polynomial, Bernstein polynomial and power law functions. The final function type used on a particular background shape is decided

[1] η is the pseudorapidity, defined as $\eta = -\ln[\tan(\theta/2)]$, θ is the angle between the particle momentum and the beam axis.

by minimising the spurious signal, which is obtained by fitting the high statistics background MC with signal plus background model. The typical diphoton mass spectrums in jet bins are shown in the Fig. 2.

All of the systematic uncertainties are considered in the analysis. The photon energy scale, photon energy resolution, the photon selection and identification efficiencies are considered. In the jet categories, the jet energy scale and resolution are dominant. All of the photon, jet and lepton uncertainties are propagated to the missing transverse energy calculation. The theoretical uncertainties like the energy scale, PDF are taken into account as well.

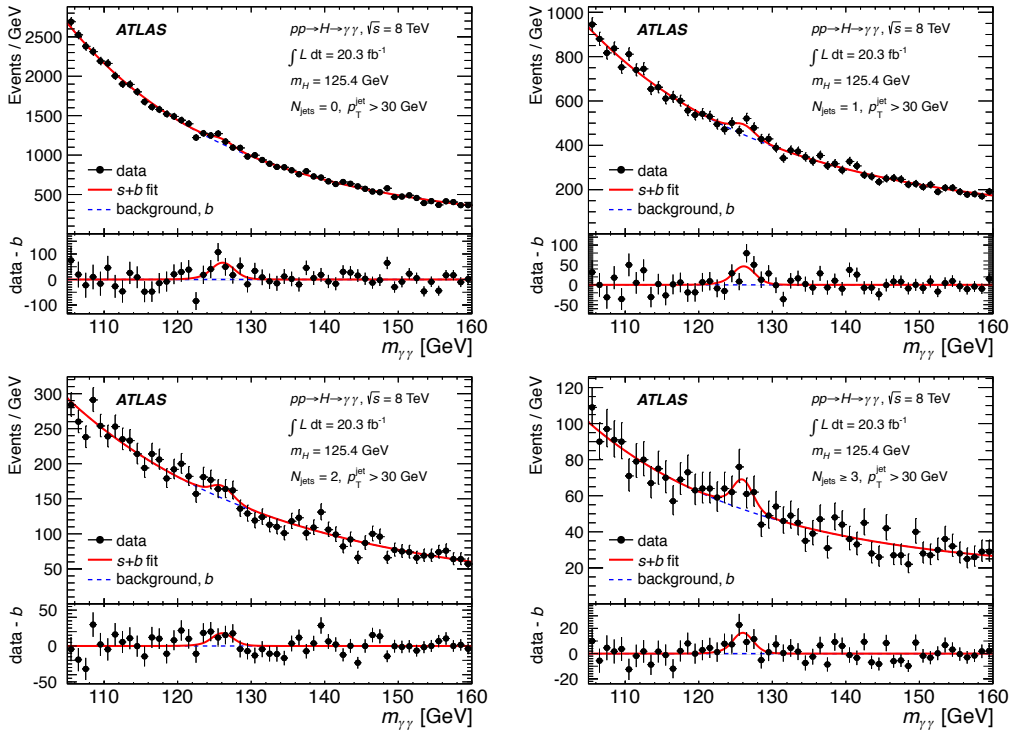


Figure 2. The diphoton invariant mass spectrum for four bins of jet multiplicity as described in the legend. The curves show the results of the single simultaneous fit to data for all multiplicity bins, where the Higgs boson mass is fixed to be $m_H = 125.4\,\text{GeV}$. The red line is the combined signal and background probability distribution functions, and the dashed line shows the background-only probability distribution function. The difference of the two curves is the extracted signal yield. The bottom inset displays the residuals of the data with respect to the fitted background component. [3]

4. The Differential and Fiducial Cross Sections

The Higgs boson event yields were measured in each bin of the differential distribution. The distributions include the jet activity, the Higgs boson kinematics, VBF sensitive variables and spin-CP variable. The variables are listed as below:

- The Higgs boson kinematics: $p^T_{\gamma\gamma}$, $|y_{\gamma\gamma}|$,
- Jet activity: Jet multiplicity N_{jets}, $p^{j_1}_T$, $|y_{j_1}|$, $p^{j_2}_T$, the scalar sum of jet transverse momenta, H_T.

- Spin-CP sensitive variables: $\Delta\phi_{jj}$, $|cos\theta^{\star}|$
- VBF-sensitive variables: $|\Delta\phi_{\gamma\gamma,jj}|$, $|\Delta y_{jj}|$

The jet multiplicity distribution is shown in the right column of Fig. 3. The signal yields were corrected for the effects of detector inefficiency and resolution. The correction were made using the bin-by-bin unfolding correction factors derived from MC samples. The results were compared with multiple theoretical predictions describing the Higgs boson+jets activities. The HRes 2.2 [7] [8] calculation was used to provide the inclusive kinematics of the diphoton system via gluon fusion. HRes is accurate to NNLO+NNLL in QCD. The BLPTW (soft-collinear effective theory, NNLO+NNLL 0-jet + NLO+NLL 1-jet cross sections), JetVHeto and MINLO [9] [10] [11] provided predictions for events associated with jets. A small distortion in the Higgs boson p_T distribution was observed by comparing the MC and the data, as shown in Fig. 3. One of the possible reason is that the Higgs boson was produced in association with the dark matter particles, which can cause the relatively higher missing transverse energy and the Higgs boson transverse momentum.

The fiducial cross section was also measured in the fiducial range $p_{T\gamma,leading(subleading)}/m_{\gamma\gamma} > 0.35(0.25)$ and $105\,\text{GeV} < m_{\gamma\gamma} < 160\,\text{GeV}$:

$$43.2 \pm 9.4(stat)^{+3.2}_{-2.9}(syst) \pm 1.2(lumi)\,\text{fb}$$

The results are shown in Fig. 4.

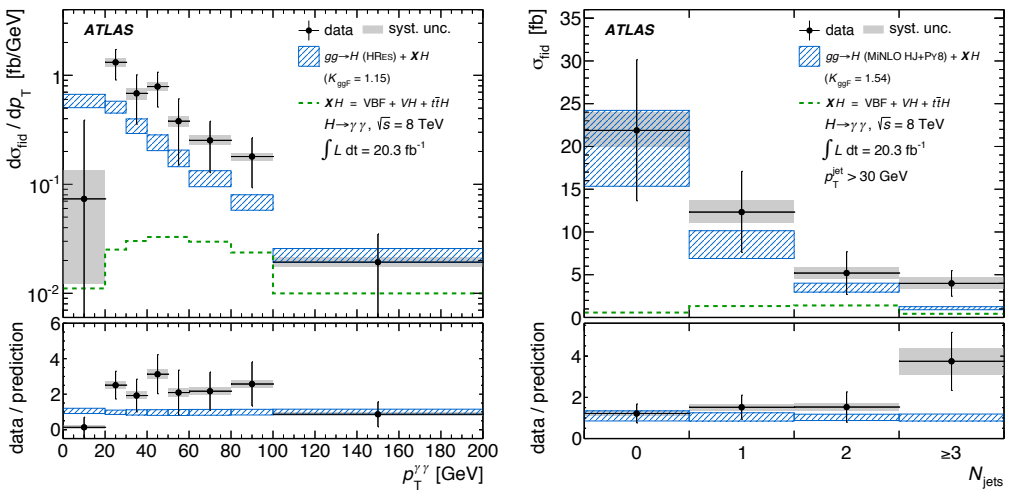


Figure 3. Left: the differential cross section of p_T^H measured in two photon final state. Right: the jet multiplicity distribution. The jets p_T threshold is 30 GeV [3]. The Hres calculation is normalised to the LHC-XS prediction using a K-factor K_{ggF}.

5. The Search for Dark Matter

The Higgs boson produced in association with dark matter candidates can explain the distortion in the diphoton p_T spectrum. Figure 1 shows a way of dark matter production, an mediator couples with two dark matter candidate and the Higgs boson. By setting up the mediator mass at 300 GeV and dark matter mass around 60 GeV, this process provides a moderate missing transverse energy (MET). The excess could be observed in both the Higgs boson p_T region and intermediate MET region. A data driven search is proposed to find the new physics. The search will be performed in two photon and MET phase space. Several categories will be used to catch

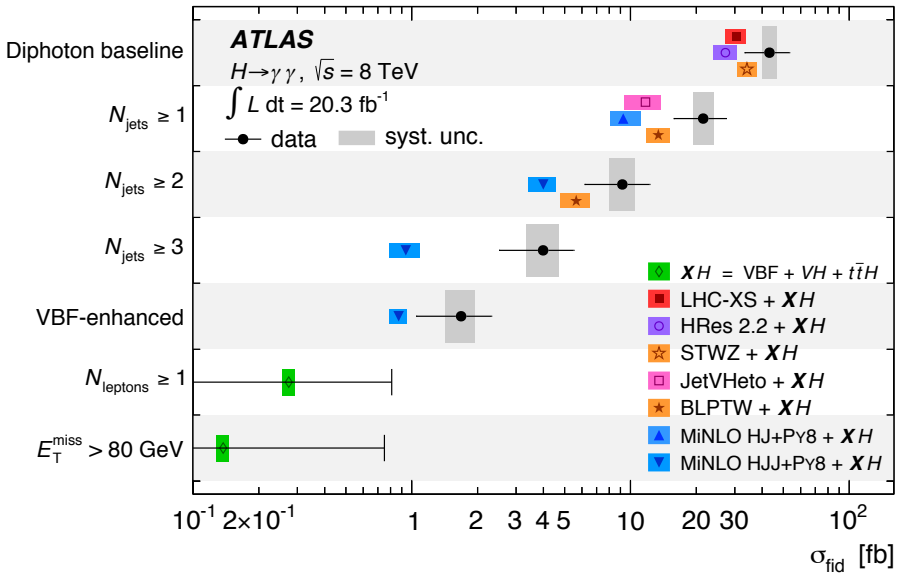


Figure 4. The measured cross sections and cross-section limits for $pp \to H$ in the seven fiducial regions. The intervals on the vertical axis each represent one of these fiducial regions. The data are shown as filled (black) circles. The error bar on each measured cross section represents the total uncertainty in the measurement, with the systematic uncertainty shown as dark grey rectangles. The error bar on each cross-section limit is shown at the 95% confidence level. The width of each theoretical prediction represents the total uncertainty in that prediction. All regions include the SM prediction arising from VBF, V H and ttH, which are collectively labelled as XH [3].

up the diphoton plus high, intermediate and low MET final states. The search will optimise the categories using 2011+2012 data and MC to find the best selection criteria through which the data has the largest excess with respect to the SM Higgs boson. The selection criteria will be applied in 2015 $\sqrt{s} = 13$ TeV, with the consideration on the changes brought from the higher centre mass energy and higher pile up contamination. Meanwhile, the MET performance of the SM Higgs boson, background and data is critical for the search. A track based MET definition will be used in 2011+2012 and 2015 data to suppress the dependence of pile up. If the excess does exist, it should be observed in other channels as well, such as ZZ to 4 leptons channel. However the diphoton selection is the most sensitive analysis in the early data.

6. Conclusions

The differential and fiducial cross section were measured after the recalibration of photon energy in ATLAS in 2014. The small excess from the Higgs boson p_T distribution could indicate a dark matter candidate produced in association with the Higgs boson. The data driven search for new physics on 2015 $\sqrt{s} = 13$ TeV data using 2011+2012 results will be performed.

References

[1] ATLAS Collaboration, Phys.Lett. B716 (2012) 1, arXiv:1207.7214[hep-ex]
[2] CMS Collaboration, Phys.Lett. B716 (2012) 30, arXiv:1207.7235 [hep-ex]
[3] ATLAS Collaboration, JHEP 1409 (2014) 112, arXiv:1407.4222 [hep-ex]
[4] ATLAS Collaboration, The ATLAS experiment at the CERN large hadron collider, J.Instrum., 3 (2008)

[5] S. Alioli, P. Nason, C. Oleari, and E. Re, JHEP 0904 (2009) 002, arXiv:0812.0578 [hep-ph]
[6] Sjostrand, Torbjrn et al. Comput.Phys.Commun. 191 (2015) 159-177 arXiv:1410.3012 [hep-ph]
[7] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini, JHEP 06 (2012) 132, [arXiv:1203.6321].
[8] M. Grazzini and H. Sargsyan, JHEP 09 (2013) 129, [arXiv:1306.4581].
[9] R. Boughezal, X. Liu, F. Petriello, F. J. Tackmann, and J. R. Walsh, Phys. Rev. D89 (2014) 074044, [arXiv:1312.4535].
[10] A. Banfi, P. F. Monni, G. P. Salam, and G. Zanderighi, Phys. Rev. Lett. 109 (2012) 202001,[arXiv:1206.4998]
[11] K. Hamilton, P. Nason, and G. Zanderighi, JHEP 10 (2012) 155, [arXiv:1206.3572].

Data processing at The South African Nuclear Energy Corporation SOC Ltd (Necsa) neutron diffraction facility

D Marais[1,2], A M Venter[1] and J Markgraaff[2]

[1] Research and Development Division, Necsa SOC Limited, PO Box 582, Pretoria, 0001, South Africa
[2] School of Mechanical and Nuclear Engineering, North-West University, Potchefstroom, 2520, South Africa

E-mail: deon.marais@necsa.co.za

Abstract. This paper describes the automated data treatment and reduction procedures developed for selected neutron diffraction instruments at Necsa as implemented in the software ScanManipulator. Examples of output data are illustrated for temperature dependent crystal structure investigations of an alumina powder, as well as depth-resolved d-spacing maps inside a welded mild steel sample.

1. Introduction

1.1. Neutron diffraction principle

Angular-dispersive neutron diffraction instruments utilize the principle that monochromatic thermal neutrons of wavelength λ are scattered coherently from the regular d_{hkl} spacing of crystalline planes in the material microstructure at angles θ_{hkl} given by the Bragg-law of diffraction, $2d_{hkl}sin\theta_{hkl} = \lambda$. Monochromatic thermal neutrons can be obtained from fission-spectrum neutrons produced in a steady-state nuclear reactor by placing neutron optical elements such as filters, monochromators and collimators in the neutron flight path. When directing a monochromatic neutron beam onto a crystalline sample, neutrons will be scattered as a distribution of Debye-Scherrer cones of varying apex angles centered on the beam axis. The apex angles can be precisely determined with neutron detectors positioned around the sample.

Efficient instrument performance is achieved through the use of neutron area detectors that enable intercepts with large sections of the Debye-Scherrer cones. This renders improved statistics and reduced measurement time. Ideally, the detector should be spherical and centered on the sample. This is however technically very challenging and financially taxing. As a compromise, arrays of flat detectors are generally used provided that relevant corrections are performed.

1.2. Data treatment and reduction software

A number of software packages are available which apply the necessary correction and reduction algorithms to raw diffraction data, GumTree [1] and Mantid [2] serving as examples. These packages tend to be large frameworks which cater for a broad spectrum of data treatment requirements, but with limited international standardization on the data acquisition and control (DAC) protocols between diffraction instruments, the software is generally not directly compatible with all instrument

modalities. This requires independent adaptation of existing third party software which could become very cumbersome.

1.3. The Necsa neutron diffraction facility(NDIFF)

The NDIFF at the SAFARI-1 research reactor of the South African Nuclear Energy Corporation (Necsa) SOC Limited has recently been upgraded with the aim to provide the African research community with neutron diffraction instrumentation of international stature. These instruments include a powder diffractometer named PITSI and a residual strain scanner named MPISI. With this upgrade, solutions were required to facilitate data acquisition, reduction and processing.

For DAC Necsa has standardized on the Australian Nuclear Science and Technology Organization (ANSTO) version of the SINQ Instrument Control Software (SICS) [3] on both instruments. Since MPISI is similar to ANSTO's Kowari neutron strain scanner, data reduction and processing can be performed with the Kowari customization of GumTree. However, since the PITSI detector geometry is not supported by GumTree, substantial modifications would be required to establish a compatible data reduction environment. With GumTree being a very expansive framework this implied a steep learning curve that had to be overcome before additions and modifications could be made. This is also true for other diffraction data processing frameworks such as Mantid. It was therefore decided to create the built-for-purpose software, ScanManipulator, to address instrument specific requirements.

Figure 1 schematically depicts the data treatment modalities implemented for PITSI to facilitate corrections, integration, normalization and stitching of data taken at a number of detector angles to render a final diffraction pattern. With the detector of PITSI only covering 20° in 2θ, a full diffraction pattern covering $9° \leq 2\theta \leq 116°$ in angular extent needs to be compiled from sequential positioning of the detector to at least seven different 2θ angles (frames) and allowing approximately 4° overlap between frames for the stitching process.

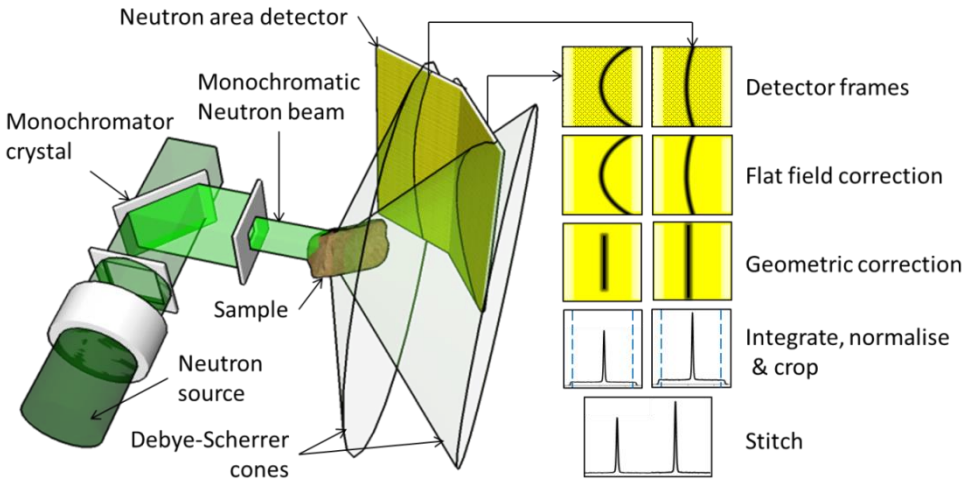


Figure 1. Illustration of an angular-dispersive neutron diffractometer geometry as well as the data correction and reduction steps required to produce a diffraction pattern from individual detector frames.

2. ScanManipulator

2.1. Implementation language.

ScanManipulator was implemented with the scripting language Python with the visual interface created with PyQt, making it a cross-platform application. It relies extensively on the scientific computing package Numpy for array manipulation and a number of third-party open-source libraries which provide the following functionalities:

- PySpec performs function fitting through regression techniques. It is used in analysis of X-ray diffraction data produced by the SPEC X-Ray Diffraction and Data Acquisition software [4].
- H5py access binary HDF5 datasets which is a primary storage mechanism of SICS [5].
- Matplotlib [6] is used to display 2D data and Mayavi [7] to display 3D data.

2.2. Data treatment and reduction functions

Although ScanManipulator was created to address PITSI's data treatment requirements, it also includes functionalities to be used with MPISI. In this respect it automatically applies corrections and function fitting either to existing data files, or in near real time to data acquired directly from the neutron detector client software. The following procedures are currently available in ScanManipulator:

2.2.1. Flat field correction. Due to various physical factors such as gas pressure, delay line resistivity variations and variable pre-amplifier gains, area detectors rarely have homogeneous sensitivity throughout its whole surface area. This can be mitigated by performing a 'flat field' correction where the detector image pixel intensities are normalized using intensity data ($a_{i,j}$) from measurements of an isotropic scatterer such as vanadium. The relative sensitivity, s, of each pixel (i,j) is calculated as $s_{i,j} = a_{i,j} / max(a)$. During physical measurements, the corrected intensity of each pixel is calculated by dividing the detected intensity with the pixel's relative sensitivity.

2.2.2. Geometric correction. A Debye-Scherrer cone at 90° diffraction angle ($2\theta_{hkl}$) intersects the surface of an area detector as a straight line. Cones at any other angle intercept the detector as a curved line where the diffraction angle is only correct on the horizontal plane of the incident beam. Randau *et al.* has derived equations to reassign each pixel to a new coordinate that 'straightens' the curved lines to their corrected 2θ angle [8].

2.2.3. Integration. After application of the corrections, the two-dimensional frame is integrated vertically to produce a horizontal (one-dimensional) peak profile along the detector length. Conversion from linear length to subtention angle θ is done through calibration. Conversion of diffraction angle to d-spacing is achieved using the Bragg law with incorporation of the monochromatic wavelength.

2.2.4. Normalization. To negate variations in the incident neutron flux during the prolonged times (hours) involved in measuring a complete diffraction pattern, intensity data are normalized with respect to a 'beam monitor count' that is directly proportional to the flux experienced by the sample. A beam monitor is a low efficiency neutron detector placed in the primary monochromatic beam path.

2.2.5. Data cropping. Area detectors are inefficient close to the perimeter of the detection surface which results in a sharp falloff in reliability. Data in this region are subsequently excluded from analysis by removing marginal sections of the integrated data which represent the left and right extremes of the neutron detector frame dataset.

2.2.6. Data stitching. As discussed, it is often unlikely to have a single large area detector covering the whole 2θ range of interest. The range can be covered by step scanning detector arrays around the instrument rotation axis to produce partially overlapped detector frames. The frames are then combined to create a single continuous diffraction pattern. This procedure involves calculating the 2θ value of each detector channel with reference to the absolute 2θ value of detector center. From this, a common 2θ axis is created which spans the total measured range and each detector frame is interpolated onto it. Regions of overlap are identified and the average values for these areas used.

2.2.7. Sample edge determination. Entry curve routines as described by Brand [9] together with additional 'z-scan' and 'wall-scan' functions are used to determine the sample's relative position with respect to the gauge volume, which is defined by the overlap of the primary and secondary apertures.

2.2.8. Peak fitting. Data from MPISI and PITSI are used to determine diffraction peak parameters, such as peak center, peak intensity, integrated intensity, full-width-half maximum, background, etc. with associated uncertainties by fitting a Gaussian function to the corrected peak data.

2.3. Input and Output

ScanManipulator is able to input a variety of data formats which include a subset of the NeXus format [10] encoded with HDF5, data files originating from the McStas neutron ray-tracing simulation package [11] as well as two columns of ASCII. A dedicated module retrieves positional neutron data directly from the DAC system. This functionality allows automated peak fitting in real time which enables measurements according to pre-defined statistical criteria, instead of time.

Complete diffraction patterns can be exported to a FullProf [12] compatible format allowing quantitative phase analysis through Rietveld-refinement [13]. Peak fitting parameters can be exported to Microsoft Excel for processing. ScanManipulator exhibits a number of windows to facilitate visual display of datasets. These graphs can be copied to the clipboard in a txt, png, emf or bmp format.

3. Examples of data treatment functionalities and representations

ScanManipulator facilitates combining different diffraction pattern data sets as function of an instrument variable. Three examples are presented to indicate typical functionalities.

3.1. Temperature dependent phase transformation study on PITSI

As an example of grouping and visualization of temperature dependent data, the high temperature phase transformations in an alumina (Al_2O_3) sample have been investigated at 11 different temperatures using neutrons with a wavelength of 1.759 Å. At each constant temperature, a full diffraction pattern range composed of 7 detector frames each was measured. The measurement scheme shown in Figure 2 was applied that resulted in 77 datasets.

Data processing comprised applying all corrections for each detector frame and stitching into a continuous diffraction pattern. By setting the appropriate grouping parameter, data could automatically be picked and combined as function of temperature, giving 11 temperature datasets each containing a full diffraction pattern.

The diffraction results of Figure 3 indicate that a phase transformation had occurred at around 700 °C from a cubic to a hexagonal phase. In order to better identify the transition temperature, a representative Bragg peak, shown in Figure 4, was fitted with a Gaussian function. The fit was automatically applied to all datasets (temperatures) to compose the graph of Figure 5.

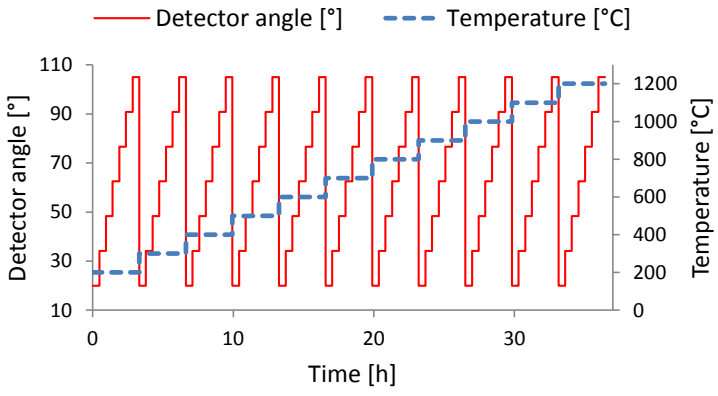


Figure 2. Graph showing the steps in detector angle and temperature versus time for a temperature dependant investigation.

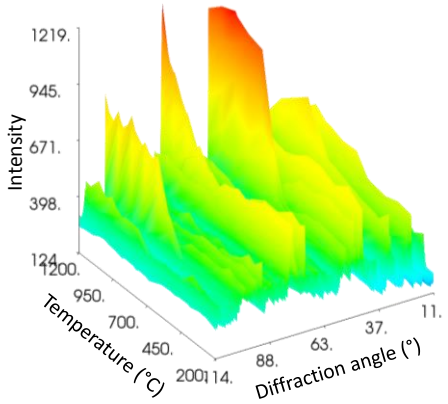


Figure 3. Waterfall plot showing diffraction pattern evolution as a function of temperature.

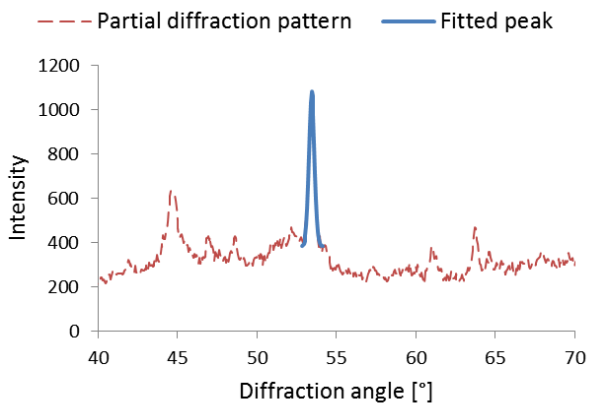


Figure 4. Diffraction pattern at 800 °C showing the peak at 2θ = 53.4° that was selected to follow its temperature dependence.

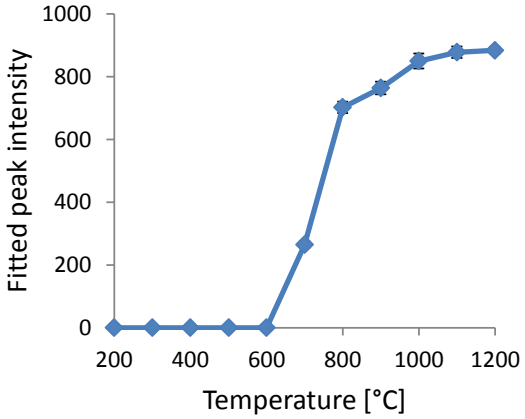


Figure 5. Graph showing the temperature dependence of the peak at 2θ = 53.4°.

3.2. Strain investigations on MPISI

Strain investigations of the interior regions of samples consist of measuring the depth and orientation dependence of a pre-selected diffraction peak. The peak angle is determined through Gaussian function fitting and is converted to d-spacing. Comparison of the d-values to an unstressed reference enables determination of the strain at that position and orientation. Figure 6 shows a typical sample setup on MPISI. By positioning the sample with respect to the gauge volume, cross-sections and even volumes of the sample can be investigated.

ScanManipulator performs the function fitting of multiple datasets automatically and exports results to Microsoft Excel where further calculations can be conducted. It can also be used as quick verification of multi-parameter trends for a specific strain direction setup as shown in Figure 7.

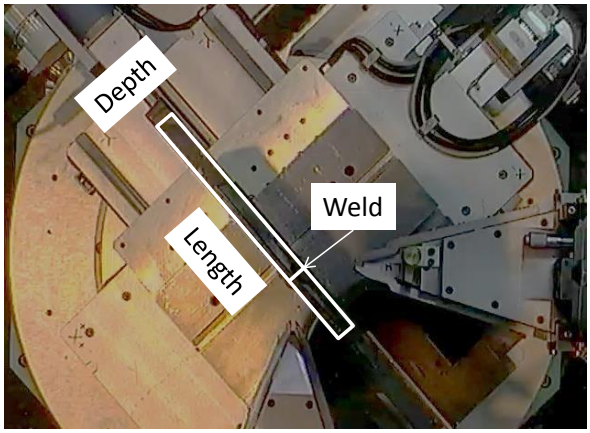


Figure 6. Photograph of a welded mild steel plate sample setup for measured on MPISI.

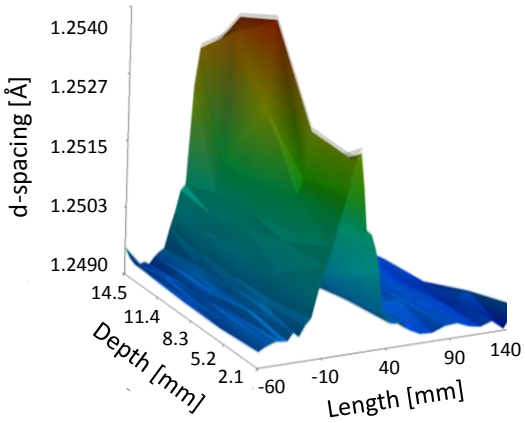


Figure 7. Graph showing the variation in (211) plane d-spacing in a welded mild steel sample as a function of multi-parameter positioning across the weld and plate depth.

3.3. Real-time peak fitting on MPISI

An instrument characterization investigation was conducted to determine the peak intensity count rate of the (211) peak of mild steel with respect to path length and gauge volume size. Figure 8 and Figure 9 summarize the multi-parameter results acquired using the real-time peak-fitting and evaluation functionality of ScanManipulator. Reaching of either of two conditions was specified, i.e. data acquisition to a maximum measurement time of 600s, or peak intensity of 40 counts. It can be seen that the required intensity could not be achieved for long path lengths with small gauge volumes within the allowed time. From the results it is evident that data acquisition against real-time evaluated set peak parameters is a viable process to be used in strain measurements.

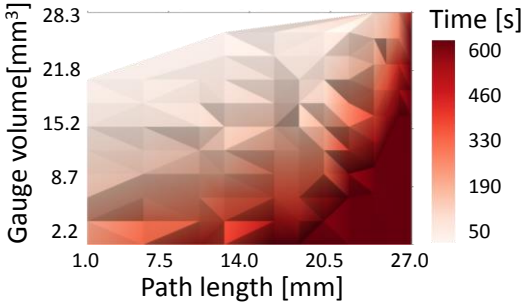


Figure 8. Surface plot showing measurement time as a function of gauge volume and path length.

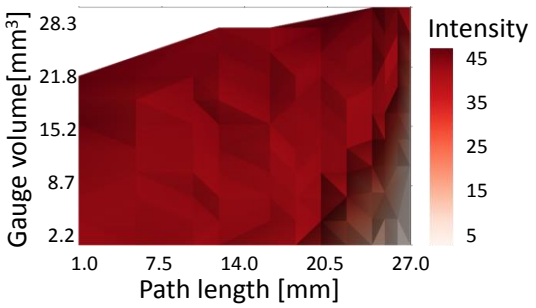


Figure 9. Surface plot showing fitted peak intensity as a function of gauge volume and path length.

4. Conclusion

A number of functionalities of ScanManipulator have been demonstrated. It provides a user friendly platform for users to automatically perform relevant corrections to neutron diffraction data originating from NDIFF instruments. Multi-parameter data can easily be displayed for visual interpretation and inspection.

Acknowledgements

This research is supported by the National Research Foundation of South Africa, grant number 84234, and Necsa SOC Limited. Any opinion, finding and conclusion or recommendation expressed in this material is that of the author(s) and the NRF does not accept any liability in this regard.

References

[1] Hugh Rayner H, Hathaway P, Hauser N, Fei Y, Franceschini F and Lam T 2006 *Physica* B **385-386** 1333-1335
[2] Arnold O *et al.* 2014 *Nucl Instrum Meth* A **764** 156-166
[3] Heer H, Könnecke M and Maden D 1998 *Physica* B **241-243** 124-126
[4] Wilkins S 2013 https://github.com/stuwilkins/pyspec Access date: 2014/05/06
[5] Collette A 2013 *Python and HDF5* (O'Reilly Media) pp 1-152
[6] Hunter J D 2007 *Comput Sci Eng* **9**(3) 90-95
[7] Ramachandran P and Varoquaux G 2011 *Comput Sci Eng* **13**(2) 40-51
[8] Randau C, Garbe U and Brokmeier H G 2011 *J. Appl. Cryst.* **44** 641-646
[9] Brand P C and Prask H J 1994 *J. Appl. Cryst* **27** 164-176
[10] Könnecke M *et al*. 2015 *J. Appl. Cryst.* **48** 301-305
[11] Lefmann K and Nielsen K 1999 *Neutron News* **10** 20
[12] Rodriguez-Carvajal J 1993 *Physica* B **192** 55-69
[13] Rietveld H M 1968 *J. Appl. Cryst.* **2** 65-71

Status of the measurements of Higgs boson properties with the ATLAS detector

Bruce Mellado[1]

School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africa

Abstract. The observation of a new particle consistent with a Higgs boson by the ATLAS and CMS experiments at the Large Hadron Collider (LHC) is now well established. With the approaching of the Run II data taking the analysis of Run I data is coming to a close. Measurements of Higgs boson properties with the ATLAS detector using Run I data are reviewed. This includes the measurement of the mass, compatibility of couplings and Spin/CP quantum numbers with the Standard Model, and the measurement of differential cross-sections. Long-term prospects of these measurements with the High-Luminosity LHC are also discussed. Prospects for Run II data taking are inferred.

1. Introduction

With the discovery of a new boson consistent with that predicted in the Standard Model (SM) by the ATLAS [1] and CMS [2] collaborations at the Large Hadron Collider (LHC), significant efforts have been made in the exploration of its properties. Here the status of these measurements is summarized, with emphasis on the results published by the ATLAS collaboration. Overall, both experiments display a consistent picture with regards to the newly discovered boson.

The exploration of the Higgs boson has branched out into a different set of relevant aspects. These include the measurements of the mass, spin/CP quantum numbers,[2] production cross-sections, and its overall compatibility with the SM. Not just for purposes of presentation, it is important to establish an order in which the results are presented. It is convenient put in place a logical succession of measurements consistent with the evolution from the most basic to the most complex. This ordering does not necessarily coincide with the order in which results are reported by the experiments.

In this light, the paper is organized as follows: Sections 2 and 3 summarize recent results of the mass and spin/CP quantum number measurements; Section 4 summarizes recent results of the total and differential cross-sections. Section 5 sums up measurements of the signal strenghts with respect to the SM. Section 6 gives a brief overview of future prospects, followed by Section 7, with summary and conclusions.

2. Mass measurement

The mass of a particle is its most basic property. The Higgs boson mass is a fundamental parameter of nature. The ATLAS and CMS collaborations have recently combined

[1] E-mail: Bruce.Mellado@wits.ac.za

[2] C stands for charge conjugation and P for parity, or space inversion.

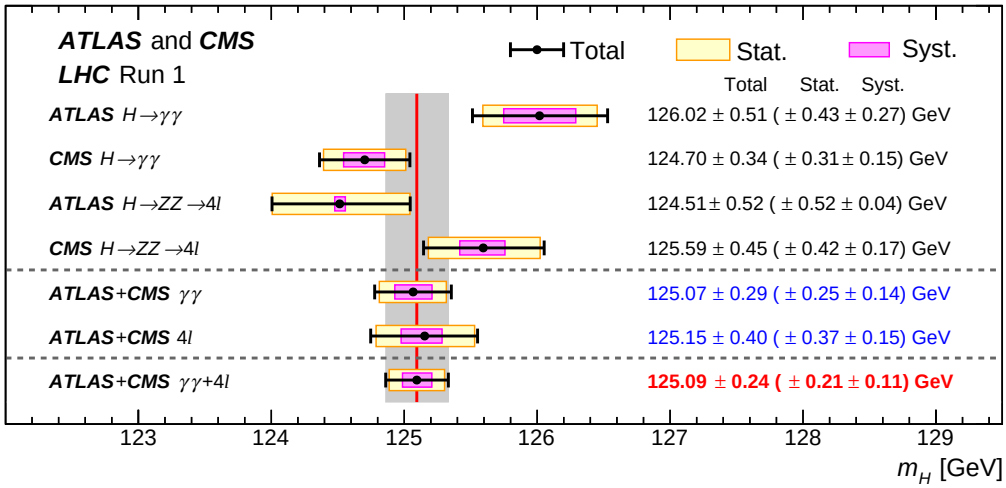


Figure 1. Summary of Higgs boson mass measurements from the individual analyses of ATLAS and CMS and from the combined analysis presented in Ref. [3]. The systematic (narrower, magenta-shaded bands), statistical (wider, yellow-shaded bands), and total (black error bars) uncertainties are indicated. The (red) vertical line and corresponding (gray) shaded column indicate the central value and the total uncertainty of the combined measurement, respectively.

measurements of the Higgs boson mass, m_H, using the most sensitive decay channels: the di-photon and $H \to ZZ^* \to 4\ell, \ell = e, \mu$ decays. With both these decays the experiments can reconstruct m_H with excellent resolution. The measurements with the di-photon and $H \to ZZ^* \to 4\ell$ channels are complementary in that the former displays the largest number of Higgs boson decay, where the latter exhibits an excellent signal-to-background ratio greater than one.

Figure 1 displays the results of the mass measurements using the two channels independently, their combination and the total combination. Results are obtained with integrated luminosities of approximately $5\,\text{fb}^{-1}$ at $\sqrt{s} = 7\,\text{TeV}$ and $20\,\text{fb}^{-1}$ at $\sqrt{s} = 8\,\text{TeV}$ of pp collision data. The measurement with the di-photon channel at ATLAS is somewhat higher than that obtained with the $H \to ZZ^* \to 4\ell$ channel. However, the discrepancy is not statistically compelling and the situation is reverse with the CMS results. The ATLAS and CMS mass measurements with the two channels are combined independently (see Fig. 1). The total combined measurement yields (in GeV) [3]:

$$m_H = 125.09 \pm 0.24 = 125.09 \pm 0.21(stat.) \pm 0.11(syst.), \quad (1)$$

where the total error is split into statistical and systematics. This measurement corresponds to a stunning accuracy of 0.2%. Because of the complexity of the measurement and its sensitivity to detector performance, it is relevant to elaborate on the main sources of systematics. Figure 2 summarizes the impact of the most important groups of systematics on the mass measurement uncertainty specified above. The mass measurement result presented here is obtained by means of a likelihood ratio that includes a number of nuisance parameters. The latter incorporate the different systematic uncertainties and are profiled. As a result the nuisance parameters obtained from the fits can somewhat differ from the expected value. The total systematic error is obtained by substracting the statistical error from the total error. The most important uncertainties are instrumental and pertain to the performance of the electromagnetic calorimeters.

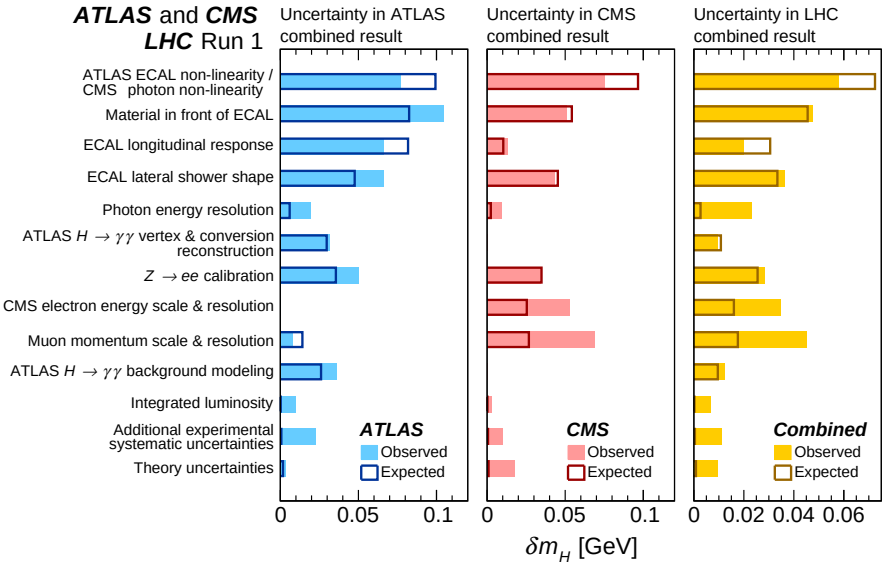


Figure 2. The impact of different groups of systematics on the ATLAS (left), CMS (center), and combined (right) mass measurement uncertainty from Ref. [3]. The observed (expected) results are shown by the solid (empty) bars (see text).

3. Spin/CP quantum numbers

The exploration of the Spin/CP quantum numbers of the newly discovered boson can be performed exploiting the properties of the decays [4] or the production mechanisms [5, 6]. These approaches are complementary. The ATLAS and CMS experiments have focused on the exploration of the decay products. This is done to understand the compability of the data with the SM hypothesis, 0^+ with respect to other hypotheses. The introduction of hypotheses other than that predicted in the SM leads to extra terms in the Lagrangian. In turn these terms generate differences in the kinematic distributions of decay products. In a recent study the ATLAS collaboration has performed a combined study of the decay kinematics of the di-photon, $H \to ZZ^* \to 4\ell$ and $H \to WW^* \to e\nu\mu\nu$ channels [7]. Here the SM is tested against two non-SM spin-0 and a simplified spin-2 hypotheses. The two non-SM spin-0 hypothesis correspond to the introduction of higher dimensional operators, including CP-conservation and CP-violation. The results exclude non-SM hypotheses considered at more than a 99% confidence level in favor of the SM. The CMS collaboration has made en extensive exploration of non-SM hypotheses with similar results in favor of the SM [8].

4. Total and differential cross-sections

The ATLAS collaboration has reported on the total and differential cross-section with the di-photon [9], the $H \to ZZ^* \to 4\ell$ [10] and their combination [11]. The results reported in Refs. [9, 10] are referred to as fiducial cross-sections. The rate of Higgs boson candidates are reported in a particular region of the phase-space defined by the decay products. Fiducial cross-sections are reported for the total phase-space explored and differentially, with respect to relevant observables (e.g. Higgs boson transverse momentum, p_T, rapidity, y, etc). Efforts are made not to bias the results with prior knowledge of Higgs boson production mechanisms in the SM. In Ref. [11] results are combined by extrapolating the fiducial cross-sections to the entire phase-space, including that where the measurement is not made and corrected by the branching fractions expected in the SM.

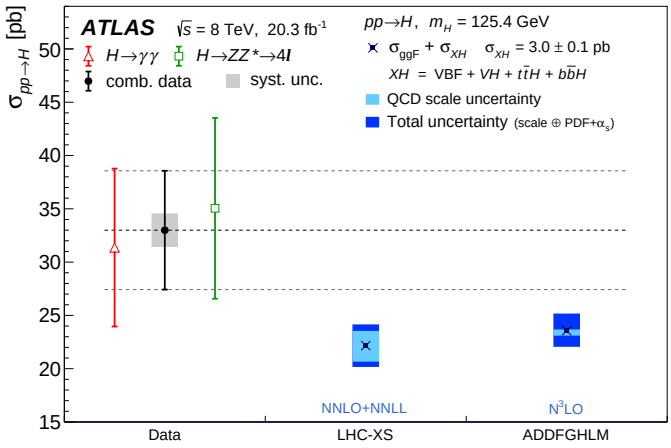


Figure 3. Measured total cross section of Higgs boson production compared to two calculations of the gluon-gluon fusion process cross section, from Ref. [11]. Contributions from other relevant Higgs boson production modes (VBF, VH, ttH, bbH) are added using cross sections and uncertainties.

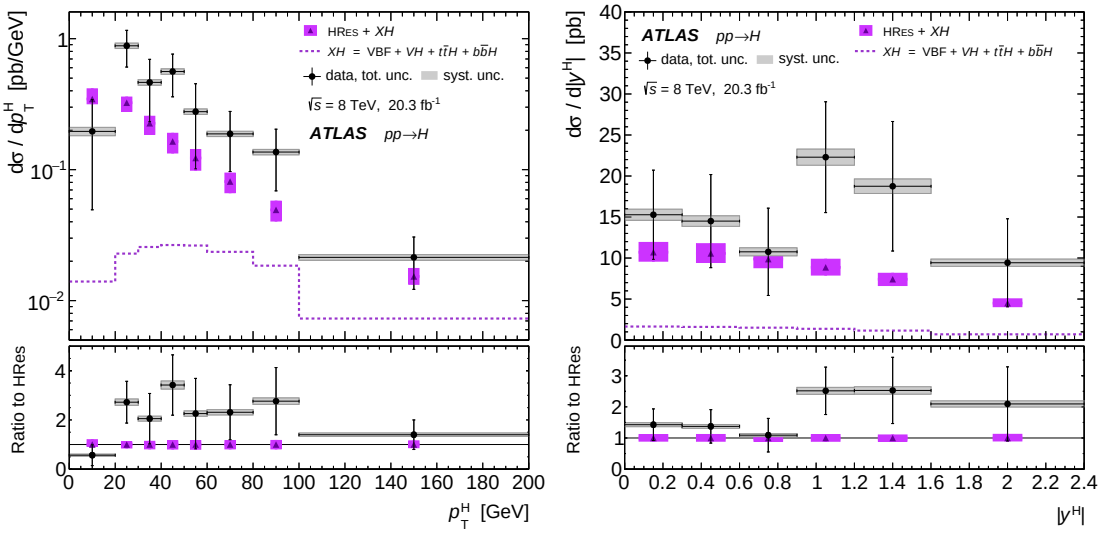


Figure 4. Differential cross-sections of the Higgs boson tranverse momentum (left) and rapidity (right) by combining the di-photon and $H \to ZZ^* \to 4\ell$ channels from Ref. [11].

Figure 3 displays the combination of the total cross-sections measured with the di-photon and the $H \to ZZ^* \to 4\ell$ channels. Results are compared to predictions existing at the time of submission to the journal. This includes an incomplete N^{3}LO cross-section for the gluon-gluon fusion production process. Since the manuscript was submitted significant progress has been made in QCD higher order corrections. The first complete gluon-gluon fusion cross-section at N^{3}LO has been made available [12]. In this calculation it is demonstrated that the correction with respect to the cross-section at NNLO evaluated at half the Higgs boson mass is about 2% and the scale uncertainty shrinks to 3%. Other production mechanisms, such as the Vector Boson Fusion, VBF, and associated production, $VH(V = Z, W)$, ttH and bbH are also taken into account.

Figure 4 displays the differential cross-sections with respect to the Higgs boson transverse

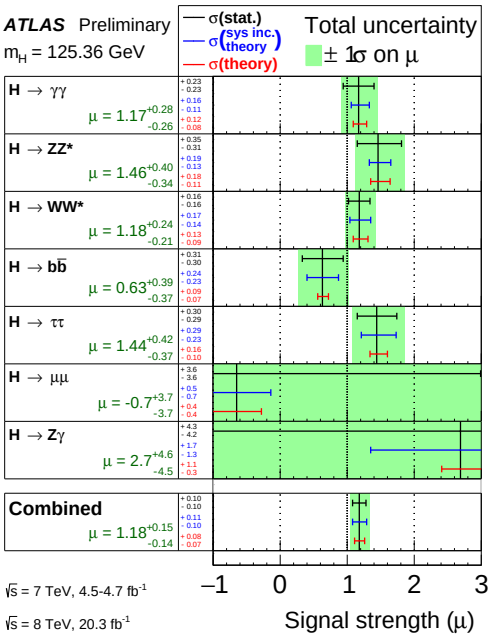


Figure 5. The observed signal strengths and uncertainties for different Higgs boson decay channels and their combination for $m_H = 125.36\,\text{GeV}$. Higgs boson signals corresponding to the same decay channel are combined together for all analyses. The best-fit values are shown by the solid vertical lines. The total $\pm 1\sigma$ uncertainties are indicated by green shaded bands, with the individual contributions from the statistical uncertainty (top), the total (experimental and theoretical) systematic uncertainty (middle), and the theory systematic uncertainty (bottom) on the signal strength shown as horizontal error bars. Results are from Ref. [15].

momentum and rapidity. These results are the combination of measurements obtained with the di-photon and $H \to ZZ^* \to 4\ell$ channels. Results are compared calculations that do not include recent progress mentioned above, nor recent calculations of gluon-gluon fusion in association with one jet at NNLO [13]. These calculations signify very important progress in QCD higher order corrections that give us renewed condifence in the robustness of theoretical predictions.

In Ref. [14] it has been pointed out that the implementation of these QCD higher order corrections do not impact significantly the compatibility of the data with the SM. Other speakers in the NPR track will discuss these results from the stand point of possible physics beyond the SM. Overall, while the level of discrepancy with the SM is not strong, the study of these observables with the new data becomes very interesting.

5. Signal strength and the Standard Model

The ATLAS collaboration has recently combined results of Higgs boson yields and constraints using the $H \to \gamma\gamma, ZZ^*, WW^*, Z\gamma, b\bar{b}, \tau\tau, \mu\mu$ decays [15]. This is a global analysis in which results optimized to observe the SM Higgs boson signal are combined in a comprehensive analysis that includes variations of SM coupling strengths. The results of the individual signal strengths, or the ratio of the observed yield to that predicted by the SM and their combination are shown in Fig. 5, in terms of μ, or the observed signal yield with respect to that predicted by the SM. The combined result is [15]:

$$\mu = 1.18 \pm 0.10(stat.) \pm 0.07(syst.) \pm 0.08(theor.), \tag{2}$$

where the errors are statistical, systematics and theoretical, respectively. Strong evidence of the Vector Boson Fusion process and coupling to down-type fermions are seen with significances of $4.3\,\sigma$ and $4.5\,\sigma$, respectively. Variations of SM-like coupling strenghts lead to conclude that, overall, the data is consistent with the SM.

It is very important to note that the results discussed here are qualitatively different from those reported in section 4. Where in the latter measurements are produced without prior assumptions on production mechanisms, results reported here have predictions from the SM embedded into them. The signal strenghts are in most of the cases obtained after the application of categories. The phase-space is weighted according to how the SM Higgs boson would be produced. As a result, a certain bias towards the SM is generated.

6. Future prospects

The LHC has made strides to provide proton-proton collisions at the energy and instantaneous luminosity frontier. The first collisions at 13 TeV have been provided and recorded by the experiments. This milestone will be followed by others. Among these is the transition from 50 ns bunch spacing to 25 ns with which to reach a record instantaneous luminosity of $1.3 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The integrated luminosity for proton-proton collisions by the end of the year is expected to be about $10\,\mathrm{fb}^{-1}$. With the total Higgs boson cross-section expected to increase by a factor of about 2.5, comparable sensitivity will be reached with respect to that achieved in Run I. By the end of Run II, sometime in 2018, about 100-120 fb^{-1} is expected to be delivered, leading to a factor of 10 more Higgs boson candidates compared to Run I. The LHC envisions to deliver $300\,\mathrm{fb}^{-1}$ by 2023, followed by a long shutdown of about 2.5 years to pursue the high luminosity upgrade. The ultimate peak luminosity is $5.7 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and $4000\,\mathrm{fb}^{-1}$ of integrated luminosity by the end of the LHC's lifetime. Coupling determination may reach accuracies of 5%-10% per experiment, depending on the coupling.

7. Summary and Conclusions

Strong progress has been made by the ATLAS and CMS collaborations in determining properties of the newly discovered scalar boson. This includes the measurement of the mass with a 0.2% accuracy, the exploration of the spin/CP quantum numbers, measurement of fiducial total and differential cross-sections, and the measurement of signal strenghts, with the exploration of couplings. Run II will be critical in elucidating if physics beyond the SM is hidden in the scalar sector.

References

[1] ATLAS Collaboration (G. Aad *et al.*) 2012 *Phys.Lett.* **B716** 1
[2] CMS Collaboration (S. Chatrchyan *et al.*) 2012 *Phys.Lett.* **B716** 30
[3] ATLAS and CMS Collaborations (G. Aad *et al.*) 2015 Phys. Rev. Lett. **114** 191803
[4] Dell Aquila J R and Nelson C A 1986 Phys. Rev. **D33** 80
[5] Englert C, Goncalves-Netto D, Mawatari K and Plehn T 2013 JHEP **1301** 148
[6] Djouadi A, Godbole R, Mellado B and Mohan K 2013 Phys. Lett. **B723** 307
[7] ATLAS Collaboration (G. Aad *et al.*) 2015 ATLAS-CONF-2015-008
[8] CMS Collaboration (V. Khachatryan *et al.*) 2014 arXiv:1411.3441 [hep-ex]
[9] ATLAS Collaboration (G. Aad *et al.*) 2014 JHEP **1409** 112
[10] ATLAS Collaboration (G. Aad *et al.*) 2014 Phys. Lett. B **738** 234
[11] ATLAS Collaboration (G. Aad *et al.*) 2015 Phys.Rev.Lett. **115** no.9, 091801
[12] Anastasiou C, Duhr C, Dulat F, Herzog F and B. Mistlberger 2015 Phys. Rev. Lett. **114** no. 21 212001
[13] Boughezal R, Caola F, Melnikov K, Petriello F and Schulze M (2015) arXiv:1504.07922 [hep-ph]
[14] von Buddenbrock S, Chakrabarty N, Cornell A S, Kar D, Kumar M, Mandal T, Mellado B, Mukhopadhyaya B and Reed R (2015) arXiv:1506.00612 [hep-ph]
[15] ATLAS Collaboration (G. Aad *et al.*) 2015 ATLAS-CONF-2015-007

Probing new physics in the Higgs sector with effective field theories at the Large Hadron Collider

Bruce Mellado[1], **Luis March**[2], **Xifeng Ruan**[3]
School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africaa

Abstract. With the discovery of a particle consistent with a Higgs boson a new window of opportunity for searches for new physics opens up. These can be performed via precision physics or direct searches. The ATLAS and CMS experiments at the Large Hadron Collider study the Higgs boson couplings to other particles assuming that their structure is the same as that predicted in the Standard Model. With the addition of new physics via higher dimensional operators, in the framework of an effective field theory, the structure of these couplings changes. The implications on the Higgs boson production rates and the differential cross-sections are discussed. Prospects for the sensitivity of the ATLAS and CMS experiments to these higher dimensional operators are also discussed. This includes measurements both at the level of decay and production.

1. Introduction

At present, all the data obtained from the many experiments in particle physics are in agreement with the Standard Model (SM). In the SM there is one particle, the Higgs boson, that is responsible for giving masses to all the elementary particles [1, 2, 3, 4]. In July 2012 the ATLAS and CMS experiments at the Large Hadron Collider (LHC) reported the discovery of a boson, a Higgs-like particle with a mass $m_H \approx 125$ GeV based on the data accumulated during 2011 and a part of 2012 period [5, 6].

The experiments at the LHC have not revealed any definitive direct signature of new physics so far.[4] However, one is led to suspect that such physics should affect the interaction Lagrangian of the Higgs boson. This generates, for example, effective operators of dimension-6 contributing to interactions of the Higgs and the electro-weak boson fields, HVV , with $V = W, Z, \gamma$. Probing such effective couplings for the recently discovered scalar is therefore tantamount to opening a gateway to fundamental physics just beyond our present reach.

Such 'effective' interaction terms need to be $SU(2) \times U(1)$ invariant if they arise from physics above the electroweak scale. Constraints on such terms have already been studied, using precision electroweak data as well as global fits of the current Higgs boson data (see Ref. [7] and references therein). Many studies have also considered anomalous Higgs boson couplings in the context of future e^+e^- and ep colliders.

[1] E-mail: Bruce.Mellado@wits.ac.za
[2] E-mail: Luis.March.Ruiz@cern.ch
[3] E-mail: Xifeng.Ruan@cern.ch

[4] It is important to note that a number of anomalies in the Higgs boson data collected so far are seen. These will be summarized by other speakers. Whether these anomalies are statistical fluctuations or are connected with new physics beyond the SM becomes a very interest prospect for the Run II data taking.

The general conclusion, based on analyses of the 8 TeV data at the LHC, is that several (though not all) of the gauge invariant, dimension-6 HVV terms can at most have coefficients ~ 5 TeV^{-2}. It still remains to be seen whether such small coefficients can be discerned with some ingeniously constructed kinematic distributions. Some work has nonetheless been done to study such distributions, in terms of either the gauge invariant operators themselves or the structures finally ensuing from them [8, 9]. At the same time, it is of interest to see if meaningful constraints do arise from the study of total rates at the LHC.

This paper is organized as follows: Section 2 briefly summarizes the effective field theory approach, Section 3 defines convenient observables based on ratios and Section 4 discusses the results. The paper is concluded with a summary in Section 5.

2. Effective field theory: Higher dimensional operators

In order to see any possible deviations from the SM in the Higgs boson sector, we will follow an effective field theory (EFT) framework. We consider $SU(2)_L \times U(1)_Y$ invariant operators of dimension up to 6, which affect Higgs boson couplings to itself and/or a pair of electroweak vector bosons. We have concentrated here on dimension-6 (D6) CP-conserving operators which affect Higgs boson phenomenology. They include [7]:

- Operators which contain the Higgs doublet Φ and its derivatives:

$$\mathcal{O}_{\Phi,1} = (D_\mu\Phi)^\dagger\Phi\Phi^\dagger(D^\mu\Phi); \quad \mathcal{O}_{\Phi,2} = \frac{1}{2}\partial_\mu(\Phi^\dagger\Phi)\partial^\mu(\Phi^\dagger\Phi); \quad \mathcal{O}_{\Phi,3} = \frac{1}{3}(\Phi^\dagger\Phi)^3 \tag{1}$$

- Those containing Φ (or its derivatives) and the bosonic field strengths:

$$\mathcal{O}_{GG} = \Phi^\dagger\Phi G^a_{\mu\nu}G^{a\,\mu\nu}; \quad \mathcal{O}_{BW} = \Phi^\dagger\hat{B}_{\mu\nu}\hat{W}^{\mu\nu}\Phi; \quad \mathcal{O}_{WW} = \Phi^\dagger\hat{W}_{\mu\nu}\hat{W}^{\mu\nu}\Phi;$$

$$\mathcal{O}_W = (D_\mu\Phi)^\dagger\hat{W}^{\mu\nu}(D_\nu\Phi); \quad \mathcal{O}_{BB} = \Phi^\dagger\hat{B}_{\mu\nu}\hat{B}^{\mu\nu}\Phi; \quad \mathcal{O}_B = (D_\mu\Phi)^\dagger\hat{B}^{\mu\nu}(D_\nu\Phi), \tag{2}$$

where

$$\hat{W}^{\mu\nu} = i\,\frac{g}{2}\sigma_a W^{a\,\mu\nu}; \quad \hat{B}^{\mu\nu} = i\,\frac{g'}{2}B^{\mu\nu}$$

and g, g' are, respectively, the $SU(2)_L$ and $U(1)_Y$ gauge couplings. $W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - g\epsilon^{abc}W^b_\mu W^c_\nu$, $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ and $G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f^{abc}G^b_\mu G^c_\nu$. The covariant derivative of Φ is given as $D_\mu\Phi = (\partial_\mu + \frac{i}{2}g'B_\mu + ig\frac{\sigma_a}{2}W^a_\mu)\Phi$. The Lagrangian, in the presence, of the above operators can be generally expressed as:

$$\mathcal{L} \supset \kappa\left(\frac{2m_W^2}{v}HW^+_\mu W^{\mu-} + \frac{m_Z^2}{v}HZ_\mu Z^\mu\right) + \sum_i \frac{f_i}{\Lambda^2}\mathcal{O}_i, \tag{3}$$

where, in addition to the D6 operators, we also allow the SM-like HWW and HZZ couplings to be scaled by a factor κ. While $\kappa \neq 1$ is indicative of certain kinds of new physics, we are especially interested in the new observable features associated with the higher dimension operators (HDOs). Therefore, we have set $\kappa = 1$ for simplicity.

No operator of the form $\mathcal{O}_{GG}$ is assumed to exist since we are presently concerned with Higgs boson interactions with a pair of electroweak vector bosons only. The operator $\mathcal{O}_{\Phi,1}$ is severely constrained by the T-parameter (or equivalently the ρ parameter), as it alters the HZZ and HWW couplings by unequal multiplicative factors. As far as HZZ and HWW interactions are concerned, $\mathcal{O}_{\Phi,2}$ only scales the SM couplings, without bringing in any new Lorentz structure. It also alters the Higgs boson self-coupling, something that is the sole consequence of $\mathcal{O}_{\Phi,3}$ as well.

In view of the above, we focus on the four operators $\mathcal{O}_{WW}$, $\mathcal{O}_{BB}$, $\mathcal{O}_W$ and $\mathcal{O}_B$. We do not include the operator $\mathcal{O}_{BW} = \Phi^\dagger \hat{B}_{\mu\nu} \hat{W}^{\mu\nu} \Phi$ in the present analysis, because it mixes the Z and γ fields at tree level, violates custodial symmetry (by contributing only to the Z-boson mass) and is, therefore, highly constrained by the S and T-parameters at the tree level.

The already existing limits on the various operators discussed above are found in numerous references. Even within their current limits, some of the operators are found to modify the efficiencies of the various kinematic cuts. The question we address in the rest of the paper is: can these limits be improved in the next run(s) of the LHC through careful measurement of the ratios of total rates in different channels? As we shall see below, the answer is in the affirmative.

3. Ratios of cross-sections as chosen observables

The four HDOs under consideration affect Higgs boson production as well as its decays, albeit to various degress. For example, HDO-dependent single Higgs boson production processes are in association with vector bosons (VH) *i.e.* $pp \to VH$ (where $V = \{W, Z\}$) and vector-boson fusion (VBF). We show the production cross-sections in these channels at 14 TeV in Fig. 1, as functions of the four operator coefficients (f_i) taken one at a time.[5] The relevant decay channels which are dependent on such operators are $H \to WW^*, ZZ^*, \gamma\gamma, Z\gamma$. Figure 2 contains these branching ratios (BR) as functions of the four coefficients under consideration.

The VBF and VH rates are sensitive to f_{WW} and f_W, but depend very weakly on f_{BB} and f_B, while the cross-section $\sigma(pp \to WH)$, is completely independent of f_{BB} and f_B. The HDO effects in $H \to \gamma\gamma$ and $H \to Z\gamma$ for $f_i \sim \mathcal{O}(1)$ is of the same order as the loop-induced SM contribution, unlike in the case of the HWW and HZZ couplings. Therefore, $\text{BR}_{H\to\gamma\gamma}$ becomes highly sensitive to f_{WW} and f_{BB}. Consequently, the 7+8 TeV data already restrict their magnitudes to small values of the order of $\lesssim 5$ TeV^{-2}. The limits on f_W and f_B, however, are relatively weaker, even after simultaneous imposition of constraints from electroweak precision data and LHC results.

Based on the above information, we set out to find observables which are sensitive to $f_i \lesssim 5$ TeV^{-2} in the high luminosity runs at the LHC. It is not completely clear yet how much statistics is required to probe such small values with various event shape variables. On the other hand, the more straightforward observables, namely, total rates in various channels, are also fraught with statistical, systematic and theoretical uncertainties which must be reduced as far as possible when precision is at a premium.

An approach that is helpful is looking at ratios of cross-sections in different channels. In this paper, we invoke two kinds of ratios. First, we take ratios of events in two different final states arising from a Higgs boson produced via the same channel (in our case, gluon fusion). Such a ratio enables one to get rid of correlated theoretical uncertainties such as those in PDF and renormalization/factorization scales. They also cancel the uncertainty in total width, which is correlated in the calculation of BRs into the two final states. Secondly, we consider the ratio of rates for the same final state for two different production channels (such as VBF and VH). Although the uncertainty in the BR cancels here, the theoretical uncertainties at the production level do not. Moreover, since the final state is the same in this case, some systematic uncertainties which are correlated (related to identification, isolation, trigger etc.) will also cancel. However, this is helpful in another manner. For some of the operators, the f_i-dependent shifts with respect to the SM are in the opposite direction for the numerator and the denominator in such ratios. We shall see that the use of both these kinds of ratios, including those involving the channel $Z\gamma$, can capture the HDO coefficients at a level unprecedented, going down to values where new physics can show up.

[5] We have used CTEQ6L1 parton distribution functions (PDFs) by setting the factorization (μ_F) and renormalization scales (μ_R) at the Higgs boson mass ($m_H = 125$ GeV).

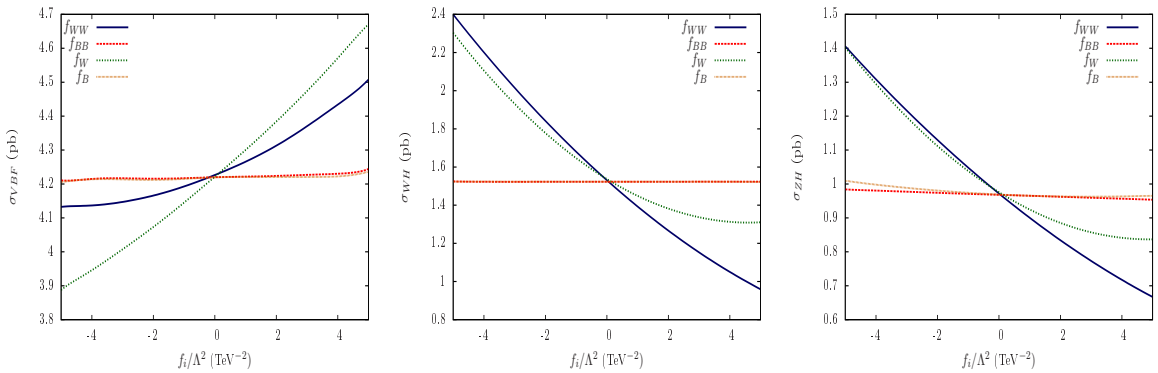


Figure 1. Higgs boson production cross-sections for the VBF and VH channels in the presence of HDOs at 14 TeV. Here the operators are varied one at a time.

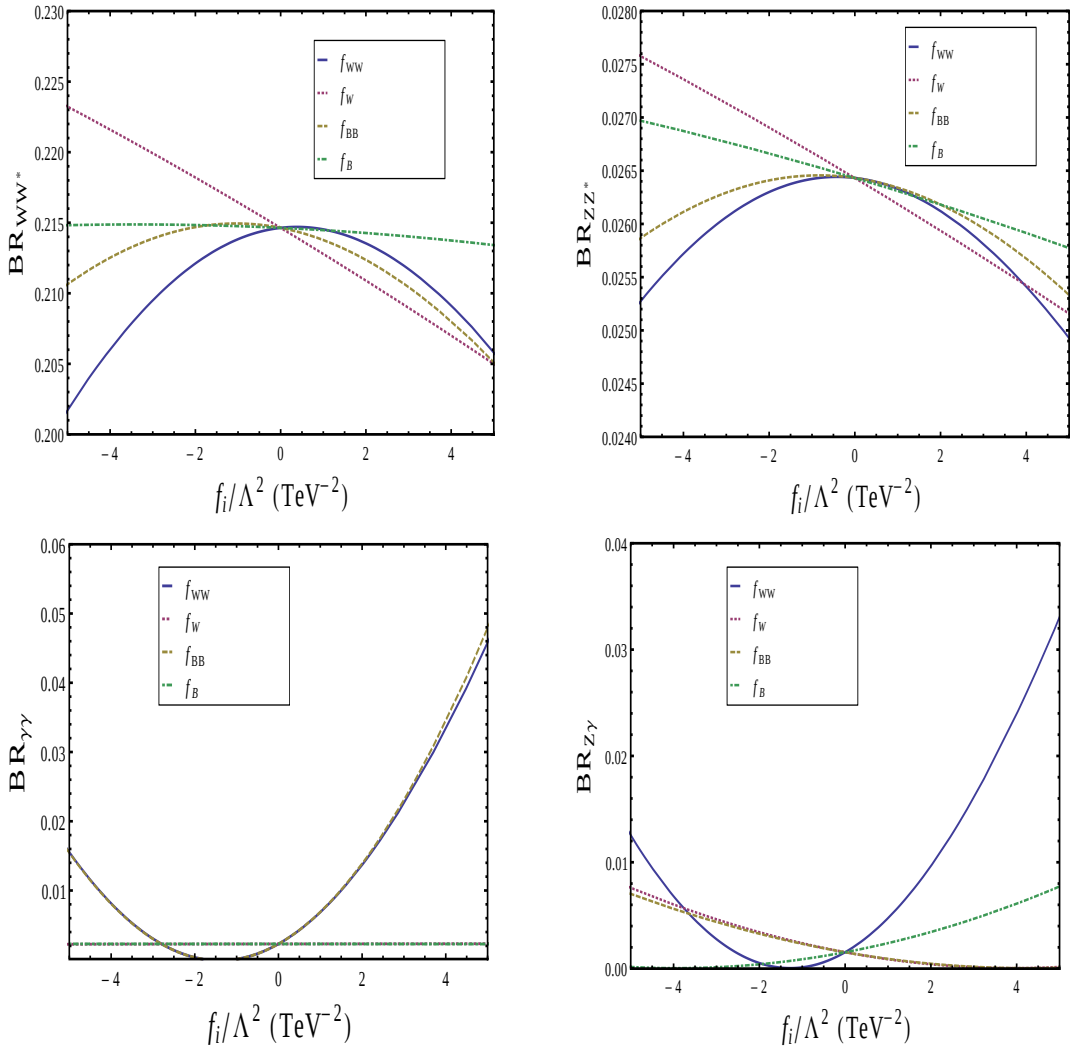


Figure 2. Branching ratios of $H \to WW^*, ZZ^*, \gamma\gamma, Z\gamma$ in presence of HDOs. The operators are varied one at a time.

Observables	$\mathcal{O}_{WW}$	$\mathcal{O}_{BB}$	$\mathcal{O}_{W}$	$\mathcal{O}_{B}$
$\mathcal{R}_1$ @ 7+8 TeV	$[-3.32,-2.91]$ $\cup$ $[+0.12,+0.57]$	$[-3.32,-2.91]$ $\cup$ $[+0.12,+0.57]$	Not bounded	Not bounded
$\mathcal{R}_1$ @ 14 TeV	$[-2.75,-2.66]$ $\cup$ $[-0.06,+0.03]$	$[-2.75,-2.66]$ $\cup$ $[-0.06,+0.03]$	Not bounded	Not bounded
$\mathcal{R}_2$ @ 14 TeV	$[-1.32,+1.21]$	Not bounded	$[-1.51,+1.72]$	Not bounded
$\mathcal{R}_3$ @ 14 TeV	Not used	Not used	Not used	$[-8.04,-7.63]$ $\cup$ $[-0.21,+0.17]$

Table 1. We summarize our obtained allowed region of the coefficients of HDOs using the three observables.

In the light of what is discussed above, we construct the following observables based on ratios [7]:

$$\mathcal{R}_1(f_i) = \frac{\sigma_{ggF} \times BR_{H\to\gamma\gamma}(f_i)}{\sigma_{ggF} \times BR_{H\to WW^*\to 2\ell 2\nu}(f_i)}, \tag{4}$$

where $\ell = e, \mu$ and f_i's are the operator coefficients;

$$\mathcal{R}_2(f_i) = \frac{\sigma_{VBF}(f_i) \times BR_{H\to\gamma\gamma}(f_i)}{\sigma_{WH}(f_i) \times BR_{H\to\gamma\gamma}(f_i) \times BR_{W\to\ell\nu}}, \tag{5}$$

and

$$\mathcal{R}_3(f_i) = \frac{\sigma_{ggF} \times BR_{H\to Z\gamma\to 2\ell\gamma}(f_i)}{\sigma_{ggF} \times BR_{H\to WW^*\to 2\ell 2\nu}(f_i)}. \tag{6}$$

Equations 4, 5 and 6 are sensitive to the operators, $\mathcal{O}_{WW}$ and $\mathcal{O}_{BB}$, $\mathcal{O}_{WW}$ and $\mathcal{O}_{W}$, and $\mathcal{O}_{B}$, respectively.

4. Results of the analysis

For our subsequent collider analysis, the chain we have used is as follows - first we have implemented the relevant dimension-6 interaction terms, as shown in Section 2, in FeynRules, and generated the Universal FeynRules Output (UFO) model files. These UFO model files have been used in the Monte-Carlo (MC) event generator MadGraph [10] to generate event samples. Next, the parton-showering and hadronization are performed using Pythia6 [11], and finally detector level analysis is carried using Delphes [12].

In Table 1 we summarize our obtained region of the parameter space allowed using three ratios, $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$. We present $\mathcal{R}_1$ using combined ATLAS+CMS data for 7+8 TeV runs. We also present a projected study for all three observables at 14 TeV with an integrated luminosity of 3000 fb^{-1}. The allowed regions on f_{WW} and f_{BB} are reduced at the 14 TeV 3000 fb^{-1} run as compared to the current data. Using the ratio $\mathcal{R}_2$ one can also put bounds on f_{WW} and f_W. There is a 'special' region of parameter space where the $\mathcal{R}_1$ mimics the SM expectation, therefore, $\mathcal{R}_2$ can also be used to infer the presence of $\mathcal{O}_{WW}$ with 'special' values of coefficient f_{WW}. The operator $\mathcal{O}_B$ does not show any aprreciable sensitivity in any production of the Higgs boson or its decay except in the $BR_{H\to Z\gamma}$. Therefore, the ratio $\mathcal{R}_3$ is constructed to constrain f_B by a significant amount, as evident from Table 1.

5. Summary and Conclusions

We have investigated how well one can constrain dimension-6 gauge-invariant operators inducing anomalous HVV interactions. Probing the gauge invariant operators individually, we feel, are important, since they can point at any new physics above the electroweak symmetry breaking scale. While the operators contributing to $H \to \gamma\gamma$ are subjected to the hitherto strongest limits using the (7+8) TeV data, the remaining ones are relatively loosely constrained, in spite of the bounds coming from precision electroweak observables. At any rate, it is necessary to reduce uncertainties as much as possible, since any realistically conceived new physics is likely to generate such operators with coefficients no greater than $\approx \mathcal{O}(1)$ TeV^{-2}. We show that a good opportunity to probe them at this level, and improve spectacularly over the existing constraints, arises if event ratios in various channels are carefully studied. These include both ratios of events in different final states with the same Higgs boson production channel and those where a Higgs boson produced by different production modes ends up decaying into the same final state. While a majority of the theoretical uncertainties cancel in the former category, the latter allows us to probe those cases where some dimension-6 operators shift the rates in the numerator and the denominator in opposite directions. We find that, after a thorough consideration of all uncertainties, all the couplings can be pinned down to intervals of width $\approx \mathcal{O}(1)$ TeV^{-2} using 3000 fb^{-1} of integrated luminosity at 14 TeV. Even with 300 fb^{-1}, the improvement over existing constraints is clearly expected, and the results are more uncertainty-free than in any other hitherto applied method. However, we must mention here that this approach should be complemented with the study of differential distributions, which is not within the scope of this paper.

References

[1] Englert F and Brout R 1964 *Phys.Rev.Lett.* **13** 321
[2] Higgs P 1964 *Phys.Lett.* **12** 132
[3] Higgs P 1964 *Phys.Rev.Lett.* **13** 508
[4] Guralnik G, Hagen C and Kibble T 1964 *Phys.Rev.Lett.* **13** 585
[5] ATLAS Collaboration (G. Aad *et al.*) 2012 *Phys.Lett.* **B716** 1
[6] CMS Collaboration (S. Chatrchyan *et al.*) 2012 *Phys.Lett.* **B716** 30
[7] S. Banerjee, T. Mandal, B. Mellado and B. Mukhopadhyaya, JHEP **1509**, 057 (2015)
[8] Englert C, Goncalves-Netto D, Mawatari K and Plehn T 2013 JHEP 1301 148
[9] Djouadi A, Godbole R, Mellado B and Mohan K 2013 Phys. Lett. B 723 307
[10] J. Alwall *et al.*, JHEP **1407**, 079 (2014)
[11] T. Sjostrand, S. Mrenna and P. Z. Skands, JHEP **0605**, 026 (2006)
[12] J. de Favereau *et al.* [DELPHES 3 Collaboration], JHEP **1402**, 057 (2014)

Efficient processing of physics quantities for the Processing Unit for the upgrade of the Tile Calorimeter of ATLAS

D Ohene-Kwofie[1], E Otoo[1] and B Mellado [2]

[1] School Of Elect. & Info Engineering ,University Of the Witwatersrand, Johannesburg, SA
[2] School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africa

E-mail: daniel.ohene-kwofie@cern.ch

Abstract. The ATLAS detector, operated at the Large Hadron Collider (LHC) records proton-proton collisions at CERN every 25ns resulting in a sustained data flow up to Pb/s. The upgraded Tile Calorimeter of the ATLAS experiment will sustain about 5PB/s of digital throughput. These massive data rates require extremely fast data capture and processing. Although there has been a steady increase in the processing speed of CPU/GPGPU assembled for high performance computing, the rate of data input and output, even under parallel I/O, has not kept up with the general increase in computing speeds. The problem then is whether one can implement an I/O subsystem infrastructure capable of meeting the computational speeds of the advanced computing systems at the petascale and exascale level. We propose a system architecture that leverages the Partitioned Global Address Space (PGAS) model of computing to maintain an in-memory data-store for the Processing Unit (PU) of the upgraded electronics of the Tile Calorimeter which is proposed to be used as a high throughput general purpose co-processor to the sROD of the upgraded Tile Calorimeter. The physical memory of the PUs are aggregated into a large global logical address space using RDMA-capable interconnects such as PCI-Express to enhance data processing throughput. Additionally, lossless (i.e. original data can be perfectly reconstructed from the compressed data) compression schemes are explored to enhance data bandwidth utilisation and provide increased throughput.

1. Introduction

The Large Hadron Collider, is the most powerful proton-proton collider ever built [1]. The discovery of the Higgs Boson [2, 3] was independently observed by the ATLAS and Compact Muon Solenoid (CMS) detectors. A Toroidal Large Hadron Collider Apparatus (ATLAS) is the largest of all the LHC detectors. It consists of a series of concentric rings: the inner detector, Tile Colorimeters and the Muon Spectrometers. TileCal is the central hadronic calorimeter of the ATLAS experiment at the LHC at CERN. It is primarily used to measure the energy and direction of hadrons and jets as they are produced. The Trigger and Data Acquisition System (TDAQ) is designed for event selection, processing and storage of the read-out data of the detector [4]. This selection mechanism is based on three trigger levels in the data flow that defines the different domains for the read-out electronics in terms of methods and rates for this selection.

Detector electronics are being upgraded (as part of scheduled upgrade phases on the accelerator and experiments components) to allow an increase in the Level-1 acceptance rate of

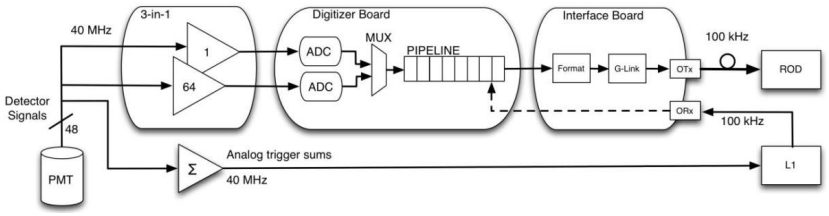


Figure 1: Current Architecture of the Tile Calorimeter.

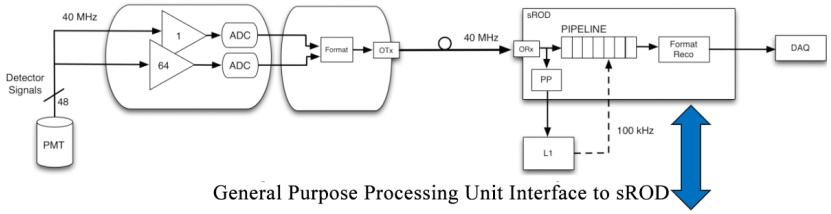


Figure 2: Upgraded Architecture of the Tile Calorimeter.

events from 70kHz to 100kHz [5]. Subsequently increasing the raw data sent for processing.

We present an architecture for on-line data processing using an ARM cluster configuration, where the data is maintained in an in-memory storage for the upgraded TileCal. The architecture consists of a cluster of ARM processing units, part of whose memories are aggregated to form a combined large logical global address space to facilitate in-memory data processing.

1.1. The Off-line Data processing Challenge

Scheduled upgrades to the ATLAS detector anticipated in 2022 will result in a much higher rate of collisions [6] at the LHC, resulting in an increase by 200 times the current rate to over 41Tb/s data output from the TileCal [4]. Storing such massive dataset for off-line processing presents a great challenge and is not desirable. The MAC project at the University Of Witwatersrand is aiming for a cost-effective, and high data throughput Processing Unit (PU), using several consumer ARM processors in a cluster configuration, as general purpose co-processor to augment the read-out system (sROD) of the upgraded TileCal.

A major challenge with processing such large volumes of data is the input/output (I/O) sub-system. The continuously growing gap between CPU and I/O speed has resulted in the conspicuous performance gap between the processor speeds and what the storage I/O subsystem can deliver. Figure 3 shows the increasing gap between CPU speed and disk storage. The solution is to reduce disk accesses and enhance in-memory data processing.

We present a brief overview of a complete in-memory storage system for on-line data processing in the ARM cluster configuration for the upgraded electronics of the TileCal. In-memory data processing provides extremely fast response time and very high throughput, with an average of about $100 - 1000$ times lower latency for a complete Random Access Memory(RAM) storage than disk-based storage systems and consequently a $100 - 1000$ times greater throughput [9].

2. PGAS Architecture for the processing of physics quantities

The architecture consists of a cluster of ARM processing units, part of whose memories are aggregated to form a combined logical global address space. The general operational architecture includes low cost ARM processing units interconnected via PCI Express interconnects (PCIe). The PCIe interconnect facilitates Remote Data Memory Access (RDMA) and offers very low-latency host-to-host transfers by copying the information directly between the host application

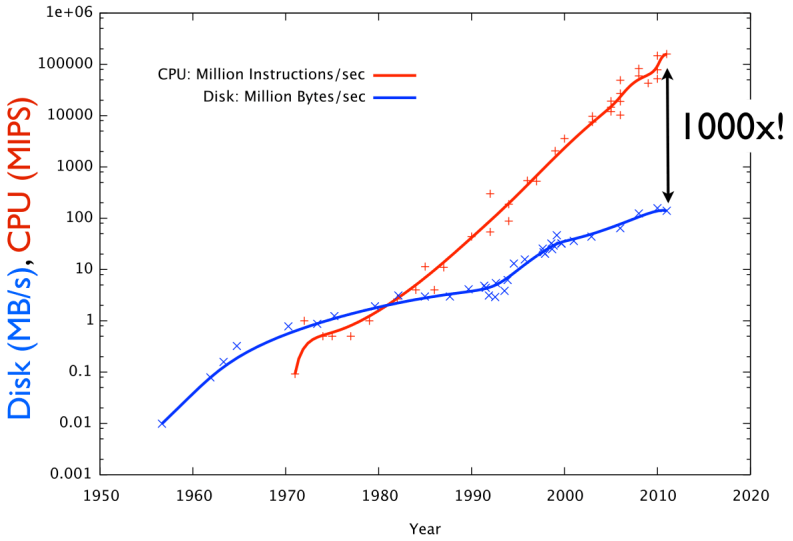


Figure 3: Increasing gap between disks and CPU speed [7].

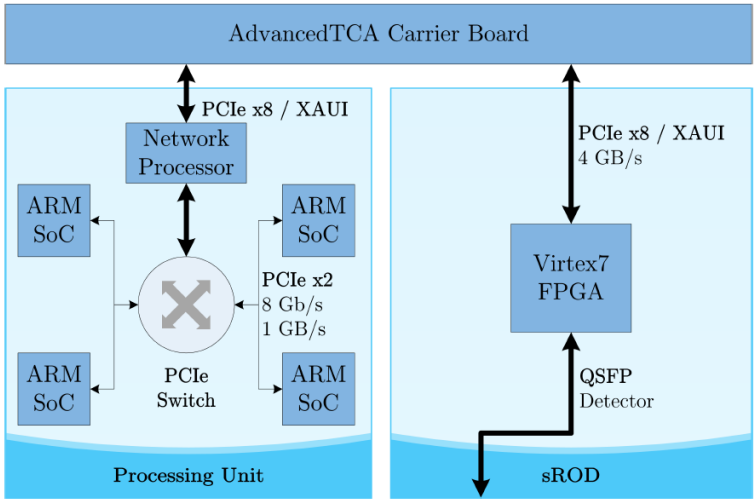


Figure 4: Schematic of the PU and SROD prototype connection [8].

memories. This enhances a seamless global logical address space for in-memory data processing as a low-overhead protocol. They are relatively affordable, low power and also provide straightforward, standards-compliant extensions that address multi-host communication and I/O sharing capabilities. Figure 5 illustrates the schematic diagram of the architecture.

Each block depicts an ARM node with 4 cores. The architecture seeks to provide low latency guarantees by ensuring effective fast access to data in-memory for application processing. Thus, minimising I/O which is the bottleneck for high throughput computing required by the ATLAS. Part of the on-going research work is to address a number of challenges related to fault-tolerance and high speed access of the memory resident data.

3. The PGAS High Data Throughput Architecture

The Figure 5 and the experiments conducted so far only demonstrates the processing capability of the PU under the PGAS computing model. This is easily transformed into a configuration

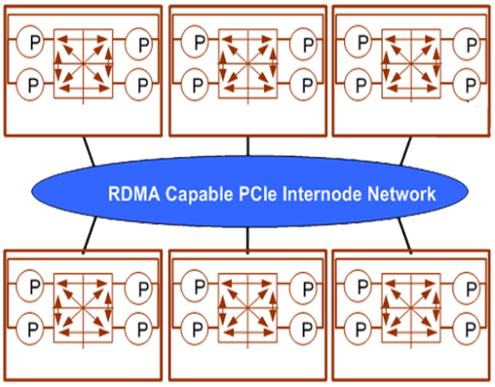


Figure 5: Schematic Diagram of the Architecture.

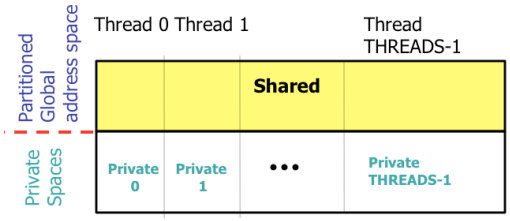


Figure 6: PGAS Memory layout.

where the aggregated memories of the ARM processors serve as a large in-memory buffer for data staging. By augmenting the ARM cluster with data input ports and data output ports to external hierarchical storage devices, data can be streamed into the ARM cluster memory using parallel I/O which can then be compressed and streamed out using parallel I/O.

Lossless data compression methods such as zlib and bzlib2, and general purpose compression schemes, present methods to encode the data with fewer bits by removing the redundancies in the data before either storage or transmission over a communication channel. Such techniques further enhance the efficiency and ensure effective use of available bandwidth of the communication channel. This results in the reduction of the amount of data injected in the channel, thereby enhancing the data throughput of the physics quantities of the upgraded ATLAS Tile Calorimeter.

4. Preliminary Evaluations

Preliminary investigations conducted show promising results. We benchmarked the system with the NASA Advanced Supercomputing (NAS) Parallel Benchmarking tool [10]. This benchmark was designed not just for parallel-aware algorithmic and software methods but also to provide an easy verifiability of correctness of results and performance figures. The evaluations are run using the Fast Fourier Transform(FFT) algorithm which solves a 3D partial differential equation using an FFT-based spectral method [10], also requiring long range communication. FFT performs three one-dimensional(1-D) FFT's, one for each dimension. The FFT benchmark is adopted since they are applicable in several signal processing algorithms, and also to the optimal filtering task of the Tilecal.

The evaluation was done on 4 nodes of the Wits High Throughput Electronic Lab (HTEL) Tegra K1 (2.3GHz Quad-Core ARM Cortex-A15) cluster with 2GB of memory each and 1Gbp Ethernet interconnect between nodes

We ran the benchmark with varying workloads as well as varying number of threads. Each workload is ran 6 times and the resulting throughput in floating point operations per second(MFLOPS) reported. A maximum of 4 threads per node are spawned and the dimensions of the FFT is varied from small ($64 \times 64 \times 64$ 3D grid) to large ($256 \times 256 \times 128$ 3D grid). The User Datagram Protocol (UDP) is used as the inter-process communication protocol between cluster nodes. UPC supports UDP, Message Passing Interface (MPI) as well as RDMA capable interconnects(e.g infiniband, PCIe, 10GbE, etc.) for inter-node communication. The current experimental setup did not have RDMA and therefore UDP was used since it is usually faster than MPI as far as inter-node communication over Ethernet is concerned.

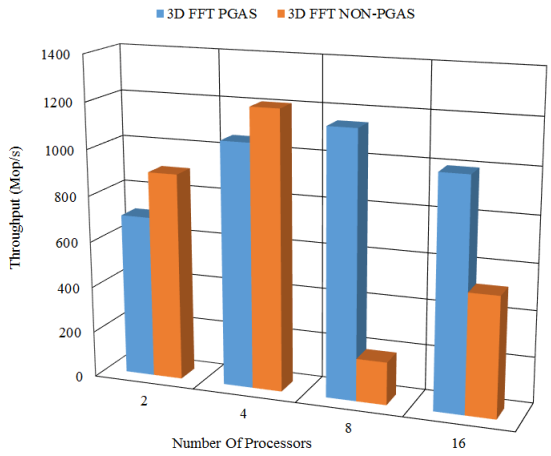


Figure 7: Throughput for small workloads.

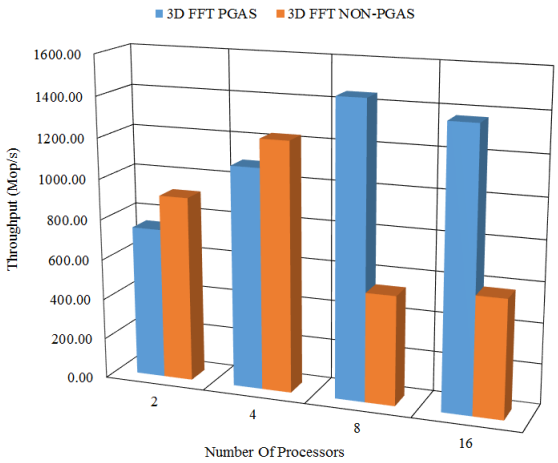


Figure 8: Throughput for large workloads.

Figures 7 and 8 show the results for the data processing throughput in MFLOPS with varying number of threads and workloads. Generally, as the number of threads increase, there is a corresponding increase in the data processing throughput since less processing is done per thread. We observe a significant and better performance increase with PGAS as depicted in Figure 8 (about 3× more throughput than the NON-PGAS FFT with 8 threads). The NON-PGAS experienced a significant drop in performance when the number of threads increased from 4 to 8. This could be due to the communication overhead as data is transferred across nodes for processing and aggregation.

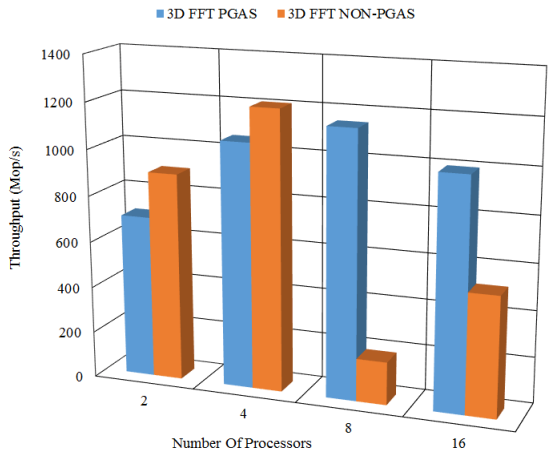


Figure 9: Average latencies for varying workloads.

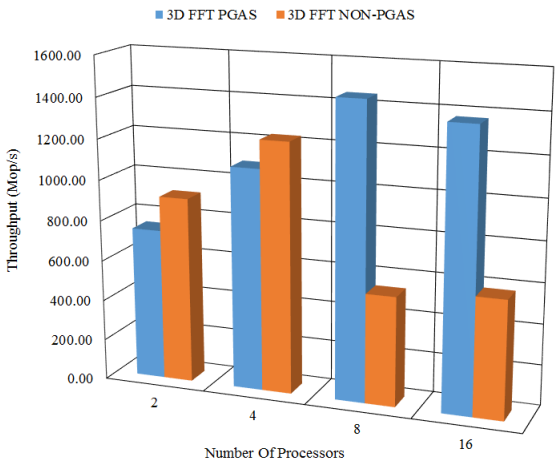


Figure 10: Latencies for large workloads.

Figure 9 shows the average latencies for each run of the experiment. There is a general increase in latency as workload size increases as expected. We also observe that as the number of threads increase latency drops a little and begins to increase after 8 threads. This is due to increase in inter-node communication. PGAS performs much better with lower latencies as compared to its NON-PGAS variant.

With a much higher bandwidth device such as 20Gb/s PCIe connectivity or a 10GbE, we expect minimal latencies and much higher throughput (GFLOPS).

5. Conclusions

Management, efficient access and analysis of the Petabytes of data, that is likely to be generated and/or used in the upgraded ATLAS TileCal present extremely challenging tasks. I/O bottlenecks in processing such huge amounts of data require techniques that utilise higher levels of the memory hierarchy to enhance data throughput.

The PGAS model presents an efficient way to process data in a distributed environment by providing a global partitioned shared address space for in-memory data. This enhances the data processing throughput. Additionally, the use of low cost CPUs such as ARM with PCIe interconnects, ensure low power consumption (high performance per watt) and thus cost effective alternative for data processing in ATLAS TileCal. Data compression frameworks are also a great strategy to both effectively utilise I/O bandwidth and thus increase data throughput. Since off-line processing of the huge data volume from the TileCal is a great challenge, on-line processing using cost effective PUs, enhanced with PGAS processing techniques, has been proposed. The strategy both ensures high throughput and efficient data bandwidth utilisation when coupled with various compression techniques.

Future work anticipated includes further rigorous experimentation using the ARM PUs with RDMA capable PCIe intra/interconnects for kernel bypass applications. Additionally, various lossless compression frameworks will be explored to further determine which scheme is effective for on-line data processing without compromising on high throughput.

Acknowledgments

We acknowledge the financial assistance of the South African National Research Foundation (NRF) and SA-CERN towards this research. We would also like to extend our gratitude to the School of Physics for making resources available for this project.

References

[1] CERN 2014 The Large Hadron Collider `http://home.web.cern.ch/topics/large-hadron-collider`
[2] Baines J T and et al 2004 *IEEE Transactions on Nuclear Science* **51** 361–366 ISSN 0018-9499
[3] Reed R and et al 2013 A Revised High Voltage Board for the Consolidation of Front End Electronics on the Tile Calorimeter of the ATLAS Detector at the LHC *SAIP 2013* (Johannesburg, South Africa)
[4] Carrioa F and et al 2014 *Journal of Instrumentation* **C02019**
[5] The ATLAS Collaboration 2012 Letter of Intent for the Phase I Upgrade of the ATLAS `https://cdsweb.cern.ch/record/1402470`
[6] The ATLAS Collaboration 2012 Letter of Intent for the Phase II Upgrade of the ATLAS `https://cdsweb.cern.ch/record/1502664`
[7] Jonathan Dursi, SciNet 2012 Parallel I/O doesnt have to be so hard: The ADIOS Library `http://wiki.scinethpc.ca/wiki/images/8/8c/Adios-techtalk-may2012.pdf`
[8] Cox M A, Reed R and Mellado B 2015 *Journal of Instrumentation* **C01007**
[9] Ousterhout J, Agrawal P, Erickson D, Kozyrakis C, Leverich J, Mazières D, Mitra S, Narayanan A, Ongaro D, Parulkar G, Rosenblum M, Rumble S M, Stratmann E and Stutsman R 2011 *Commun. ACM* **54** 121–130 ISSN 0001-0782
[10] Serres O, Andreev N, Francois C, Abhishek A, Smita A, Veysel B, Yiyi Y, Chauvin S, Vroman F and El-Ghazawi T 2011 UPC NAS Parallel Benchmarks `http://threads.hpcl.gwu.edu/sites/npb-upc`
[11] Phoboo A E 2014 Dealing With Data `http://atlas.ch/news/2014/dealing-with-data.html`
[12] ATLAS Collaboration 2008 *Journal of Instrumentation* **S08003**
[13] George M 2014 *Journal of Instrumentation* **C05004**
[14] Baines J T, Bee C P, Bogaerts A, Bosman M, Botterill D, Caron B, Dos Anjos A, Etienne F, Gonzalez S, Karr K, Li W, Meessen C, Merino G, Negri A, Pinfold J L, Pinto P, Qian Z, Touchard F, Werner P, Wheeler S, Wickens F, Wiedenmann W and Zobernig G 2004 *IEEE Transactions on Nuclear Science* **51** 361–366 ISSN 0018-9499
[15] El-Ghazawi T, Carlson W, Sterling T and Yelick K 2005 *UPC:Distributed Shared Memory Programming* (Hoboken, New Jersey: John Wiley & Sons Inc.)

A study of radiation damage in plastic scintillators using magnetic resonance techniques for the upgrade of the ATLAS detector

C. Pelwan[1], H. Jivan[1], D. Joubert[1], J. Keartland[1], M. Madhuku[2], B. Mellado[1], G. Peters[1], K. Sekonya[2], E. Sideras-Haddad[1]

[1]School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africa
[2]iThemba LABS (Gauteng), Johannesburg 2050, South Africa.

E-mail: chad.dean.pelwan@cern.ch

Abstract. During the phase two upgrade of the Large Hadron Collider (LHC), various components of the accelerator and ATLAS detector are due to be replaced or upgraded to withstand the increase in instantaneous luminosity. The minimum bias trigger scintillator (MBTS) plastics, located at $2.09 \leq |\eta| \leq 3.84$, on the EndCaps of the Tile Calorimeter (TileCal) in the ATLAS detector were subjected to ionizing radiation that allows them to track the trajectories and measure the energies of energetic particles. However, it is this interaction that causes structural damage within the polystyrene based MBTS plastics. The 6 MeV proton tandem accelerator at iThemba LABS, Gauteng is used to simulate the damage that the MBTS plastics are subjected to in the ATLAS detector in order to find a plastic scintillator type that could replace the one currently used. In order to understand structural damage, electron paramagnetic resonance (EPR) is employed to detect structural defects in two grades of polystyrene based plastic scintillators and three grades of polyvinyl based plastic scintillators. A replication of the spectra seen by the EPR technique will be attempted using a computational ab-intio approach. This should offer insight into the electronic structure of the plastic scintillators and how ionizing radiation causes structural damage to them. In this paper we present results from two Polyvinyl toluene (PVT) based plastic scintillators.

1. Introduction

With the second run of the Large Hadron Collider (LHC) set to begin in July this year, the shut down after the first run called for various components of the two ring, particle accelerator and its four detectors to be upgraded or replaced [1]. The minimum bias trigger scintillator (MBTS) plastics are an example of components that were to be replaced once their efficacy was depleted due to the exposure to energetic particles they interact with [2].

In total, there are 16 MBTS plastics situated around the beam pipe within the largest of the four detectors, the ATLAS detector, at a pseudo-rapidity between $2.07 \leq |\eta| \leq 3.42$ on the EndCaps of the Tile Calorimeter (TileCal) in the ATLAS detector. Before nominal operation of the LHC, the polystyrene based, 2 cm by 2 cm plastic scintillators were essential in vetoing background and finding an effective mean for the proton-proton collisions. When the MBTS plastics interact with energetic particles, usually ionizing radiation, there is an exchange of energy from the particles to the plastics scintillators. The plastic scintillators produce light, through the scintillation mechanism, which is then passed down via wavelength shifting fibres to

photo-multiplier tubes (PMTs) and the signal is digitized and can be analysed. However, this energetic interaction causes damage to the benzene molecule in the plastic scintillator and, in turn, the scintillation mechanism which is responsible for luminescence [3].

Understanding the damaged caused by ionizing radiation in plastic scintillators is non-trivial because of the amorphous nature of the materials, however, it is vital since the MBTS plastics are mainly subjected to low to medium energies (in the GeV range) in the ATLAS detector. Therefore, the study focuses on the interaction of particles with the plastics at relatively lower energies (in the MeV region) first hypothesising that it is possible to characterize damage in the plastic scintillators and, more specifically, to the benzene molecule. Damage has been induced to a plastic scintillator identical to that used for the MBTS plastic along with five other plastic scintillator grades. Two plastic scintillators will be investigated: polyvinyl toluene and polyester over a range of six doses. The two polystyrene samples are of a Dubna and Protvino make and the three polyvinyl toluene samples where called EJ200, EJ208, and EJ260. Electron paramagnetic resonance (EPR) be employed to characterize the damage seen in the various plastic scintillators and a computational ab-intio method will be used in an attempt to recreate the spectra [4]. In this paper we present preliminary results from EJ208 and EJ60, the two PVT based plastics. The calculation employs density functional theory (DFT) to build up the bulk material plastic from the benzene molecule. This should give us insight into which bonds are broken or formed when the particles interact with ionizing radiation [5].

2. Material Preparation and Irradiation

The damage to the plastic scintillators was induced by 6 MeV protons and SRIM (Stopping and Range of Ions in Matter) was used to model this interaction. Using SRIM, we found that the 6 MeV protons stopped at a distance of around 470 μm in the plastic scintillators. Therefore, the plastics needed to be cut and polished to a width thinner than that for ionization to occur. An extension of SRIM is TRIM (Transport of Ions in Matter), it was used to calculate the average energy of a proton, was approximately 2.07 MeV. This value added in calculating the dose each sample was subjected to. The samples were cut and polished to width of about 250 μ m with the volume that was around 1 cm $\times$ 1 cm $\times$ 250 μm. The samples were cut and polished in the Materials Preparation Room at the University of Witwatersrand.

Irradiation of samples took place at iThemba LABS, Gauteng using the tandem accelerator which is capable of accelerating protons to an energy of 6 MeVs. The beam is thenincident on the samples which are placed on a rotating carousel at the end of the beam line. In order to investigate the effect of radiation damage on the plastic scintillators, a range of doses were. Six doses were chosen, namely, 0.8 MGy, 2.5 MGy, 8.0 MGy, 25.0 MGy, 50.0 MGy, and 80.0 MGy. Each sample of each type of plastic scintillator was subjected to the different doses and analysed along with an un-irradiated sample. The exact dosage was calculated using the following equation

$$\mathcal{D} = \frac{I_B \; A \; E_{ave} \; \Delta t}{Q \; m_{Ir}}, \tag{1}$$

where I_B is the current of the beam incident on the sample with an irradiated are A for a time Δt. E_{ave} is the average energy calculated using the TRIM calculation and q is its charge (that of the proton). Once the samples are irradiated, they are then ready to be tested using the magnetic resonance spectroscopy techniques.

3. Electron Paramagnetic Resonance

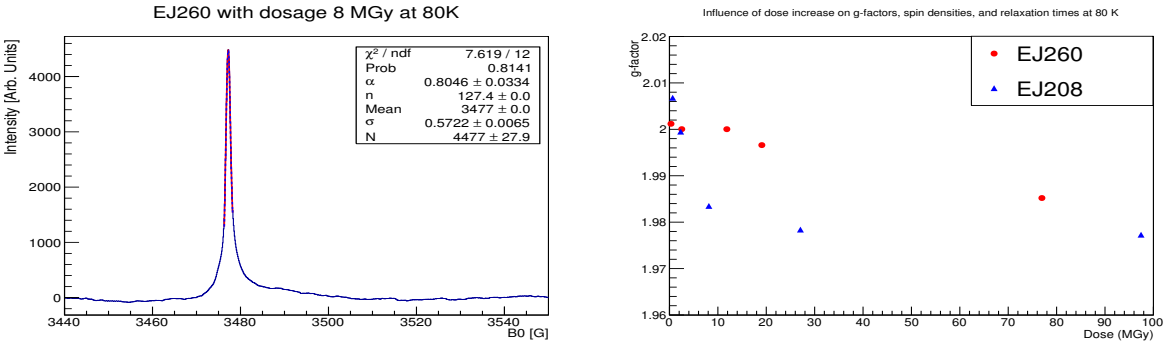
EPR is a spectroscopy technique that will be used to probe the electronic structure. The X-band, continuous wave (CW) ESP380 spectrometer was used to study how the increase in dosage effects the EPR spectrum produced for each sample. This is done by studying unpaired electrons and

ions in the samples when placed in a large magnetic field and a resonant frequency is applied. The following resonance condition summarizes this

$$\Delta E = h\omega_0 = g\mu_B B_0. \tag{2}$$

Equation (2) tells us that we can characterize a difference between two energy quantum states with a particular angular resonant frequency, ω_0. This is proportional to the magnetic field, B_0, the Bohr magneton, μ_B, and the g-factor of the sample, g. This should give us more insight into the environment in each sample and how it differs with an increase in dose. From the EPR spectra seen it is possible to obtain the g-factor of each sample, the relaxation time of the sample and the spin density. It is then possible to see how these parameters change with an increase of dose.

4. Preliminary EPR Results

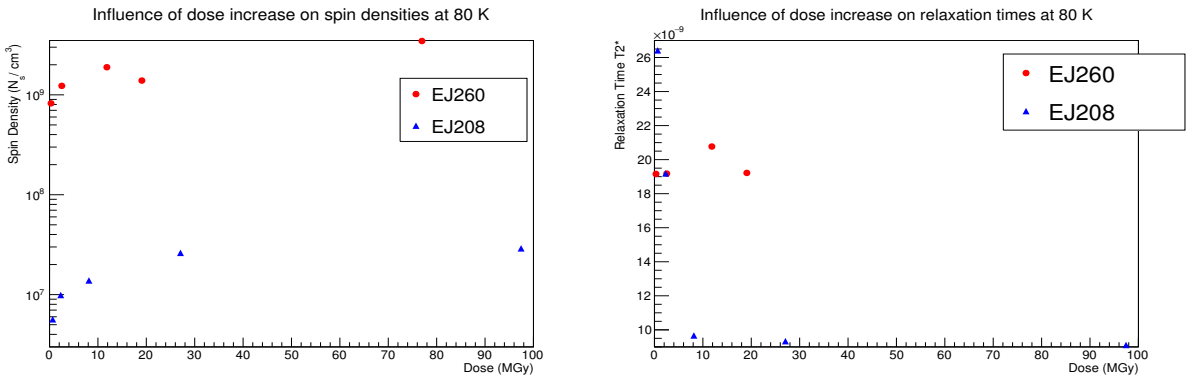


1 (a) 1 (b)

Figure 1: Integrated EPR spectra of irradiated EJ260 at 80 K (left) change in g-factor with dose for EJ260 and EJ208 (right)

Figure 1 (a) shows the spectra of EJ260 irradiated to a dose of 8 MGy. A CrystalBall distribution was used to fit the shape of the signal. This fit allows us to obtain information about various parameters like the g-factor seen in figure 1 (b). The asymmetric nature of the spectrum could arise from the anisotropy of the g-factor. The g-factor gives us an indication of the environment seen by the unpaired electrons in each sample. As the dose increases, we see that the g-factor decreases away from that of an un-irradiated reference sample. This could be due to the fact that defects in the samples caused by the ionizing radiation effect the potential the unpaired electrons see and thus change the g-factor.

Figure 2 (a) shows the unpaired spin concentrations, N_s, seen in both samples. This concentration is an indication of the number of unpaired electrons seen in a sample. N_s is obtained by integrating over the EPR absorption spectrum and dividing by the volume of the irradiated spot on the sample [6]. In general, the number of unpaired electrons increases as dosage increases. The increase is different for EJ260 than it is for EJ208. The reason for this is not yet know.



2 (a) 2 (b)

Figure 2: Influence of dose increase on unpaired spin concentration (left) and relaxation time T_2^* (right) at at 80 K

The relaxation time, T_2^*, is an indication of the spin-spin relaxation time for amorphous structures. It is calculated using full width at half the maximum of each integrated EPR peak.

$$T_2^* = \frac{2}{\sqrt{3}} \frac{\hbar}{g\mu_B B_{\mathrm{FWHM}}}. \tag{3}$$

It indicates the average time it takes for the spins to transition to a lower energy state. As the dose increases and more damage is induced, the relaxation time decreases for both samples.

5. Conclusion

The damage induced by ionizing radiation changes the structure in the plastic scintillators. It causes defects in the sample that shifts the g-factors in the samples and increase the number of unpaired electrons in them. This result is seen in both the experimental and computational work. The unpaired electrons also lose energy to other unpaired electrons and the defects in the material and there is a broadening in the EPR peak. This is due to the fact that the hyperfine parameters change when the samples are damaged by irradiation.

Future work will investigate radiation damage in one other PVT based plastic scintillator and three other polystyrene types.

6. References

[1] G Aad and Collaboration A 2008 *INSTITUTE OF PHYSICS PUBLISHING AND SISSA* **08003**

[2] Kwee R, Kolanoski H and Ellis N 2012 *Development and Deployment of an Inner Detector Minimum Bias Trigger and Analysis of Minimum Bias Data of the ATLAS Experiment at the Large Hadron Collider* Ph.D. thesis Humboldt-Universitat zu Berlin

[3] Torrisi L 2002 *Radiation Physics and Chemistry* **63** 89–92 ISSN 0969806X

[4] Marx D and Hutter J 2009 *Ab Initio Molecular Dynamics* (CAMBRIDGE UNIVERSITY PRESS) ISBN 9780521898638 URL `www.cambridge.org/9780521898638`

[5] Martin R M 2006 *Understanding Carbon Nanotubes* **1914** 199–276 ISSN 0163-1829 URL `http://dx.doi.org/10.1007/3-540-37586-4_4`

[6] Jones B, Barklie R, Khan R, Carey J and Silva S 2001 *Diamond and Related Materials* **10** 993–997 ISSN 09259635

A Di-Higgs Search in the $\gamma\gamma b\bar{b}$ Decay Channel Using the ATLAS Detector

Robert G. Reed, Stefan Von Bruddenbrock, Deepak Kar, Bruce Mellado

School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africa

E-mail: robert.reed@cern.ch

Abstract. The Higgs boson was discovered on July 4th 2012. The Higgs boson now offers a portal into physics beyond the Standard Model (BSM). The search for a di-Higgs production in the $\gamma\gamma b\bar{b}$ decay channel is sensitive to extensions in the Standard Model and provides a way of probing new physics. In the $\gamma\gamma b\bar{b}$ search using 2012 data an excess was seen around the 300 GeV mass of the four object ($m_{\gamma\gamma b\bar{b}}$), although differing in the significance, for both ATLAS and CMS. The $\gamma\gamma b\bar{b}$ channel is an early analysis which means the strategy and all associated frameworks must be ready as soon as possible. The first look at data will be to verify the excess seen in 2012 followed by a further investigation for BSM physics. This proceedings describes the implementation and overall status of the analysis framework developed for the $\gamma\gamma b\bar{b}$ decay channel.

1. Introduction

The ATLAS and CMS Collaborations announced the discovery of a Higgs-like boson, on July 4^{th} 2012, with a mass of 125 GeV [1,2]. After further studies using the full datasets available during Run I [3,4] this boson was observed to be consistent with the Standard Model (SM) Higgs Boson with a significance of 10σ as seen in Fig. 1 [5]. The Standard Model predicts the Higgs boson as a consequence to spontaneous symmetry breaking which gives mass to elementary particles. The discovery and the measurement of its mass provides new means to probe physics that is beyond the Standard Model.

2. The ATLAS detector

The ATLAS detector is one of two main general purpose detectors at the LHC. The ATLAS detector is different in many respects to the CMS detector most noticeable is the air-core superconducting magnets that produce the toroidal field for the muon spectrometer. The detector is cylindrical in shape and has a forward-backward symmetry. The detector consists of concentric sub detectors. Starting from the inside we have the inner tracker and silicon micro-strip detector located inside the transition radiation tracker which provide tracking in the pseudorapidity[1] range $|\eta| < 2.5$, electromagnetic calorimeter and lead liquid-argon sampling device is divided into one barrel ($|\eta| < 1.475$)) and two end cap regions ($1.375 < |\eta| < 3.2$), the

[1] In experimental particle physics, pseudorapidity, η, is a commonly used spatial coordinate describing the angle of a particle relative to the beam axis. It is define by $\eta \equiv -ln\left[tan\left(\frac{\theta}{2}\right)\right]$ where θ is the angle between the particles three-momentum and the positive direction of the beam axis.

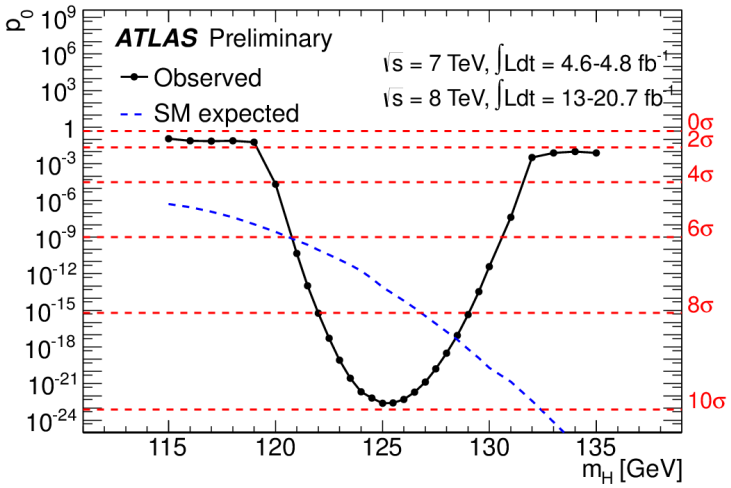


Figure 1: Combined search showing the local probability in the search for the standard Higgs boson [5]. The dashed line represents the Standard Model expected local probability.

hadron calorimeter is also divided into a barrel ($|\eta| < 0.8$) and extended barrel ($0.8 < |\eta| < 1.7$) both of which consist of steel and plastic scintillators. Each sub detector is designed to identify different particles, their energies and trajectories. All signals from the sub-detectors are used to reconstruct the event and all particles that are produced in the collision. The reconstructed data is then processed by various analysis groups to look for interesting physics such as that found beyond the SM. The Higgs boson opens a portal to study physics beyond the SM by making use of particular decay channels which are sensitive to extensions of the SM theory. Such a channel is the di-Higgs to $\gamma\gamma b\bar{b}$ which is a sub analysis of the Higgs to $\gamma\gamma$ group.

3. The $\gamma\gamma b\bar{b}$ run I results

The $\gamma\gamma b\bar{b}$ final state search is an excellent channel for Higgs boson pair production searches due to the large branching ratio of the $h \to b\bar{b}$ process [6]. Combining the large branching ratio with the clean di-photon trigger, di-photon mass resolution and low background we get a very clean analysis. Figure 2 shows the 95% Confidence Limit (CL) on the cross section times the branching ratios of a narrow resonance decaying into two Higgs bosons [7]. It is immediately clear there is an excess around the $m_X \approx 300$ GeV area, where m_X is the mass of the heavy mediator. The CMS detector also observed an excess in a similar region but with a smaller significance [8]. The excess in itself warrants further investigation, due to the presence in both CMS and ATLAS, which will be the driving factor in the early data collection of Run 2 for this decay channel. The framework described in following sections will attempt to verify if this excess still exists in the 2015 data taking period (Run II) using the new data formats.

4. Run II $\gamma\gamma b\bar{b}$ analysis framework

The $\gamma\gamma b\bar{b}$ framework inherits its core functionality from its parent analysis group. This means the calibration and event selection is largely pre-defined by the $\gamma\gamma$ group. I will focus on how the sub analysis differs from the di-photon analysis which is described in Ref. [1,4]. All analysis groups are using the new xAOD data format that is the standard in ATLAS. The analysis code is written in c++11 using an object orientated approach. This allows a central repository of tools which can be used by many groups allowing more eyes to debug the same code. The $\gamma\gamma$ group has written a core analysis code which sub groups can inherit from allowing centralised

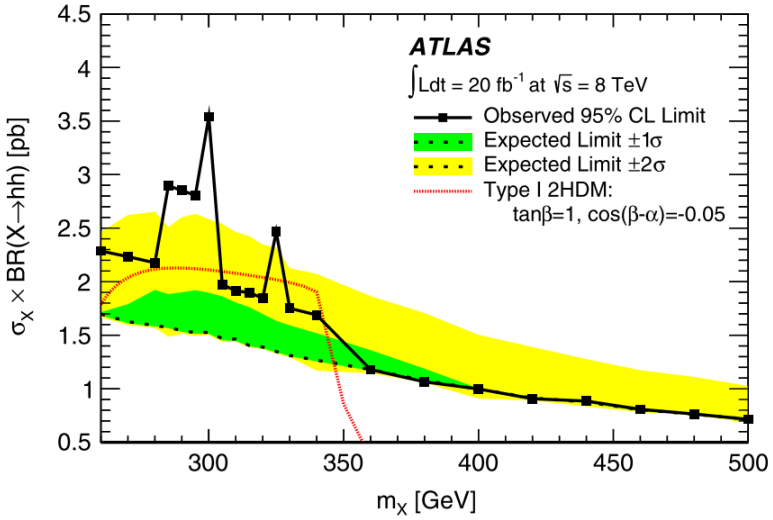


Figure 2: A 95 % CL upper limit on the cross section times branching ratios of a narrow resonance decaying to pairs of Higgs bosons as a function of m_X [7].

calibration and selections across all sub groups. The difference in the object selection between the $\gamma\gamma b\bar{b}$ and $\gamma\gamma$ groups can be described on an object by object basis.

4.1. Photons

There is no difference in photon selection from the $\gamma\gamma$ group as the $\gamma\gamma b\bar{b}$ analysis is an extension including two jets (a narrow cone of hadronic material and possibly other particles produced by the hadronisation of a quark or gluon during a collision) formed from b-quarks these are flavored jets called BJets.

4.2. Electrons

Since electrons are not of interest we follow the same recommendations as the $\gamma\gamma$ group.

4.3. Jets

The primary focus in the $\gamma\gamma b\bar{b}$ channel is the identification of jets formed by b quarks. Jets are required to be in the central region $|\eta| < 2.5$ for better b-tagging efficiency. A tighter cut on the Jet Vertex Fraction (JVF) of 0.5 is applied which is different to the $\gamma\gamma$ cut of 0.25. The JVF discriminant measures the probability that a jet originated from a particular vertex. This is important for b-tagging reasons due to the displaced secondary vertex caused by b-mesons in the b-jet identification. In the latest derivation of data production the JVF will be replaced by a new Jet Vertex Tagger (JVT) which is expected to perform better. Finally, the jet is required to have a transverse momentum greater than 25 GeV.

4.4. Muons

Muons are selected in a similar way to the $\gamma\gamma$ analysis. Muons are rejected if they overlap any jet unless that jet is in fact a b-tagged jet. In the case a muon overlaps a b-tagged jet we add the four momentum of the muon to the jet.

4.5. b-jets

Jets are tagged by flavour using multivariate analysis techniques. A new tagging tool is used to perform this classification but an additional variable is saved using the Run I tool for comparison.

A 70% working point is used for b-tagged jet classification. Events are rejected if the leading (sub) b-jets have transverse momentum less than 55 (35) GeV. The invariant mass of the dijet system must be within the mass window 95 GeV - 135 GeV. The value of the b-tagging working point is currently the same as Run I for comparisons but this has yet to be optimised.

5. Configuring the framework

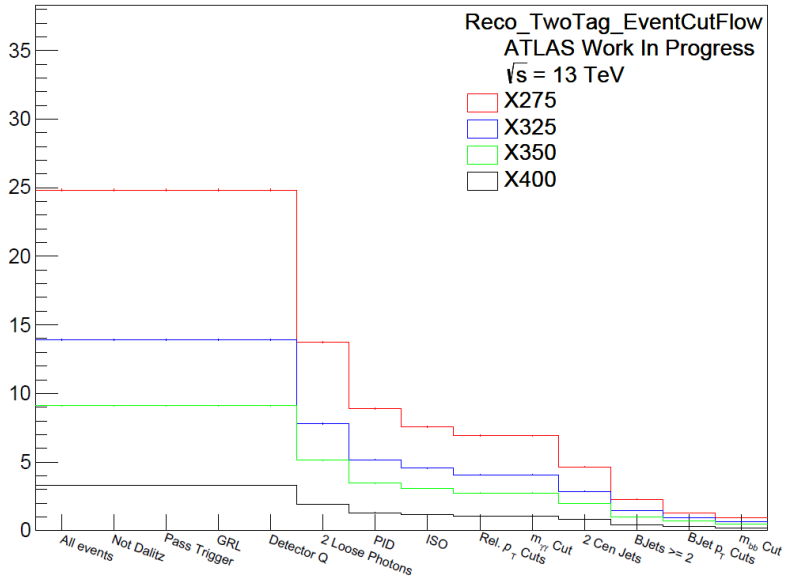
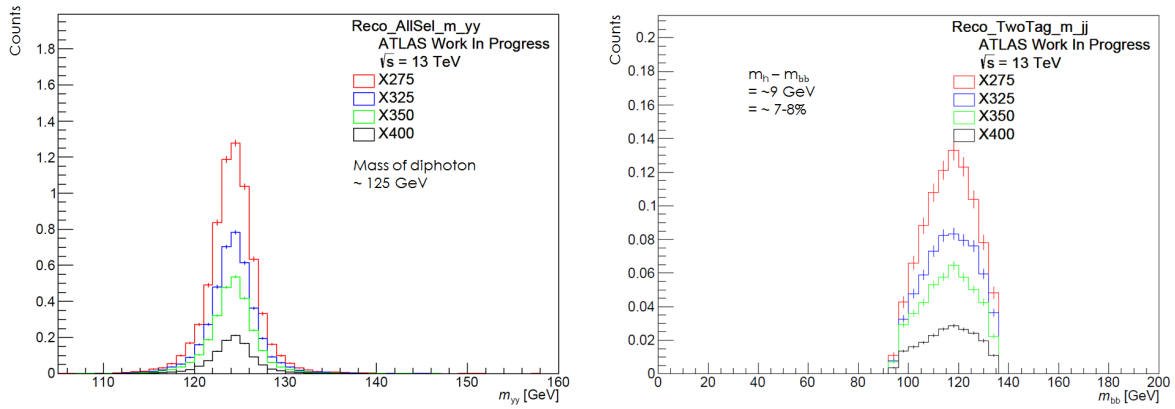


Figure 3: The cutflow produced by the $\gamma\gamma b\bar{b}$ framework.

In order to steer the analysis we use configuration files which are used to setup the analysis code to perform the desired function. Figure 4 shows the different options available to the user. The benefit of using such configuration files is that the user can change cut values or tagging working points without having to edit the source code and recompile. This feature makes the framework user friendly and flexible to different needs. In the latest release there are over 58 different parameters that can be tuned in the configuration file and this is likely to grow as we keep adding features.

Flag	Default	Description
#Photon Selection		
HH2yybb.PhotonSelection.ApplyFakePhotonWeights	NO	Add fake photon event weights (Ongoing)
# BTagging Selection		
JetHandler.EnableBTagging:	YES	Turn Btagging on/off
JetHandler.Selection.RemoveBJets:	NO	Remove all bjets
JetHandler.Selection.SelectBJets:	NO	Keep only bjets
JetHandler.Selection.BTagOperatingPointNames:	MV1_60 ...	MV1 working point names
JetHandler.Selection.BTagOperatingPointCuts:	0.991565 ...	Actual MV1 cut value
HH2yybb.Selection.UseWorkingPoint:	MV1_70	Selected working point to use
JetHandler.Selection.BTagAbsRapidity:	2.5	Cut on btagging range

Figure 4: An example of some options available to the user in the configuration file. Over 60 different features are available.



(a) (b)

Figure 5: a) Invariant mass of di-photons and b) di-b-jets for four different signal samples.

6. Preliminary results

6.1. Cutflow

Figure 3 shows the cutflow produced by the framework for four different Monte-Carlo signal simulations. Each simulation has a different mass for the heavy scalar (X). Events are rejected at each cut level. The final cut is the mass window of the $b\bar{b}$ object. Any event passing this cut will be a candidate for the four object $\gamma\gamma b\bar{b}$.

6.2. Invarient mass of $\gamma\gamma$ and $b\bar{b}$

Figure 5 a) shows the invariant mass of the selected di-photons and di-b-jets. The $m_{\gamma\gamma} \approx 125\ GeV$ which indicates the framework is selecting the correct photons from the decay of one of the Higgs. In the case of the di-b-jet mass, Fig. 5 b), we see a slight underestimate ranging between $7-8\%$ with a value of $m_{b\bar{b}} \approx 120\ GeV$. The slight under estimate can be studied by looking at the ratio between the reconstructed energy, E_{Reco}, to the energy of the b-quark, E_{bquark}. This is called the jet response.

6.3. b-jet response

The b-jet response, defined by $Response = (E_{RecoBJet} - E_{bquark})/E_{bquark}$, is a relative indication of how far above or below the reconstructed b-jet energy is to the truth energy of the b-quark. This is useful for calibrations specific to b-jets for the $\gamma\gamma b\bar{b}$ analysis. Its only applicable when looking at simulations since the truth information is not available in actual data. Figure 6 shows the response calculated for all b-jets matched to their respective b-quark. The mean value of -0.078 reflects the negative $7-8\%$ we saw in Fig. 5 b). An asymmetric Gaussian is fitted which predicts the mean values of the peak as well as the width above and below the mean. It turns out that the response is dependant on the transverse momentum p_T and the pseudorapidity η. Further studies need to be done to investigate this dependence.

7. Discussion and Conclusions

This framework will provide the platform necessary to perform statistical procedures to verify the existence of the excess seen in Run I. Physics data is expected to be produced at 25 ns by August 2015 which means the framework and all associated statistical machinery must be ready and in place. The analysis framework is in a mature stage of development and the results shown provide verification that it is performing as expected. Work that will follow this advancement

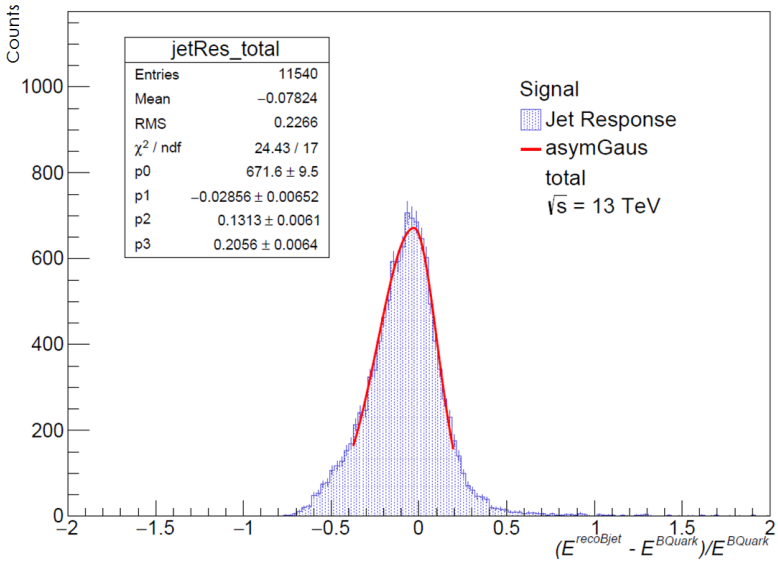


Figure 6: Total b-jet response with asymmetric Gaussian fit

will be the understanding and development of the statistical machinery that will be used for limit setting and local probability calculations. Further studied will be done on the b-jet response and the impact it will have on the $m_{\gamma\gamma b\bar{b}}$ distribution. The framework as it stands is largely ready for the Run 2 data taking period. This framework will be used by members of the $\gamma\gamma$ group and any sub group which may benefit from b-tagging functionality.

8. Acknowledgements

The National Research Foundation for the bursary they provided. The Department of Science and Technology and the SA-CERN program who provided financial assistance for research visits to CERN. We 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 *Physics Letters B* **716** 1–29
[2] The CMS Collaboration 2012 *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics* **716** 30–61
[3] The CMS Collaboration 2014 *The European Physical Journal C* **74** 3076
[4] The ATLAS Collaboration 2012 *Phys. Rev. Lett. 108* **111803** 1–19
[5] The ATLAS Collaboration 2013 *ATLAS-CONF-2013-034*
[6] Heinemeyer S, *et al.* 2013 *CERN-2013-004, FERMILAB-CONF-13-667-T*
[7] The ATLAS Collaboration 2015 *Physical Review Letters* **114** 081802
[8] CMS Collaboration 2014 *CMS PAS HIG-13-032*

A Framework Tool for Integrating the Back-End Infrastructure in the ATLAS Detector Control System

Robert Graham Reed, Bruce Mellado

School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africa

E-mail: robert.reed@cern.ch

Abstract. The Large Hadron Collider at CERN is scheduled for a Phase II upgrade in 2022. The ATLAS collaboration is required to upgrade the detector to handle the increased centre of mass collision energy. The overall amount of data transmitted off the detector will also increase which requires a complete redesign of the front-end electronics. A large proportion of these electronics will be relocated to the back-end infrastructure. A Demonstrator program has been established as a proof of principle which will be house in a new rack mounted setup. The proposed solution will be an Advanced Telecommunication Computing Architecture which will not only house but also allow advanced management features and control at a hardware level. The integration of this new architecture into the ATLAS Detector Control System will be discussed.

1. Introduction

The Large Hadron Collider (LHC), a 27 km accelerator ring containing two beams of protons, collides protons at extremely high energies in order to help answer questions about the sub atomic universe. The ATLAS detector [1] is one of two general purpose detectors which record collisions at a rate of 20 interactions every 50 ns. The ATLAS detector is scheduled to undergo a major upgrade in the year 2022 [2]. 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, the central region of the hadronic calorimeter, is used to measure energies and directions of hadrons and leptons. Plastic scintillators which, when particles pass through, emit light collected by photo-multiplier tubes (PMTs) to measure energy and direction of hadrons. These light signals are directed to front-end electronics where they are digitised and processed for the first level of triggering.

Data from the front-end electronics is transferred off the detector to the back-end electronics where further processing occurs. These electronics are house in Versa Module Europa (VME) crates. These systems are over 30 years old and will not be able to accommodate the required upgrade to handle the increased data rates that will be delivered by the LHC in the year 2022. As a result the entire front and back-end electronics will be upgraded for the Phase II upgrade [2].

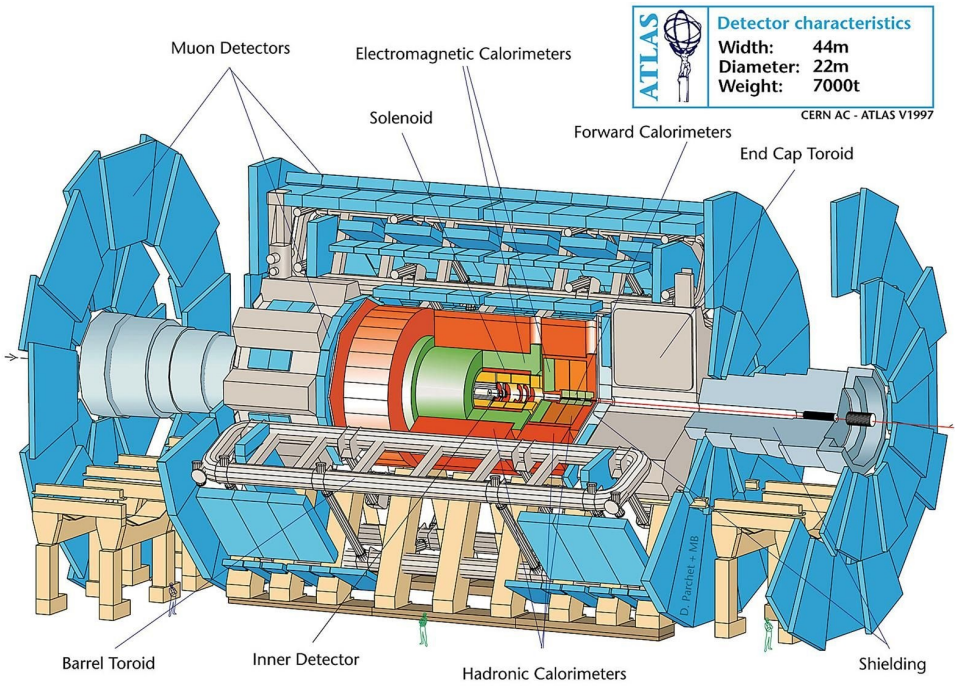


Figure 1: Artistic representation of the ATLAS detector.

2. Trigger Pre Processor

A part of the Phase II upgrade is the redesign of the Read Out Driver (ROD) of the Tile Calorimeter. This will involve electronics from both the front and back-end system. This new improved board, called the Trigger Pre Processor (PPr), will contain the sensitive electronics that are currently housed in the front-end area. This will allow the full data volume digitized at 40 MHz to exit the detector and be processed in the back-end area. Figure 2 shows the schematics of the upgrade with the pipelines and triggering contained in the new Trigger Pre Processor. Since the PPr 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 PPr and the Tile Calorimeter as a whole [3].

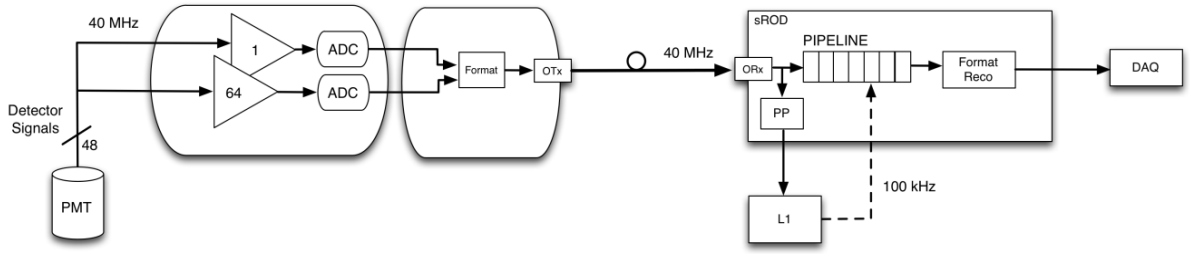


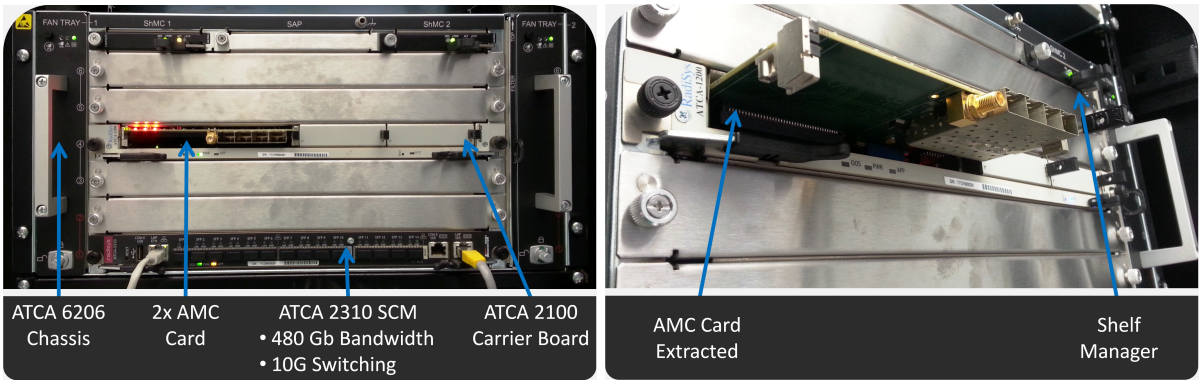
Figure 2: Upgraded read out electronics.

The PPr will be housed in an Advanced Telecommunications Computing Architecture (ATCA) chassis. The ATCA standard will be replacing the previous VME standard as it is superior in many aspects [4]. Once the PPr prototypes have been tested the designs will be brought to South Africa where the PCB will be manufactured and all components mounted. The PPr will be inserted into the ATCA system at the University of the Witwatersrand and

tested.

3. ATCA Framework

The ATCA chassis allows intelligent monitoring and control while offering high speed connectivity via a 40 Gbps backplane. Figure 3 shows a front view of the chassis that is currently installed at the University of the Witwatersrand. This model has six slots in a dual star topology offering redundancy. A 10 Gbps switch module provides the routing and switching configuration. A carrier board allows smaller cards, called Advanced Mezzanine Cards (AMC), to be inserted. The PPr is a doubled sized AMC card that can be inserted as shown in Figure 3 (b). All boards can be inserted and extracted while the system is powered on. This is called Hot Swapping. The Shelf Manager is responsible for the all the auxiliary services such as power control, fan speeds, temperature readings and voltage readings among others. Communication to the Chassis is done via a Simple Network Management Protocol (SNMP). This protocol is used since the Detector Control Software, WinCC, has a built in driver to manage the connection. 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." [5].



(a) (b)

Figure 3: a) Front of ATCA chassis and b) AMC extracted from chassis at the University of the Witwatersrand.

Table 1: Tile Calorimeter Bandwidth Upgrade.

Phase	Present	Upgrade
Number of fibers	256	4096
Fiber Bandwidth	800 Mbps	10 Gbps
Total Bandwidth	205 Gbps	41 Tbps

4. The ATLAS Detector Control System

The Phase II upgrade will issue in new electronics such as the ATCA back-end infrastructure. This new hardware needs to be integrated into the existing Detector Control System (DCS) [6]. The DCS is a highly parallel system with distributed nodes that monitor and control 12

individual sub detectors in ATLASin a redundant fashion. The DCS has the task of bringing the detector and all components into various desired operational states while monitoring and performing corrective actions in case of abnormal behaviour, all in an automated fashion.

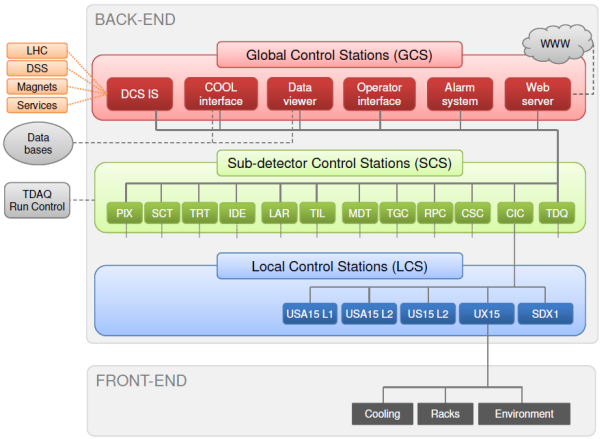


Figure 4: Schematic of the ATLAS Detector Control System [6].

Figure 4 shows the distributed nature of the DCS and the logical separation of the controlling nodes. It consists of various drivers arranged in tiers according to geographic, function or dependency groups. The software is called WinCC Open Architecture which is a product of Siemens [7].

All sensors and variables of interest in the ATCA system need to be polled in a regular fashion to allow the DCS to correctly deduce the state of all devices. This requires the creation of data points in the WinCC software that is very specific to the hardware in question. A framework tool for the ATLAS DCS has been designed which provides the expert the means to automatically generate the required data points. It does this with all the correct configurations such as the SNMP address, formatting, descriptions, polling intervals, access permission and more. This framework tool is called fwATCA.

5. fwATCA Framework Tool

The fwATCA framework tool is required to perform four functionalities:

(i) Search - This is the most important. The framework needs to search through the ATCA Shelf and determine what it contains and find all the sensors that correspond to each device.

(ii) Sort - Once all the devices have been found, along with their sensors, the information needs to be sorted in a meaningful and clear way.

(iii) Create - The sorted information must then be generated into data points that are accessible by the Detector Control System.

(iv) Configure - After creation one may want to configure a selected data point (or all) with descriptions, alerts, formatting, polling intervals or simply to delete them.

Figure 5 shows the framework panel that has been able to search through the ATCA shelf. Sort the information into two main categories, with sub categories, and then create the data points (The shaded colour around the sensor name indicates that the sensor has not been created). The sorting of the data points is defined by two categories: Standard Data Points and Custom Data Points. The first being sensors that are populated by the Shelf Manager. Standard sensors are always available as they belong to the chassis itself and are always present. The Standard Data Point category is then broken into smaller groups depending on the physical

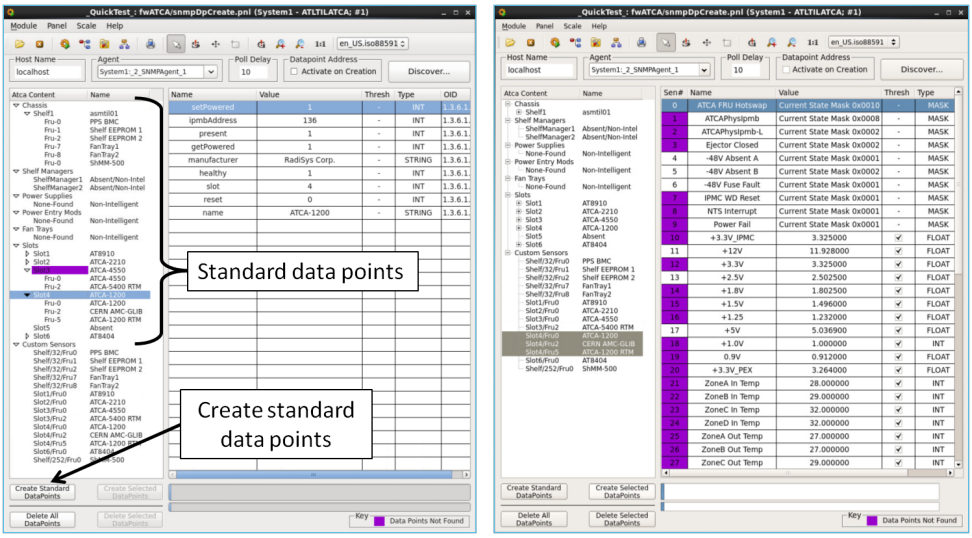


Figure 5: Framework panel that allows data point creation.

devices itself such as fan trays, power modules, shelf managers or physical slot bays. The number of devices in each category will depend on the chassis type since some chassis only have six slots while others may have fourteen. The second category, Custom Data Points, contains the sensors that belong to the boards that are inserted. Naturally these sensors are only available if the board is present. The custom sensors are provided by the board manufacturer and must meet the PICMG specifications. The information these sensors provide is different for each board which means they are harder to integrate in an automated fashion. However, this has been achieved. The custom sensors are processed after creation allowing them to be categorised according to the information they provide.

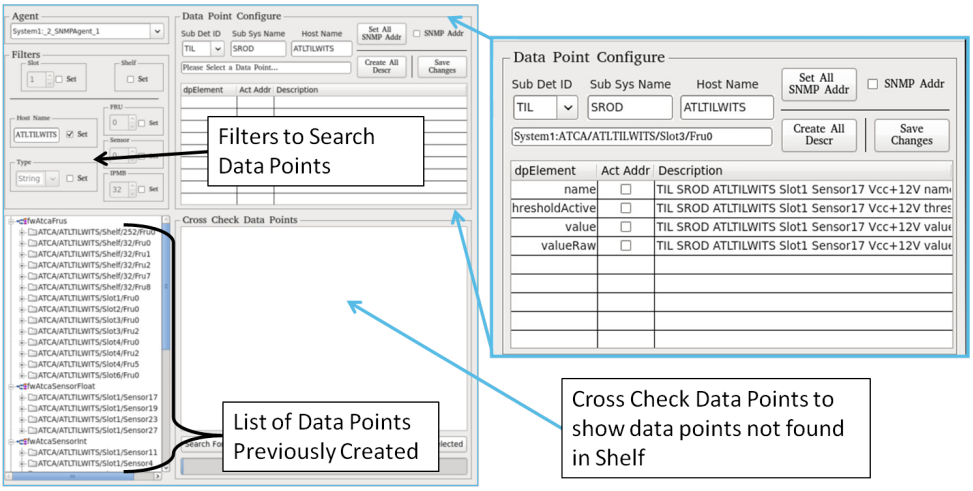


Figure 6: Framework panel that allows data point configuration.

Figure 6 provides the expert with the tools requires to configure the data points after creation. This involves creating descriptions, viewing (or deleting) data points that are no longer needed or have been removed from the chassis, searching data points with filters and controlling the SNMP polling.

6. Conclusions

The CERN community will be upgrading to the new ATCA standards as a replacement of the current VME technology. The work up to now has facilitated the understanding of these new systems. The framework described in this proceedings is an expert tool that will assist in the integration of new hardware into the current DCS. The tool provides a platform that allows the development of custom monitoring scripts and panels that can be used by any sub-detector in ATLAS. The tool automates searching, sorting and creation of large complicated back-end infrastructure and provides a scalable framework to assist in the integration efforts.

7. Acknowledgements

The University of the Witwatersrand has facilitated the efforts by purchasing the required hardware to allow hands on development for this work and for other big science applications. The National Research Foundation (NRF) for the bursary they provided. The SA-CERN program provided financial assistance for research visits to CERN. The authors 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 2008 *Journal of Instrumentation* **3** S08003–S08003 ISSN 1748-0221 URL `http://stacks.iop.org/1748-0221/3/i=08/a=S08003`

[2] ATLAS C 2012 Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment URL `http://cds.cern.ch/record/1502664`

[3] 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* URL `http://cds.cern.ch/record/1628753`

[4] Ballestreroa S *et al.* 2014 ATCA in ATLAS Tech. rep. URL `https://edms.cern.ch/file/1304001/1/ATCA-Backupdoc-rev2.pdf`

[5] PICMG About Us URL `http://www.picmg.org/v2internal/aboutus.htm`

[6] Lantzsch K *et al.* 2012 *Journal of Physics: Conference Series* **396** 012028 ISSN 1742-6588 URL `http://stacks.iop.org/1742-6596/396/i=1/a=012028`

[7] 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`

GPU-based Computation of Energy and Time for the Upgrade of the Tile Calorimeter of the ATLAS Detector

Marc Sacks, Bruce Mellado

School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africa

E-mail: marc.sacks@cern.ch

Abstract. After the expected 2022 upgrade of the Large Hadron Collider, increased running luminosity will necessitate the redesign of the front-end and back-end detector electronics. The massive increase in raw data generated by the ATLAS Tile Calorimeter (TileCal) as well as the increased out of time pile-up will require more sophisticated hardware and signal processing algorithms for accurate energy reconstruction. ARM processors are common in mobile devices due to their low cost, low energy consumption and good performance. Coupled with high performance general purpose GPU processing, a cost-effective, high data throughput Processing Unit (PU) is being developed at the University of the Witwatersrand. This PU, working in tandem with the planned Tile Pre-Processor (TilePPr), will enable computationally complex signal processing on the TileCal raw data while maintaining minimal software design difficulty for the end-user. The GPU will allow for parallel processing, which can be used effectively in the context of the detector which has nearly 10000 independently read-out channels. An overview of the planned PU design is given. Performance and throughput test results for several implementations of algorithms on the NVIDIA TK1 GPU are given.

1. Introduction

A Toiroidal LHC ApparatuS (ATLAS) is a detector in the Large Hadron Collider (LHC) at CERN [1]. It is a general purpose detector used in the search for Standard Model and beyond the Standard Model physics, comprising several sub-detectors; namely the inner detector (ID), the electromagnetic and hadronic calorimeters, and the muon spectrometer. These detectors measure the energy and trajectories of the particles formed within it due to high-energy proton-proton (p-p) and heavy ion collisions. The calorimeter has two subsystems, namely the electromagnetic (EM) and hadronic calorimeters. The EM detector is a liquid argon based detector, while the hadronic calorimeter is composed of sheets or tiles of plastic scintillators sandwiched between inactive plates of steel. The tiles of plastic scintillator have led to the hadronic calorimeter being referred to as the Tile Calorimeter or TileCal [2]. Proton bunches collide in ATLAS approximately every 50 ns. This bunch crossing rate produces hundreds of millions of collisions per second, which the TileCal read-out system reads out at a total bandwidth of about 205 Gbps [3]. After the upgrade of TileCal, in 2022, this will increase to 41 Tbps [3]. Storing this amount of data is neither possible nor desirable, as such a trigger system is used to separate rare physical processes from background events. The approach is to detect events of interest in real-time and to store those events for analysis; all other data

is discarded. The trigger system must also deal with increased pile-up, which will occur as a function of increased luminosity. To accommodate for the increased bandwidth and pile-up, new trigger systems are being developed [4]. The current and future trigger systems are composed of 3 and 4 separate trigger levels respectively, with a read-out driver (ROD) forming the interface between the Level-1 and Level-2 triggers [4]. The ROD is being upgraded to the so-called Tile Pre-Processor (TilePPr). This paper will describe a proposed co-processing unit (PU) for the TilePPrs based on ARM and graphic processing unit (GPU) technology. First a more complete description of the trigger system and associated hardware is given in section 2, followed by a description of GPUs and parallel processing systems in section 3, and finally a discussion of the TilePPr co-processing unit and results to some preliminary tests of trigger algorithms implemented on GPUs are given in section 4.

2. ATLAS Trigger System

On the boundary between the Level-1 and Level-2 triggers are the read-out drivers (ROD) [2]. The RODs are responsible for reading in data from the Level-1 trigger upon a Level-1 accept and transmitting this data to the next level of the trigger [3]. The RODs are also responsible for the reconstruction of the time-energy profiles of collision events, calculating the quality-factor of these reconstructions, synchronising data and trigger, compressing digital samples, calculating transverse energy, and other maintenance activities [3]. In Phase-II, the Level-1 trigger will be implemented at the level of the TilePPr [4].

2.1. Trigger Hardware

Current hardware designs are largely based on Field Programmable Gate Arrays (FPGA) [3]. FPGAs are unique pieces of hardware in that they behave somewhat like a blank slate of an integrated chip (IC). The user of an FPGA is in control of an array of programmable logic blocks and can configure them in many ways. For instance, an FPGA can be programmed to behave like a collection of operational amplifiers, or even to behave like something as complex as a CPU. In many ways an FPGA can be seen as a formattable ASIC and indeed can be used to emulate an ASIC before it is physically created. The current Level-1 trigger algorithm has been implemented as a massively parallel system on FPGAs and the entire TileCal processing chain relies on its rapid latency [3]. Parallel systems are described in the following section. The RODs and TilePPrs also make use of FPGAs, and additionally DSPs [3]. The HLT uses computer farms consisting of x86 CPU architecture, however research into the use of GPUs to speed up processing time is being investigated [5]. The versatility, customizability and performance of FPGAs are diminished by their high cost and difficulty of use. FPGA programming is low-level and unfamiliar to many programmers, requiring a knowledge of electrical circuits, and therefore they do not lend themselves to rapid prototyping and alteration of code.

3. Parallel Systems and Parallel Processing

Parallel processing refers to the concurrent processing of data by two or more processors [6]. Concurrent processing usually implies a speed-up as compared to the serial processing of the same system. That multiple processors can function simultaneously usually implies that they can function independently of one another. This is not always the case; processors may need to communicate with one another, and the extent of their independence can be used as an indicator of the degree of parallelism of the system being analysed. The degree of parallelism can in turn be used to estimate the amount of speed-up that can be expected from the parallel processing of a system. A simple example of a parallel system is the process of matrix multiplication, where entries of the output matrix can be calculated separately and simultaneously. TileCal can be viewed as a parallel system of 10000 channels before the trigger selection algorithm, as

the trigger might need to access one or many channels at once As such, the hardware used in the trigger system takes advantage of this parallelism.

3.1. Parallel Processing with Graphic Processing Units

Graphic Processing Units (GPU) are responsible for generating images for display on screen. This can be understood as a parallel process: pixels on a screen can be thought of as entries in a matrix and a GPU can process these elements independently. The parallel processing capabilities of GPUs has been extended into fields other than graphics processing and this is referred to as general purpose graphic processing units (GPGPU) [7]. GPUs are typically less expensive than FPGAs and they can be programmed in a variety of languages, for instance in a C/C++ platform developed by NVIDIA called Compute Unified Device Architecture (CUDA) [7]. The programming languages C, C++, and their derivatives are widely known and as such GPGPU makes it possible for non-specialists to program parallel code rapidly, as well as to understand, use, and alter existing code with ease.

4. TilePPr Co-Processing Unit

An ARM-based processing unit (PU) which will operate in conjunction with the TilePPr is being developed at the University of the Witwatersrand [8]. ARM processors are low-cost, energy efficient alternatives to traditional x86 architecture processors. As such they are found in many mobile devices (cellular phone, tablets) since the 1990s. ARM processors are performant in their own right, but to take advantage of the massively parallel environment of TileCal, the introduction of GPGPU into the PU is being investigated. Details such as the number and configuration of the GPUs in the PU are at this stage unclear. In the early testing phases of the PU, the GPU(s) will most likely interface with ARM processors with PCIe x3.

The PU is intended to offer the necessary computational power to deal with the increased pile-up of the HL-LHC. Algorithmic methods aimed at providing an improved calculation of quality-factor as well as improved resolution of events suffering from pile-up are under investigation [9]. These signal-processing techniques could be implemented partially or wholly on the PU. To determine what techniques could be implemented on the PU, further investigation into the performance of GPGPU in conjunction with ARM processors must be conducted.

Two algorithms were implemented here; the Optimal Filtering (OF) algorithm which is currently implemented on the RODs and used for online energy reconstruction, and an algorithm which reconstructs pulse shapes based on a deconvolution technique. OF calculates the amplitude, baseline, and phase shift of a given pulse using three sets of weights combined with a set of samples in a dot product operation. The deconvolution technique approximates the transfer function of the detector system and, knowing the output, can calculate the input. The deconvolution algorithm has shown promising results with regard to energy estimation in high pile-up environments especially when used in conjunction with OF. Details of the OF and deconvolution algorithms can be found in [9] and [10], for the purposes of this text it is sufficient to understand them as simple matrix multiplication operations.

The results in table 1 are based on the implementation of these two algorithms on the NVIDIA TK1 GPU. The table gives results of the performance of the OF algorithm alone and OF running in parallel with the deconvolution algorithm.

The algorithms were run on randomly generated static data; future implementations will use streamed simulated detector data. The following observations can be made regarding the results shown in table 1. Computation time increases linearly with number of events when running both algorithms. However, it can be seen that when only the OF algorithm is performed on the events, the performance is drastically reduced. Performing only OF on the data results in a computation time up to 7.5× slower than with the combined algorithms. Because the GPU is performing both algorithms in parallel, it is more efficient to run multiple algorithms on the same data set

Table 1. GPU Performance for OF Only and OF and Deconvolution Combined

Number of Events	Computation Time (μs)	GFlop/s[†]
96 (OF Only)	565	0.09
96	81	0.61
128	167	0.39
256 (OF only)	2136	0.06
256	295	0.44
512	555	0.47
512 (OF only)	2826	0.06
1024	1085	0.48
4096	4221	0.50

[†]b of floating point operations per second

simultaneously. This is desirable in the context of the PU where running as many algorithms as possible on a given data set will enable the passing of richer results to the Level-2 trigger. The poor performance of the GPU when running OF only is also due to the way hardware manages memory. The layout of GPU memory is more difficult to manage than a typical CPU and performance is heavily reliant on this. To illustrate this consider the results shown in table 2.

Table 2. GPU Performance for Alternate Implementation of OF and Convolution

Number of Events	Computation Time (μs)	GFlop/s
128	86	0.74
512	163	1.61
4096	1066	0.61

This alternate implementation of the algorithms is achieved by varying the dimensions of the matrices undergoing multiplication, and results in a much reduced computation time. The implementation of each individual algorithm as well as the interactions of the algorithms with each other will be the subject of future investigation.

5. Conclusion

The HL-LHC will require an upgraded trigger system to deal with increased data production and increased pile-up. A co-processing unit for the TilePPr based on ARM and GPU technology is under investigation. The parallel nature of the energy reconstruction in the TileCal detector makes the use of parallel processing technology an attractive option. It is possible to implement a parallel system with FPGAs, however they are expensive and difficult to program. GPGPU provides a less expensive alternative and one which is also easier to program and update. Preliminary test results indicate that GPGPU could be used in the high throughput environment of TileCal as part of a PU operating in assistance to the TilePPr.

References

[1] The ATLAS Collaboration *et al* 2008 *JINST* **3** S08003
[2] ATLAS Trigger Group *et al* 2008 *JINST* **3** P03001
[3] Carrió F, Mellado B, Reed R *et al* 2014 *JINST* **9** C02019
[4] ATLAS Collaboration 2012 CERN-LHCC-2012-022. LHCC-I-023
[5] Halyo V, Hunt A, Jindal P, LeGresley P, Lujan P 2013 *JINST* **3** P10005
[6] Pacheco P S 2011 *An Introduction to Parallel Programming* (USA: Elsevier Inc)
[7] Sanders J, Kandrot E 2010 *CUDA by Example* (Addison-Wesley Professional) 2-7
[8] Cox M, Reed R, Mellado B 2015 *JINST* **10** C01007
[9] Balabram L E, Seixas J M, Filho L M 2014 Quality Factor for the Hadronic Calorimeter in High Luminosity Conditions *J. Phys.: Conf. Ser.* 608 012044 http://iopscience.iop.org/1742-6596/608/1/012044)
[10] Stärz S 2012 *JINST* **7** C12017

Division C – Photonics

Gene expression changes in diabetic wound healing as induced by photobiostimulation *in vitro*

S M Ayuk, H Abrahamse and N N Houreld
Laser Research Centre, Faculty of Health Sciences, University of Johannesburg, P.O. Box 17011, Doornfontein, Johannesburg, 2028, South Africa
Tel: +27 11 559-6833 Fax +27 11 559-6884

Email: nhoureld@uj.ac.za

Abstract. Diabetes Mellitus (DM) is a complex metabolic disorder resulting in hyperglycaemia. Impaired wound healing is a serious complication of diabetes, and is a severe public health problem. Photobiostimulation is a non-invasive form of treatment known to enhance healing of such wounds using low energy lasers. This study investigated the changes in extracellular matrix (ECM) gene expression in diabetic wounded fibroblasts *in vitro* after photobiostimulation at 830 nm. Normal (N-unstressed), normal wounded (NW-stressed) and diabetic wounded (DW-stressed) fibroblasts were incubated for 48 h after irradiation using a continuous wave diode laser at a wavelength of 830 nm with 5 J/cm^2. Non-irradiated cells (0 J/cm^2) were used as controls. The gene expression profile (84 genes) was assessed using an ECM real-time reverse transcription polymerase chain reaction (RT-PCR) array with the appropriate controls included. Sixty one genes were significantly regulated (55 up-regulated and 6 down-regulated) in N-cells; 40 genes (20 up-regulated, and 20 down-regulated) in NW-cells; and 42 genes (9 up-regulated and 33 down-regulated) in DW-cells. Several genes were down-regulated in DW-cells as compared to N- and NW-cells. Photobiostimulation modulated the expression of important genes in wound healing, including cell adhesion molecules, integrins, ECM proteins, proteases, and inhibitors involved in the ECM. An in depth comprehension of the molecular and biological processes may create an improved therapeutic protocol for the treatment of diabetic wounds.

1. Introduction

Diabetes Mellitus (DM) is a complex metabolic disorder resulting in hyperglycaemia (high glucose). Complications of diabetes and other non-healing ulcers remain a major concern in public health. Global statistics show that millions of people suffer from this disease, with a prevalence of 15% with diabetic foot ulcers and 3% with lower limb amputation [1]. Wound healing involves a series of biological processes that must occur appropriately within a given time, and necessitates the interaction of various cells and growth factors to ensure proper healing [2-4]. Fibroblasts are very important in wound healing; they maintain homeostatic balance, and through cell proliferation, differentiation and extracellular matrix (ECM) development stimulates the process of tissue repair by synthesising collagen around the matrix [5,6]. Chronic ulcers may result from an imbalance between ECM formation and degradation [7]. Hyperglycaemia can affect these cellular and biological activities through oxidative stress or alkaline

glycation end (AGE) products which in turn affect changes in genes encoding for ECM proteins and proteases [8], leading to impair wound healing [9]. Several investigations have shown diabetes to affect gene expression as well as wound healing [10,11]. Photobiostimulation is a non-invasive form of light treatment known to enhance healing of wounds. Photobiostimulation at various wavelengths and dosages is known for its stimulatory effect, and enhances wound healing in animal and clinical studies, as well as cell cultures [12-14]. Literature has shown that photobiostimulation is able to effect the expression of genes in animal and cell culture models [15-17]. In order for these biological processes to be achieved, wavelengths within the visible and Near Infra-Red (NIR) spectrum are implemented during treatment with a power output of 10-200 mW. This treatment is acceptable in many clinical practices however disputes on protocol and treatment specifications, for different models, create room for more research [13]. The purpose of this study was to evaluate the expression of extracellular matrix genes in diabetic wounded cells after photobiostimulation with 830 nm light.

2. Materials and methods

2.1. Cell Culture and Laser Irradiation

Isolated human skin fibroblasts (Academic ethics Committee, Clearance Reference Number: 01/06, University of Johannesburg) were used to create various models, namely normal (N-unstressed, Figure 1A), normal wound (NW-stressed, Figure 1B) and diabetic wounded (DW-stressed). Cells were cultured via standardised protocols [18]. Approximately 6 x 10^5 cells was re-suspended in growth media and seeded in 3.4 cm diameter culture dishes. Using media with a basal concentration of 5.6 mM/L D-glucose, an *in vitro* diabetic model was established by adding 17 mM/L D-glucose [19]. The Council for Scientific and Industrial Research (CSIR) / National Laser Center (NLC) of South Africa provided the lasers. The stressed wounded models were created by introducing a central scratch on a confluent monolayer of cells using a 1 ml sterile disposable pipette (Figure 1B) [20,21] and left to stabilise at 37^0C in 5% CO_2 for 30 min, after which laser irradiation was performed at 830 nm (Fremont, California, USA, RGBlase, TECIRL-70G-830SMA or FC-830-300-mm2-sma-1-0) (continuous wave; 98 mW; 9.1 cm^2; 10.76 mW/cm^2; 5 J/cm^2; duration of irradiation 7 min 43 s). Non-irradiated (untreated, 0 J/cm^2) N-cells were used as controls for irradiated N- and NW-cells, while non-irradiated (untreated, 0 J/cm^2) DW-cells were used as controls for irradiated DW-cells. This was due to significant differences in gene expression between non-irradiated NW- and DW-cells. Post irradiation cells were incubated for 48 h. Total RNA was isolated and quantified, and 1 µg was reverse transcribed into cDNA and the gene profile was determined making use of the human ECM and adhesion molecules RT² Profiler™ real-time reverse transcriptase polymerase chain reaction (RT-PCR) Array (SA Biosciences, PAHS-0132Z) [15]. The statistical analysis was achieved on three repeats of each sample.

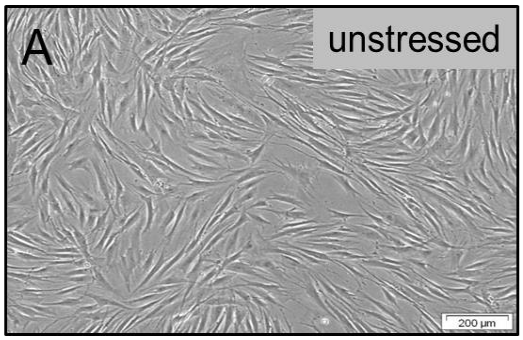


Figure 1A. Micrograph of unstressed fibroblasts seeded to 70-80% confluency.

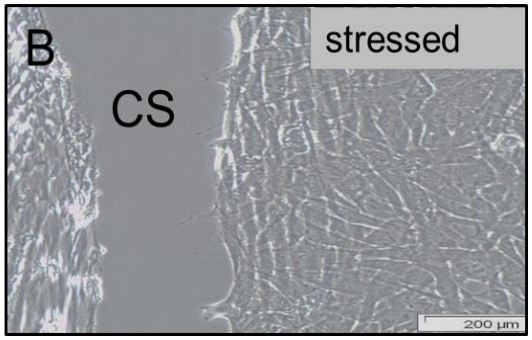


Figure 1B. Stressed fibroblasts created with central scratch (CS).

3. Results The results showed that laser photobiostimulation mediated the expression of 84 genes. Sixty one genes were significantly regulated (55 up-regulated and 6 down-regulated) in N-cells; 40 genes (20 up-regulated, and 20 down-regulated) in NW-cells and 42 genes (9 up-regulated and 33 down-regulated) in DW-cells. There was a significant change in gene expression as compared to the respective controls (Table 1).

Table 1. Summarises the gene expression of N-, NW- and DW-cells. The total RNA from non-irradiated (N- and DW-) and irradiated (N-, NW- and DW-) cells were characterised in triplicates. The gene expression profile is denoted as gene up-regulation in black or gene down-regulation in bold.

Key (Black-Up-regulation and Bold-Down-regulation)	**Normal**	**Normal Wounded**	**Diabetic Wounded**
Cell adhesion molecules participates in cell migration, cell proliferation, survival and differentiation			
Transmembrane Molecules	CD44, CDH1, HAS1, ICAM1, ITGA1, ITGA2, ITGA3, ITGA5, ITGA6, ITGA7, ITGA8, ITGAL, ITGAM, ITGB2, ITGB3, ITGB4, MMP14, MMP15, MMP16, NCAM1, PECAM1, SELE, SELL, SELP, SPG7, **ITGB5**	ITGA5, ITGA6, ITGA7, ITGA8, ITGAV, ITGB2, ITGB4, MMP15, PECAM1, SELP, **HAS1, ITGA2, ITGAL, ITGB1, SGCE, VCAM1**	MMP15, SELE, SELL, SGCE, VCAM1, **CD44, HAS1, ICAM1, ITGA1, ITGA2, ITGA3, ITGA5, ITGAV, ITGB1, ITGB3, MMP14, SPG7**
Cell-Cell Adhesion	CD44, CDH1, COL14A1, COL6A2, ICAM1, ITGA8, **COL11A1**	ITGA8, **COL14A1, VCAM1**	COL11A1, COL14A1, VCAM1, **CD44, COL6A2, CTNND1**
Cell-Matrix Adhesion	ADAMTS13, ITGA1, ITGA2, ITGA3, ITGA5, ITGA6, ITGA7, ITGA8, ITGAL, ITGAM, ITGB2, ITGB3, ITGB4, THBS3, **ITGB5**	ITGA5, ITGA6, ITGA7, ITGA8, ITGAV, ITGB2, ITGB4**, ITGA2, ITGAL, ITGB1, SGCE, SPP1**	**CD44, ITGA1, ITGA2, ITGA3, ITGA5, ITGAV, ITGB1, ITGB3, SPP1, THBS3**
Other Adhesion Molecules	CNTN1, COL5A1, COL6A1, COL7A1, COL8A1, CTNNA1, FN1, KAL1, LAMA1, LAMA2, LAMA3, LAMC1, THBS2, CLEC3B, VTN, **VCAN, CTGF**	COL12A1, COL5A1, KAL1, LAMA3, CLEC3B, **VCAN, CTGF, CTNND2, LAMA1, LAMA2, LAMC1, TNC**	VCAN, **COL12A1, COL16A1, COL5A1, COL6A1, COL7A1, CTNNB1, FN1, KAL1, LAMA1, LAMA3, LAMB3, THBS1**
Extracellular Matrix maintains skin integrity and homeostasis and relates with several structural and extracellular proteins			
Basement Membrane Constituents	LAMA1, LAMA2, LAMA3, LAMC1	COL7A1, SPARC**, LAMA1, LAMA2, LAMC1**	**COL7A1, LAMA1, LAMA3, LAMB3**
Collagens and ECM Structural Constituents (Synthesis of ECM)	COL5A1, COL6A1, COL7A1, COL8A1, FN1, KAL1, **COL11A1, COL1A1**	COL12A1, COL1A1, COL5A1, KAL1, **COL14A1**	COL11A1, COL14A1, **COL12A1, COL16A1, COL5A1, COL6A1, COL6A2, COL7A1, COL14A1, FN1, KAL1**
ECM Proteases (MMPs responsible for ECM degradation)	ADAMTS1, ADAMTS13, ADAMTS8, MMP11, MMP12, MMP13, MMP14, MMP15, MMP16, MMP2, MMP7, MMP8, MMP9, **MMP1**	MMP15, MMP2, MMP8, TIMP1, **MMP1, MMP10, MMP12, MMP3**	ADAMST8, MMP15, **ADAMST1, MMP1, MMP14, MMP2, MMP3, MMP8, SPG7**
ECM Protease Inhibitors	COL7A1, KAL1	KAL1	**COL7A1, KAL1, THBS1**
Other ECM Molecules	HAS1, THBS2, CLEC3B, VTN, **VCAN, CTGF**	CLEC3B, **VCAN, CTGF, ECM, SPP1**	**HAS1, SPP1**

4. Discussion and Conclusion

Photobiostimulation affects various cellular responses both *in vivo* and *in vitro*, as well as gene expression [15-17,22,23] This study focused on the gene expression profile of normal, normal wounded and diabetic wounded human skin fibroblast cell models. The explanation behind these findings could be the change in the respiratory chain of mitochondria to release energy in the form of adenosine tri phosphate (ATP) as well as reactive oxygen species (ROS) leading to the release of transcription factors stimulating gene transcription and eventually ECM formation, cell proliferation, collagen production, growth factor production, as well as the expression of various proteins [24]. Wound healing is impaired in diabetes due to the poor synthesis of growth factors and proteins [11,25,26], as well as cell-ECM interaction [27]. In this study, the main components affected by photobiostimulation in the ECM and adhesion molecules are the cell adhesion molecules (Integrins, Cadherins, Immunoglobulins and Selectins) as well as ECM molecules (Fibronectin, Collagen, Laminin, Matrix Metalloproteinases (MMPs) and their inhibitors). This study showed that 55 genes in N-cells were significantly up-regulated, while only 20 genes in NW-cells and nine genes in DW-cells were up-regulated. Genes in DW-cells were significantly down-regulated possibly as a result of the ECM dysregulation due to hyperglycaemia. In addition, the inflammatory process is delayed enhancing the release of proteases such as MMPs promoting ECM degradation and vital growth factors and receptors responsible for wound healing. The dysregulation may also prevent integrins from binding with fibronectin, and hence impairing migration [28,29]. According to Matsumoto and colleagues [30] a concentration of 1.5% glucose impaired wound healing in mesothelial cells. Their study showed that photobiostimulation in a paracrine or autocrine fashion regulates gene expression in the ECM in isolated skin fibroblasts *in vitro*. Similarly, Peplow and colleagues [23] reviewed 17 papers on gene expression in human and animal cell cultures and confirmed that photobiostimulation at green, red, and NIR stimulated gene expression, even though further work needs to be done to elucidate its cellular effects. In addition, McDaniel *et al.,* [31] exposed human skin fibroblasts to multiple wavelengths (590/870 nm LED array) and concluded that simultaneous exposure can affect cell metabolism as well as gene expression. They suggested that gene expression patterns in fibroblasts could be influenced by adjusting the wavelength intensities in both visible and NIR treatments.

In conclusion, photobiostimulation modulates gene expression of cell adhesion molecules, integrins, ECM proteins, proteases, and inhibitors involved in the ECM at a wavelength of 830 nm. This study confirms that hyperglycaemia is responsible for impaired wound healing in cell cultures *in vitro*. Furthermore an in depth comprehension of the molecular and biological processes may create an improved therapeutic protocol for diabetic wounds and future studies conducted on gene modulation and their receptors for improve understanding.

Acknowledgement

Appreciation goes to the CSIR National Laser Centre, African Laser Centre, and the University of Johannesburg, South Africa for their endless support.

References

[1] International Diabetes Federation, IDF 2014 *Diabetes Atlas*, 6th ed. Brussels, Belgium.
[2] Koh T J. and DiPietro L A. 2011 *Expert Rev Mol Med.* **13** e23.
[3] Grieb G, Steffens G, Pallua N, Bernhagen J and Bucala R. 2011 *Int Rev Cell Mol Biol.* **291** 1
[4] Penn J W, Grobbelaar A O. and Rolfe K J. 2012 *Int J Burn Trauma* **2** 18
[5] Rhee S. 2009 *Exp Mol Med*. **41** 858.
[6] Broughton G 2nd, Janis J.E. and Attinger C E. 2006 *Plast Reconstr Surg.* **117** 12S
[7] Li Z, Guo S, Yao F, Zhang Y. and Li T. 2013 *J Diabetes Complications*. **27** 380
[8] Mason R M. and Wahab N A. 2003 *J Am Soc Nephrol*. **14** 1358

[9] Guo, S. and DiPietro, L.A. *J Dent Res*. **89** 219
[10] Ge K, Wu J J, Qian L, Wu M J, Wang F L, Xu B and Xie T 2015 *Genet Mol Res.* **14** 4802
[11] Sun H, Mi X, Gao N, Yan C and Yu F S. 2015 *Invest Ophthalmol Vis Sci.* **56** 3383
[12] Dawood M S. and Salman S D. 2013 *Lasers Med Sci.* **28** 941
[13] Huang Y Y, Sharma S K, Carroll J. and Hamblin M R. 2011 *Dose Response.* **9** 602
[14] Silveira P C L, Silva L A, Freitas P T, Latini A, Pinho R A 2011 *Lasers Med Sci* **26** 125
[15] Ayuk S M, Houreld N N and Abrahamse H. 2014 *Int J Photoenergy*, **2014** 17.
[16] Houreld N N, Ayuk S M and Abrahamse H. 2014 *J Photochem Photobiol B: Biology* **140** 146
[17] Rodrigues N C, Brunelli R, de Araújo H S, Parizotto N A. and Renno A C. 2013 *J Photochem Photobiol B. Biology* **120** 29
[18] Hawkins D. and Abrahamse H. 2005 *Photomed Laser Surg*. **23** 251
[19] Houreld, N. and Abrahamse, H. (2007) *Diabetes Technol Ther.* **9** 451
[20] Cory, G. Scratch‐wound assay.2011 *Methods Mol Biol* **769** 25
[21] Goetsch K P. and Niesler C U. 2011 *Anal Biochem*. **411** 158
[22] Ayuk S M, Houreld N N. and Abrahamse H. 2012 *Diabetes Technol and Therapeut.* **14** 1110
[23] Peplow P V, Chung T-Y, Ryan B, Baxter G D. 2011 *Lasers Surg Med.* **43** 843
[24] Karu T. *Biomedical Photonics Handbook* 2003 **48.** 1
[25] de Loura Santana C, de Fátima Teixeira Silva D, Deana A M, Prates R A, Souza A P, Gomes M T, de Azevedo Sampaio B P, Shibuya J F, Bussadori S K, Mesquita-Ferrari R A, Fernandes K P. and França C M. 2015 *PLoS One.* **10** e0122042.
[26] Mohd Hilmi A B, Hassan A. and Halim A S. 2015 *Adv Wound Care* **4** 312
[27] Watt F M. and Fujiwara H. 2011 *Cold Spring Harb Perspect Biol*. **3** pii: a005124
[28] Muller M, Trocme C, Lardy B, Morel F, Halimi, S. and Benhamou, P Y. 2008 *Diabet Med*, **25** 419
[29] Ongenae K C, Phillips T J. and Park H-Y. 2000 *Dermatol Surg* **26** 447
[30] Matsumoto M, Tamura M, Miyamoto T, Furuno Y, Kabashima N, Serino R, Shibata T. Kanegae K, Takeuchi M, Abe H, Okazaki M. and Otsuji,Y. 2012 *Life Sci.* 90
[31] McDaniel D H, Weiss R A, Geronemus R G, Mazur C, Wilson S. and Weiss M A. 2010 *Lasers Surg Med.***42** 540

Synchrotron modeling of the gamma-ray to optical afterglow of GRB 130427A and expected neutrino flux

Jessymol K Thomas, Reetanjali Moharana and Soebur Razzaque

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

E-mail: jessymolkt@uj.ac.za, reetanjalim@uj.ac.za, srazzaque@uj.ac.za

Abstract. GRB 130427A, at redshift 0.34, is one of the brightest and most energetic long duration gamma-ray bursts (GRBs) ever detected. A 95 GeV photon, the highest energy ever detected from a GRB, has also been reported by the Large Area Telescope (LAT) on board the Fermi Gamma ray Space Telescope. Simultaneous observations in the gamma ray, X-ray and ultraviolet/optical ranges, make this GRB one of the most well studied in history. We have modeled temporal the temporal evolution of flux in different frequencies and the broadband spectral energy distribution at different time intervals by using optically thin synchrotron radiation from a relativistic blast wave expanding in a constant density interstellar medium and in a wind-type medium with density gradient. We find that the afterglow of GRB 130427A is better described in case of a wind-type medium. We also calculate the expected neutrino flux from this GRB, if protons are accelerated to ultrahigh energies in the blast wave and interact with afterglow photons. Neutrino telescopes which are currently operating and those being planned for the future should be able to either detect this flux or, at least, constrain ultrahigh energy cosmic ray acceleration in GRBs.

1. Introduction

The Gamma Ray Burst detected on 27th April 2013 (GRB 130427A) is one of the most energetic explosions, with redshift z=0.34. The burst took place only a quarter of the way across the universe and it is one of the brightest gamma ray bursts (GRBs) ever recorded. GRB 130427A has the highest fluence recorded and has the longest gamma ray emission in the 100 MeV energy range for a duration of 20 hours [1] . It was such a hyper energetic burst that we could record a 95-GeV photon at 244 seconds and a 32-GeV photon after 34366 seconds [1]. NASA's Fermi Gamma ray Space Telescope with the Gamma ray Burst Monitor (GBM) and the Large Area Telescope (LAT) and the Swift Gamma Ray Burst explorer, the multi wavelength space observatory with all its three instruments, the Burst Alert Telescope (BAT),the X-ray Telescope (XRT) and the Ultraviolet/Optical Telescope (UVOT), have observed GRB 130427A and its afterglows in the gamma ray, X-ray, ultraviolet, and optical energy ranges. The emission was also detected at radio, infrared and visible wavelengths with ground based telescopes. The X-ray afterglow is recorded by the Swift at photon energies 0.3-10 keV and by the NuSTAR at 3-80 keV, for more than a week [2]. The RAPTOR telescopes have also observed the direct optical and infrared afterglows for a few hours [3].

Most of the models of GRBs and the afterglows are based on the synchrotron emission from the relativistic electrons accelerated within the forward shock [4, 5]. The ejecta of the GRBs

interact with the interstellar medium and produces afterglow radiation. This radiation can be described by the synchrotron emission of electrons in the blast wave. With time the intrinsic properties of the blast wave like Lorentz boost factor Γ, the internal radius R, in turn the energy density and magnetic field changes. The synchrotron spectrum of electron is affected by the minimum Lorentz boost factor $\gamma_{e,min}$ and cooling Lorentz boost factor $\gamma_{e,c}$ of electron, where it starts to produce Synchrotron radiation. These Lorentz factors of electron depends on the intrinsic parameters of the blast wave. So we expect a temporal evolution of the radiation flux which can be attained by modeling of the blast wave. In case of , $\gamma_{e,c} < \gamma_{e,min}$, all the electrons cool to be in fast cooling regime. Whereas for $\gamma_{e,c} > \gamma_{e,min}$, the Synchrotron spectrum is in slow cooling regime. We have modeled the broadband spectral energy distribution (SED) at different time intervals and the temporal evolution of the flux for different frequencies, by using optically thin synchrotron radiation from a relativistic blast wave expanding in a constant density interstellar medium (ISM) and in a wind-type medium in which the density decreases as $1/R^2$, where R is the radial distance from the center of the GRB. The data for the spectral energy distribution and light curves of GRB 130427A are better explained with wind environment compared to the ISM environment. GRB spectra are of significant importance in the study of the source environment and the formation of the jet.

2. Synchrotron spectrum model and light curves of GRB 130427A

To get the overall instantaneous synchrotron spectrum of electrons we need to model the blast wave evolution. The synchrotron frequencies, ν_m, ν_c and ν_s correspond to the minimum energy, cooling energy and the saturation energy of the electrons respectively. ν_a is the self absorption frequency.

The instantaneous synchrotron spectrum follows a power law with index p for an electron energy E is, $N(E) = N_0 E^{-p}$. The observed photon flux in the slow cooling regime, as given in [5] is:

$$F_\nu = F_{\nu,max} \begin{cases} (\frac{\nu}{\nu_a})^2 (\frac{\nu_a}{\nu_m})^{\frac{1}{3}};\ \nu < \nu_a, \\ (\frac{\nu}{\nu_m})^{\frac{1}{3}};\ \nu_a < \nu < \nu_m, \\ (\frac{\nu}{\nu_m})^{\frac{-(p-1)}{2}};\ \nu_m < \nu < \nu_c, \\ (\frac{\nu_c}{\nu_m})^{\frac{-(p-1)}{2}} (\frac{\nu}{\nu_c})^{\frac{-p}{2}};\ \nu_c < \nu_s. \end{cases} \tag{1}$$

The equations for the synchrotron frequencies in wind environment are given as [6, 7]:

$$\nu_m = 9.5 \times 10^{13} \epsilon_{b,0.1}^{1/2} \epsilon_{e,0.1}^{2} E_{55}^{1/2} (1+z)^{1/2} t_d^{-3/2}\ Hz, \tag{2}$$

$$\nu_c = 2.1 \times 10^{15} \epsilon_{b,0.1}^{-3/2} E_{55}^{1/2} (1+z)^{-3/2} t_d^{1/2} A_*^{-2}\ Hz, \tag{3}$$

$$\nu_a = 8.3 \times 10^{9} \epsilon_{b,0.1}^{1/5} \epsilon_{e,0.1}^{-1} E_{55}^{-2/5} (1+z)^{-2/5} t_d^{-3/5} A_*^{6/5}\ Hz, \tag{4}$$

$$\nu_s = 8.22 \times 10^{22} A_*^{-1/4} E_{55}^{1/4} (1+z)^{-3/4} t_d^{-1/4} \phi_1^{-1}\ Hz, \tag{5}$$

and $F_{\nu,max}$, the observed peak flux, is given as:

$$F_{\nu,max} = 3.53 \times 10^{-24} \epsilon_{b,0.1}^{1/2} E_{55}^{1/2} (1+z)^{-1/2} t_d^{-1/2} d_{l,28}^{-2}\ erg/cm^2/s/Hz. \tag{6}$$

Here E_{55} is the initial kinetic energy of the blast wave in units of 10^{55} ergs, ϵ_e is the electron equipartition fraction or the fraction of energy going to the relativistic electrons, $\epsilon_{e,0.1} = \epsilon_e/0.1$, ϵ_b is the fraction of energy going to the magnetic energy, $\epsilon_{b,0.1} = \epsilon_b/0.1$, $d_{l,28}$ is the luminosity

distance in units of 10^{28} cm unit, t_d is the time after the prompt emission in days, ϕ is the number of gyroradii needed for the electron acceleration in the magnetic field and $\phi_1 = \phi/10$, $A_* \equiv \dot{M}_{-5}/v_8$ corresponding to a mass-loss rate of $\dot{M}_w = 10^{-5}\dot{M}_{-5}M_\odot yr^{-1}$ in wind by the progenitor star with velocity $v_w = 10^8 v_8$ cm/s.

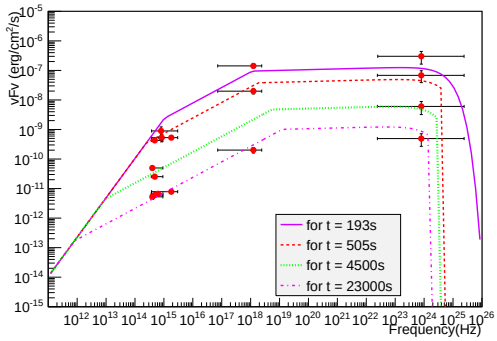


Figure 1. Spectral energy distribution in wind medium. Here the red points are the observed data with errors.

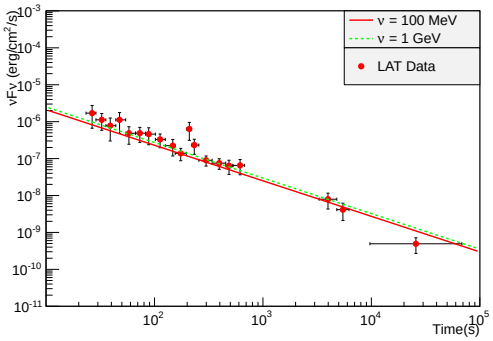


Figure 2. Light curve for LAT data in wind medium.

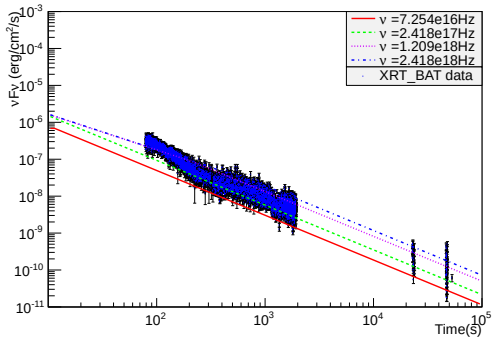


Figure 3. Light curve for XRT-BAT data in wind medium.

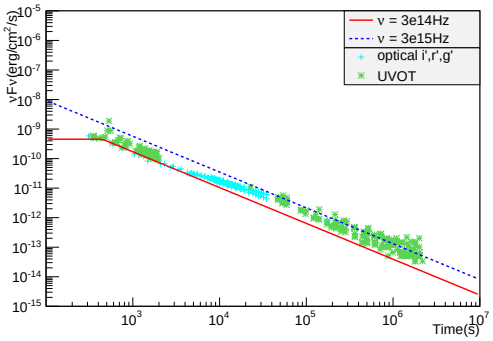


Figure 4. Light curve for UVOT/optical data in wind medium.

Figure 1 shows observed LAT, XRT-BAT and UVOT/optical data from GRB 130427A at different times and corresponding synchrotron model fluxes at those times. The data points are average fluxes in time intervals around the reported times. Here the model parameters are, $\epsilon_{b,0.1} = 0.025$, $\epsilon_{e,0.1} = 0.095$, $E_{55} = 0.5$, $p = 1.95$, $A_* = 0.1$, $z = 0.34$, $d_{l,28} = 0.56$ and $\phi_1 = 0.1$. In Fig. 2, Fig. 3 and Fig. 4, the energy fluxes are plotted as a function of time starting from the trigger of GRB 130427A and fitted with the model parameters for different frequencies using LAT, XRT-BAT and UVOT/optical data, respectively.

We do the same synchrotron modeling of GRB 130427A afterglow data for interstellar medium. In ISM, the different frequencies and the observed peak flux are:

$$\nu_m = 1.644 \times 10^{14} \epsilon_{b,0.1}^{1/2} \epsilon_{e,0.1}^{2} E_{55}^{1/2} (1+z)^{1/2} t_d^{-3/2} \; Hz, \tag{7}$$

$$\nu_c = 1.931 \times 10^{13} \epsilon_{b,0.1}^{-3/2} E_{55}^{-1/2} (1+z)^{-1/2} t_d^{-1/2} n^{-1} \; Hz, \tag{8}$$

$$\nu_a = 5.53 \times 10^9 \epsilon_{b,0.1}^{1/5} \epsilon_{e,0.1}^{-1} E_{55}^{1/5} (1+z)^{-1} n^{3/5} \ Hz, \tag{9}$$

$$\nu_s = 5.334 \times 10^{22} n^{-1/8} E_{55}^{1/8} (1+z)^{-5/8} t_d^{-3/8} \phi_1^{-1} \ Hz, \tag{10}$$

$$F_{\nu,max} = 7.95 \times 10^{-23} (1+z)^{-1} n_0^{1/2} d_{l,28}^{-2} E_{55} \epsilon_b^{1/2} \ erg/cm^2/s/Hz. \tag{11}$$

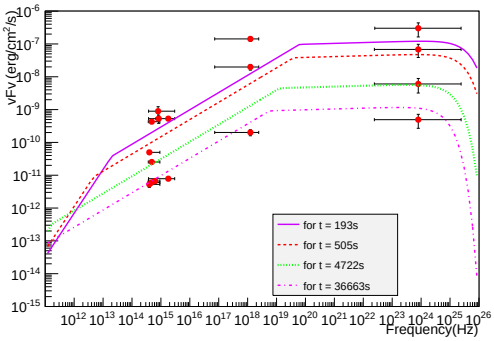


Figure 5. Spectral energy distribution in ISM.

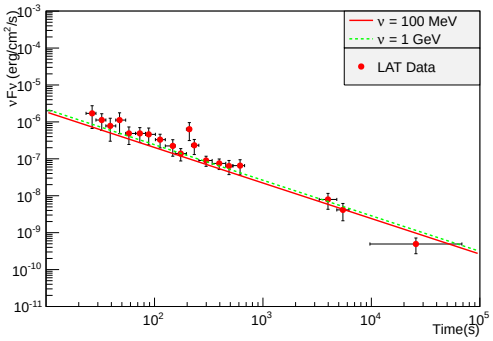


Figure 6. Light curve for LAT data in ISM.

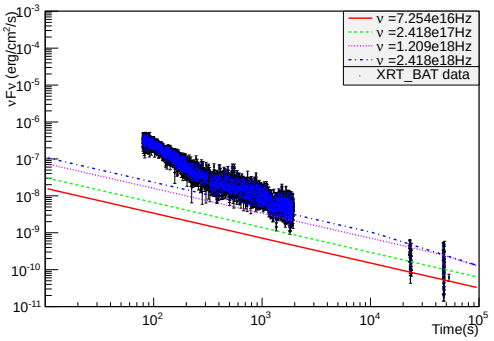


Figure 7. Light curve for XRT-BAT data in ISM.

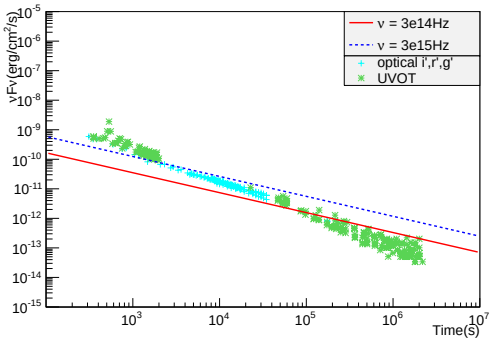


Figure 8. Light curve for UVOT/optical data in ISM.

Fig. 5 shows the spectral energy distribution for the GRB 130427A in ISM. The fitting of data for ISM is not as good as we obtained for wind environment. The model parameters used for the spectrum fit are: $\epsilon_{b,0.1} = 4 \times 10^{-4}$, $\epsilon_{e,0.1} = 0.028$, $E_{55} = 0.5$, $z = 0.34$ and $d_{l,28} = 0.56$, $p = 1.95$, $n = 1$ and $\phi_1 = 0.1$. Fig. 6, Fig. 7 and Fig. 8 are the light curves for LAT, XRT-BAT and UVOT/optical data, respectively.

3. Neutrino flux calculation for GRB 130427A

Neutrinos are the only high energy particles which can directly convey astronomical information from far and from the most high energy processes in the universe. Long-duration GRBs like GRB 130427A are the candidate sources of Ultra High Energy Cosmic Rays (UHECRs). The interaction of these cosmic rays with synchrotron radiated photons can produce neutrinos [6, 8]. The calculation of the flux for these neutrinos is done here using the parameters obtained from the synchrotron emission model fit for GRB 130427A.

In GRBs, interaction of shock accelerated relativistic protons with low energy photons can produce ultra high energy muons, pions, neutrons and kaons. The interactions are as follows [6, 9, 10]:

$$p\gamma \rightarrow X\pi^+\pi^-, XK^+K^- \tag{12}$$

where, X can be protons.

$$p\gamma \rightarrow \Delta^+ \rightarrow n\pi^+ \tag{13}$$

Here we calculated the neutrino flux from pion and muon decay, where pions are produced via Δ^+ resonance from $p\gamma$ interaction in the GRB 130427A blast wave.

$$\pi^+ \rightarrow \mu^+ + \nu_\mu \rightarrow e^+ + \nu_e + \nu_\mu + \bar{\nu}_\mu \ . \tag{14}$$

To calculate the neutrino flux, the photon spectrum obtained from the synchrotron spectrum can be expressed in terms of proper density of the synchrotron photon $n'_\gamma(E')$ [5, 6]. Here $E' = \frac{h\nu}{\Gamma}(1+z)$ is the photon energy in the co-moving frame, where ν is defined in the equation(1), Γ is the bulk Lorentz factor and R is the radius of the blast wave. For the slow cooling:

$$n'_\gamma(E') = \frac{2d_L^2(1+z)F_{\nu,max}}{R^2c\Gamma E'(m)} \times \begin{cases} (\frac{E'}{E'_a})^2(\frac{E'_a}{E'_m})^{\frac{1}{3}}\frac{E'_m}{E'};\ E' < E'_a, \\ (\frac{E'}{E'_m})^{\frac{-2}{3}};\ E'_a < E' < E'_m, \\ (\frac{E'}{E'_m})^{\frac{-(p-1)}{2}};\ E'_m < E' < E'c \\ (\frac{E'_c}{E'_m})^{\frac{-(p-1)}{2}}(\frac{E'}{E'_c})^{-p/2}\frac{E'_m}{E'}e^{\frac{-E'}{E'_s}};\ E'_c < E' < E'_s. \end{cases} \tag{15}$$

Here we consider the slow cooling afterglow synchrotron spectrum and the $p\gamma$ opacity scales with proton energy in the slow cooling regime as [11] :

$$\tau_{p\gamma}(E_p) = \tau_{p\gamma}(E_{pl}) \times \begin{cases} (\frac{E_{ph}}{E_{pl}})^{\frac{k}{2}-\frac{1}{2}}(\frac{E_p}{E'_{Ph}})^{\frac{k}{2}};\ E_p < E_{ph}, \\ (\frac{E'_p}{E_{pl}})^{\frac{k}{2}-\frac{1}{2}};\ E_{ph} < E_p < E_{pl}, \\ 1;\ E_p > E_{pl}. \end{cases} \tag{16}$$

Where E_{pl} is the minimum energy corresponds to the break energy $h\nu_m$ in the afterglow synchrotron spectrum, E_{ph} is the energy corresponding to the break energy $h\nu_c$ and E_{ps} is the highest energy the protons. For details, see Ref. [6].

In our calculation, we have put a condition on opacity: $\tau_{p\gamma} \leq 3$. In order to calculate the neutrino flux, the proton flux $J_P(E_p)$ and the intermediate pion flux $J_\pi(E_\pi)$ need to be calculated first and finally the neutrino flux is calculated. For details, see Ref. [6, 11].

Expected neutrino fluxes from GRB 130427A at different time intervals (same as we considered in modeling) in a wind environment and in ISM are plotted in figures 9 and 10 respectively. The wind environment and ISM model parameters were used for the calculation of the respective neutrino fluxes.

4. Conclusions

In this work, we have done synchrotron modeling of GRB 130427A in a wind environment and in ISM. We studied the spectral energy distribution of the GRB at different time intervals and the temporal evolution of the flux for different frequencies, using data in gamma-ray, X-ray and optical ranges. This was done considering a slow cooling regime for the electrons, both in a wind environment and constant density ISM. The wind environment model described in this work seems to fit the data of GRB 130427A better than the ISM model, if applied with reasonable parameters. The GRB 130427A is expected to be a powerful source of ultra high

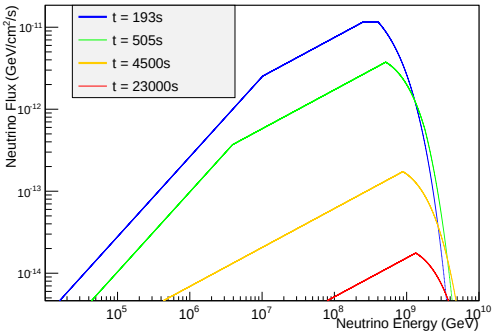


Figure 9. Neutrino flux for wind

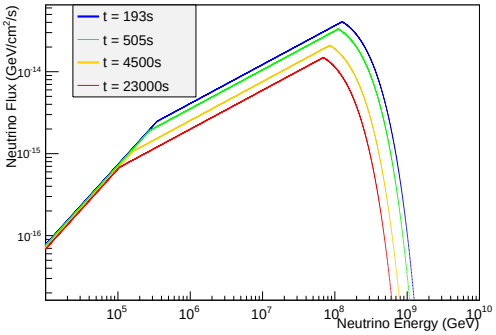


Figure 10. Neutrino flux for ISM

energy cosmic rays (UHECRs) and the interaction of these UHCRs with the burst's afterglow could produce neutrinos. We have calculated the expected neutrino flux from the GRB taking the model parameters at different time intervals used in the modeling (reference section 2), in both ISM and a wind environment.

The detailed study of this extremely energetic GRB and the calculation of the associated neutrino flux from it will give more insight in to the possible sources and the environments of the GRBs. Hopefully, the neutrino telescope currently in operation and those being planned for the future will be able to detect the neutrino flux calculated in this work, or a non-detection will help in putting an upper limit on UHECRs acceleration and neutrino production in GRBs.

5. References

[1] Ackermann M et al. 2013 *Science* **343** 42
[2] Kouveliotou C et al. 2013 *Astrophys.J.Lett* **779** 1
[3] Vestrand W T 2014 *Science* **343** 38
[4] Meszaros P and Rees M J 1997 *Astrophys.J.* **476** 23
[5] Piran T 2005 *Rev. Mod. Phys.* **76** 1143
[6] Razzaque S 2013, *Phys. Rev.* D **88** 103003
[7] Granot J, Piran T, Sari R 2010 *Astrophys.J.* **527** 236
[8] Waxman E and Bahcall J N 1997 *Phys. Rev. Lett.* **78** 2292
[9] Guetta D, Hooper D, Alvarez-Muniz J, Halzen F, and Reuveni E 2004 *Astropart.Phys.* **20** 429
[10] Gupta N and Zhang B 2007 *Astropart.Phys.* **27** 386
[11] Razzaque S, Yang L 2015 *Phys.Rev.* D **91** 043003

Novel zincate phosphors: A new red emitting phosphor for LED applications

Vijay Kumar[#], Sudipta Som, Somrita Dutta, Hendrik C Swart

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

[#]Email: vj.physics@gmail.com

Abstract. A series of $CaLa_2ZnO_5$ (CLZ):xEu^{3+} (x = 0, 1, 2, 3, 4, 5, 6 and 7 mol %) have been synthesized by the combustion route. X-ray diffraction results revealed the formation of a single-phase CLZ with a tetragonal structure. The CLZ: Eu^{3+} phosphors irradiated with ultraviolet light exhibited the red photoluminescence (PL) owing to the $^5D_0 \rightarrow {}^7F_J$ transitions of Eu^{3+}. The PL emission intensity first increased up to 5 mol% of Eu^{3+} concentration and then decreased. The reduction in emission intensity was attributed to a concentration quenching phenomenon that is explained on the basis of the different coupling mechanisms. The International Commission on Illumination diagram analysis results directed the appropriateness of the studied phosphor for pure red emission in light emitting diode applications.

1. Introduction

Rare earth (RE) doped phosphors, through many years, have been explored as the most studied materials for solid state lighting applications, as replicated from the large number of spectra in the existing literature [1-5]. Currently, the requirement of materials with improved luminescent properties in multidisciplinary areas has speed up research endeavours in this area focussed towards attaining superior material properties. RE ions are considered as prominent luminescent centers in various inorganic host lattices and, categorized by a partially filled 4f shell that is enclosed by $5s^2$ and $5p^6$ electrons, are broadly integrated to produce high quantum efficiency phosphors [5]. The addition of small concentration of RE ions into host matrix came into existence as a promising method for regulating the optical and PL properties [5]. Significant volume of research has also been carried out on various materials doped with Eu^{3+} so that they can be used in solid-state lighting applications. The beauty of Eu^{3+} ions doped phosphors has attracted the research community due to their importance in red emitting diodes [6-7].

Recently, ternary oxides XY_2ZO_5 (Y_2= rare-earth, X=Ca, Ba, Z=Zn, Cu) have attracted significant attention because of their motivating structural, optical and superconducting properties along with the exceptional physical and chemical stability [8-9]. The research interest has also been directed towards the ternary oxides XY_2ZO_5 based upconversion (UC) and down conversion materials. The luminescence properties of La_2BaZnO_5 and Gd_2BaZnO_5 activated with Eu^{3+}, Tb^{3+} and Tm^{3+} were reported for the first time in 1985 [10]. After that, significant research efforts have been directed towards the luminescent properties of RE doped XY_2ZnO_5 phosphors synthesized by various methods [11]. Recent results suggest that the sol-gel derived $CaLa_2ZnO_5$ (CLZ): Er^{3+}-Yb^{3+} phosphor is an efficient green UC phosphor [12]. The researchers have also studied optical thermometry in the temperature range of 298-513 K using the FIR technique. The optical and luminescent properties of

citric based sol-gel derived CLZ: Eu^{3+} phosphors have been investigated by Bandi et al. [13]. They concluded that the studied system with a higher Eu^{3+} content under the excitation of blue light could be a good orange-red-emitting phosphor for its possible utility in white LED devices.

A large number of synthesis methods have been employed in the preparation of Eu^{3+} doped phosphors. It is a well-established fact that heat treatment of light emitting materials have a straightforward effect on the luminescence properties due to the variation in the crystallite size, crystallinity, surface morphology, shape and size of particles. We have synthesized CLZ powder by the solid state reaction method by optimizing the reaction temperature to ensure the maximum luminescent intensity of the synthesized host. Later, the host matrix was doped with a different concentration of Eu ions at an optimized reaction temperature. The influence of Eu^{3+} concentrations on the photoluminescence (PL), structural and morphological properties has been investigated and explained within the scope of this article.

2. Experimental details

The ternary oxide based red phosphors namely $CaLa_2ZnO_5$ (CLZ):xEu^{3+} (x = 0, 1, 2, 3, 4, 5, 6 and 7 mol %) were synthesized by the combustion route. Analytical reagent (AR) grade lanthanum oxide (La_2O_3, 99.9%), calcium carbonate ($CaCO_3$, >99%), zinc oxide (ZnO, 97%) and europium oxide (Eu_2O_3, 99.98%) obtained from Sigma-Aldrich were used as the starting materials. Details of synthesis procedure are reported in our previous work [14].

X-ray diffraction (XRD) patterns of the as-prepared phosphors were recorded using a Bruker D8 focus X-ray diffractometer. The PL studies were performed on a Hitachi Fluorescence Spectrometer F-2500 with a 150W Xe lamp as an excitation source in the range 220-750 nm.

3. Results and discussion

3.1. Effect of Eu^{3+} doping on the crystal structure of $CaLa_2ZnO_5$ red phosphors

To investigate the structure of undoped and Eu^{3+} (1-7 mol %) doped phosphors, the influence of different Eu^{3+} doping concentrations in the crystal lattice and to detect the phase, XRD patterns of the synthesized CLZ:Eu^{3+} phosphors were recorded. XRD patterns of undoped, and Eu^{3+} (1-7 mol. %) doped CLZ phosphors are shown in Fig. 1. The diffraction peaks were indexed to a tetragonal structure of CLZ with major (h k l) peaks at (103), (113), (022), (004), (023), (114), (221), (030), (303), (134), (026), (401), (042) and (216) and lattice constants a = b = 6.547 Å and c = 3.854 Å were obtained [3]. The sharp XRD patterns of the CLZ phosphors indicate the presence of a single phase structure. Similar XRD patterns were observed for the different Eu^{3+} concentrations. The dopant concentrations do not lead to significant changes in the host structure. On the other hand, the relative intensities of the (022) and (113) diffraction peaks of the prepared phosphors has been increased with the rise in the Eu^{3+} concentration up to 5 mol% and decreased at higher doping content. The enhancement in peak intensity may be due to the improvement of the crystallinity. The maximum intensity was observed at 5 mol% Eu^{3+} concentration for the (022) diffraction peak. The highest intensity of the CLZ at 5 mol% Eu^{3+} also exhibited maximum PL emission intensity when compared to the other doping contents [13]. The XRD analysis results are in good agreement with the PL results. A small shift was also observed in the (022) peak with an increase in the Eu^{3+} concentration. This observation specifies that the Eu^{3+} ions were totally combined into the CLZ host lattice. It may be due to the addition of Eu^{3+} ions in the CLZ lattice owing to the migration of the atoms and rearranged the density of states in the CLZ unit cell. It is also evident that the peak shifts towards a higher angle up to 5 mol% of Eu^{3+} doping, subsequently no considerable peak displacement were detected for higher concentrations. This result indicates that a marginal reduction of the lattice parameters has been found after the introduction of 1 to 5 mol % of Eu^{3+} in the CLZ lattice and then saturates.

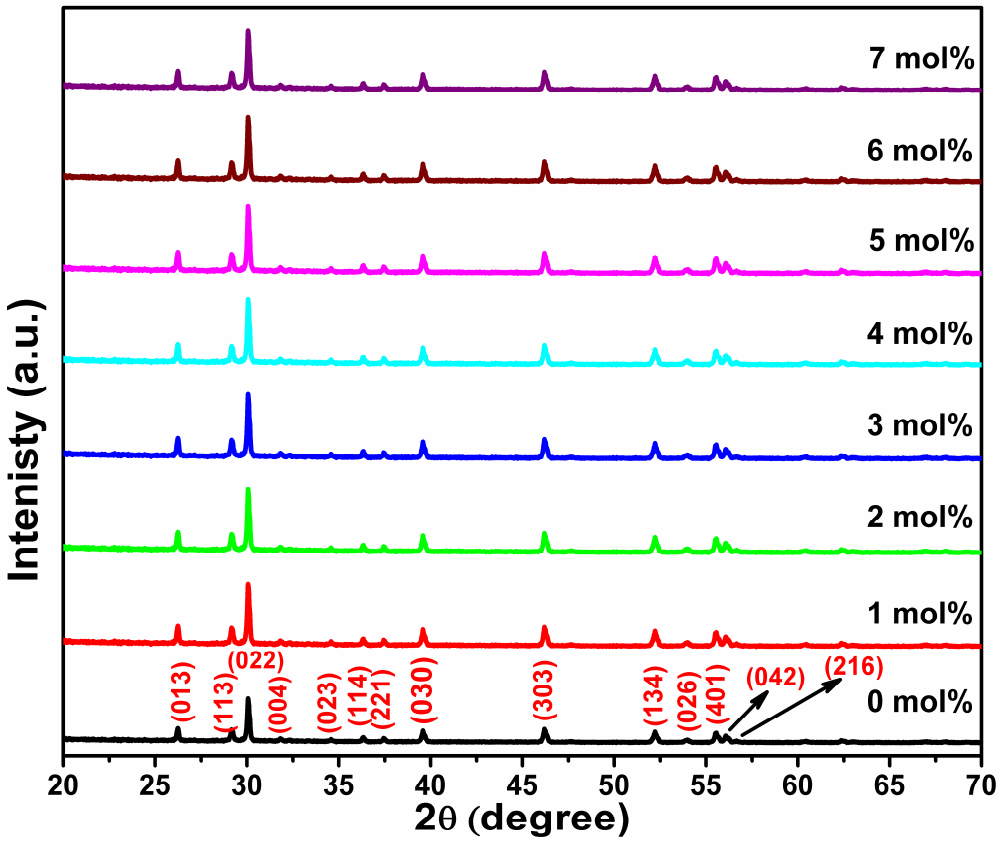


Fig.1: XRD patterns of the $CaLa_2ZnO_5$ phosphors containing different Eu^{3+} concentrations.

3.2. Photoluminescence study of $CaLa_2ZnO_5:Eu^{3+}$ phosphors

The concentration of dopant has an influence on the performance of a phosphor. Hence, it is essential to ascertain the composition of CLZ: Eu^{3+} phosphors with optimum PL emission intensity. A series of CLZ: Eu^{3+} phosphors with different concentration of Eu^{3+} (1-7 mol %) ions were prepared. The emission spectra of CLZ: Eu^{3+} phosphors with different Eu^{3+} concentration are recorded at an excitation wavelength of 456 nm and are shown in Fig. 2(a). A broad emission band in the undoped sample around ~ 400 to 500 nm (as shown in the inset of Fig. 2(a)) was observed due to the presence of different intrinsic defects such as Zn vacancy or Zn interstitial in the host matrix [15-16]. After doping of Eu^{3+} this defect emission diminishes at the cost of the increment of the characteristics peaks of Eu^{3+}. The origin of the defect related emission in this host is under investigation. However, seven different emission bands extending between 500 - 750 nm were detected due to transitions from the excited $^5D_2$ to $^7F_3$ (512 nm), $^5D_1$ to $^7F_1$ (539 nm) and $^5D_0$ to the ^7F_j (j = 0-4) levels of Eu^{3+} ions [13, 17]. The presence of the luminescence from the higher excited states, such as $^5D_1$ and $^5D_2$, indicates that the nonradiative relaxation to the $^5D_0$ level was not very efficient. This may be due to the existence of stronger electron-phonon coupling. The observed bands were found to be narrow in nature which is attributed to the characteristic electron shielding effect in Eu^{3+} ions. The most intense emission peak detected at around 624 nm, which is ascribed to the electric-dipole transition ($^5D_0 \rightarrow {}^7F_2$) and found to be hypersensitive to the host structure symmetry. Most of the f–f transitions of the rare earth ions are less affected by the environment. But some are quite susceptible to the crystal environment and turn out to be more powerful. These types of transitions are termed as hypersensitive transitions. Another emission in the vicinity of 595 nm was ascribed to the magnetic dipole transition of $^5D_0 \rightarrow {}^7F_1$ and is

insensitive to the site symmetry, as they are parity-allowed. It is to be noted that if the Eu^{3+} ions occupy the inversion symmetry center, only the magnetic dipole transition $^5D_0 \rightarrow {}^7F_1$ is expected in the emission spectrum unless the electric-dipole transition $^5D_0 \rightarrow {}^7F_2$ will dominate. In the present case the dominance of the electric-dipole transition $^5D_0 \rightarrow {}^7F_2$ was observed. This indicates that the location of Eu^{3+} deviates from the inversion symmetry, i.e. at low balance positions with the increase in Eu^{3+} concentration [17]. A considerable enhancement in emission intensity of the phosphor with the increase in Eu^{3+} concentration was observed, and it approached the maximum at 5 mol% and then decreased. The reduction in emission intensity at a higher Eu^{3+} concentration may be due to the concentration quenching phenomena [17].

The CIE (International Commission on Illumination) chromaticity diagram of CLZ: Eu^{3+} phosphors for different Eu^{3+} content is displayed in Fig. 2(b). CIE parameters such as colour coordinates (x, y) and correlated colour temperature (CCT) were calculated to describe the colour emission of the prepared phosphors. The colour coordinates were determined using the Colour Calculator software. The colour correlated temperature (CCT) is given by the McCamy empirical formula [17].

$$CCT = -437n^3 + 3601n^2 - 6861n + 5514.31 \quad (1)$$

where $n = (x-x_e) / (y-y_e)$ and chromaticity epicentre is at $x_e = 0.3320$ and $y_e = 0.1858$.

It is evident from the CIE diagram that the colour coordinates shifted from a light blue to a dark red region with the increase in the Eu^{3+} concentrations. The experimental results reflected that the red emission colour can be adjusted by simply varying the concentration of the dopant ions. A red emission (0.64, 0.34) was observed for the 5 mol % concentration of Eu^{3+}. The ideal red chromaticity value given by National Television Standard Committee is (0.67, 0.33), which is very close to the experiential value in the present case for the 5 mol% of Eu^{3+} [17]. At a higher Eu^{3+} concentration, the position of the colour coordinates shifted towards faded red light emission. The CCT value of the synthesized phosphors was found to vary from 1716 to 3847 K. It is well quoted in the literature that the warm white light employed for home appliances should have a CCT value less than 5000K [17]. Thus, the synthesized phosphors may be suitable for ideal red emission for home appliances. The different CIE parameters are calculated with respect to the PL data of the undoped and Eu^{3+} doped CLZ phosphors and depicted in Table 1. The ideal phosphor materials can be categorized by their most prominent wavelength and colour clarity of the radiated colour. This idea is mainly used for the narrow banded light sources. The CCT calculation is a crucial factor to be used in broad banded light sources. Thus, it is very necessary also to evaluate the colour purity of the emitted red light. To study the effect of the varied Eu^{3+} concentration on the colour purity, the colour purity was calculated using the relation given below [17]:

$$ColourPurity = \frac{\sqrt{(x_s - x_i)^2 + (y_s - y_i)^2}}{\sqrt{(x_d - x_i)^2 + (y_d - y_i)^2}} \times 100\% \quad (2)$$

where (x_s, y_s) are the coordinates of a sample point, (x_d, y_d) are the coordinates of the dominant wavelength and (x_i, y_i) are the coordinates of the illuminated point. The calculated results are presented in Table 1. It is clearly seen from the table that the color purity increased with the rise in Eu^{3+} content and was maximum at 5 mol% of Eu^{3+}. All the emission transitions are shown in the energy level diagram of Eu^{3+} in this host as shown in Fig. 2(c).

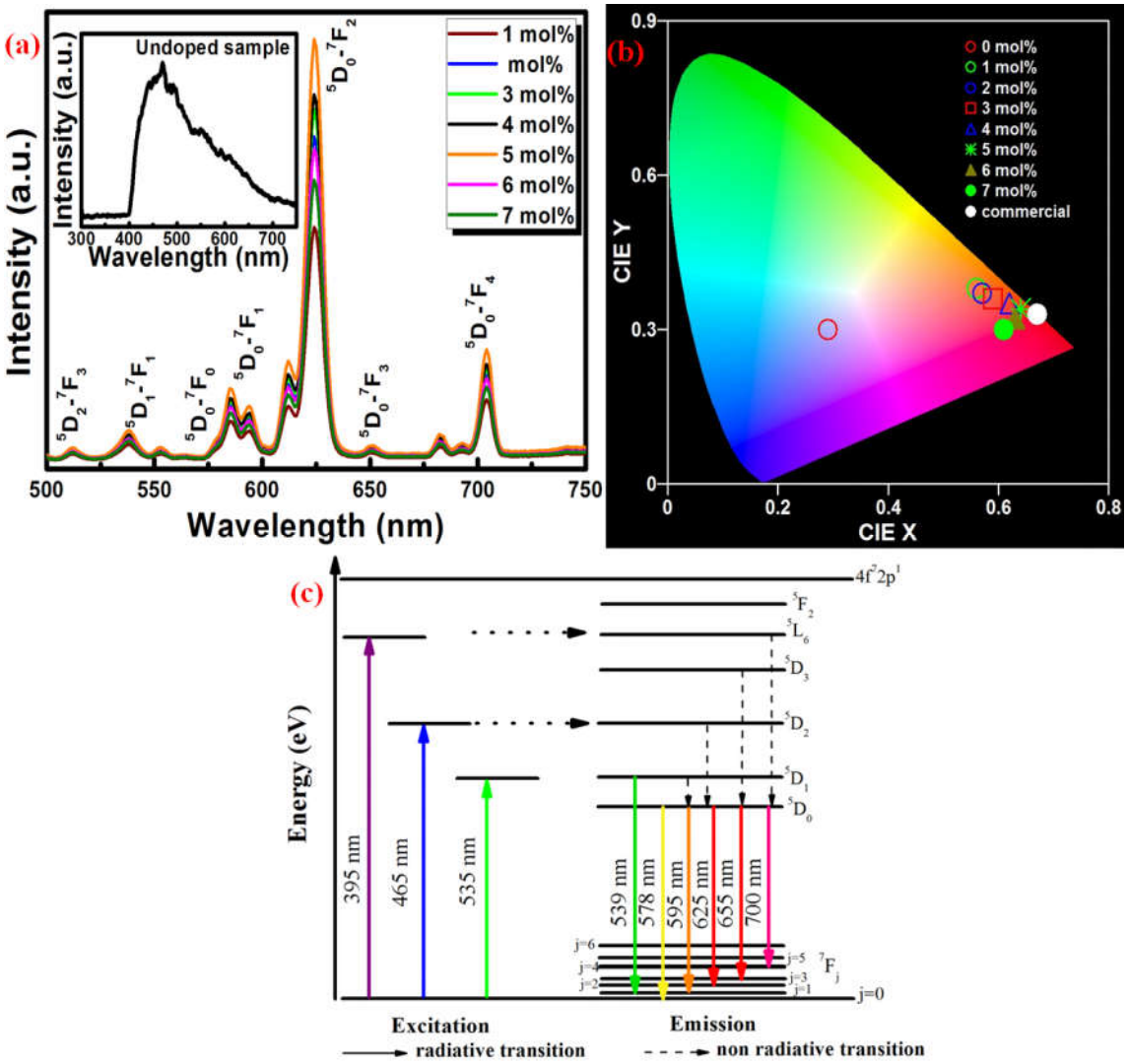


Fig. 2: (a) Effect of doping concentration on the PL emission of $CaLa_2ZnO_5:Eu^{3+}$ phosphors. **(b)** The colour coordinates diagrams for $CaLa_2ZnO_5:Eu^{3+}$ phosphors for different Eu^{3+}. **(c)** The energy level diagram of Eu^{3+} in the $CaLa_2ZnO_5$ host matrix.

Table 1: CIE parameters of CLZ: Eu^{3+} with different Eu^{3+} concentration.

Phosphor	Concentration (mol%)	Colour Coordinates		CCT (K)	Colour Purity (%)
		x	y		
La_2CaZnO_5: Eu^{3+}	0	0.28	0.29	-	20
	1	0.56	0.38	1716	69
	2	0.57	0.37	1718	71
	3	0.59	0.36	1831	77
	4	0.62	0.35	2200	85
	5	0.64	0.34	2694	91
	6	0.63	0.32	3250	88
	7	0.61	0.30	3847	83

4. Conclusion

In summary, a series of Eu^{3+} doped CLZ phosphors were prepared by the solid-state reaction method. The crystal system of the prepared samples was identified as a tetragonal structure, and all the peaks were indexed correctly. The Eu^{3+} doped CLZ phosphors showed a dominant red emission peak centered at 625 nm under 395 nm UV light excitation. The luminescence intensity increased with Eu^{3+} concentration up to 5 mol% and then decreased. The reduction of emission intensity with increasing Eu^{3+} is associated with a concentration quenching phenomena. The results have confirmed that the synthesized phosphors may be considered as a potential red phosphor candidate that can be utilized in solid-state lighting and display devices when combined with UV or blue chips.

Acknowledgment

The research is also supported by the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation of South Africa. The financial assistance from the University of the Free State is highly appreciated.

References

[1] Daicho H, Iwasaki T, Enomoto K, Sasaki Y, Maeno Y, Shinomiya Y, Aoyagi S, Nishibori E, Sakata M, Sawa H, Matsuishi S, and Hosono H 2011 *Nat. Commun.* **3** 1132.
[2] Luminescence. From Theory to Applications (ed Ronda, C. R.) ISBN: 978-3-527-31402-7 (Wiley-VCH, 2007).
[3] Pavitra E, Seeta R R G, Park J Y, Wang L, Moon B K and SuYu J 2015 *Sci Rep* **5** 10296.
[4] Pust P, Weiler V, Hecht C, Tücks A, Wochnik A S, Hen A K, Wiechert D, Scheu C, Schmidt P J and Schnick W 2014 *Nat Mater* **13** 891.
[5] Kenyon A J 2002 *Prog Quant Electron* **26** 225.
[6] Barnes M D, Mehta A, Thundat T, Bhargava R N, Chabra V, Kulkarni B 2002 *J Phys Chem B* **104** 6099.
[7] Gedanken A, Reisfeld R, Sominski L, Zhong Z, Koltypin Y, Panczer G, Gaft M, Minti H 2000 *Appl Phys Lett* **77** 945.
[8] Etchart I, Berard M, Laroche M, Huignard A, Hernandez I, Gillin W P, Curry R J, Cheetham A K 2011 *Chem Commun* **47** 6263.
[9] Jaffres A, Viana B, van der Kolk E 2012 *Chem Phys Lett* **527** 42.
[10] Lammers M J J, Donker H, Blasse G 1985 *Mat Chem Phys* **13** 527.
[11] Yoshida A, Ogawa H, Kan A, Kondo T 2005 *J Eur Cer Soc* **25** 2897.
[12] Li L, Guo C, Jiang S, Agrawal D K, Li T 2014 *RSC Adv* **4** 6391.
[13] Bandi V R, Grandhe B K, Jang K, Lee H S, Shin D S, Yi S S, Jeong J H 2012 *J Alloys Comp* **512** 264.
[14] Kumar V, Som S, Dutta S, Das S, Swart H C 2016 *Mater. Design* **93** 203.
[15] Pereira A S, Peres M, Soares M J, Alves E, Neves A, Monteiro T, Trindade T 2006 *Nanotechnology* **17** 834.
[16] Husain S, Alkhtaby L A, Bhat I, Giorgetti E, Zoppi A, Miranda M M 2014 *J Lumin* **154** 430.
[17] Som S, Kunti A K, Kumar V, Kumar V, Dutta S, Chowdhury, Sharma S K, Terblans J J, Swart H C 2014 *J Appl Phys* **115** 193101.

Investigated charge carrier effects in silicon membranes using a femtosecond laser

WI Ndebeka[1], PH Neethling[1], CM Steenkamp[1], H Stafast[1,2], EG Rohwer[1]

[1] Laser Research Institute, Physics Department, Stellenbosch University, PO Bag X1, Matieland 7602, South Africa

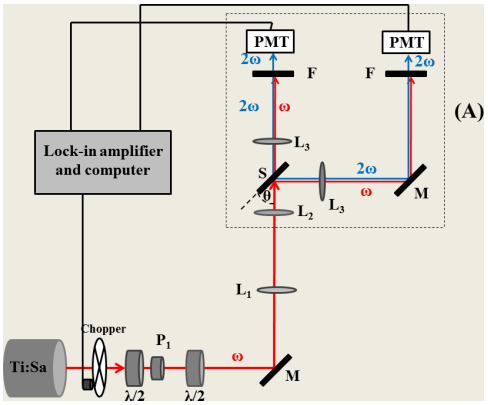
[2] Leibniz Institute of Photonic Technology, Albert-Einstein-Str. 9, 07745 Jena, Germany

E-mail: ndebeka@sun.ac.za

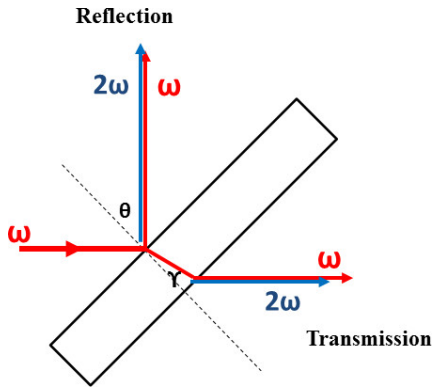
Abstract. The second harmonic (SH) generated at the Si/SiO_2 interface varies on a time scale of several seconds when illuminated with high intensity near infrared laser pulse trains (λ = 800 nm, 80 fs, 80 MHz repetition rate, $E_{pulse} \leq 10$ nJ). The temporal behaviour arises from generation of trap sites and subsequent trapping of charges at the interface via multi-photon processes. These trapped charges create an interfacial electric field which influences the nonlinear properties of the Si/SiO_2 interface and leads to a time dependent second harmonic (TDSH) signal under continuous irradiation. This is known as electric field induced second harmonic (EFISH) generation. In this work, measurements are focused on the simultaneous measurements of EFISH signal from a free standing oxidized Si membrane both in reflection and transmission as a function of the irradiation time. Results show that the transmission of the fundamental irradiation as well as the transmitted SH signal generated from the Si membrane increases, reaches a maximum, and then decreases again as the input intensity is increased. The nonlinear behaviour of the transmitted signal is explained using free charge carrier absorption (FCA) in silicon.

1. Introduction

The silicon-silicon dioxide (Si/SiO_2) interface has been one of the most studied systems in the realms of materials and condensed matter physics. Despite the great outpouring of research and the huge commercial success, there has been little effort to construct a comprehensive, unified microphysical model of the Si/SiO_2 system [1-3]. The optical properties, both linear and nonlinear, of the interface region which are of interest, since this transition region varies considerably from those of the adjacent bulk phases. The second harmonic (SH) generation is used as a versatile tool to investigate buried surfaces and interfaces of centrosymmetric material, in which the electric field induced second harmonic (EFISH) generation processes is used as a valuable technique to investigate the mechanisms of charge transfer and trap generation [5 - 7]. The time dependence of the resulting EFISH signal is a measure of the rate of trap site generation and population, by both electrons and holes being pumped across the interface [8]. In this work, simultaneous time-dependent second harmonic (TDSH) measurements in reflection from and transmission through a Si membrane are presented. The transmission and reflection of the fundamental laser power are also measured and discussed using free carrier absorption (FCA) in silicon.



(a)



(b)

Figure 1: (a) Schematic diagram of the EFISH setup used during this study. (b) Laser beam reflection from and transmission through thin Si membranes; external angle of incidence $\theta = 45\,^{\circ}$; within silicon $\gamma \approx 11\,^{\circ}$. ω and 2ω are the fundamental and SH signals, respectively.

2. Experiments

The optical source used in this work is a titanium-doped sapphire femtosecond laser oscillator (Spectra Physics, 3941-M3S, tsunami) capable of tunable laser operation over a broad range of near infrared (NIR) wavelengths. The femtosecond laser is pumped by a frequency doubled Nd:YVO_2 laser (Spectra Physics, Millenia V) with 5.5 W continuous wave (cw) maximum power output at a wavelength of 532 nm [9]. Figure 1a shows the schematic diagram of the experimental setup used in this work. The p-polarized beam from the Ti:Sapphire laser ($\lambda_c = 800\,\text{nm}$, $\tau = 80\,\text{fs}$ at 80 MHz repetition rate) passes through a half wave plate, $\lambda/2$, and a polarizer for power adjustment. The polarized beam is then incident on the sample, after being guided by a set of silver mirrors, collimated and focused by two lenses, L_1 and L_2, respectively. Light is incident on the sample at an external angle of $45\,^{\circ}$. The incident laser light is focused down to 11 μm diameter on the sample surface. The focal plane position is determined by a z-scan like measurement meaning that the sample is moved along the incident laser beam through the beam focus and the generated SH intensity is measured as a function of the sample position z. The fundamental laser beam has a Gaussian spatial beam profile. The Rayleigh length of the focus is $z_R = 5.57\,\mu\text{m}$ is obtained by fitting the SH intensity with the following expression [10]:

$$P_{2\omega} = \frac{K}{1 + z^2/z_R^2}, \tag{1}$$

with $K = (\eta P_\omega^2)/(\pi w_0^2)$. Here P_ω and $P_{2\omega}$ are the fundamental incident and SH converted powers, respectively, η is the SH conversion efficiency constant and $w_0 = w(z = 0)$, with $w(z)$ being the radius of the incident beam of frequency w at a position z along the propagation direction. For both the transmitted and reflected beams, the fundamental and SH light are collimated by identical lenses L_3. A power meter is used to measure the fundamental transmitted and reflected powers. For the power measurements, no filter is used to block the SH light generated on the sample as its contribution to the measurements is negligibly small. For the time-dependent second harmonic (TDSH) measurements, two identical filters, F were used to allow the SH light to pass through while blocking the transmitted and reflected fundamental beams.

The detection of the SH signals were done by use of two identical photomultiplier tubes, PMT. To enhance the signal-to-noise ratio, the SH signals measured in transmission and in reflection are amplified by a digital lock-in amplifier, with 2 physical units, in combination with a 500 Hz light

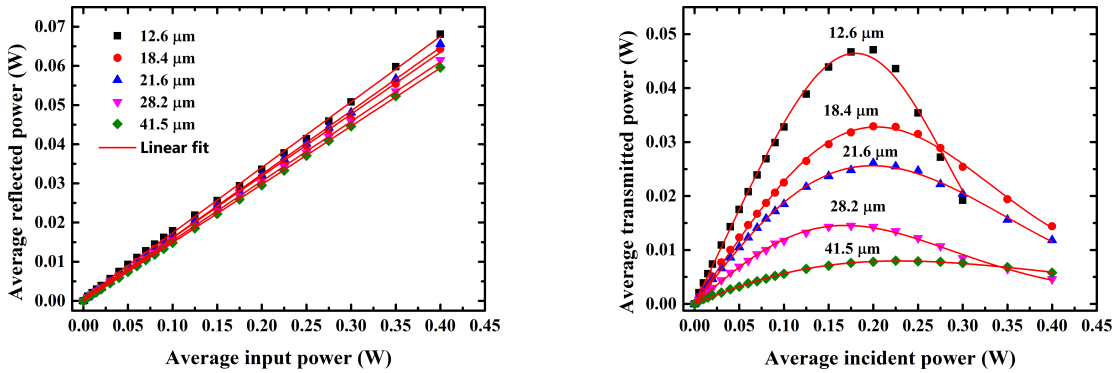
chopper. The lock-in amplifier, with LabView software, was connected to a computer for data acquisition. The samples were prepared from commercial ⟨100⟩-Si wafers by chemical etching using tetramethylammonium hydroxide (TMAH) to produce thin membranes of $3 \times 3\,\mathrm{mm}^2$ area with different thicknesses [13]. The wafers are slightly p-doped ($3 - 6 \times 10^{14}\,/\mathrm{cm}^3$). Before the investigations, the samples are cleaned using acetone and hydrofluoric acid to remove dirt and old oxide layers, respectively. In contact with air, the clean surfaces oxidize under dark room conditions reaching an equilibrium thickness of ($< 5 \pm 0.5$) nm within 48 hours [11].

3. Results

The naturally oxidized Si membrane samples, with different thicknesses, were irradiated with a femtosecond pulsed laser ($\lambda_c = 800\,\mathrm{nm}$, $\tau = 80\,\mathrm{fs}$) at an external angle of incidence of 45 °. Both the fundamental and second harmonic (SH) signals were recorded for transmission through the membranes and reflection from the membranes as shown figure 1b.

3.1. Fundamental measurements

The beam from the oscillator is collimated, passes through a chopper of 50 % transmission and is focused by two lenses, L_1 and L_2 ($f_1 = -50\,\mathrm{mm}$ and $f_2 = 35\,\mathrm{mm}$, respectively) to diameter of ($11 \pm 2\,\mu\mathrm{m}$). The transmitted and reflected beams are collimated by lenses L_3($f_3 = 120\,\mathrm{mm}$). One power meter is used to measure the transmitted and reflected average power.



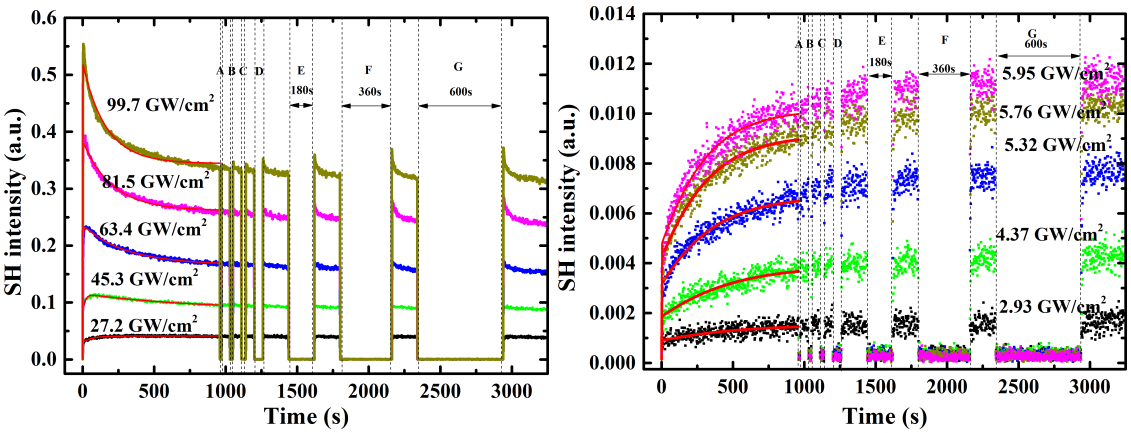
(a) (b)

Figure 2: Fundamental measurements of the average laser power reflected from (a) and transmitted through (b) a Si membranes with different thicknesses. The solid curves are fits; cf. details in the text

Figure 2 shows the results of the fundamental beam reflection from (figure 2a) and transmission through (figure 2b) Si membranes, with different thicknesses, as a function of the incident average laser power, P_{in} up to $\approx 400\,\mathrm{mW}$. For reflection, the average laser reflected power P_{re} displays a linear dependence on the incident power. This is in agreement with previous findings and confirms the validity of the Fresnel equations using the optical constants of silicon samples under ambient conditions [9]. However, the transmitted laser power P_{tr} shows a different behaviour; starting with a linear increase, it reaches a maximum around $0.15 - 0.25\,\mathrm{W}$ and decreases again as the average incident laser power is increased. Contributions from internally reflected beams to the transmitted laser power P_{tr} are negligible ($\leq$ about 2 %) because of the high reflection losses at the interfaces, even without any account for laser beam attenuation by absorption and/or scattering [12].

3.2. Time-dependent second harmonic measurements

For the time-dependent second harmonic (TDSH) generation, the p-polarized beam out of the laser (800 nm, 80 fs, 80 MHz) was incident on the Si sample, with thickness d , at an external angle 45 °. The pulse energy of the laser is < 10 nJ and the central wavelength was monitored by a spectrometer. Detection of the SH signals was made by use identical photomultiplier tubes (PMT) after blocking fundamental beams with filters.



(a) (b)

Figure 3: Time-dependent second harmonic measurements in reflection (a) and in transmission (b) with dark periods: A (10 s), B (20 s), C (30 s), D (60 s), E (180 s), F (360 s) and G (600 s). The intensity value given for each curve corresponds to the local intensity of the fundamental laser beam just inside the front interface for reflections, and just inside the back interface for transmission. The membrane was $\approx 33.6\,\mu$m thick.

Figure 3 displays simultaneous measurements of second harmonic signals reflected from (figure 3a) and transmitted through (figure 3b) a silicon membrane, with thickness $d \approx 33.6\,\mu$m as a function of irradiation time, with every measurement starting on a virgin sample spot.

In reflection measurements (figure 3a), for peak intensities $\leq 27.2\,\mathrm{GW/cm^2}$, the SH signal increases on a time scale of minutes before reaching a steady state (equilibrium) up to 960 s. This effect is known as electric field induced second harmonic (EFISH) generation [11-13]. The temporal evolution can be reproduced using two exponential functions and the fit is represented by the solid curve in figure 3a. For peak intensities $> 27.2\,\mathrm{GW/cm^2}$, the SH signal show a different behaviour; first it increases on a time scale of seconds, reaches a maximum and then decreases on a time scale of minutes before reaching equilibrium. In this case the temporal evolution is reproduced using three exponential functions (see equation 3). It is observed that upon re-irradiation the SH intensity increases above the level before the dark period but less than the initial SH intensity upon onset of laser illumination. The decline after each new SH maximum occurs on a significantly shorter time scale than the slow decline after the initial maximum of the virgin sample spot.

In transmission measurements (figure 3b), the SH signal increases on a time scale of several minutes before reaching an equilibrium value. The transmitted intensities for the TDSH in transmission were calculated from the transmitted fundamental powers and correspond to the intensities listed for the reflection measurements. The dark periods seem to not affect the SH signals upon re-irradiation: the SH signal starts at the value it had when the signal was blocked.

4. Discussion and conclusion

4.1. Fundamental measurements

The absorption of Ti:sapphire laser light by slightly doped silicon membranes (in the 12 μm to 40 μm range) was investigated with a femtosecond pulsed laser. The measurements revealed a strong nonlinear power dependence of the fundamental beam transmission through the sample. The nonlinearity observed with the laser beam transmission is not attributed to photo-damage of the Si membranes because the process is reversible, i.e. it gives the same trend if the measurements start with high incident powers, but it is attributed to free carrier absorption (FCA) in silicon [13]. Laser induced sample heating and coherent two- and/or multi-photon absorption are of minor importance and can in any case not explain the observed decrease of transmitted power at high incidence powers [11]. Similar results obtained by Yamada et al. [12] with a pulsed dye laser at 590 nm irradiating a 0.6 μm thin Si layer on sapphire show a pronounced maximum followed by a lower constant transmission for further increasing input intensity. This complex transmission behaviour could nicely be fitted by a simple FCA model (equation 2 of [13])

$$\begin{aligned} I_{trans}(I_{in}, z_m) &= (I_{in} - I_c)\exp(-\alpha_1 z_m)\exp(-\alpha_{FCA} z_m) + I_c \exp(-\alpha_1 z_m) \\ &= F\left[(I_{in} - I_c)\exp(-E_n I_{in}^n) + I_c\right], \end{aligned} \tag{2}$$

where $F = \exp(-\alpha_1 z_m)$. I_c is defined as the I_0 threshold value below which electron-hole pairs are generated but do not contribute to FCA. I_{in} and I_{trans} are the average intensities just inside the front and back Si/SiO_2, respectively. α_1 is the linear absorption coefficient of Si, $z_m = d/\cos\theta$ is the optical path length in silicon with thickness d. If α_{FCA} is the FCA absorption coefficient, n_{eh} and σ_{eh} the concentration and absorption cross section of eh pairs, respectively, $\alpha_{FCA} \approx n_{eh}\sigma_{eh}$ with $n_{eh}\sigma_{eh} = \sigma_0 n_{eh}^n$ (for $n \approx 2$) [13] and $n_{eh} = \kappa I_{in}$ as the eh pairs are generated by one-photon absorption. Therefore $e^{-\alpha_{FCA} z_m} = e^{-\kappa_n z_m I_{in}^n} = e^{-E_n I_{in}^n}$. The application of equation (2) to figure 2b is shown by the solid curves. Since the original submission of this paper the model has been applied to a different dataset and values for the fitting parameters have been obtained [14]

4.2. Time-dependent second harmonic

The temporal evolution of the experimental data is reproduced, in reflection (for peak intensities $\leq 27.2\,GW/cm^2$) and in transmission geometries by fitting two exponential functions given by [15-17]. For peak intensities $> 27.2\,GW/cm^2$ (reflection), the temporal evolution of the SH response (figure 3a) cannot be explained only with electron effects. An additional effect is added to describe hole dynamics, which requires an additional two exponential functions. However, upon fitting the experimental data with four exponential functions, the fourth time constant gives unrealistic values. The experimental data were then reproduced using three exponential functions given by:

$$I^{(2\omega)}(t) \propto \sum_{i=1}^{3} \left[1 + a_i \exp(-t/\tau_i)\right]^2, \tag{3}$$

where $a_1, a_2 < 0$ and $a_3 > 0$ and $\tau_i > 0$. The time constants τ_i are dependent on the incident laser peak intensity $I_{l,p}$ [6 - 8] and follow the power law $1/\tau_i(I_{l,p}) \propto I_{l,p}^{n_i}$. In reflection geometry, for peak intensities $\leq 27.2\,GW/cm^2$, it was found that $n_1 = 2.7$ and $n_2 = 3.2$. The interpretation of these results is based on previous work done on bulk Si/SiO_2 [6], [17 - 20]. The explanation is that laser irradiation leads to electron injection and trapping in the ultrathin SiO_2 [8] and to a permanent modification of the SH response due to photoinduced trap site creation. The

superposition of electron and hole processes leads to a time-dependent EFISH signal. The energy barrier for electron injection into the SiO_2 amounts to 4.3 eV [17, 18]. Electrons surmount this interface in the bulk or at Si/SiO_2 interface by direct three-photon absorption explaining why n_1 and n_2 are close to 3 or cascade processes. For laser peak intensities $> 27.2\,GW/cm^2$ (reflection geometry), the temporal evolution of equation (3) gives the characteristic time constant τ_3, and following the power law it is found that $n_3 = 6.59$. The extracted time constant intervals are given in Table 1.

For the EFISH measurements in transmission, the extracted time constant do not follow a simple power law dependence on the incidence intensity. This would seem to indicate that FCA plays a significant role in the establishment of the interfacial electric field at the back interface. Further investigations are necessary to pinpoint this role.

Table 1: Time constants extracted from the graphs using two- and three-exponential fits

	τ_1	τ_2	τ_3
Transmission	$14 - 68\,s$	$150 - 305\,s$	
Reflection	$0.005 - 21\,s$	$0.75 - 321\,s$	$373 - 6316\,s$

In summary, the application of fs Ti: sapphire laser for the investigation of Si membrane samples in the $12\,\mu m$ to $40\,\mu m$ thickness range reveals a strong nonlinear dependence of the transmitted laser powers while the reflected powers show a linear dependence. The nonlinearity observed is attributed to FCA in Si. A FCA model (equation (2)) is used to fit the transmitted laser power data. Simultaneous measurements of the SH signals in reflection from and transmission through Si membranes show that the observed phenomena are results of the electron injection and trapping inside the ultrathin SiO_2.

5. References

[1] Helms C R and Poindexter E H 1994 *Report on Progress in Physics* **57(8)** 791
[2] Lüpke G 1999 *Surface Science Reports* **35** 75-161
[3] Lu Zhong-Yi, Nicklaw C J, Fleetwood D M, Schrimpf R D and Pantelides S T 2002 *Phys. Rev. Lett.* **89** 285505
[4] Neethling P H 2008 *Electric field induced second harmonic (EFISH) measurements of highly boron doped p^+-type Si/SiO_2*. PhD Thesis, University of Stellenbosch
[5] Nyamuda G P, Steenkamp C M, Stafast H and Rohwer E G 2011 *Appl. Phys.* B **104** 735
[6] Scheidt T, Rohwer E G, von Bergmann H M and Stafast H 2004 *European Phys. Journal : Appl. Phys.* **27** 393
[7] Scheidt T, Neethling P H, Rohwer E G, von Bergmann H M and Stafast H 2008 *Appl. Phys.* **104** 083712
[8] Scheidt T 2005 *Charge carrier dynamics and defect generation at the Si/SiO_2 interface probed by femtosecond optical second harmonic generation.* PhD Thesis, Friedrich-Schiller University Jena- Germany
[9] Sokolowski-Tinten K, Bilakowski J. D and von der Linde D. 1995 *Phys. Rev. B* **51** 14186
[10] Bing Gu and Jing Chen and Ya-Xian Fan and Jianping Ding and Hui-Tian Wang 2005 *J. Opt. Soc. Am. B* **22** 2651–2659
[11] Mihaychuk J C, Bloch J and van Driel H M 1995 *Opt. Lett.* **20** 2063
[12] Yamada M, Kotani H, Yamamoto K and Abe K 1981 *Phys. Lett.* **85A** 191
[13] Heisel P -C, Ndebeka W I, Neethling P H, Paa W, Rohwer E G, Steenkamp C M and Stafast H 2016 *Appl. Phys. B: Lasers and Optics* **122** 60
[14] Ndebeka W I, and Neethling P H, and Steenkamp C M, and Bergmann J, and Rohwer E G and Stafast H. June 2016 Interband and free charge carrier absorption in silicon at 800 nm: experiments and model calculations, manuscript in preparation
[15] Aktsipetrov O A, Fedyanin A A, Golovkina V N and Murzina T V 1994 *Opt. Lett.* **19** 1450
[16] Bloch J, Mihaychuk J G and van Driel H M 1996 *Phys. Rev. Lett.* **77** 620
[17] Mihaychuk J G, Shamir N and van Driel H M 1999 *Phys. Rev.* B **59** 2164
[18] Cernusca M, Heer R and Reider G A 1998 *Appl. Phys. B: Lasers and Opt* B **66** 367
[19] Balk P 1988 *Material Science Monographs* **32**
[20] Scheidt T, Rohwer E G, von Bergmann H M and Stafast H 2004 *Phys. Rev. B* **69** 155314

Fluorescence behaviour of Eu doped Gd_2O_3 nanosheets via CuO incorporation

Anurag Pandey, R. E. Kroon and H. C. Swart
Department of Physics, University of the Free State, P.O. Box 339, Bloemfontein-9300, South Africa
E-mail: anuragpandey439@gmail.com, swarthc@ufs.ac.za

Abstract. The present research deals with the variation in fluorescence emission of Eu doped Gd_2O_3 phosphor in the presence of different CuO concentrations. Sheet-like phosphor materials have been successfully synthesized by a simple solution combustion route. The crystalline nature, phase formation and presence of elements in the prepared materials were identified by X-ray diffraction and energy dispersive X-ray spectroscopy, respectively. Microstructure analysis has been performed by scanning electron microscopy. The photoluminescence emission spectra have been recorded to observe the fluorescence behaviour upon 325 nm laser excitation and the corresponding colour coordinates were also calculated.

1. Introduction

Lanthanide activated phosphors have been studied extensively since the last few decades owing to their strong luminescence behaviour useful in multifunctional applications e.g. fluorescent lamps, display devices, solar cells, light emitting diodes, optical and temperature sensing, etc. [1-3]. The selection of materials (host and dopants) andthe synthesis technique play a crucial role in the development of such phosphors. Gadolinium oxide(Gd_2O_3) is an interesting material for luminescence investigation due to its characteristics of serving both as host as well as doping elements for different applications purposes. It gives a suitable environment for doping elements as a host because of its high band gap, low phonon frequency,and good thermal and chemical stability [2, 4, 5]. Numerous studies have been performed by researchers with lanthanide doped/codoped gadolinium oxide for strong multicolor emissions via both the upconversion and downconversion processes [2, 5-7]. Europium (Eu) is very popular as doping element among the lanthanides for its strong red emission through 4f-4f transitions and the Eu^{3+} to O^{2-} charge transfer excitation band [8].Several articles have been reported on Eudoped Gd_2O_3 nano-phosphor/crystals for optical/structural investigations [5, 7, 9].Incorporation of transition metals may affect the structural and optical properties of the materials. CuO is one of these materials containing transition metals that can be considered. Severalphosphors have been reported via codoping ofcopper and lanthanide combinations[10, 11] but no such report is available with Cu/Eucodoping.Therefore, it is interesting to see the effect of the incorporation of CuO into the Gd_2O_3:Eusystem. Control over shape and size of the phosphors makes it more advantageous from an applications point of view and influences the optical, electrical and magnetic properties of the materials [9].Various synthesis techniques have been reported for producing materials of different size and shapes but the solution combustion method is found to be betterdue to its low processing temperature, homogeneous mixing and smaller particle formation [12].

The formation of sheet-like structuresand the influence on the luminescence emission upon adding CuO to the Gd_2O_3:Eusystemwas the aim of the present study. The structural and optical characterizations wereperformed by X-ray diffraction (XRD), scanning electron microscopy (SEM), energy dispersive X-ray spectroscopy (EDS) and photoluminescence (PL)measurementsand explained in detail.

2. Material synthesis and characterization

Eudoped Gd_2O_3 phosphor powder has been prepared by the low temperature solution combustion route [12]. Gd_2O_3,Eu_2O_3 and CuO were used as starting materials, HNO_3 as solvent and urea as organic fuel. All the chemicals were of highly pure (99.90-99.99%) analytical grade purchased from Sigma Aldrich. A number of samples have been prepared according to the following composition

$$(100\text{-}x\text{-}y)\ Gd_2O_3 + x\ Eu_2O_3 + y\ CuO \qquad (1)$$

where$x$ = 5.0 mol% and y =0.0, 1.0, 3.0, 5.0 mol%.

The oxide materialswere dissolved in concentrated HNO_3 to convert them to the form of nitrates. The nitrate forms of the host and dopants were mixed together and dissolved until transparent. Then the urea (organic fuel) was added to the final transparent solution and stirred about an hour at 60 ^{0}C till a transparent gel was obtained. The formed gel-like solution was taken in an alumina crucible and placed inside an electrical furnace preheated to 500 ^{0}C where combustion took place. The obtained samples were grounded to obtain fine and homogeneous powders and then annealed at 800 ^{0}C for 2 h. The heat-treated samples have been used for further characterization.

The XRD patterns were measured by a Bruker-D8 Advance diffractometer with a Cu-target radiation (λ= 0.154 nm), SEM and EDS images have been taken by using a JSM-7800F analytical field emission scanning electron microscope.A325 nm He-Cd laser was used to excite the samples and the emission was measured using an Ocean Optics USB2000+ spectrometer. The colour coordinates were calculated by using the 1931 Commission Internationale de L'Eclairage (CIE) software.

3. Results and discussion

3.1. Structural analysis

The XRD spectra of the Gd_2O_3:Eu/CuOnanosheets recorded in the 15 to 65 degree range are shown in Fig. 1. The patterns matched suitably with JCPDS card no. 12-0797 of the cubic Gd_2O_3 with space group Ia3. All the observed peaks were well indexed and the sharp diffraction peaks indicating the crystalline nature of the prepared phosphors. Interestingly, some extra peaks were detected on incorporation of the CuO and seemed to increase on increasing the concentration of CuO.These impurity peaks were identified for different orientation of pure monoclinic CuO (as assigned in Fig. 1) JCPDS card no. 48-1548. The crystallite size of the developed phosphors has been calculated by using the Scherrer equation [13]

$$D = \frac{k\lambda}{\beta Cos\theta} \qquad (2)$$

where$D$ is the crystallite size,k is the shape factor (assumed to be 0.9),λ is wavelength of the X-ray radiation (0.154056 nm for Cu Ka radiation).$\theta$is Bragg diffraction angle and β is the full width of the diffraction peak measured at half of its maximum intensity (in radians).The values of the average crystallite size were found to be 34 nm and did not changed significantly on incorporation of CuO (from 37 to 41 nm).

To observe the surface morphology of the prepared phosphors, the SEM images of the Gd_2O_3:Euwithout and with 3 mol% CuO codoped samples have been obtained (Fig. 2). Fig. 2(a) shows the SEM image of Gd_2O_3:Euat1µm scale. The image displays a sheet-like particle structure of different sizes. The SEM image with CuO incorporated into the sample also demonstrates a sheet like structure as observed from Fig. 2(b). The thickness of the sheets was found to be about 100 nm as indicated by the red circles of Fig. 2(c). Thus the formation of sheet-like structures with nano thicknesses (i.e. nano-sheets) was confirmed by the SEM characterization.

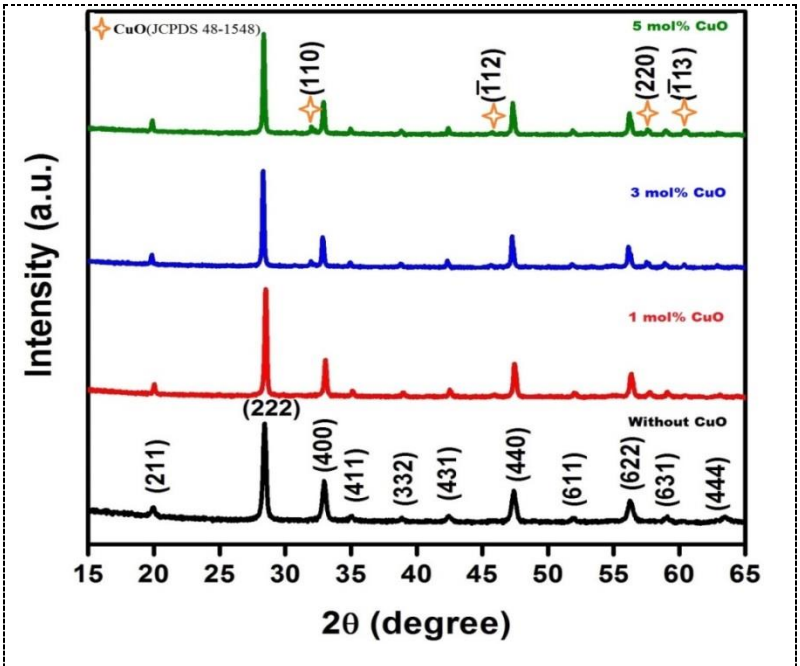


Figure 1. XRD patterns of the Gd_2O_3:Eunanosheets with different CuO concentrations.

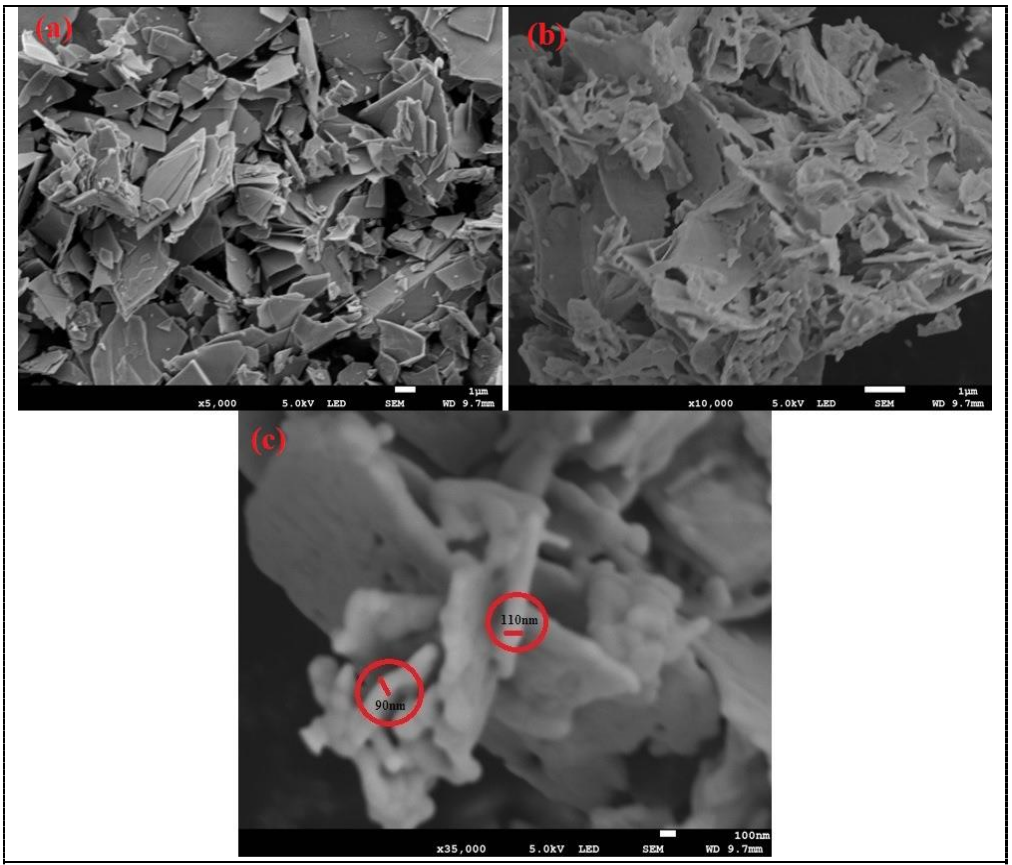


Figure 2. SEM images of the Gd_2O_3:Eunanosheets (a) without and, (b) and (c) with CuO incorporation.

Fig. 3 shows the EDS spectra of the prepared materials corresponding to that in Fig. 2(a) and (b). The EDS patterns shown in Fig. 3(a) and (b) confirmed the elemental presence of all the components of the prepared phosphors. The peak of carbon is found in both cases due to the use of carbon tape to stick the sample during characterization. Also a weak peak of aluminum detected at about 1.5 keV might be due to contamination from the use of the alumina crucible during the preparation of the samples.

The optical photographs of the phosphor samples are given as insets in Fig. 3(b). From the photographs we can see that the colour of the material has changed from white to brown/black on adding the amount of CuO from 1.0 to 5.0 mol%.

3.2 Luminescence study

The room temperature PL emission spectra of the Eudoped and CuO codoped Gd_2O_3 phosphor excited by a 325 nm He-Cd laser and recorded in the 375-750 nm range are shown in Fig. 4. The whole emission spectra comprised of two parts, one is a broad band ranging from 375-550 nm and the other consisting of several sharp peaks ranging from 550-750 nm. The broad band centered at about 440 nm is assigned to the Eu^{2+} emission through the $^4f_6{}^5d_1$ (5d) excited state to the $^8S_{7/2}$ ($^4f_7$, i.e. 4f) ground state transition [14]. A few sharp emission peaks are detected in the orange-red region which is attributed to Eu^{3+} emission via the excited $^5D_0$ state to different ground states $^7F_{0,\ 1,\ 2,\ 3,\ 4}$ transitions [14, 15]. On comparing the emission spectra it was found that the emission band that corresponded to Eu^{3+} has decreased whereas that which corresponded to the Eu^{2+} has increased drastically on the incorporation of 1 mol% CuO. As the Eu^{3+} ions can occupy the Gd^{3+} sites without inversion symmetry therefore the electric dipole transition $^5D_0 \rightarrow {}^7F_2$ is dominant [14] and the formation of moreEu^{2+} active centres in presence of CuO ions is the result of capturing unoccupied Cu^{2+} s-d states[16]. The total emission intensity throughout the recorded region was suppressed on addition of more CuO (in the case of 3 and 5 mol% CuO) because of Cu-Cu clustering inside host lattice [17].

The emission peak at about 612 nm was dominant over the others in the case of the singly Eu doped materials which resulted in a strong red emission from the material. On the other hand the broad blue band peaking at 440 nm has been dominant over the red in the case of the CuO doping and hence an intense blue emission was observed from the sample. This change in colour of the emitted light from the samples was confirmed by the corresponding calculated colour coordinates. The CIE chromaticity diagram with the indicated coordinate positions is given as an inset of Fig. 4.

4. Conclusion

Strong red and blue light emitting sheet-like Eu/CuO doped/codoped Gd_2O_3phosphors have been successfully synthesized via a low temperature combustion process. The XRD analysis confirmed the cubic phase Gd_2O_3 with some monoclinic CuO impurities. SEM images have shown sheet-like structures of the prepared phosphors with thickness in the nanometer range. Optical spectroscopy of the synthesized phosphors resulted in intense red and blue emissions upon a 325 nm laser excitation in the case of without and with CuO incorporation, respectively. The PL emission spectra have shown the emissions corresponding to Eu^{2+} and Eu^{3+} both which was affected by the CuO incorporation. The calculated colour coordinates of emitted radiation were found to be (0.60, 0.32), (0.21, 0.15), (0.18, 0.13) and (0.19, 0.14) in the case of without, 1 mol%, 3 mol% and 5 mol% CuO doping.

Acknowledgements

This research is supported by the South African Research Chairs Initiative of the Department of Science and Technology (84415) and National Research Foundation of South Africa. The financial support from the University of the Free State is also acknowledged. The SEM-EDS measurements were performed using Jeol JSM-7800 Field Emission SEM equipped with Oxford Aztec EDS.

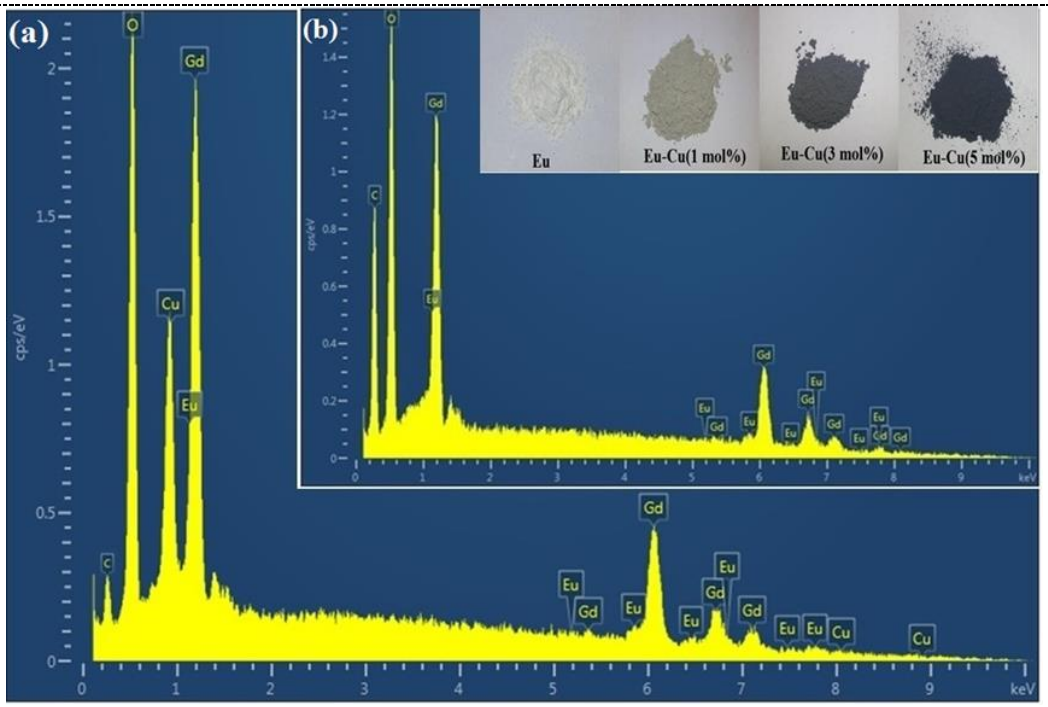


Figure 3. EDS images of the Gd_2O_3:Eu with (a) CuOand (b) without CuO nanosheets with their optical photographs as indicated.

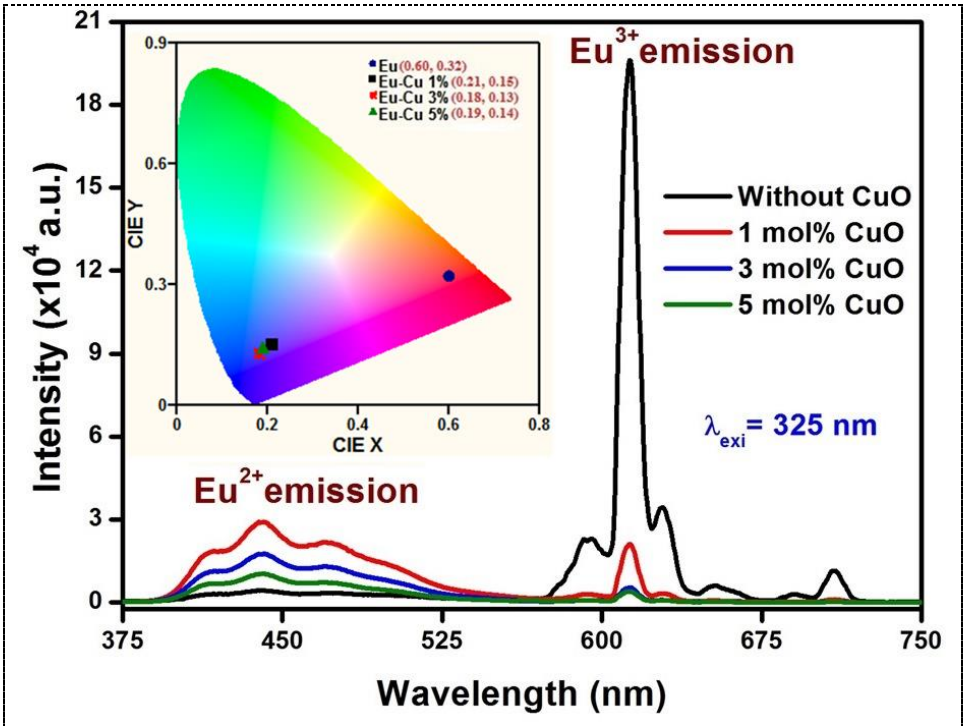


Figure 4. PL emission spectra of the Gd_2O_3:Eunanosheets with different CuO concentration and corresponding colour coordinates (inset).

References

[1] Fischer S., Goldschmidt J. C., LoperP., Bauer G. H., BruggemannR., Kramer K., BinerD., HermleM. andGlunzS. W. 2010*J. Appl. Phys.***108**044912.

[2] Singh S. K., Kumar K. andRai S. B. 2009*Appl. Phys. B*, **94** 165.

[3] Pandey A., Rai V. K., Kumar V., Kumar V. andSwart H. C. 2015*Sensor Actuat. B-Chem.***209** 352.

[4] HiralT. andOrikoshiT. 2004*J. Colloidal Interface Sci.***269** 103.

[5] MaalejN. B., QurashiA., AssadiA. A., MaalejR., Shaikh M. N., LlyasM. andGondalM. A. 2015*Nanoscale Res. Lett.* **10** 215.

[6] Dwivedi Y., BahadurA. andRai S. B. 2011*J. Appl. Phys.*, **110** 143103.

[7] SeoS., Yang H. and Holloway P. H. 2009*J. Colloidal Interface Sci.***331** 236.

[8] Li Y., Zhao Z., Song Z., Wan R., Qui J., Yang Z., Yin Z., Liu X., Liu Q. andZhou Y. 2015*J. Am. Ceram. Soc.*1-7 [DOI: 10.1111/jace.13589].

[9] Lee K. H., Bae Y. J. andByeonS. H. 2008*Bull. Korean Chem. Soc.* **29** 2161.

[10] Thomas R. andNampoori V. P. N. 1988*Solid State Commun.***68**821.

[11] Sharma S. K., Pitale S. S., Malik M. M., Gundu Rao T. K., Chawla S., Qureshi M. S. andDubey R. N. 2010*J. of Lumin.***130** 240.

[12] Pandey S., Pandey A. and Rai V. K. 2013*J. Display Technol.***9** 989.

[13] Kroon R. E. 2013 *S. Afr. J. Sci.* **109 (5/6)** 2 pages [http://dx.doi.org/10.1590/sajs.2013/a0019].

[14] Zhang J.-C., Long Y.-Z., Zhang H.-D., Sun B., Han W. P. and Sun X. Y. 2014 *J. Mater. Chem. C* **2** 312.

[15] Kumar V., Kumar V., Som S., Duvenhage M. M., Ntwaeaborwa O. M., Swart H. C. 2014 *Appl. Surf. Sci.***308** 419.

[16] Choudhury B., Dey M., Choudhury A. 2013 *Int. Nano Lett.***3** 25.

[17] Koao L. F., Dejene B. F., Swart H. C. 2014 *Physica B***439** 173.

A Nonlinear Optical loop Mirror enhanced three wavelength oscillations Erbium doped fiber laser source based on Fiber Bragg Grating reflectors

S. Qhumayo and R. Martinez

Photonics Research Group, Department of Electronic and Electrical Engineering
University of Johannesburg, PO Box524, Auckland Park, 2006
siyaz.qhumayo@gmail.com

Abstract. A three wavelengths erbium doped fiber laser source is developed. The configuration is a ring cavity comprises a single mode, single-clad 2.5 meter of Erbium doped fiber as a gain medium, an optical circulator, optical isolator for unidirectional operation and a 980/1550 nm wavelength division multiplexing (WDM) optical coupler for coupling the laser diode pump source and the optical feedback of the ring cavity laser source. The broadband spontaneous emission generated in the gain medium is coupled out of the ring cavity via port 1 through port 2 of the optical circulator. On this port, printed on a photosensitive single mode fiber, the Bragg gratings with 20% reflectivity at 1540nm, 1547nm and 1555nm are connected and their function is to select the oscillation wavelengths and to serve as the laser output coupler. To suppress the wavelength competition caused by homogeneous broadening of the Erbium doped fiber at room temperature, thereby enhancing the laser wavelength stability, a Nonlinear Optical loop Mirror with 2.5 km of single mode standard telecommunication fiber is used. Pumped by the 980nm laser diode at 80mW, the laser source has emission at 1540nm, 1547nnm and 1555nm with an average of 1 mW maximum optical power level in each oscillation wavelength.

1. Introduction

Multi-wavelength Erbium doped fiber lasers in recent years have attracted a great deal of research interest because of their applicability to telecommunications, fiber sensors and spectroscopy. For telecommunications, in particular, Erbium doped fiber lasers exhibit excellent characteristics that include their ability to emit simultaneous wavelength oscillations making them multi-wavelength optical sources that are potentially applied in wavelength division multiplexing (WDM) systems. In addition, these fiber lasers have emission in the optical fiber low loss wavelength region, commonly known as Telecommunications C-band, over which Erbium doped fiber Amplifiers (EDFA) operate.

The well-known pitfall of Multi-wavelength Erbium doped fiber lasers (MW-EDFL) is that, at room temperature, EDFL cavity tends to oscillate one wavelength at any particular time. This is caused by the wavelength competition that arises from the cross-gain saturation, existing as a result of the predominantly homogeneously broadened EDFA gain medium at room temperature, leading to unstable oscillations [1]-[2]. Much of the previous and current research in this area has focused on developing techniques to overcome this challenge. As such, several techniques to mitigate the effect of cross-gain saturation have been proposed and implemented with varying degrees of success. For example, it has been observed that the homogeneous line width of Erbium doped fiber exceeds 10nm at 290K and decreases with temperature, down to 1nm at 77K [2].

Hence, cooling the EDFA in liquid nitrogen at 77K to reduce the homogeneous line width and inserting the intracavity Fabry-Perot etalon as oscillation wavelength filter have been used [2]. Exploiting the intracavity polarization evolution for emission wavelengths, the similar cooling mechanism proved effective [3]. However, this technique is suitable for laboratory environment and inconvenient for field applications due to its bulkiness. The simultaneous lasing was achieved by using the inhomogeneous gain medium provided by a twin-core Erbium doped fiber [4]. Complexity is still an issue in this design. Some of the required characteristics of MW-EDFL are that they should meet the International Telecommunication union standard ITU-T which specifies the frequency spacing between adjacent channels to be in the order of 1.6 nm, 0.8nm, 0.4 nm and 0.2nm [5][6]; and that the power distribution between channels be uniform. From a design perspective, that emphasizes the importance of having control over which wavelengths to oscillate in a laser system. Using separate gain media for each oscillation wavelength guarantees simultaneous wavelength oscillation at any given time [7]. However meeting ITU-T standards using this configuration will be a challenge. Incorporating Raman amplifier or semiconductor optical amplifier in the laser cavity can also mitigate the homogeneous broadening in MW-EDFL [8][9][10], however, this is not a pure Erbium doped fiber laser system, but a hybrid of gain media. That might need additional design consideration since the gain media behavior is different.

An intensity dependent loss mechanism introduced by a nonlinear loop mirror is by far proving to be the simplest method of mitigating the cross-gain saturation in multi-wavelength Erbium doped fiber lasers [11][12][13][14]. In all the above reported work, the focus has been on multiwavelength generation on an Erbium doped fiber gain medium, with little or no intention to gain control on the wavelengths oscillation in the laser cavity. In this work, we exploit the cavity intensity dependent loss ability of the nonlinear loop mirror to stabilize the power and oscillation wavelength of the fiber laser with complete control over laser oscillation wavelengths. This is achieved by making use of the laboratory designed single mode fiber printed Fiber Bragg gratings as the wavelength selecting filters. Thus, the wavelength spacing, the number of wavelength and the exact center wavelength oscillation can be controlled using our laser configuration. The ability to control the above laser parameters is important for the ITU-T standardization of laser sources for WDM systems [6]. Experimental characterization have been performed, and a three wavelength laser with oscillations at 1540nm, 1547nm and 1555nm have been achieved. This comes after adjusting the length of the NOLM to an optimal value convenient for stability in our configuration.

2. Experimental Setup and principle of operation

A schematic of the experimental setup is shown on figure 1. A laser diode pump of 980nm wavelength and 80 mW of pump power is used. This pump power is coupled to the 2.5 m Erbium doped fiber gain medium by a WDM coupler. To ensure the uni-directionality of the laser, an optical isolator is inserted in the fiber ring cavity of the laser. A nonlinear Optical loop Mirror (NOLM) with a 70:30 loop power splitting ratio is inserted in the ring cavity to provide the intensity depend loss (IDL) for the lasing wavelengths. An optical circulator is used to couple out of the cavity, the amplified spontaneous emission (ASE) generated in the EDF from which the oscillation wavelengths are filtered by series configured Fiber Bragg gratings with 20 % reflectivity at 1540nm, 1547nm and 1555nm. The FBG also serve as the output coupler of the laser in which the feedback power is 20% of the cavity power and the output is 80% of the cavity power.

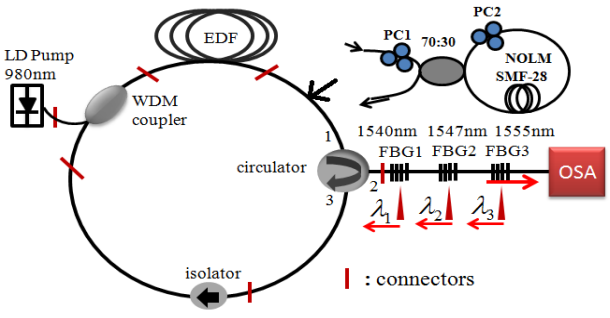


Figure 1: Schematic diagram of the experimental setup of the proposed three wavelength laser and the NOLM

The output of the laser was interrogated by a Optical spectrum analyzer (OSA) with resolution set to 0.05nm.The experiment was run at room temperature, with no loop mirror first and with loop mirror. The length of the loop mirror was varied. A loop mirror with lengths 500m, 1250m, 1700m, 2500m and 3000m was used.

3. Results and discussion

In a series of experiment the configuration in figure 1 was used, without the loop in the first place. The output of the laser was monitored over an approximately 3 hour period using the optical spectrum analyser and the laser spectrum in 3D view is shown in figure 2.

As can be seen in figure 2, there is only one wavelength at 1540 nm oscillating in the laser cavity suggesting a huge instability and strong cross-gain saturation. Then the NOLM was introduced with varying the length of the loop from 500 m -3000m. When the length of 500m was used, after careful adjustment of the polarization controllers, there was no change in the lasing wavelength and the power fluctuations remained between -5dBm and 5dBm as shown in figure 3 after a 3 hour monitoring. This suggests the persistence of the gain depletion by a dominant wavelength at 1540nm.

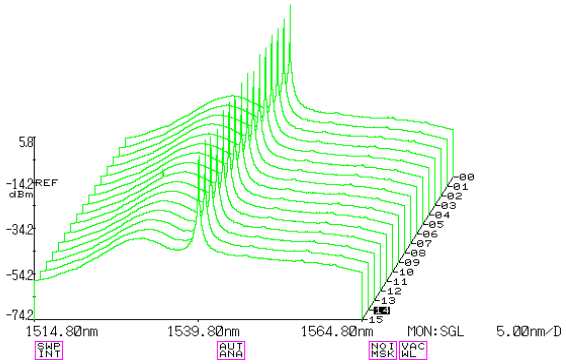


Figure 2: Spectrum of the laser without NOLM

For convenience, we have shown on figure 4, the change in the lasing wavelength and power stability improvement with the increase in the NOLM length. It was observed that when the loop length was 1250m, a change in the lasing wavelength was introduced, although it was only an addition of one more wavelength, suggesting the weakening of the gain depletion by a dominant wavelength, as shown on figure 4.

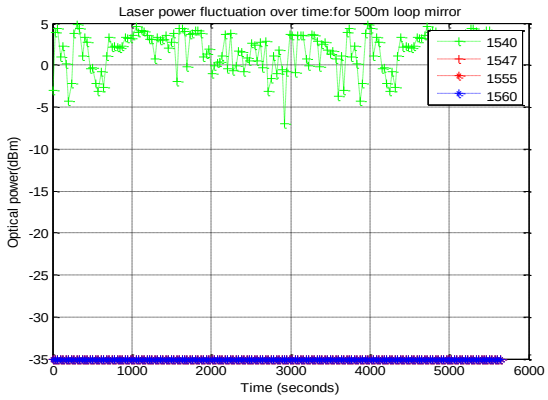


Figure 3:Wavelength and laser power stability of 1aser without 500m of NOLM length

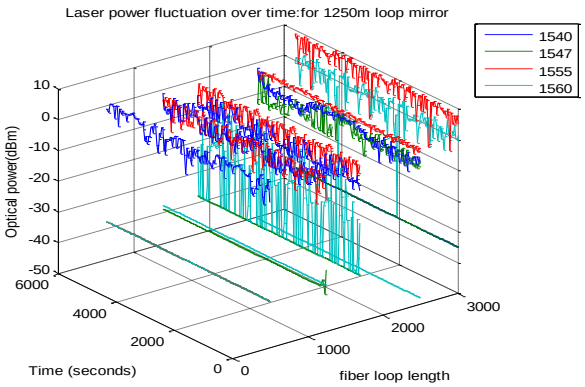


Figure 4:Laser power stability and lasing wavelength as a function of NOLM length.

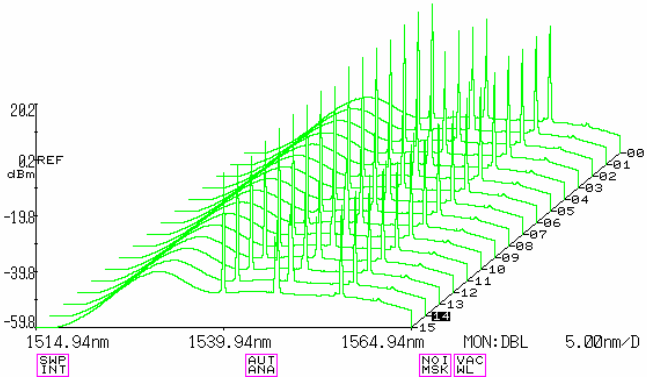


Figure 5:Emission spectrum of a three wavelengths laser with 2.5km Of NOLM length

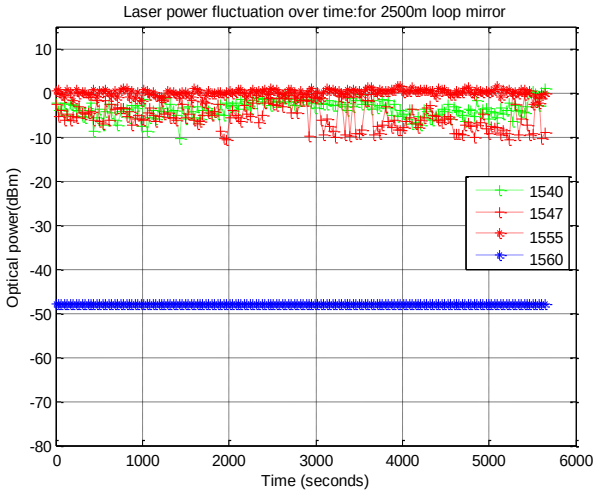


Figure 6: Laser power and wavelength stability of a three wavelength laser with 2.5km of NOLM length

From figure 4, it can be seen that the increase in the length of the loop NOLM increase the number of wavelengths oscilatting the cavity. However, it was also obseverved that there is a length above which the further increase in the loop length affects both the stability and the lasing wavelength of the laser. In our experiments, the length above which the satability start to decrease is 2500m, as seen in figure 4. This suggest an optimal length of 2500m is apropriate for both the maximum laser power stability and the lasing wavelength under this configuration. In figure 5, a spectrum of a three wavelength fiber laser with a NOLM length of 2500m is shown.

Division D1 – Astrophysics

Very-high energy emission from pulsars

M Breed[1], C Venter[1] and A K Harding[2]

[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

E-mail: 20574266@nwu.co.za

Abstract. The vast majority of pulsars detected by the *Fermi* Large Area Telescope (LAT) display exponentially cutoff spectra with cutoffs falling in a narrow band around a few GeV. Early spectral modelling predicted spectral cutoffs at energies of up to 100 GeV, assuming curvature radiation. It was therefore not expected that pulsars would be visible in the very-high-energy (VHE) regime ($>$ 100 GeV). The *VERITAS* announcement of the detection of pulsed emission from the Crab pulsar at energies up to 400 GeV (and now up to 1.5 TeV as detected by *MAGIC*) therefore raised important questions about our understanding of the electrodynamics and local environment of pulsars. *H.E.S.S.* has now detected pulsed emission from the Vela pulsar down to tens of GeV, making this the second pulsar detected by a ground-based Cherenkov telescope. Deep upper limits have also been obtained by *VERITAS* and *MAGIC* for the Geminga pulsar. We will review the latest developments in VHE pulsar science, including an overview of the latest observations, refinements, and extensions to radiation models and magnetic field structures, and the implementation of new radiation mechanisms. This will assist us in understanding the VHE emission detected from the Crab pulsar, and predicting the level of VHE emission expected from other pulsars, which is very important for the upcoming *CTA*.

1. Introduction

Since the launch in June 2008 of *Fermi* LAT [12], a high-energy (HE) satellite measuring γ-rays in the range 20 MeV$-$300 GeV, two pulsar catalogues (1PC, [1]; 2PC, [2]) discussing the light curve and spectral properties of 117 pulsars have been released. The vast majority of the *Fermi*-detected pulsars display exponentially cutoff spectra with cutoffs around a few GeV. These spectra are believed to be due to curvature radiation (CR), which is assumed to be the dominating emission process in the GeV band (see Section 1.2).

1.1. Standard pulsar emission models

There exist several physical radiation models that can be used to study HE emission from pulsars. These include the polar cap (PC; [15]), slot gap (SG; [11]), outer gap (OG; [14]), and the pair-starved polar cap (PSPC; [20]) models, which can be distinguished from each other based on the different assumptions of the geometry and location of the 'gap regions'. The 'gap region' is where particle acceleration takes place due to an unscreened, rotation-induced E-field parallel to the local B-field, as well as subsequent emission by these particles.

In PC models emission from HE particles is assumed to originate close to the neutron star (NS) surface. These particles are accelerated by large E-fields near the magnetic poles (known

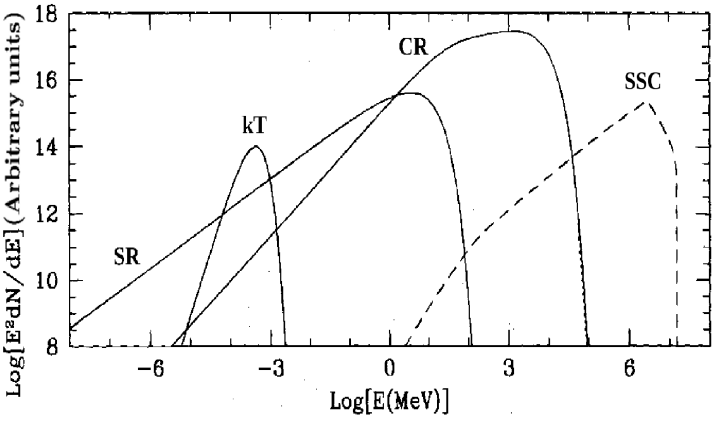


Figure 1. Early prediction of the phase-averaged spectrum for the Vela pulsar. The solid lines represent the spectral components for CR, SR, and the thermal surface flux (kT). The dashed curve represents the TeV pulsed spectral component associated with the ICS of SR of the primary $e^{\pm}$ (SSC). Adapted from [29].

as the magnetic PCs) only up to a few stellar radii. In SG models, the radiation comes from narrow gaps close to the last open field lines (the field lines that are tangent to the light cylinder where the corotation speed equals the speed of light c), with the gaps extending from the NS surface up to high altitudes. In the OG model, the gap region extends from the null-charge surface, where the Goldreich-Julian charge density is zero [17] up to high altitudes, also close to the last open field lines. The PSPC model involves a gap region that extends from the NS surface to the light cylinder over the full open volume [20], since the potential is unscreened in this case, so that there are not enough pairs to fully screen the E-field.

1.2. Radiation and pair creation processes

To explain HE emission in the standard models, one has to take detailed particle transport and radiation mechanisms into account. These mechanisms include CR, synchrotron radiation (SR), and inverse Compton scattering (ICS). CR occurs whenever charged particles are constrained to move along curved paths, e.g., along curved B-field lines (e.g., [18]), therefore involving a change in their longitudinal kinetic energy. When the emitted CR photon energy and the local B-field are high enough, magnetic pair production may occur (where an HE photon converts into an electron-positron pair, $e^{\pm}$), leading to a cascade of $e^{\pm}$ pairs which may screen the parallel E-field outside the gaps. The pair cascade is characterized by the so-called multiplicity, i.e., the number of pairs spawned by a single primary. The pairs may radiate SR if they have velocity components perpendicular to the local B-field so that this process involves a change in the particles' transverse kinetic energy. Also, ICS occurs due to the relativistic particles which upscatter soft photons (e.g., originating at a heated PC), which results in the "boosting" of the photon energies up to very high energies. ICS photons may also be converted into $e^{\pm}$ pairs. Two-photon pair creation ($\gamma\gamma$-absorption) may also occur, in particular in OG models. Cyclotron emission combined with subsequent ICS has also been considered by [28] to explain the broadband spectrum of the Crab pulsar.

1.3. Historic perspective of VHE spectral modelling

Early modelling, assuming the standard OG model, predicted spectral components in the VHE regime when estimating the ICS of primary electrons on SR or soft photons. This resulted in a natural bump around a few TeV (involving $\sim$ 10 TeV particles) in the extreme Klein-

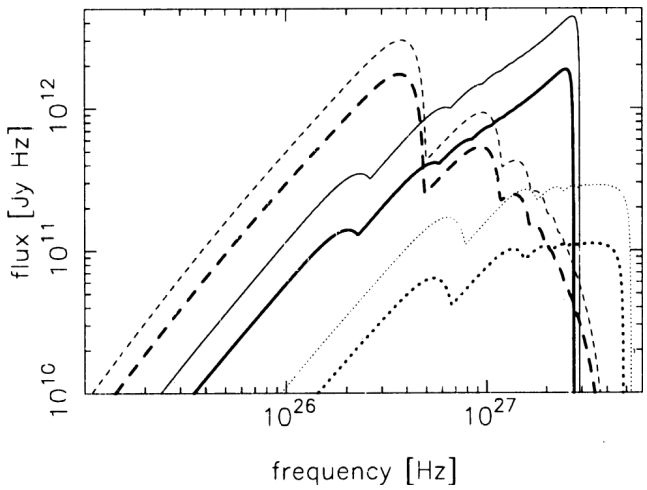


Figure 2. Expected TeV spectra for three bright pulsars including the Crab (solid lines), PSR B0656+14 (dashed lines), and PSR B1509−58 (dotted lines). The thick and thin curves represent inclination angles (between the spin and magnetic axis) of $\alpha = 30°$ and 45°, respectively [23].

Nishina limit as seen in figure 1 and 2. However, these components may not survive up to the light cylinder and beyond [14, 29, 23], since $\gamma\gamma$ pair creation leads to absorption of the TeV γ-ray flux. Other studies assumed CR to be the dominant radiation mechanism producing γ-ray emission when performing spectral modelling and found spectral cutoffs of up to 50 GeV. For example, [13] modelled the cutoffs of millisecond pulsars (MSPs). These are pulsars which possess relatively low B-fields and short periods. Their model assumed a static dipole B-field and a PC geometry, and predicted CR from the primary electrons that are released from the PC. Their predicted CR spectral component cut off at $\sim$ 100 GeV. The CR photons may undergo pair production in the intense low-altitude B-fields, and the newly formed electron-positron secondaries will emit SR in the optical and X-ray band (see [19]). Therefore, they concluded that the HE CR from MSPs occurred in an energy band that was above the detection range of satellite detectors like *Energetic Gamma-Ray Experiment Telescope (EGRET)* and below that of ground-based Cherenkov detectors such as the *High Energy Stereoscopic System (H.E.S.S. I)*.

Later studies investigated the X-ray and γ-ray spectrum of rotation-powered MSPs using a PSPC model [20], and found CR cutoffs of $\sim 10-50$ GeV (see [16, 32]). Optical to γ-ray spectra were also modelled by [21] assuming an SG accelerator and a retarded vacuum dipole (RVD) B-field, for the Crab pulsar. They found spectral cutoffs of up to a few GeV. Another study modelled the phase-resolved spectra of the Crab pulsar using the OG and SG models, and found HE cutoffs of up to $\sim$ 25 GeV [24]. Cutoffs around $\sim$ 10 GeV were found for the OG model using the RVD B-field [31].

2. Observational revolution

In view of the above theoretical paradigm it was not expected that a pulsar should be visible in the VHE regime. It was therefore surprising when the *Very Energetic Radiation Imaging Telescope Array System (VERITAS)* announced the detection of pulsed emission from the Crab pulsar above $\sim$ 100 GeV [8], followed by the detection by the *Major Atmospheric Gamma-Ray Imaging Cherenkov (MAGIC)* of emission up to $\sim$ 400 GeV (soon after their initial detection of emission at $\sim$ 25 GeV) [6, 7]. The *MAGIC* Collaboration has since reported the detection of pulsed photons with energies up to 1.5 TeV [10]. Ground-based Cherenkov telescopes are now searching for more examples of VHE pulsars.

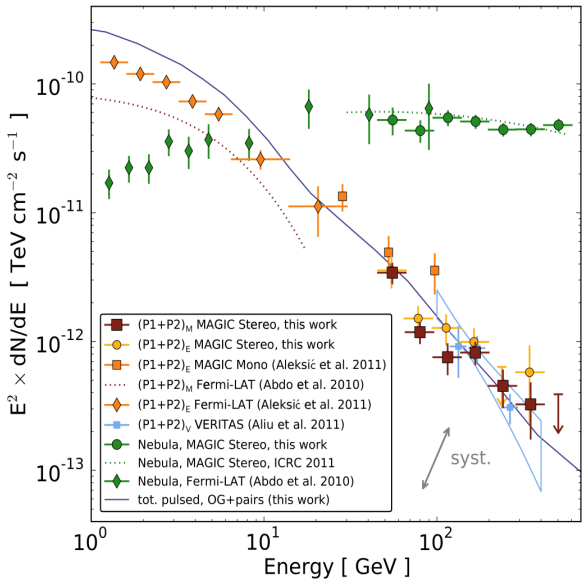


Figure 3. The observed and modelled spectra for phase-resolved γ-ray emission, including both emission peaks P1 and P2, from the Crab pulsar as measured by *MAGIC* (dark red squares). The plot also contains measurements from *Fermi* LAT (orange diamonds) and *VERITAS* (light blue squares and solid line). An OG model including emission from pairs are assumed. The systematic error of the *MAGIC*-Stereo measurement is shown and corresponds to a shift of $\pm 17\%$ in energy and $\pm 19\%$ in flux. From [6].

In figure 3 the phase-resolved γ-ray spectrum for the Crab pulsar for both light curve peaks P1 and P2, as measured by *MAGIC*, is shown. The spectrum also contains measurements by *Fermi* LAT and *VERITAS*. The observations by *MAGIC* are in good agreement with those by *VERITAS*. The *VERITAS* data fit a broken power law, and raised important questions whether the CR component is extended or if a second component is required to explain the observed spectra.

The detection of the Crab pulsar above several GeV prompted *Fermi* to search for pulsed emission at HEs. They detected significant pulsations above 10 GeV from 20 pulsars and above 25 GeV from 12 pulsars [3]. The Crab pulsar is the first source which have been detected over almost all energies ranging from radio to VHE γ-rays. More recently, pulsed emission was detected from the Vela pulsar above 30 GeV with the *H.E.S.S.* [30] and up to ~ 80 GeV (at the 4σ-level) with the *Fermi* LAT [26]. *MAGIC* furthermore detected no emission from Geminga above 50 GeV [4]; neither did *VERITAS* above 100 GeV [9]. A stacking analysis involving 115 *Fermi*-detected pulsars (excluding the Crab pulsar) was performed by [25] using *Fermi* data. However no emission above 50 GeV was detected, implying that VHE pulsar detections may be rare, given current telescope sensitivities. From all these observations there are three effects visible in the energy-dependent pulse profiles: the peaks remain at the same phase, the P1/P2 ratio decreases as energy increases, and the pulse width decreases with increasing energy. More VHE pulsars may be found by the *Cherenkov Telescope Array (CTA)* which will have a ten-fold increase in sensitivity compared to present-day Cherenkov telescopes.

3. Theoretical ideas

All of the standard pulsar emission models (see Section 1.1) predicted HE spectral cutoffs between a few GeV and up to ~ 100 GeV, assuming B-fields such as the static dipole and RVD solutions. Clearly, refinements to these radiation models and B-fields are needed to explain the

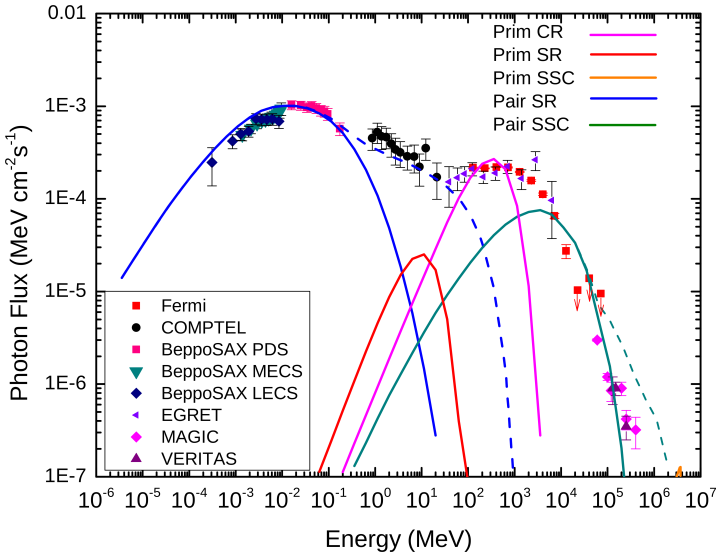


Figure 4. Modelled spectrum of phase-averaged pulsed emission from the Crab pulsar. The spectral components from primary electrons and pairs (as labelled) are for a magnetic inclination angle $\alpha = 45°$, observer angle $\zeta = 60°$ and pair multiplicity $M_+ = 3 \times 10^5$. From [22].

observed VHE emission from the Crab. There are a few ideas for such refinements. One is a revised OG model by [24] which can produce IC radiation of up to ~ 400 GeV due to secondary and tertiary pairs upscattering infrared to ultraviolet photons [6]. In this OG model the IC flux depends sensitively on the B-field structure near the light cylinder.

Another idea was proposed by [27], invoking the SSC radiation process. This is indeed a promising radiation mechanism, where relativistic particles upscatter the SR photons emitted by the same population. The SSC radiation mechanism was applied by [22] to predict optical to X-ray, and γ-ray spectra (see figure 4) assuming an SG model and a force-free B-field. This process relies critically on the assumed electrodynamics and the magnetospheric structure. They performed simulations for the Crab and Vela pulsars, as well as two MSPs, i.e., B1821−24 and B1937+21. However, the only significant predicted SSC component was for the Crab pulsar. They also found that the pair SR matched the observed X-ray spectrum of the MSPs. They furthermore tested the addition of an HE power law extension to the pair spectrum (dashed lines in figure 4) whose SR spectrum would account for the observed emission in the $1 - 100$ MeV range. However, the resulting SSC component exceeded the observed *MAGIC* and *VERITAS* points, implying that the observed $1 - 100$ MeV emission is not produced by the same particles that produce the SSC emission.

4. Conclusions and future prospects

The abovementioned detections of the VHE pulsed emission from pulsars and the explanation thereof implies that this emission may yield strong constraints on γ-ray radiation mechanisms, the location of acceleration regions, and the B-field structure. There is thus an urgent need for refinements and extensions of standard pulsar emission models and radiation mechanisms, including more realistic B-fields. Some examples are discussed above, but there are many more.

SG model refinements include photon-photon pair production attenuation within the model and also more realistic B-fields such as dissipative magnetospheric solutions. One could also model the emission pulse profiles as a function of energy.

Ground-based Cherenkov telescopes are now searching for more examples of VHE pulsars.

New pulsar models will assist us in predicting the level of VHE emission expected from them, which would be very important for the upcoming *CTA*. The low threshold energy of *CTA* will provide an overlap with the *Fermi* energy range and will help to discriminate between CR and a potentially new spectral component.

Acknowledgments
This work is based on research supported by the National Research Foundation (NRF) of South Africa (Grant Numbers 90822, 93278, and 99072). Any opinions, findings, and conclusions or recommendations expressed are that of the authors, and the NRF accepts no liability whatsoever in this regard. A.K.H. acknowledges the support from the NASA Astrophysics Theory Program.

References

[1] Abdo A A *et al.* 2010 *Astrophys. J. Supp.* **187** 460
[2] Abdo A A *et al.* 2013 *Astrophys. J. Supp.* **208** 17
[3] Ackermann M *et al.* 2013 *Astrophys. J. Supp.* **209** 34
[4] Ahnen, M L *et al.* 2016 *ArXiv:1603.00730*
[5] Aleksić J *et al.* 2011 *Astrophys. J.* **742** 43
[6] Aleksić J *et al.* 2012 *Astrophys. Astr.* **540** A69
[7] Aliu E *et al.* 2008 *Science* **322** 1221
[8] Aliu E *et al.* 2011 *Science* **334** 69
[9] Aliu E *et al.* 2015 *ApJ* **800** 61
[10] Ansoldi S *et al.* 2016 *Astron. & Astrophys.* **585** A133
[11] Arons J 1983 *Astrophys. J.* **266** 251
[12] Atwood W B *et al.* 2009 *Astrophys. J.* **697** 1071
[13] Bulik T, Rudak B and Dyks J 2000 *MNRAS* **317**, 97
[14] Cheng K S, Ho C and Ruderman M 1986 *Astrophys. J.* **300** 522
[15] Daugherty J K and Harding A K 1982 *Astrophys. J.* **252** 337
[16] Frąckowiak M and Rudak B 2005 *Adv. Space Res.* **35** 1152
[17] Goldreich P and Julian W H 1969 *Astrophys. J.* **157** 869
[18] Harding A K 1981 *Astrophys. J.* **245** 267
[19] Harding A K, Muslimov A G and Zhang B 2002 *Astrophys. J.* **576** 366
[20] Harding A K, Usov V V and Muslimov A G 2005 *Astrophys. J.* **622** 531
[21] Harding A K, Stern J V, Dyks J and Frackowiak M 2008 *Astrophys. J.* **680** 1378
[22] Harding A K and Kalapotharakos C 2015 *Astrophys. J.* **811** 63
[23] Hirotani K 2001 *Astrophys. J.* **549** 495
[24] Hirotani K 2008 (arXiv:0809.1283)
[25] McCann A *et al.* 2015 *Astrophys. J.* **804** 86
[26] Leung C Y *et al.* 2014 *Astrophys. J.* **797** L13
[27] Lyutikov M, Otte N and McCann A 2012 *Astrophys. J.* **754** 33
[28] Lyutikov M 2013 *MNRAS* **431** 2580
[29] Romani R W 1996 *Astrophys. J.* **470** 469
[30] Stegmann C 2014 talk at "Astroparticle Physics 2014" Amsterdam
[31] Tang A S P *et al.* 2008 *Astrophys. J.* **676** 562
[32] Venter C and de Jager O C 2005 *Astrophys. J.* **619** L167

Spectral studies of flaring quasar PKS 1424-418 above 100 MeV with *Fermi*-LAT

Feraol F. Dirirsa[1,*], Richard J. Britto[1] and Soebur Razzaque[1]
on behalf of the *Fermi*-LAT Collaboration

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

E-mail: *fdirirsa@uj.ac.za

Abstract. The Flat Spectrum Radio Quasar (FSRQ) PKS 1424-418 is an Active Galactic Nucleus (AGN) located at a redshift z = 1.522. This source has shown several flaring episodes through the whole electromagnetic spectrum in recent years. The *Fermi*-Large Area Telescope (*Fermi*-LAT), a space-based gamma-ray detector, has detected four outbursts during the 2012 October to 2013 September period, which were also followed up by the Hartebeesthoek Radio Astronomy Observatory (HartRAO). We present an analysis of *Fermi*-LAT data on PKS 1424-418 during this period. Studies of the flaring pattern of FSRQs, such as PKS 1424-418, can provide interesting constraints related to the physics of the gamma-ray production in these objects.

1. Introduction

The *Fermi* Gamma-Ray Space Telescope has been collecting data from the whole sky since 2008 [1]. The majority of the extragalactic objects observed by *Fermi*'s main instrument—the Large Area Telescope (LAT), are classified as blazars [2]. They constitute a large group of active galaxies that host an active galactic nucleus (AGN). In the popular unification scenario of AGNs, blazars are understood as sources in which a relativistic jet is produced by the central engine and aligned with the observer's line of sight [3]. These objects are marked by a high flux from the radio to the gamma-ray band and their spectral energy distributions (SEDs) are characterized by a double-peak structure [4], one peaking between the Infrared and the X-ray region, the other one at gamma-ray energies. The low energy peak can be widely interpreted as synchrotron radiation from relativistic electrons in the jets, while the second high frequency bump is presumably due to inverse Compton scattering between the same electron population interacting either with the synchrotron soft photons (synchrotron self-Compton process) or with other photons (external Compton process) that originated in the local environment [4].

Currently LAT onboard *Fermi* permits an unprecedented view of the SED and variability of blazars, especially the most powerful flat spectrum radio quasars (FSRQs). In this work, we study spectra of four prominent outburst episodes of the FSRQ PKS 1424-418 ($z = 1.522$) that were detected by *Fermi*-LAT from 2012 October to 2013 September [5]. The Automatic Telescope for Optical Monitoring (ATOM) also provided a dense optical light curve coverage for this source [6]. PKS 1424-418 is a bright, compact radio source and a highly polarized optical quasar. The bulk emission of this radiation may occur in the innermost regions of the AGN,

within the broad-line region (BLR). This source was in an active phase in 2012–2013, during which the large flux allowed us to track the spectral behaviour of the source. Therefore, we study this flaring pattern in order to investigate possible constraints related to the physics of the gamma-ray production in FSRQs.

2. Method and Data processing

We analysed the *Fermi*-LAT data between 2012 October 01 and 2013 September 30, corresponding to the Modified Julian Days (MJD) 56201.0 to 56565.0, using the likelihood analysis package of the standard *Fermi Science Tools* v9r33p0, publicly available at the *Fermi* mission website[1]. We used the *Pass 7 reprocessed* data representation, the "source" event class with energy between 100 MeV and 300 GeV, and the P7REP_SOURCE_V15 instrument response functions. The maximum zenith angle was set to 100° to avoid contamination from the Earth limb. Photons were selected in a circular region of interest (ROI) of 10° radius, centered at the position of the source of interest. We took into account the diffuse backgrounds (modeled using "gll_iem_v05_rev1.fit" for the Galactic diffuse emission and "iso_source_v05.txt" for the isotropic background model).

The Enrico software, an user contributed package built upon the python backbone of the science tools, is also used to analyse the spectral energy distribution of the source of interest. We have performed an unbinned likelihood analysis (using the "gtlike" *Science Tool* [7]) of the *Fermi*-LAT data by using a source model in which the spectrum of all the sources of the third *Fermi*-LAT source Catalog (3FGL) located within a 20° radius from PKS 1424-418 are modeled. The likelihood analysis consist in optimising the value of the spectral parameters of the modeled sources by a maximum likelihood statistics procedure.

In the following sections where our results are presented, the "PRELIMINARY" label was added to the figures, since our results constitute a work in progress that requires further validation within the *Fermi*-LAT Collaboration.

2.1. Spectral energy distribution analysis

SEDs of blazars are well modeled by a log-parabola function. This model has been already used with success to fit the spectra of blazars in the catalog of 157 X-ray SEDs observed with BeppoSAX [8]. The log-parabolic model of equation (1) is one of the simplest ways to represent curved spectra when they do not show a sharp high energy cut-off like that of an exponential. We write the log parabola function as

$$dN(E)/dE = N_0' \left(\frac{E}{E_b}\right)^{-\left(\alpha+\beta\log\left(\frac{E}{E_b}\right)\right)} \text{ ph/(cm}^2 \text{ s keV)}, \quad (1)$$

where N_0' is a prefactor, the α index denotes the slope at $E = E_b$ and the β index models the curvature of the log-parabola[2]. We fixed the reference energy E_b at 415.8×10^3 keV as in 3FGL and therefore the spectrum is completely determined only by the three parameters N_0', α and β. The energy flux is given by

$$F(E) = EdN(E)/dE = EN_0' \left(\frac{E}{E_b}\right)^{-\left(\alpha+\beta\log\left(\frac{E}{E_b}\right)\right)}. \quad (2)$$

[1] http://fermi.gsfc.nasa.gov/ssc/data/analysis/software/

[2] When used as a unit, we abbreviate "photon" as "ph"

Thus to see the energy at which the source luminosity peaks on a logarithmic frequency scale, we have to plot $\log EF(E)(= \nu F(\nu))$ versus $\log E$, *i.e.* the energy flux is then simply written as

$$E^2 dN(E)/dE = (1.60 \times 10^{-9}) E^2 N_0' \left(\frac{E}{E_b}\right)^{-\left(\alpha+\beta \log\left(\frac{E}{E_b}\right)\right)} \text{ erg/(cm}^2 \text{ s)}, \quad (3)$$

where the numerical constant is the energy conversion factor from keV to erg and N_0' is a prefactor, such as $N_0' = N_0/(1.60 \times 10^{-9})$ and N_0 being the prefactor of the spectrum in this rescaled model. A power law with an exponential cut-off (PLExpCutoff) model has also been used in our unbinned likelihood analysis, given by

$$E^2 dN(E)/dE = (1.60 \times 10^{-9}) E^2 N_0' \left(\frac{E}{E_0}\right)^{\gamma_1} \exp\left(-\left(\frac{E}{E_c}\right)\right) \text{erg/(cm}^2 \text{ s)}, \quad (4)$$

where E_0 is the pivot energy at which the error in the differential photon flux is minimal, E_c is the cut-off energy, and γ_1 is the photon spectral index. E_0 was arbitrary fixed at 100 MeV.

2.2. Light curve

Variations in flux of PKS 1424-418 are represented by light curves that we draw in a one-day time binning. Though the spectral shape of this source exhibits a definite curvature, we chose to model it with a simple power law, in order to minimize statistical errors on the photon indices. Since the differential flux of a power law model can be represented by $dN(E)/dE = N_0 (E/E_0)^{-\gamma}$, the integral flux was then computed between 100 MeV and 300 GeV, such as:

$$F(\text{E} > 100 \text{ MeV}) = \int_{100 \text{ MeV}}^{300 \text{ GeV}} N_0 \left(\frac{E}{E_0}\right)^{-\gamma} dE. \quad (5)$$

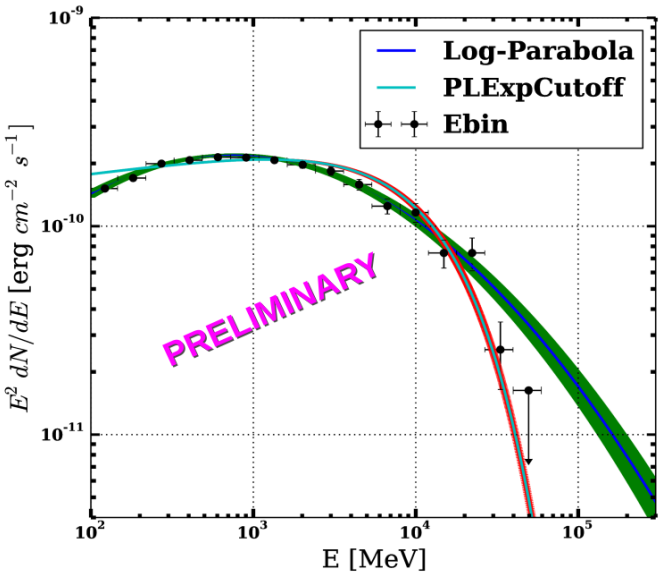


Figure 1. SED of PKS 1424-418, from 2012 October 01 to 2013 September 30 (MJD 56201–56565) fitted with a LP function (fitting parameters: $N_0 = 7.63 \times 10^{-10^{+8.6\times10^{-12}}_{-8.7\times10^{-12}}}$ erg^{-1}cm^{-2} s^{-1}, $\alpha = 1.87 \pm 0.01$, $\beta = 0.11 \pm 0.01$, $E_b = 415.8$ MeV) and a PLExpCutoff function (with fitting parameters: $N_0 = 1.12 \times 10^{-8}$ erg^{-1}cm^{-2} s^{-1}, $E_0 = 100$ MeV, $E_c = 11688^{+865}_{-1146}$ MeV, $\gamma_1 = -2.0$.) The likelihoods of the fitting procedures are given by the values $-loglike = 150379.3$ and 150394.0 for the LP and PLExpCutoff fits respectively.

Table 1. Flaring activity of PKS 1424-418 from 2012 October until 2013 September.

Source periods	Start & Stop date(MJD)	Start & Stop date (UTC)	Duration (days)
Period A	56201.0 – 56257.0	2012 Oct. 01 – 2012 Nov. 26	57
Period B	56257.0 – 56362.0	2012 Nov. 26 – 2013 Mar. 11	105
Period C	56362.0 – 56461.0	2013 Mar. 11 – 2013 Jun. 18	99
Period D_1	56461.0 – 56504.0	2013 Jun. 18 – 2013 Jul. 31	43
Period D_2	56504.0 – 56565.0	2013 Jul. 31 – 2013 Sep. 30	61

Hard spectra have $\gamma < 2$ with peak at high energies, soft spectra have $\gamma > 2$ with peak at low energies and flat spectra with $\gamma = 2$ means the spectrum has a similar amount of power in all spectral bands. In this analysis, the power law model has been used to analyze the light curves that were produced using the gtlike tool.

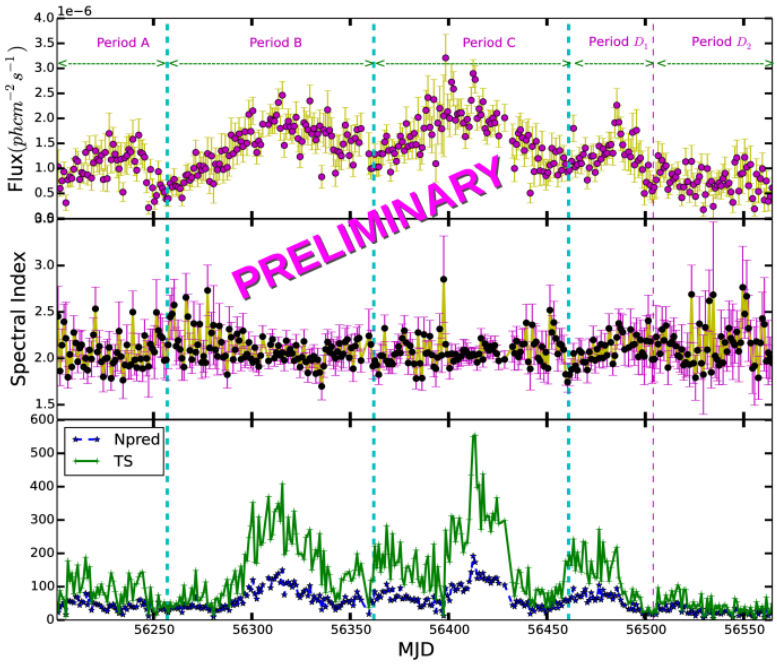


Figure 2. Top panel: Light curve of daily flux (periods categorized according to Table 1). Middle panel: Photon spectral index as a function of time. Bottom panel: TS and N_{pred} versus time.

3. Data analysis

Fig. 1 shows the 12-month averaged SED of PKS 1424-418 (corresponding to the whole data set shown in Table 1), along with the log-parabola and PLExpCutoff functions whose parameters were obtained by the unbinned likelihood analysis.

In Fig. 2, the top panel displays the light curve of the four flaring states of PKS 1424-418 (labeled as A, B, C, D_1) and post-flare (labeled as D_2), over a one-year period of one-day time binning. It exhibits flux variability across a wide time range. A detailed study of these flux variations can provide information on the physics and dynamics of the emitting region. In this

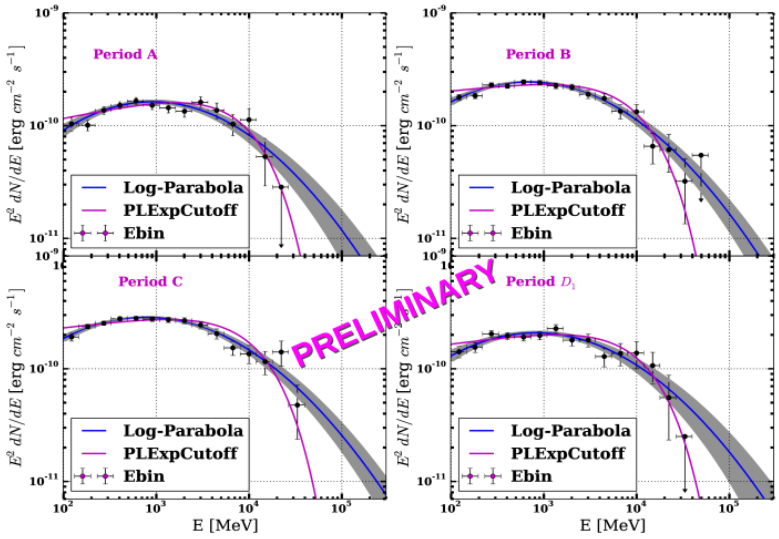


Figure 3. SEDs of four flaring period fitted by LP and PLExpCutoff models from energy range 100 MeV to 300 GeV. Fitting parameters of both models are given in Tables 2 and 3.

Table 2. Unbinned likelihood spectral fitting parameters and derived qualities of fit ($-loglike$) for PKS 1424-418, using the LogParabola (LP) model.

Time Range	N_0 ($\mathrm{erg^{-1}cm^{-2}\ s^{-1}}$)	α	β	E_b (MeV)	$-$loglike	χ^2(ndf)	$\chi^2_{\rm red}$
Period A	$(5.47\times10^{-10})^{+1.85\times10^{-11}}_{-1.81\times10^{-11}}$	1.81±0.04	0.12 ±0.02	415.8	46443.3	12.01(14)	1.33
Period B	$(8.58\times10^{-10})^{+1.67\times10^{-11}}_{-1.64\times10^{-11}}$	1.89±0.02	0.11±0.02	415.8	72051.0	10.45(15)	1.05
Period C	$(9.93\times10^{-10})^{+1.84\times10^{-11}}_{-1.83\times10^{-11}}$	1.87±0.02	0.10±0.01	415.8	67736.0	7.36(15)	0.74
Period D_1	$(5.55\times10^{-10})^{+1.48\times10^{-11}}_{-1.45\times10^{-11}}$	1.85±0.03	0.12±0.02	415.8	38742.6	9.61(15)	0.96

Table 3. Unbinned likelihood spectral fitting parameters and derived qualities of fit ($-loglike$) for PKS 1424-418, using the PLExpCutoff model.

Time Range	N_0 ($\mathrm{erg^{-1}cm^{-2}\ s^{-1}}$)	E_0 (MeV)	E_c (MeV)	γ_1	$-$loglike	χ^2(ndf)	$\chi^2_{\rm red}$
Period A	$(7.30\times10^{-09})^{+5.07\times10^{-10}}_{-5.00\times10^{-10}}$	100.00	9298^{+2291}_{-1716}	$-1.83^{+0.04}_{0.00}$	46442.5	14.65 (14)	1.63
Period B	$(1.28\times10^{-08})^{+4.72\times10^{-10}}_{-4.63\times10^{-10}}$	100.00	11087^{-1938}_{-1534}	-1.91 ±0.02	72058.9	25.20 (15)	2.52
Period C	$(1.45\times10^{-08})^{+5.18\times10^{-10}}_{-5.30\times10^{-10}}$	100.00	12694^{+1990}_{-1711}	-1.90 ±0.02	67742.6	27.83(15)	2.78
Period D_1	$(8.04\times10^{-09})^{+4.37\times10^{-10}}_{-4.30\times10^{-10}}$	100.00	10821^{+2302}_{-1765}	-1.89 ±0.03	38744.2	14.77(15)	1.48

figure, the vertical dashed lines shows the range of the five periods that we identified for an analysis based upon a high flux observed, and the test statistic sketched at the bottom panel in the figure is given by $TS = -2\log L_{\rm max,0}/L_{\rm max,1}$, where $L_{\rm max,0}$ and $L_{\rm max,1}$ are the maximum likelihoods of the model that includes or not the source of interest, respectively, and N_{pred} is the number of photon counts predicted by the optimised model.

As shown in Fig 2, a first flare was observed from the beginning of 2012 October and lasted for almost two months (period A). The second flare was observed in 2013 January, in the middle of period B. It was significantly brighter than the flare detected in 2012 October. The third

flare was observed in 2013 April (middle of period C) with a very high daily flux. The TS and N_{pred} were higher than the values of the previous recorded ones as shown in the bottom panel of Fig. 2. The fourth flaring episode (period D_1) was observed in 2013 July with a high daily flux—around 1×10^{-6} ph cm^{-2} s^{-1}, almost every day of this month. During the post-flare observations (period D_2), the source was fainter, with a photon index ranging between about 1.8 and 2.8. Flaring episodes often spanned from weeks to months, though rapid and intense sub-flares are observed in between.

Fig. 3 shows the SED distributions of PKS 1424-418, along with both the log-parabola (blue line) and PLExpCutoff (marble line) functions whose parameters where obtained by the unbinned likelihood analysis. The parameters of the different fits can be found in the Tables 2 and 3, respectively. The SED shapes have not changed significantly, though flux varied from one flaring episode to another. A particular hint of flattening of the SED was observed when the source became brighter (as shown in Fig. 2, period C). However, the SED is well fitted by a log-parabola function with $\chi^2_{\rm red} = 0.73$ over the whole energy range covered (E > 100 MeV).

Error bars represent the statistical errors while the vertical arrows show upper limits. Upper limits on the flux were calculated when the Test Statistic (TS) value for the source was lower than 9 (significance lower than $\sim 3\sigma$) or when N_{pred} was lower than 3. Results from the photon spectral index study suggest that hardening did not happen when the source was brighter (as it would be if following the general trend of bright FSRQs). There is rather a hint of hardening after the flaring peak of period B. However, statistical uncertainties are too large to clearly exhibit this feature.

4. Summary

The analysis of four flaring periods of PKS 1424-418 during its 2012 October–2013 September outburst was presented. The spectral shapes of this source during the whole period of flaring activities and the post-flaring episode were analysed using two spectral models, namely LP and PLExpCutoff. The spectral analysis were performed using the unbinned maximum likelihood estimator (gtlike science tool). The SEDs of individual flares were plotted and show that the spectrum was also curved during the single flaring episodes. The spectral shapes do not vary significantly between the four flaring episodes despite dramatic flux variations. On the light curve, we do not see higher flux coinciding with the hardening of the photon spectral index.

Acknowledgments

The *Fermi*-LAT Collaboration acknowledges support for LAT development, operation and data analysis from NASA and DOE (United States), CEA/Irfuand IN2P3/CNRS (France), ASI and INFN (Italy), MEXT, KEK, and JAXA (Japan), and the K.A. Wallenberg Foundation, the Swedish Research Counciland the National Space Board (Sweden). Science analysis support in the operations phase from INAF (Italy) and CNES (France) is also gratefully acknowledged. The work presented in this paper was supported in part by an MWGR 2015 grant from the National Research Foundation with Grant No. 93273.

References

[1] Atwood W. B., et al., 2009, *ApJ*, **697**, 1071
[2] Abdo A. A., et al., 2010, *ApJS*, **188**, 405
[3] Dermer C. D., 2014, *Astro-Ph.HE*, **1**, 1408
[4] Massaro E., et al., 2006, *A&A*, **448**, 861-71
[5] Stefano C., et al., 2013, *The Astronomer's Telegram*, **4770**, 1
[6] Buson S., et al., 2014, *A&A*, **40**, A569
[7] http://fermi.gsfc.nasa.gov/ssc/data/analysis/documentation/Cicerone/
[8] Massaro E., et al., 2004, *A&A*, **413**, 489-503.
[9] D'Ammando F., et al., 2013, *The Astronomer's Telegram*, **4714**, 1

The Vela Supercluster - does it provide the missing link to explain the local flow fields?

A Elagali, R C Kraan-Korteweg and M E Cluver

Astrophysics, Cosmology and Gravity Centre (ACGC), Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, SA

E-mail: ahmedagali70@gmail.com

Abstract. As part of a larger effort to uncover the structures hidden behind the Milky Way, we analyse 4,756 optical redshifts of galaxies in the Hydra/Antlia and Vela region ($245^\circ \leq l \leq 295^\circ, |b| \leq 10^\circ$). These redshifts were mainly obtained with the 2dF+AAOmega spectrograph at the Anglo-Australian Telescope (88% of the redshifts), the remaining data are obtained with other telescopes as well as taken from literature. This analysis is suggestive of the existence of a supercluster in this region, hereafter called the Vela Supercluster (Vela SCL). The Vela SCL is at a mean redshift of $cz \sim 18,000\,\mathrm{km\ s^{-1}}$, and extends about $70 \times 70\,\mathrm{Mpc}$ ($20^\circ \times 20^\circ$) on the sky. We use a nearest neighbour algorithm to identify the galaxy clusters and groups members, to determine their velocity dispersions and the corresponding virial masses. Although the Vela SCL is sparsely sampled, we find 13 galaxy clusters and 19 galaxy groups contained within it. The masses of these galaxy clusters/groups range from a few times $10^{14} - 10^{15}\,\mathrm{M}_\odot$. Taking account of the sparse sampling, the Vela SCL seems comparable to the Shapley Supercluster (Shapley SCL). This may have considerable implications on the bulk flow and peculiar velocity of the Local Group (LG), since the Shapley SCL contribution to the LG motion is $\sim 9\%$ (Muñoz & Loeb 2008).

1. Introduction

In this paper, we investigate an excess of galaxies discovered in the Zone of Avoidance (ZoA) in the Hydra/Antlia and Vela regions ($245^\circ \leq l \leq 295^\circ, |b| \leq 10^\circ$) at approximately $cz \sim 18,000\,\mathrm{km\ s^{-1}}$. This overdensity seems to extend over a large area of the sky, straddling the Galactic Plane. Based upon the velocity distribution of the galaxies and the number of galaxy clusters and groups found, we argue that this overdensity has the signature of a massive supercluster, hidden by the dust of the Milky Way. To put our results in context to known superclusters, we first give a short overview on the typical properties of superclusters at similar distances.

Superclusters are the largest, most massive and extended extragalactic systems in the Universe. These structures can extend over $\sim 30 - 100\,\mathrm{Mpc}$, and have masses of a few $10^{16}\,\mathrm{M}_\odot$ [1]. In general, superclusters tend to associate with regions of comparable size and very low density called "voids" [1]. In most cases they are connected through filaments and sheets of galaxies [2]. Studying the dynamics of these structures can provide us with an enhanced picture for structure formation theories [3, 4] as well as the distribution of dark matter in the universe [5, 6].

During the past few decades, large-scale surveys have revealed relatively nearby superclusters. One of the best studied superclusters in the local Universe is the Shapley Supercluster (Shapley SCL). It was first noted by Shapley [7], it has been explored spectroscopically [8], later in X-ray

[9], followed by multi-wavelength imaging [10]. This structure is centred at approximately (RA, Dec)=($13^{\rm h}$ $25^{\rm m}$,$-30°$). At redshift of $z \sim 0.05$, it extends over an area of $12° \times 30°$ and contains 33 galaxy clusters. The Shapley SCL occupies a spatial volume of $\sim 10^{15}\,{\rm Mpc}^3$, and has a mass of about $\sim 5 \times 10^{16}\,{\rm M}_\odot$ [10, 11]. In comparison with the Shapley SCL, the Horologium-Reticulum Supercluster (HRSC) is considered as the second most massive supercluster within 300 Mpc [12, 13]. The HRSC is centred at approximately (RA, Dec)= ($3^{\rm h}$ $19^{\rm m}$,$-50°$), extends over an area of $12° \times 12°$ at mean redshift $z \sim 0.066$, and contains 21 galaxy clusters [13]. Another massive supercluster within 300 Mpc is the the Corona Borealis Supercluster (CSC) [1, 14]. The CSC is centred at approximately (RA, Dec)= ($15^{\rm h}23^{\rm m}$, $30°$), extends over an area of $6° \times 6°$ at mean redshift $z \sim 0.07$, contains 12 galaxy clusters, and has a mass of about $\sim 12 \times 10^{16}\,{\rm M}_\odot$ [14]. In summary, the Shapley SCL stands out more prominently compared to the other two massive superclusters in the nearby universe because of its significant extent and unusual massive structures.

2. The Redshift Catalogue

Our redshift catalogue is based on deep optical galaxy catalogues in the Hydra/Antlia and Vela regions ($245° \leq l \leq 295°, |b| \leq 10°$) [15, 16], and the near infrared 2MASS Extended Source Catalogue(XSC) [17]. The major part of the redshift data is based on recent observations by two of us (Kraan-Korteweg and Cluver) using the 2dF+AAOmega spectrograph on the Anglo Australian Telescope in February 2014. A total of $4,344$ galaxy spectra were obtained in 25 fields in the Hydra/Antlia and Vela regions. The majority of these redshifts, 95.8% yield reliable galaxy redshifts, while the remaining spectra were dominated by a foreground star for which it was not possible to extract the spectrum of the targeted galaxy. These remaining 4.2% redshifts were not included in our analysis.

Other redshift data are the result of earlier unpublished and complementary work by Kraan-Korteweg and collaborators. These are from the 6dF instrument on the Australian Astronomical Observatory's UK Schmidt Telescope, the Optopus multi-object spectrograph facility on the European Southern Observatory's 3.6 m Cassegrain focus telescope, the Southern African Large Telescope (SALT), and the South African Astronomical Observatory (SAAO) using the 1.9 m telescope. We also searched the extragalactic database HYPERLEDA for redshifts in our ZoA survey region. Where crossmatches existed, we compared the literature values with the AAOmega observations. After we compiled all the above redshifts, we eliminated the doubly observed galaxies. In our subsequent analysis we only used the redshifts with higher reliability. We have a total of $4,756$ galaxies redshifts for the galaxies within the Hydra/Antlia and Vela regions ($245° \leq l \leq 295°, |b| \leq 10°$). The majority of these redshifts (88%) result from the AAOmega February Observations.

3. The Velocity Distribution and Structure of the Vela Overdensity

In Fig. 1 we present the on-sky distribution of all the galaxies with redshifts in the direction of the Vela overdensity. These galaxies are colour-coded according to different velocity ranges. Black are galaxies with velocities $v < 15000\,{\rm km\ s^{-1}}$, red are galaxies in the velocities range $15000 < v < 25000\,{\rm km\ s^{-1}}$, and cyan are galaxies with velocities $v > 25000\,{\rm km\ s^{-1}}$. The black circles mark the AAOmega observation fields from which the majority of our data originate. The Vela overdensity (red dots) is conspicuous in 20 of the 25 AAOmega fields and can be traced for $\Delta l \sim 25°$ above the plane in the range($285° > l > 260°$), and for $\Delta l \sim 20°$ below the plane in the range ($280° > l > 260°$). Moreover, the velocity histograms on both side of the plane are remarkably similar, which brings to perspective that the Vela overdensity might be connected across the plane (despite the current lack of redshift data in the mostly obscured ZoA) and consequently having a width of about $\Delta b \sim 20°$.

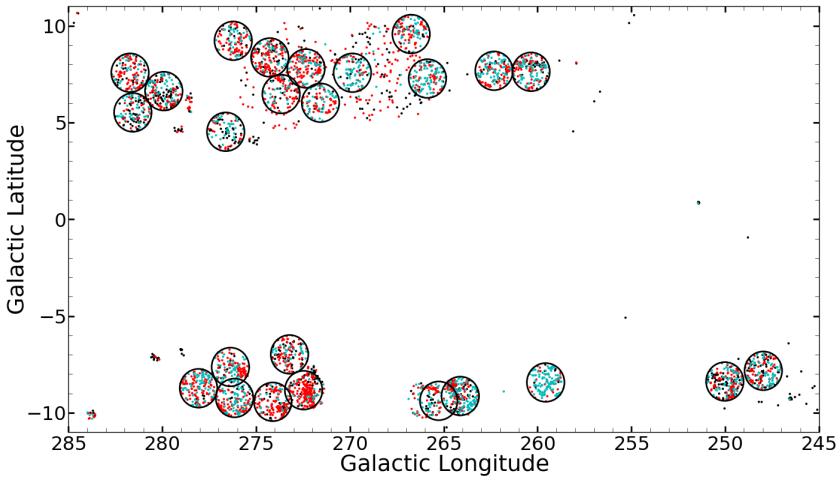


Figure 1. The on-sky distribution of all the galaxies available in the direction of the Vela overdensity, these galaxies are colour-coded according to there different velocities. Black ($v < 15000\,\mathrm{km\ s^{-1}}$), red ($15000 < v < 25000\,\mathrm{km\ s^{-1}}$), and cyan ($v > 25000\,\mathrm{km\ s^{-1}}$). The black circles delimit the AAOmega observation fields.

In Fig. 2, we show the velocity distribution of the galaxies in the direction of the Vela overdensity with velocity range from $0 < v < 50000\,\mathrm{km\ s^{-1}}$. In this histogram one can identify three peaks. The main body of the Vela overdensity is represented by the prominent peak centred at $v \sim 18000 \pm 1000\,\mathrm{km\ s^{-1}}$ and associated extended shoulders that range from $15000 - 22000\,\mathrm{km\ s^{-1}}$. This velocity distribution is very similar to that of Shapley SCL, which extends from $13000 - 18000\,\mathrm{km\ s^{-1}}$, based on a full systematic map of the Shapley SCL (refer to Fig.2 in [8]). Although, Vela is a bit further away on average (by a factor f = 1.3). This is the first indication that the Vela overdensity likely is a massive supercluster. The other two velocity peaks likely belong to further large scale structures hidden within the Zone of Avoidance.

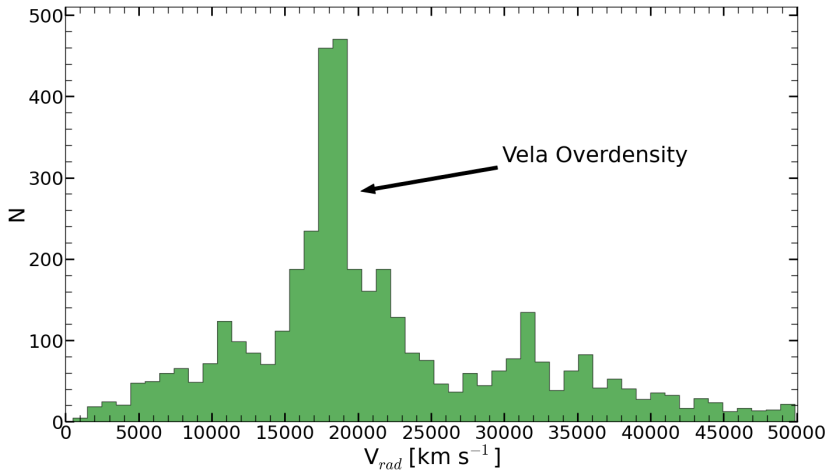


Figure 2. Velocity distribution of the galaxies in the direction of Vela overdensity for the survey area $245° \leq l \leq 295°; |b| \leq 10°$, most of these redshifts are from the AAOmega observations.

Figure 3 shows a wedge diagram out to velocities of $v = 30000\,\mathrm{km\ s^{-1}}$ in the general Vela SCL region, covering longitude range $255° \leq l \leq 285°$ and $\Delta b = 20°$, hence covering the ZoA within $|b| \leq 10°$. Based on this figure, the Vela SCL shows up as a prominent wall-like structure. It actually consists of two walls, the main wall (Wall One) at $v \sim 18000\,\mathrm{km\ s^{-1}}$ and a smaller one (Wall Two) at slightly higher velocity. Tracers of Wall One appear both above and below the Galactic plane. The second wall is present only above the plane. This is a second indicator that the Vela overdensity may form a massive supercluster. Based on Fig. 3, the Vela overdensity occupies the region from $l = 262°$ to $l = 282°$ and in velocity space from approximately $v \sim 16000\,\mathrm{km\ s^{-1}}$ extending to $v \sim 25000$ km s^{-1}. This indicates a very broad and extremely massive extended object of size of at least 20 degrees both in Galactic longitude and Galactic latitude, which corresponds to $70 \times 70\,\mathrm{Mpc}$ on the sky (assuming a value of $\mathrm{H_0} = 75\,\mathrm{km\ s^{-1}Mpc^{-1}}$).

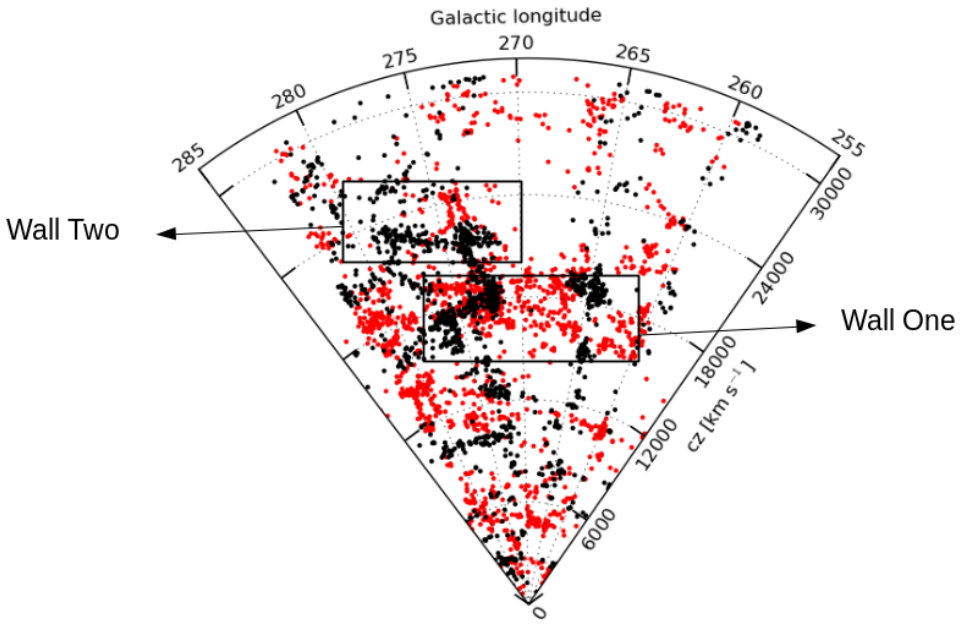


Figure 3. Velocity wedge diagram out to velocities of $v = 30000\,\mathrm{km\ s^{-1}}$ in the general Vela SCL longitude range $255° \leq l \leq 285°$ with a width of $\Delta b = 20°$, centred around the ZoA $|b| \leq 10°$. Black and red dots denote galaxies above and below the Galactic plane, respectively.

4. Clusters/Groups Results

To locate potential galaxy clusters and groups within the Vela overdensity, we used a nearest neighbour algorithm that searches for galaxies with the highest number of neighbours in velocity space. These neighbours must lie within one Abell radius at that specific redshift from the respective centre [19]. We then examined each galaxy cluster/group candidate by producing a velocity histogram to study the velocity dispersion, a velocity distribution as a function of cluster-centric distance from the optimized centre of the candidate, a K_s-band luminosity function, and a B-band luminosity function for the candidate members. From these figures, the dispersion velocities, and the masses of each candidate, we were able to approximately classify these structures. A total of 13 galaxy clusters with masses in the range $3 - 36 \times 10^{14}\,\mathrm{M_\odot}$, and 19 galaxy groups with masses in the range $0.2 - 20 \times 10^{14}\,\mathrm{M_\odot}$ were identified. This is a typical profile of a supercluster, for instance, the Shapley SCL contains 33 clusters and galaxy groups, with masses in the range of $8.3 \times 10^{14} - 3.3 \times 10^{15}\,\mathrm{M_\odot}$ [8]. Figure 4 shows the velocity distribution as a function of the distance from the centre for one of the prominent cluster candidates. It lies at $(l, b) \sim (272.2°, -8.9°)$ and has an average extinction of $E(B-V) \sim 0.2$ [15]. The red dots are

the galaxies that belong to the cluster and lie within one Abell radius from the optimized centre, and the blue dots are the galaxies in the background and foreground of the cluster. There are 108 galaxies that belong to this cluster candidate and have a velocity dispersion of $\sigma \sim 500\,\mathrm{km\ s^{-1}}$. Figure 5 displays the on-sky distribution of the galaxies in this cluster candidate. The black dot represents the optimized central galaxy, the red points shows the galaxies that belong to the cluster and lie within one Abell radius, and the blue represents the foreground and background surrounding galaxies. The many number of clusters/groups found in the sparsely sampled Vela overdensity (see Fig. 1) is a further strong indicators that this overdensity is a supercluster.

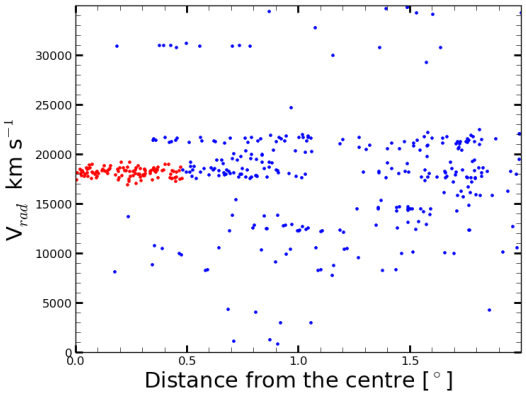


Figure 4. Velocity distribution as a function of distance from the optimized centre of one of the cluster candidates. The red dots are the galaxies that lie within one Abell radius from the optimized centre around the mean velocity of the cluster, while the blue are the galaxies in the central area of this cluster.

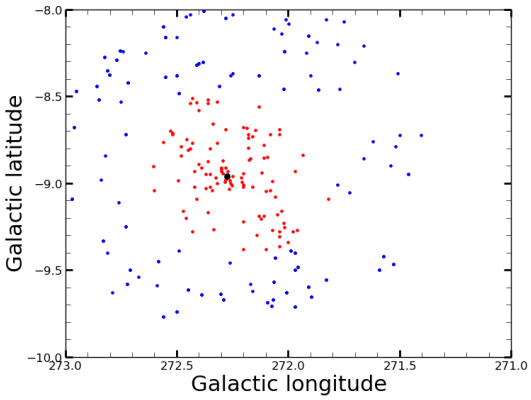


Figure 5. On-sky distribution of the galaxies in the same cluster candidate. The red dots again mark the galaxies that lie within one Abell radius from the optimized centre (the black dot).

5. Conclusions

Based on the prominent peak in the velocity distribution of the galaxies in the Vela overdensity above/below the Galactic plane at $v \sim 18,000\,\mathrm{km\ s^{-1}}$, the main broad wall that can be traced on either side of the plane (Fig.3), and the many massive galaxy clusters/groups found in the Vela overdensity, the evidence for the existence of a massive supercluster in the Hydra/Antlia and Vela region is strong. We call it Vela SCL. The extent of Vela SCL, as derived from the distribution of these galaxies, is $70 \times 70\,\mathrm{Mpc}$, which is comparable at $v \sim 18,000\,\mathrm{km\ s^{-1}}$ to the extent of the Shapley SCL $30 \times 75\,\mathrm{Mpc}$ [8]. It is 1.3 times further away than Shapley SCL. Given the fact that the Shapley SCL contributes about 9% of the Local Group motion [11] (if not more; see [18]), the proximity of a further supercluster (Vela) in its vicinity, might have implications on bulk flows studies. Consequently, the discovery of this supercluster might play a role in resolving the long-enduring bulk flow controversies and the misalignment of flows with the Cosmic Microwave Background (CMB) measurement.

Acknowledgments

Acknowledgements We gratefully acknowledge the support by the South African National Research Foundation. AE is very grateful to Dr.Christina Magoulas for her reading and

comments on the paper. This publication makes use of data from the Two Micron All Sky Survey (2MASS). The 2MASS is a joint project between the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

References

[1] Batiste M and Batuski D J 2013 *MNRAS* **436** 3331
[2] Einasto J, Joeveer M and Saar E 1980 *MNRAS* **193** 353
[3] Araya-Melo P A, Reisenegger A, Meza A, Van de weygaert R, Dünner R and Quintana H 2009 *MNRAS* **399** 97
[4] Sheth R K and Diaferio A 2011 *MNRAS* **417** 2938
[5] Cen R 1994 *APJ* **424** 22
[6] Quintana H, Ramirez A, Melnick J, Raychaudhury S, and Slezak E 1995 *APJ* **110** 463
[7] Shapley H 1930 *HarO.* **874** 9
[8] Proust D *et al* 2006 *A&A* **447** 133
[9] Mitsuishi I, Gupta A, Yamasaki N Y, Takei Y, Ohashi T, Sato K, Galeazzi M, Henry J P and Kelley R L 2012 *AIPC* **1427** 344
[10] De Filippis E, Schindler S and Erben T 2005 *A&A* **444** 387
[11] Muñoz J A and Loeb A 2008 *MNRAS* **391** 134
[12] Lucey J R, Dickens R J, Mitchell R J and Dawe J A 1983 *MNRAS* **203** 545
[13] Fleenor M C, Rose J A, Christiansen W A, Johnston-Hollitt M, Hunstead R W, Drinkwater M J and Saunders W 2006 *APJ* **131** 1280
[14] Pearson D W, Batiste M and Batuski D J 2014 *MNRAS* **441** 1601
[15] Kraan-Korteweg R C 2000 *A&AS* **141** 123
[16] Von Maltitz K 2012 Master's thesis Cape Town University
[17] Jarrett T H *et al* 2000 *AJ* **119** 2498
[18] Kocevski D D and Ebeling H 2006 *ApJ* **645** 1043
[19] Abell G O 1958 *ApJS* **3** 211

A 'road test' of ANOVA versus DFT and LS as a period-finding algorithm

C A Engelbrecht[1] and F A M Frescura[2]
[1]Department of Physics, University of Johannesburg, Kingsway, Auckland Park, Johannesburg 2092, South Africa
[2]School 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 brought Fourier-based algorithms into popular use as period-finding tools in astronomy, specifically in asteroseismology. Recently published work has indicated that the ANOVA approach could outperform these traditional approaches in certain cases. Very little practical application of ANOVA has appeared in the literature, though. Given the rapidly growing body of time-domain data in astronomy and the considerable importance of some of the conclusions that have been reached on the basis of these data, the recently published results prompt a closer look at ANOVA as a competitor to DFT and Lomb-Scargle (LS)-based methods in asteroseismology. In this paper, we present and analyse the main findings of a comparative 'road-test' of ANOVA, DFT and Lomb-Scargle algorithms, applied to a variety of observing scenarios, including typical ground-based, time-domain data of pulsating stars.

1. Introduction

Aided comprehensively by the data obtained from various space-based missions, the techniques of asteroseismology have profoundly affected our understanding of stellar structure and evolution. The accurate determination of pulsation frequencies, amplitudes and phases is essential for the successful application of asteroseismology to this end. The most commonly used techniques for period determination in variable stars are the Discrete Fourier Transform (DFT) adapted for non-equally-spaced time series [1], and the Lomb-Scargle (LS) periodogram [2]. Both of these techniques rely on certain mathematical properties of the harmonic functions for their efficiency as period-finding tools. There are various well-developed alternatives to Fourier-based methods. These have been summarised in a recent review [3]. The most prominent alternative period-finding strategy identifies the most probable periodicity occurring in a time series by minimising the dispersion in phase bins associated with the respective periods inspected (instead of calculating transforms of time series or fitting functions to them, as the Fourier-based methods do). These are called 'dispersion-based methods' or, occasionally, 'entropy-based methods'. The original dispersion/entropy-based method is the 'String-length' method introduced by Lafler & Kinman [4] for the purpose of dealing with light curves which deviate substantially from sinusoidal shapes. Schwarzenberg-Czerny [5] pioneered a thorough series of investigations into the statistical techniques embodied in the analysis of variance and their potential for the accurate detection of periodic variations in time series. We shall refer to Schwarzenberg-Czerny's formulation of these techniques as 'ANOVA' in what follows.

A recent study by Graham et al. [6] found that the ANOVA method, as described in [5], was exceptionally successful at retrieving periods in time series generated by large astronomical surveys, specifically: the MACHO, CRTS and ASAS surveys. Since the authors of the present paper probe the detection of low-amplitude pulsations in main-sequence stars, an exercise where it is essential to optimise the period-finding techniques employed, the findings reported in [6] have prompted us to investigate the efficacy of ANOVA in our own time series. We decided to measure this efficacy by exposing various sets of artificial data – with properties that are similar to the various datasets that we have at our disposal for our studies of pulsating stars – to three period-finding methods: the ANOVA method, the standard DFT algorithm described in [1] and the LS described in [2]. The following sections describe our simulated time series and the results of applying the three period-finding methods to these time series.

2. Simulated light curves

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 are typical of actual observing scenarios that produce useful asteroseismological results for pulsating early B stars. A previous study [7] by the present authors compared the efficacy of the Generalised Lomb-Scargle (GLS) method for period-finding, with DFT and LS methods. For ease of comparison, we have retained (except for computer-generated random errors) the simulated time series used in [7].

The following three realistic observing scenarios were chosen:

i) Scenario 1: A strictly equally-spaced time sampling over the 110 days, typical of what is achieved with a space-based telescope making uninterrupted observations 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 at least as well as the LS and ANOVA periodograms in this scenario (if not better).

ii) Scenario 2: A typical ground-based observing campaign at a single site, consisting of five full weeks spread over the 110-day time-span and consisting of varying numbers of photometric nights per week. This time series typifies the result of small-telescope observing allocations of the South African Astronomical Observatory at its Sutherland site in the Northern Cape. In this simulation, a total of 17 photometric nights were simulated among the 35 days of allocated observing time. This ~50% duty cycle is typical of actual conditions at Sutherland. The actual sampling times were adapted to accommodate the advance of sidereal time over the 110 days of calendar time, which is accompanied by an earlier rising time of target stars each night. 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 early 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. The method used to construct *signal strengths* is described after the figure.

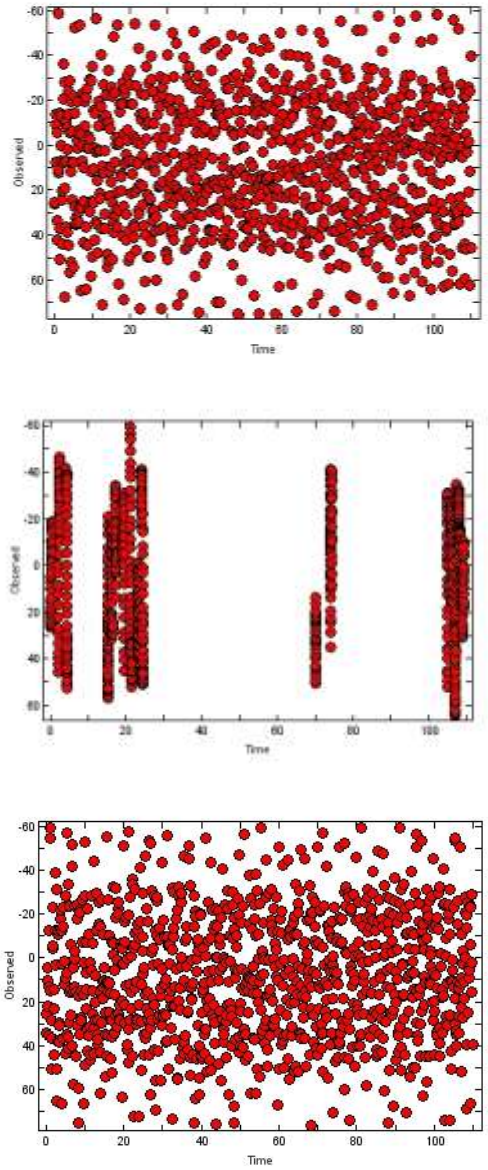


Figure 1. A sample of light curves generated for Scenarios 1, 2 and 3 (from top to bottom) respectively. Since the signal has a random noise component, the light curves will appear slightly differently every time they are calculated.

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, we added three sinusoids with respective amplitudes, frequencies and phase differences typical of low-amplitude p-modes in pulsating B stars. 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 correspond with the typical mean errors and error variances found in small-telescope photometry at Sutherland. The calculated light curves were then fed into the mathematical algorithms for calculating DFT, LS and ANOVA periodograms and the results were compared. These results are discussed in the following section.

3. Retrieving periods

The ability of the DFT, LS and ANOVA periodograms to retrieve the input frequencies and amplitudes 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 our results is presented in this section.

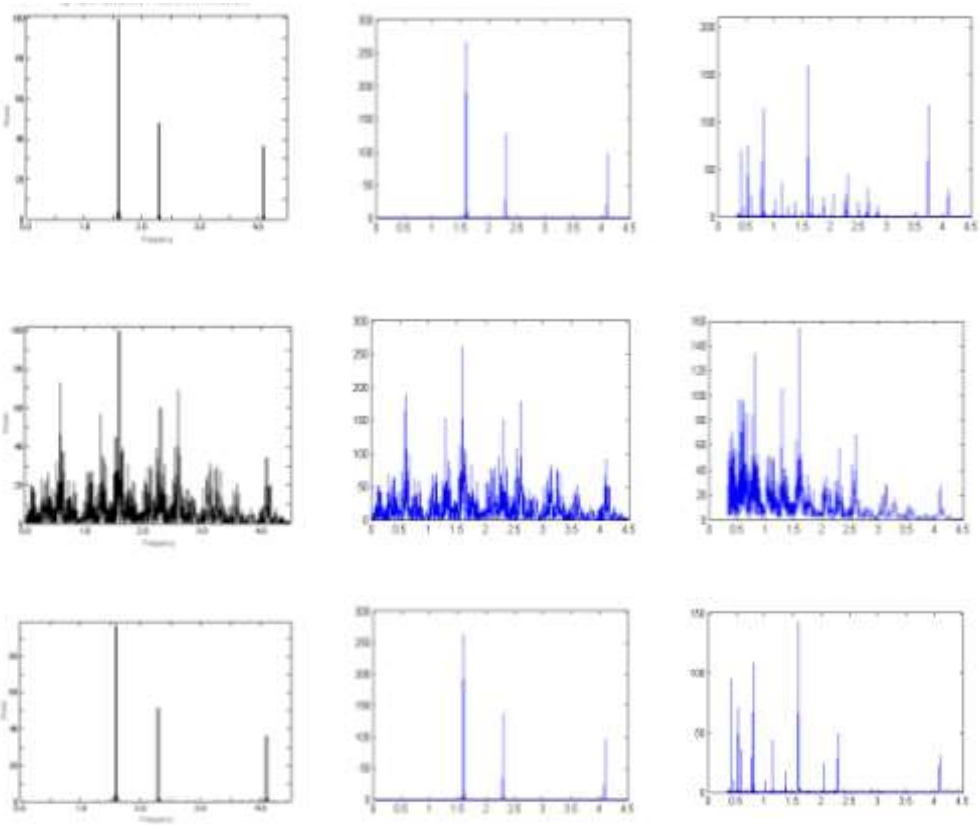


Figure 2. Periodograms obtained with DFT, LS and ANOVA algorithms (from left to right in each row), for result A. First row: scenario 1 (equal time spacing); second row: scenario 2 (ground-based observer); third row: scenario 3 (ground-based survey).

Result A: Firstly, we considered SNR values which are typical of ground-based asteroseismology of pulsating B stars, with a homogeneous error distribution across the entire observing epoch of 110 days. Initial periodograms (i.e. before any prewhitening) for Result A are shown in figure 2. For equal time spacing (scenario 1: the first row of periodograms in the figure), the DFT and LS periodograms deliver quite clear-cut results: they retrieve the three input frequencies (and their associated amplitudes) with ease and without any ambiguity. In comparison, the periodogram produced with the ANOVA algorithm appears beset with a multitude of spurious peaks. However, closer inspection

reveals that these are the sub-harmonics of the three primary frequencies from which the signal is constructed. Measures can be taken to account for the sub-harmonics when searching for periods with the ANOVA algorithm. Still, the DFT and LS methods appear simpler and less prone to mis-interpretation in the equally-spaced scenario.

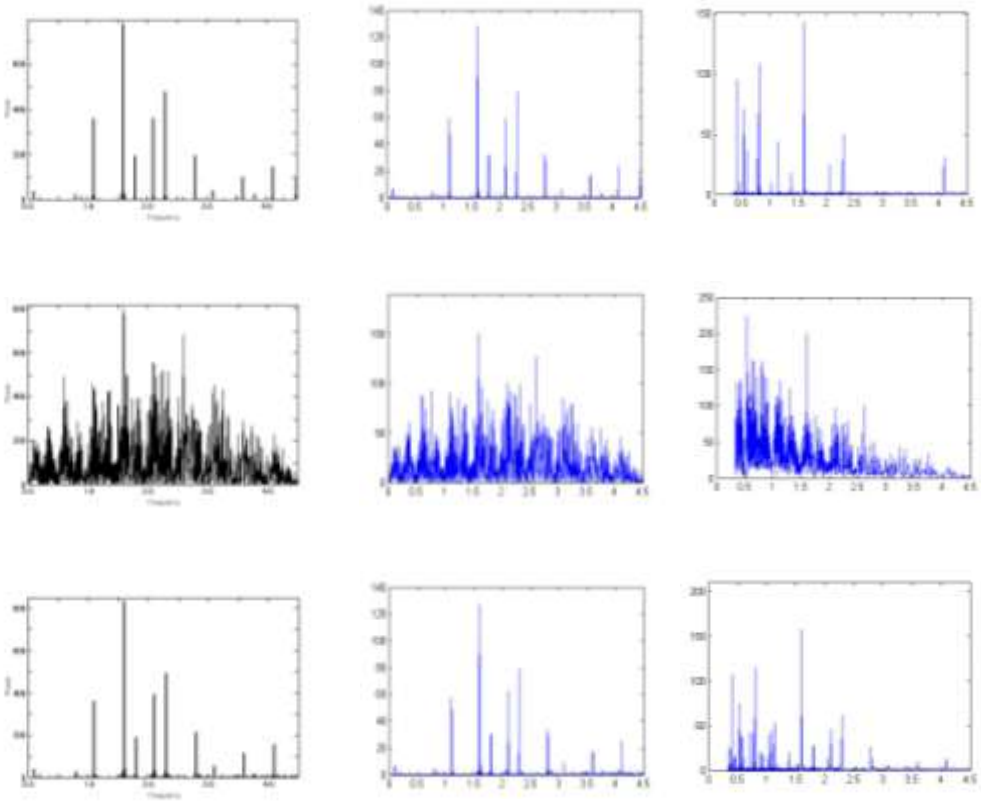


Figure 3. Periodograms obtained with DFT, LS and ANOVA algorithms respectively, for Result B. The arrows in the third column of graphs indicate the positions of the periods corresponding with the input signals, as for figure 2. The three rows correspond to scenarios 1, 2 and 3 respectively, as for figure 2.

The simulation of single-site observer-controlled observing produces the much more noisy periodograms in row 2 of figure 2. The one-day aliasing peaks are clearly seen in the DFT and LS periodograms, and a prewhitening procedure (not shown, but performed by the authors) succeeds in retrieving all three input periods without any ambiguity, removing practically all the noise in the periodogram as well. It is far more problematic extracting sub-harmonics from the ANOVA plots in this scenario. It does not appear useful to employ ANOVA in this context. Turning to row 3 of figure 2, which shows the periodograms for the scenario 3 (ground-based survey), it is surprising that the ANOVA periodogram fares better for this non-equally-spaced scenario, compared to scenario 1. This seems to be the context for which ANOVA is best suited as an alternative to DFT and LS methods.

Result B: The vulnerability of the respective period-finding methods to heavy aliasing effects was tested by removing every second day's data from the equally-spaced time series (scenario 1) and similarly thinning out the time series for the other two scenarios by 50% in a quasi-regular way. Initial periodograms (i.e. before prewhitening) for Result B are shown in figure 3. When compared with figure 2, it is clear that all nine periodograms are detrimentally affected by the 'aliasing noise' introduced by the deliberate, 'gap-generating', thinning procedure used. The primary input signal is still unambiguously identified in eight of the nine periodograms. However, for scenario 2 (middle

row), the ANOVA periodogram fails to identify even the strongest input signal as the primary peak. ANOVA appears at least as vulnerable to aliasing effects as the Fourier-based methods are, if not more so.

Result C: Vulnerability to variable background noise levels was tested by dividing the time series 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 (before prewhitening) for Result C are shown in figure 4 (for scenario 2 only). Comparison with figure 2 indicates that none of the three period-finding algorithms shows any significant vulnerability to *variable* background noise levels across an observing season. However, note that the *signal-to-noise ratios* of the dominant peaks in the plots in figure 4 are indeed lower than in their counterparts in figure 2.

4. Discussion and conclusions

The three experiments described as Result A, Result B and Result C in the previous section lend a measure of support to the period-finding performance attributed to ANOVA in [6]. It would appear prudent to apply this period-finding algorithm in parallel to the DFT and LS algorithms, particularly in situations where aliasing effects might be compromising the latter methods. Also, ANOVA might be of particular value when analysing sparsely time series (corresponding to scenario 3 in this paper). However, the ANOVA method does not appear to be a viable alternative for finding periods in time series that are typical of the single-site, observer-controlled example (scenario 2) discussed in this work.

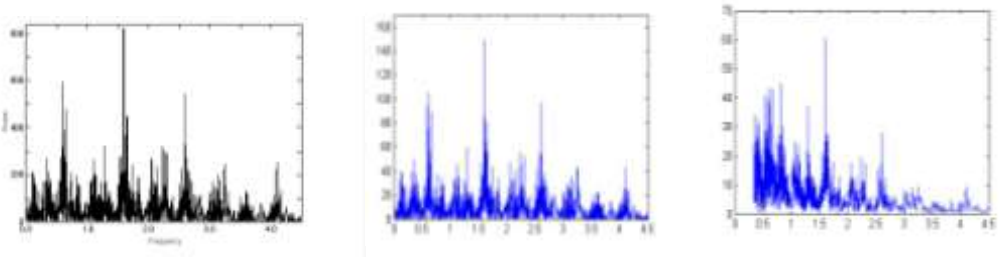


Figure 4. Periodograms obtained with DFT, LS and ANOVA algorithms respectively, for Result C. Only scenario 2 is shown, as the other scenarios add nothing new.

References

[1] Deeming T J 1975 *Ap&SS* **36** 137
[2] Scargle J D 1982 *ApJ* **263** 835
[3] Engelbrecht C A 2014 *Precision Asteroseismology: Proc. IAU Symposium 301 (Wroclaw, Poland, 19-23 August 2013)* eds J A Guzik, W J Chaplin et al (Cambridge: CUP) pp 77-84
[4] Lafler J and Kinman T D 1965 *ApJS* **11** 216
[5] Schwarzenberg-Czerny A 1989 *MNRAS* **241** 153
[6] Graham M J, Drake A J, Djorgovski S G, Mahabal A A, Donalek C, Duan V, Maker A 2013 *MNRAS* **434**, 3423
[7] Engelbrecht C A and Frescura F A M 2015 *Proc. 59th Annual Conference of the SA Institute of Physics (Johannesburg, South Africa, 7-11 July 2014)* eds C A Engelbrecht and S Karataglidis, pp 324-329

Optical spectroscopic observations of unclassified Active Galactic Nuclei in the Fermi-2LAC catalogue

L Klindt[1], B van Soelen[1], PJ Meintjes[1] and P Väisänen[2]

[1]Department of Physics, University of the Free State, Bloemfontein, 9301, South Africa,
[2]South African Astronomical Observatory, PO Box 9, Observatory 7935, Cape Town, South Africa

E-mail: lizelkeklindt@gmail.com

Abstract. Blazars constitute the most violent astronomical objects with relativistic jets emitting radiation at all frequencies. We report on optical spectroscopic observations of unclassified Active Galactic Nuclei (AGN) in the 2nd *Fermi*-LAT catalogue of AGN which exhibit blazar-like characteristics. The total target sample comprises of 19 AGN of unknown type (AGU) which are all at high galactic latitudes ($|b| > 10^{\circ}$) with optical and/or radio counterparts within the *Fermi*-LAT 95% error circle. We are commencing with a multi-wavelength campaign to classify the candidate sources by establishing Spectral Energy Distributions (SEDs) and to search for potential Very High Energy (VHE) candidates. Preliminary low-resolution spectra, obtained with the SAAO 1.9-m telescope and the Southern African Large Telescope (SALT), are mainly featureless as is expected for BL Lac objects, however Ca II K&H, MgIb and/or NaD absorption lines appear to be present, allowing for a first estimation of the redshifts which correlates favourably with that of known blazars. We report on the details of two of these sources. The shallow Ca II depression indicates that strong non-thermal emission is present in these targets and they are therefore good potential blazar sources.

1. Introduction

Active Galactic Nuclei (AGN) are galaxies that are believed to be powered by accretion onto a supermassive black hole ($M_{\rm BH} = 10^6 - 10^{10} M_{\odot}$) located at its centre. AGN can produce large relativistic jets, which emit multi-wavelength non-thermal radiation. When the relativistic jet propagates close to our line of sight the AGN is classified as a blazar, displaying rapid multi-wavelength variability, strong polarization, non-thermal emission and high apparent luminosities over the entire electromagnetic spectrum from the Doppler-boosted jet (see e.g. [1, 2, 3]). Depending on the strength of the optical emission features two distinct groups exist, namely BL Lacartae objects (BL Lacs) and Flat Spectrum Radio Quasars (FSRQs). BL Lacs are well known for their weak or absent spectral lines which is a result of the non-thermal emission in the jet that outshines the host galaxy emission. BL Lacs are optically defined as sources with emission and/or absorption lines with equivalent widths of $W_{\lambda} \leq 5$ Å, and a diluted Ca II break value less than 40% (see e.g. [4], [5], [6]). FSRQs on the other hand display both strong narrow and broad emission lines [7].

Spectral Energy Distributions (SEDs) of AGN 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 peak is due to synchrotron emission of relativistic electrons, whereas the processes producing the high-energy peak are still under debate, with models suggesting either leptonic or hadronic processes (see e.g. [8]). The leptonic model suggests that the high energy component is produced by inverse Compton (IC) scattering off either external photons from the accretion disc, dust torus or Broad Line region (External Compton - EC), or off photons produced by the synchrotron emission (Synchrotron Self Compton - SSC). In the hadronic model the emission of the high energy component is dominated by π^0 decay photons, pair cascades or synchrotron and Compton emission from secondary charged pions [9, 10, 11, 12].

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 [13], with peak frequencies of $\nu_{\rm peak} < 10^{14}$ Hz, $10^{14} \leq \nu_{\rm peak} < 10^{15}$ Hz, and $\nu_{\rm peak} > 10^{15}$ Hz, respectively. Blazars exhibit diverse properties which have been studied for years, however, many questions still remain about the physical processes involved and the demographics of these sources which therefore motivates the importance of classifying and modelling these sources over multi-wavelengths.

Based on the above mentioned observational characteristics of blazars, we have identified an original target sample of 19 Active Galactic Nuclei of Unknown type (AGU) in the 2 year *Fermi*-LAT AGN Catalogue (*Fermi*-2LAC), and place particular focus on 10 targets, with the aim to classify these systems through multi-wavelength observations, and to contribute to our search for Very High Energy (VHE) sources. We have based the sample selection on criteria similar to those previously employed by [14] in a study of unidentified sources listed in the Energetic Gamma Ray Experiment Telescope (EGRET) catalogue. Onboard the *Fermi* Gamma-ray Space Telescope spacecraft is the Large Area Telescope (LAT) which has observed every region of the sky for 30 minutes every 3 hours since August 2008. The energy range covered by LAT is 20 MeV to 300 GeV [15]. In the *Fermi*-LAT 2-year Source catalogue (866 sources in the clean sample) 81% of the sources are classified as blazars and 18% as AGU which exhibit blazar characteristics [16]. We have focused on classifying a subset of these AGUs through multi-wavelength analysis which includes optical spectroscopic and photometric observations and single dish radio observations. Here we report on the optical spectroscopic results obtained for two targets namely 2FGL J1154.1-3242 and 2FGL J0730.6-6607.

2. Sample selection

The original sample we chose from the 157 AGU sources listed in the *Fermi*-2LAC included 19 targets which are observable from Southern Africa, and were selected for multi-wavelength follow-up observations (radio to gamma-rays) in order to determine their unknown properties. In the 2LAC catalogue the most likely radio and/or X-ray counterparts to the Fermi sources were established using three different statistical tests which also considered the source density in the region surrounding the LAT source [16]. We have proposed to observe the established optical/IR counterpart to the best candidate radio sources. With a few exceptions the following criteria apply to our sample (see full discussion [17]):

- Observability from Southern latitudes: to enable the utilization of the optical telescopes we used a declination limit of $-90^\circ \leq \delta \leq +20^\circ$. Two sources with declinations exceeding this limit were still included in our sample due to their high radio brightness and were consequently candidate sources for radio observations with the HartRAO 26-m telescope which can observe up to $\delta < 45^\circ$.
- High Galactic latitude sources $|b| > 10^\circ$.
- Gamma-ray photon spectral index: $1.2 < \Gamma < 3$ where $dN/dE = N_0(E/E_0)^{-\Gamma}$ (see e.g. figure 12 in [16]).

Table 1. Candidate AGUs in the *Fermi*-2LAC catalogue with blazar characteristics. The radio counterparts within the 95% error circle are listed with their corresponding V magnitude, target position, gamma-ray spectral index (Γ), radio flux density at 4.85 GHz ($S_{4.85}$) and variability index (VI). The preliminary estimated redshifts (z), obtained from the observations reported here, are also listed.

2LAC Name	Vmag	RA_{J2000}	Dec_{J2000}	Γ	$S_{4.85}$	VI	z
J0044.7-3702	19.60	00:45:12	-37:05:49	2.57	330	92.7	
J0201.5-6626	20.56	02:01:08	-66:38:13	2.25	168	39.8	
J0644.2-6713	20.69	06:44:28	-67:12:57	2.16	218	99.6	
J0730.6-6607	15.13	07:30:50	-66:02:19	1.34	82	26.8	0.106 ± 0.001
J0855.1-0712	19.78	08:55:09	-07:15:03	2.62	1157	31.2	
J1106.3-3643	17.41	11:06:24	-36:46:59	2.20	92	24.1	
J1154.1-3242	18.88	11:54:06	-32:42:43	2.03	212	20.5	0.224 ± 0.019
J1218.8-4827	17.53	12:19:02	-48:26:28	2.40	65	26.6	
J1407.5-4257	17.47	14:07:39	-43:02:34	1.91	149	22.9	
J1955.0-5639	17.25	19:55:03	-56:40:29	1.88	9	23.7	

- Radio bright sources: $S_{4.85\ GHz} \geq 100$ mJy.
- $V_{mag} \leq 21$; fainter targets are excluded from the sample due to observational limitations.
- No determined redshift (z).
- We also considered the *Fermi*-LAT variability index (VI) to determine the probability for a source to be variable, though this was not used as a strict selection criteria; sources with a variability index of $VI > 41.6$ have a 99% chance to be variable over the two year observation period. [14, 16, 18].

Based on the above mentioned criteria we are focusing on observations of 10 targets from our original sample of 19 (see table 1). These selected sources are all still identified as AGU in the very recently released 3FGL catalogue [19]. The updated positions correspond to those previously reported in the 2LAC catalogue. (In the 3LAC catalogue the number of sources has increased to 402 Blazar Candidates of Unknown type (BCU) in the Clean Sample [20]).

3. Optical observations and data analysis

The optical observations and data analysis undertaken with the SAAO 1.9-m telescope and the Southern African Large Telescope (SALT) are summarized below.

3.1. SAAO 1.9-m telescope

Optical spectra were obtained for 8 targets of the original sample (19 AGUs) during May/June 2014 with the SAAO 1.9-m telescope. We report specifically on the results obtained for one of the targets, 2FGL J1154.1-3242, on 22 and 25 May 2014. In order to search for potential lines we have undertaken broadband spectroscopic observations utilizing the SpCCD spectrograph with grating 7 in order to achieve a wavelength range of 4200 Å (spectral resolution of ~ 5 Å). The broad wavelength range allows for a wavelength coverage from approximately 3200 Å to 8000 Å. In order to achieve a signal-to-noise ratio (SNR) ~ 10, five exposures of 3600 s each were acquired and averaged. The standard spectroscopic reduction steps in IRAF were used to reduce the two-dimensional data using copper-argon comparison spectra taken after each target spectrum.

3.2. SALT

Five sources have been observed with SALT during Semester I 2015 (December 2014 – March 2015), using the Robert Stobie Spectrograph (RSS) with the pg0300 grating and with two different Camera Stations, namely 10.0° and 11.5°. We report on observations obtained for 2FGL J0730.6-6607 on 31 December 2015. The chosen configurations allowed us to access the 3200 Å − 9500 Å visible wavelength range with a resolving power of $R = 530$. We used 2 spatial dither positions (two separate spectra at different locations of the slit, shifted in the vertical axis) with an offset of 10 arcsec in order to correct for fringing effects at the longer wavelengths. In total four exposures of 120 s and 126 s each were obtained for the two configurations, respectively. The data were reduced using IRAF packages with the standard steps and by combining the one-dimensional spectra.

4. Results and Discussion

Here we only report on the results obtained for two targets, however spectroscopic data have thus far been obtained for 9 of the targets with the SAAO 1.9-m telescope and SALT, for which reductions are currently on-going. A report on the spectra of 4 of the targets is shown in [17]. A comparison between the estimated redshifts and the gamma-ray photon spectral indices of the sources, places them within the blazar range as shown in figure 19 of [16].

4.1. SAAO 1.9-m telescope observations of 2FGL J1154.1-3242

Ca II H&K and G-band absorption lines appear to be present at 4860, 4894 and 5174 Å, respectively, in the target spectrum of 2FGL J1154.1-3242 (V = 18.88 mag.) as shown in figure 1. This gives a redshift measurement of $z = 0.224 \pm 0.019$. The shallow Ca II break indicates that strong non-thermal emission is present in this target and it is therefore a potential blazar source. Similar results have been reported for a FSRQ (PKS J0820-5705) by [14].

4.2. SALT observations of 2FGL J0730.6-6607

The RSS spectrum obtained for 2FGL J0730.6-6607 is shown in figure 2. These results show absorption features, namely Ca II H&K, MgIb and NaD, which led to a redshift measurement of $z = 0.106 \pm 0.001$. We have obtained a Ca II break value of 16% and, therefore, it can be concluded that strong non-thermal emission dominates the host galaxy emission. We have obtained similar results from data undertaken with the SAAO 1.9-m telescope which is reported in [17].

5. Multi-wavelength follow-up studies

In order to successfully classify the identified sources in our sample, multi-wavelength data are required. In the optical regime we have obtained spectra for 5 targets with the SAAO 1.9-m telescope and SALT (reductions currently on-going). Six targets have been accepted for RSS observations with SALT during 2015 Semester II. Photometric observations with the Sutherland High-speed Optical Camera (SHOC) on the SAAO 1.9-m telescope were also undertaken during December 2014. These complementary observations focused on searches for rapid photometric variability which is related to shocks that propagate down the relativistic jet and interact with the surrounding medium [21].

Further multi-wavelength observations are planned, in particular radio observations with the HartRAO 26-m telescope, as well as the search of archival data in, for example, ASI Science Data Centre (ASDC)[1]. An Example of the SED of 2FGL J0730.6-6607 obtained in ASDC is shown in figure 3. The SED of this target clearly indicates two peaks at frequencies that potentially

[1] http://www.asdc.asi.it/fermi2lac/

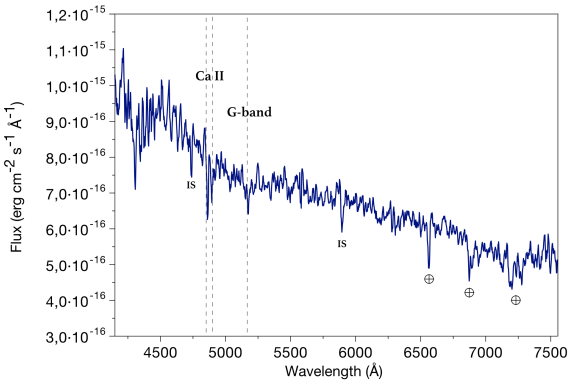


Figure 1. Optical spectrum of 2FGL J1154.1-3242 (V∼18.88 mag.) obtained with the SAAO 1.9-m telescope during May 2014. A redshift of $z = 0.224 \pm 0.019$ was obtained from the potential Ca II H&K and G-band absorption lines. Telluric lines are indicated with crossed circles.

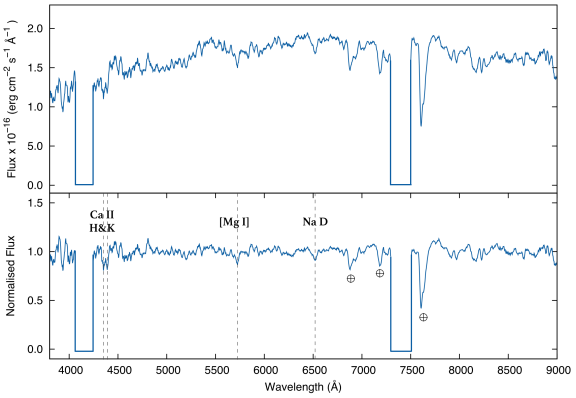


Figure 2. Flux calibrated RSS spectrum of 2FGL J0730.6-6607 (V∼15.13 mag.) obtained with SALT on 31 December 2015. A redshift of $z = 0.106 \pm 0.001$ was obtained from the Ca II H&K, MgIb and NaD lines. Top panel: flux calibrated spectrum. Bottom panel: normalized spectrum.

resemble that of HSP sources, making 2FGL J0730.6-6607 a candidate target for VHE follow-up studies.

6. Conclusion

The optical spectra obtained with the SAAO 1.9-m telescope and SALT during 2014 show promising results with the absorption features of the host galaxy present, namely Ca II H&K, MgIb and NaD lines. This allowed for estimated redshift values of $z = 0.224 \pm 0.019$ and $z = 0.106 \pm 0.001$ for 2FGL J1154.1-3242 and 2FGL J0730.6-6607, respectively. The shallow Ca II break ($< 40\%$) indicates that the targets are potentially blazars which exhibit multi-frequency non-thermal emission. In BL Lacs specifically the non-thermal emission from the jet decreases the Ca II break since the break is related to the viewing angle of the system (e.g. [5]). The double-humped shape of the SEDs with a low-energy synchrytron peak and high-energy peak resemble that of AGNs as expected, which is our first step towards classifying the identified sources in the *Fermi*-2LAC with the aim to search for VHE candidate sources. A few targets are potential HSPs and are, therefore, worth following up with TeV observations.

Acknowledgments

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. This work is based on the research supported in part by the National Research Foundation of South Africa for the grant 87919. 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. This paper uses observations made with the Southern African Large Telescope (SALT), and at the South African Astronomical Observatory.

References

[1] Ghisellini G, Padovani P, Celotti A and Maraschi L 1993 *The Astrophysical Journal* **407** 65
[2] Ghisellini G 2013 Lecture Notes in Physics Berlin Springer Verlag 873
[3] Falomo R, Pian E and Treves A 2014 arXiv:1407.7615v1

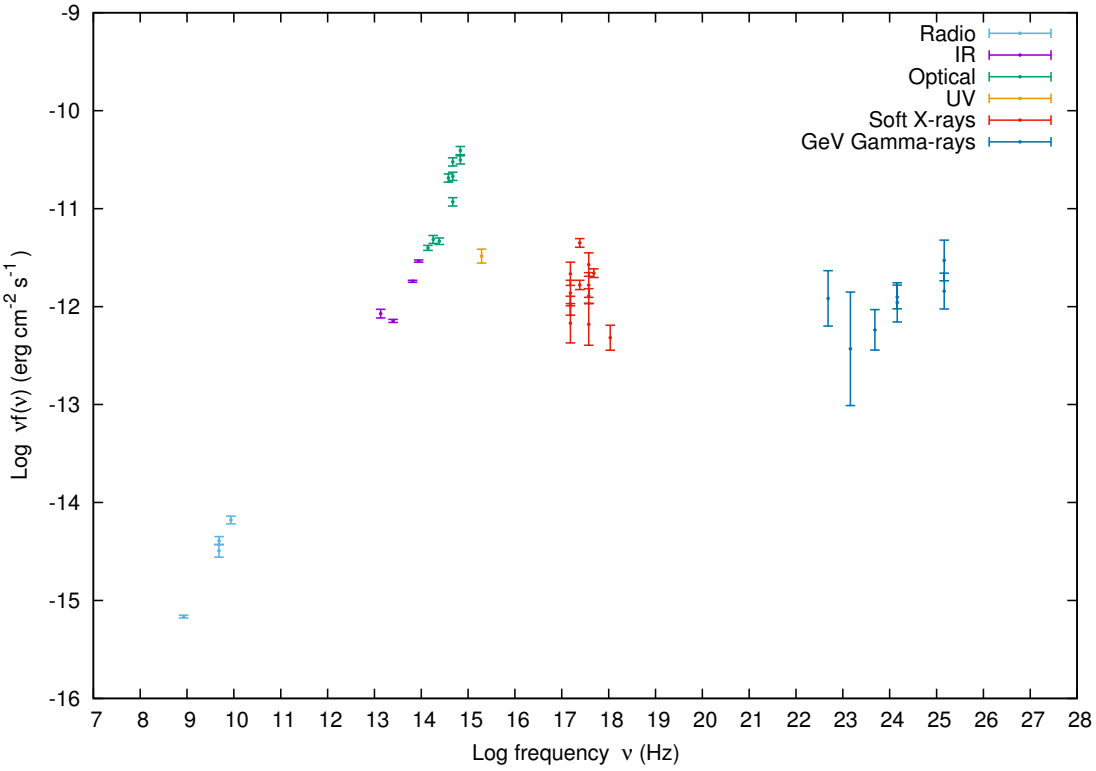


Figure 3. A Spectral Energy Distribution (SED) of 2FGL J0730.6-6607 obtained from the available multi-wavelength archived data in ASDC. The SED clearly shows the double-humped shape which is characteristic of blazars. The high-energy component shows a hard spectral index, which makes this target a VHE candidate source.

[4] Marchã MJM, Browne IWA, Impey CD and Smith PS 1996 *Monthly Notices of the Royal Astronomical Society* **281** 425
[5] Landt H, Padovani P and Giommi P 2002 *Monthly Notices of the Royal Astronomical Society* **336** 945
[6] Galbiati E *et al* 2005 *Astronomy & Astrophysics* **430** 927
[7] Urry CM and Padovani P 1995 *Publication of the Astronomical Society of the Pacific* **107** 803
[8] Böttcher M, Reimer A, Sweeney K and Prakash A 2013 *The Astrophysical Journal* **768** 54
[9] Mannheim K and Biermann PL 1992 *Astronomy & Astrophysics* **253** L21
[10] Aharonian FA 2000 *New Astronomy* **5** 377
[11] Mücke A and Protheroe RJ 2001 *Astroparticle Physics* **15** 121
[12] Mücke A, Protheroe RJ, Engel R, Rachen JP and Stanev T 2003 *Astroparticle Physics* **18** 593
[13] Padovani P and Giommi P 1996 *Monthly Notices of the Royal Astronomical Society* **279** 526
[14] Nkundabakura P and Meintjes P J 2012 *Monthly Notices of the Royal Astronomical Society* **427** 859
[15] Atwood WB *et al* 2009 *The Astrophysical Journal* **697** 071
[16] Ackermann M *et al* 2011 *The Astrophysical Journal*, **171** 37
[17] Klindt L, Meintjes PJ and van Soelen B 2015 *Multi-wavelength classification of unidentified AGN in the Fermi 2LAC catalogue: Proceedings of SAIP2014 (Johannesburg, South Africa, 7–11 July 2014)* ed C Engelbrecht and S Karataglidis (University of Johannesburg) pp 341-346
[18] Nolan PL *et al* 2011 *The Astrophysical Journal* **199** 31
[19] Acero F *et al* 2015 *The Astrophysical Journal* arXiv:1501.02003v1
[20] Ackermann M *et al* 2015 *The Astrophysical Journal* arXiv:1501.06054
[21] Marscher AP *et al* 1991 *The Astrophysical Journal* **371** 491

A Timing Noise Analysis Pipeline for HartRAO pulsars applied to PSR J1326-5859

Jacques Maritz[1*], Pieter Meintjes[1], Sarah Buchner[2], Natalia Lewandowska[3]

[1]Physics Department, University of the Free State, 9300, South Africa
[2]SKA, The Park, Park Road, Pinelands, South Africa
[3]Hartebeesthoek Radio Astronomy Observatory, P.O.Box 443, Krugersdorp 1740, South Africa

E-mail: *maritzjm@ufs.ac.za

Abstract. Timing noise in long-term pulsar timing residuals is a challenge to our understanding of present pulsar models. The quest to understand on the emission mechanisms of radio pulsars has led to various key science projects ranging from the search for gravitational waves to the development of next-generation instruments for pulsar astronomy. Timing noise studies present an essential cornerstone in these projects. One possible explanation for the existence of timing noise is mode switching in the magnetosphere of the pulsar. In a long-term study we have investigated the timing noise phenomena seen in PSR J1326-5859. It was observed with the 26 m radio telescope of the Hartebeesthoek Radio Astronomical observatory (HartRAO) for several decades and exhibits a large variety of timing noise characteristics. In this paper we review the preliminary timing noise analysis pipeline used for PSR J1326-5859 which can be also used for other southern hemisphere pulsars observed with HartRAO.

1. Introduction

Pulsars can be considered as stable cosmic clocks that serve the purpose of testing fundamental theories and advancing computing technology [1]. Integrated pulses of pulsars can be used for timing purposes. However, careful inspection of the single pulses reveal a rather unstable picture regarding their emission modes [2]. Inspecting the integrated pulses and the timing residuals over a long time span (months to years) allow us to see possible signatures of mode switching [3], nulling [4], glitching [5], precession [6], timing noise [7] and the phenomenon of intermittent pulsars [1]. Anomalous pulse emission modes have unique longitudinal positions that require sophisticated observational hardware [8] and algorithms [9].

Most of these emission modes (or signatures in the residuals of the pulsar) can be linked to changes in the magnetosphere, plasma conditions and/or the pulse profile of the pulsar itself [3]. Theoretically, any changes in the magnetosphere of the pulsar also influence the spin-down ($\dot{\nu}$). Therefore, if the spin-down of the pulsar can be computed accurately over the total time span of the observed data, it will reveal the history of the spin-down and the magnetospheric conditions of the pulsar. One can then fit different models to the spin-down of the pulsar with the hope of linking the conditions of the magnetosphere (emission history) to changes in the pulse profiles. Lyne et al. [3] observed a correlation between the spin-down and pulse profile changes in several pulsars. These correlations are in the form of mode switching in the pulsar

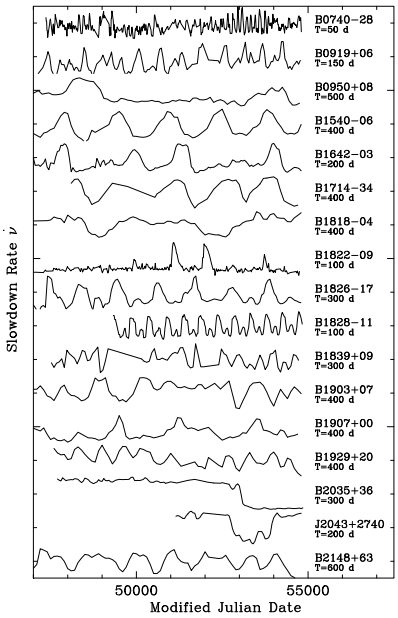


Figure 1. The spin-down history of 17 pulsars adopted from [3]. Clear mode switching between several states can be seen in the spin-down evolution of the pulsars. [3] also linked the mode switching seen in the spin-down evolution of the pulsars to pulse profile changes.

magnetosphere and could be a possible explanation for the observed timing noise signature seen in the timing residuals of several pulsars (Fig. 1).

In this paper we investigate the possible methods to link the observed timing noise signatures with possible changes in the magnetosphere and pulse profiles of the Hartebeeshoek Radio Astronomy Observatory (HartRAO) long term pulsar data sets. Timing noise analysis proves important for gravitational wave (GW) detection through the use of pulsar timing arrays, since both timing noise and GW signatures are considered to be red noise [10]. This paper is structured as follows: HartRAO Pulsar Data (Section 2), General Analysis Method (Section 3), Computing Algorithms (Section 4), Preliminary Results of spin-down Evolution of PSR J1326-5859 (Section 5) and an Discussion (Section 6).

2. HartRAO pulsar data

We investigate the timing noise signature seen in the 21 year long data set of PSR J1326-5859. This is a 478 ms [11] pulsar with a dispersion measure of $DM = 287\,\mathrm{pc\,cm^{-3}}$ [12] and an average spin-down of $-1.415 \times 10^{-14}\ \mathrm{s^{-2}}$ [13]. The data contains both 18cm and 13cm observations that were performed with the 26m HartRAO radio telescope. Timing of PSR J1326-5859 was done with the help of the packages TEMPO2 and PSRCHIVE [10,14]. Each time of arrival (and subsequently every residual) was produced via single polarization observations with an integration time of 20 minutes (Fig. 2).

3. General method of analysis

Timing residuals are produced by comparing the observed times of arrival (TOA) to the predicted ones, this process can be done with a combination of software: DSPSR (folding of raw data) [15], PSRCHIVE (manipulating of folded data) and TEMPO2 (timing analysis). Raw data are folded according to a pulsar ephemeris and the integrated pulse is used as a standard pulse to determine the TOA. A timing model is then fitted to the list of TOA and factors that are not

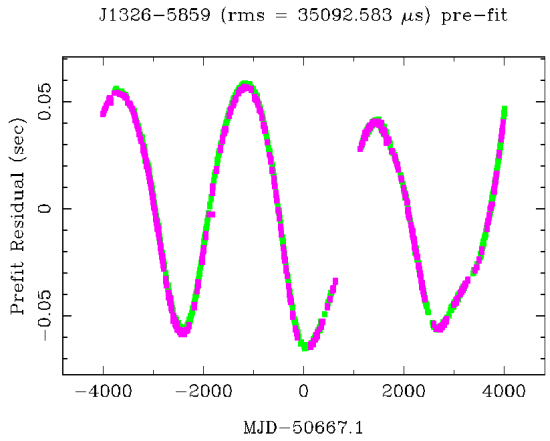


Figure 2. Timing residuals of J1326-5859 showing a timing noise signature that is quasi-periodic. The break in the residuals was due to telescope maintenance. Purple and green points represent observations at 13cm and 18cm.

predicted in the model will produce a unique signature in the residuals. In PSR J1326-5859 the signature is in the form of long time scale quasi-periodic timing noise.

When trying to connect the observed timing noise signature (or residuals) to the spin-down evolution of the pulsar, the immediate complexity that arises is that of computing the second derivative of the signature. This is done by reconstructing a function (F) that fits the residual signature. If the function F is not fitted in an accurate way, then the second derivative ($\ddot{F}$) will resemble false history of the spin-down of the pulsar. Any behaviour that is not considered as normal mode of emission, will be reflected in the spin-down history, like for example: glitches, mode switching, encounters with secondary masses and nulling. In the next section we investigate the possible methods to reconstruct the function F.

[16] developed novel methods and techniques to correlate the short term variability seen in both the pulsar spin-down and pulse profile variations by implementing Gaussian processes. We closely follow their recipe for determining the spin-down evolution using Gaussian process reconstructions of the residuals.

4. Algorithms

There are several methods to compute the spin-down evolution by using the timing residuals of a pulsar. One such method is to use the glitch plugin provided with the TEMPO2 code [10]. This plugin accepts hand-selected regions of residuals and computes the spin-down for each of these regions. However, the second derivative ($\ddot{F}$) is sensitive to the size of the chosen regions. Consequently, this method is restricted by the prerequisites of knowing what type of emission history to look for in the spin-down history (such as mode switching).

Another popular method to reconstruct the function F is the use of Gaussian fits. The main benefit of this fitting process is that the sizes of regions are optimized and the error in the reconstructed function decreases if the residuals are locally predictable. The Gaussian fitting process requires no assumptions of whether F is related to a specific model. The value of F when evaluated at any point x is a Gaussian random variable with mean $\mu(x)$ and variance $Var(x)$. The function values at a point x and some other point x' are related by the covariant function, $k(x, x')$. Thus, the latter indirectly implies that fitting a semi-predictable signature will produce a good fit to the residuals. The GaPP (Gaussian process in python) code [17] is

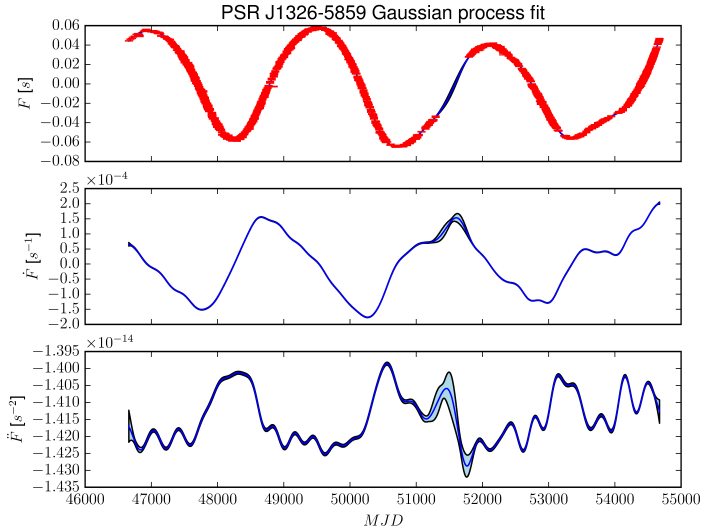


Figure 3. Spin-down evolution of PSR J1326-5859. Upper panel: Fitted timing noise residual. Middle Panel: Spin-down ($\dot{F}$). Lower Panel: Spin-down evolution ($\ddot{F}$). Errors of the fit are indicated by the shaded regions. The scaling of $\ddot{F}$ arises through the process of partial derivatives with time, see [16].

used to perform the Guassian process regression on the residuals. The GaPP is used for the reconstruction of dark energy data.

After computing the spin-down history of the pulsar, one can look at the history to identify certain signatures worth investigating and link that to any pulse profile changes. This can be done by a self developed python code that loads, fits and cleans the integrated pulses that were produced for each observation. One can afterwards plot, manipulate or stack pulses in certain regions of time. The only limitation of this process is the quality of the pulse profile itself.

5. Preliminary spin-down evolution of PSR J1326-5859

A list of residuals and TOA errors can be produced for the total data set using the TEMPO2 code. These residuals and TOA errors serve as input for the GaPP code. The GaPP code initializes the Gaussian process with initial parameters for the mean and the variance according a kernel:

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{(x - x')^2}{2\ell^2}\right). \quad (1)$$

One important part of the process is the optimization of the step length (l). Lastly follow the training of the hyper-parameters, reconstruction of the function F and its second derivative $\ddot{F}$ (Fig. 3). The parameters for the Gaussian process were optimized to be $l = 232$ days and $\sigma_f = 2.8 \times 10^{-2}$.

We used the pipeline to stack pulses in two temporal regions (48000 − 48500 MJD and 49000 − 49500 MJD) of different spin-down values (Fig. 4). A constant number of pulses were added in both regions to ensure a consistent S/N. In this paper the stacking was done for only the first part of the data set. The investigation of the second half will be left for future work.

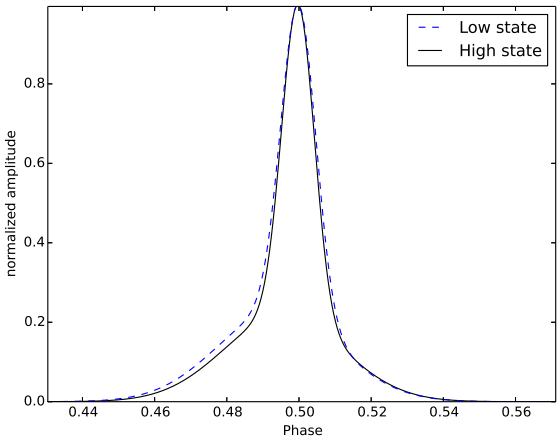


Figure 4. Stacking of 1644 MHz integrated pulses of PSR J1326-5859 in two different regions. The high and low state correspond to the largest and lowest spin-down observed in PSR J1326-5859.

6. Discussion

Switching in the spin-down evolution of PSR J1326-5859 is clearly visible (Fig. 3). The observed switching reveals itself as changing between two extreme states in the first half of the data set and becomes less predictable in second half. The stacking of pulses (of equal stacking length) in different states of the spin-down evolution proves to be insufficient for correlation searches. Thus, correlating the observed spin-down evolution to the pulse profiles, will require more sophisticated techniques that will be done in future work. In addition to the switching, we also observe an increase in the mean of the spin-down from MJD = 53000. Spin-down activity in the second half of the data becomes less clear. Similar results can be seen in the spin-down evolution of several pulsars reported in [3] (see Fig. 1-3 in their paper) and [16]. Recently a numerical model was developed for the evolution of non-spherical pulsar parameters (period and inclination angle) with plasma filled magnetospheres [18]. This MHD model was used to fit the observed residuals of PSR B1828-11 and the Crab pulsar according to a precession model. As future work we will be fitting different models to the spin-down evolution of PSR J1326-5859 (especially the second part of the data set) to test possible signatures such as MHD precession or encounters with massive objects.

7. Acknowledgments

We would like to acknowledge the SKA project for the financial support of this work.

References

[1] Lorimer D R 2005 *Handbook of pulsar astronomy* (Cambridge, UK New York: Cambridge University Press) ISBN 978-0-521-82823-9

[2] Osłowski S, van Straten W, Hobbs G B, Bailes M and Demorest P 2011 *Monthly Notices of the Royal Astronomical Society* **418** 1258–1271 (*Preprint* 1108.0812)

[3] Lyne A, Hobbs G, Kramer M, Stairs I and Stappers B 2010 *Science* **329** 408–412 (*Preprint* 1006.5184)

[4] Backer D C 1970 *Nature* **228** 42–43

[5] Lyne A G, Smith F G and Pritchard R S 1992 *Nature* **359** 706–707

[6] Frescura F and Flanagan C S 2003 *Radio Pulsars* (*Astronomical Society of the Pacific Conference Series* vol 302) ed Bailes M, Nice D J and Thorsett S E p 237

[7] Chukwude A E 2007 *Chinese Journal of Astronomy and Astrophysics* **7** 521–530

[8] Karuppusamy R, Stappers B and van Straten W 2008 *Publications of the Astronomical Society of the Pacific* **120** 191–202

[9] Jones D I 2012 *Monthly Notices of the Royal Astronomical Society* **420** 2325–2338 (*Preprint* `http://mnras.oxfordjournals.org/content/420/3/2325.full.pdf+html`) URL `http://mnras.oxfordjournals.org/content/420/3/2325.abstract`

[10] Hobbs G, Edwards R and Manchester R 2006 *Chinese Journal of Astronomy and Astrophysics Supplement* **6** 020000–192

[11] D'Alessandro F, McCulloch P M, King E A, Hamilton P A and McConnell D 1993 *Monthly Notices of the Royal Astronomical Society* **261** 883–894

[12] Hobbs G, Faulkner A, Stairs I H, Camilo F, Manchester R N, Lyne A G, Kramer M, D'Amico N, Kaspi V M, Possenti A, McLaughlin M A, Lorimer D R, Burgay M, Joshi B C and Crawford F 2004 *Monthly Notices of the Royal Astronomical Society* **352** 1439–1472 (*Preprint* `astro-ph/0405364`)

[13] Manchester R N, Hobbs G B, Teoh A and Hobbs M 2005 *VizieR Online Data Catalog* **7245**

[14] Hotan A W, van Straten W and Manchester R N 2004 *Publications of the Astronomical Society of Australia* **21** 302–309 (*Preprint* `astro-ph/0404549`)

[15] van Straten W and Bailes M 2010 (*Preprint* `1010.006`)

[16] Brook P R, Karastergiou A, Johnston S, Kerr M, Shannon R M and Roberts S J 2016 *Monthly Notices of the Royal Astronomical Society* **456** 1374–1393 (*Preprint* `1511.05481`)

[17] Seikel M, Clarkson C and Smith M 2012 *Journal of Cosmology and Astroparticle Physics* **6** 36 (*Preprint* `1204.2832`)

[18] Arzamasskiy L, Philippov A and Tchekhovskoy A 2015 *ArXiv e-prints,1504.06626*

Possible extragalactic astrophysical counterparts of IceCube neutrino events

Reetanjali Moharana, Richard J. Britto, Soebur Razzaque

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

E-mail: reetanjalim@uj.ac.za

Abstract. Sources of the IceCube observed neutrino events within energy range 30TeV-2PeV, is now an open question in astronomy and astrophysics. The dominant shower-type neutrino events have large errors in measuring their directions, hence it is difficult to identify their astrophysical sources. These neutrinos can have counterparts in non-thermal X-rays and gamma rays. So a cross-correlation study of IceCube neutrino events with extragalactic candidate sources using X-ray and gamma-ray selected source catalogs such as Swift-BAT, 3LAC and TeV-Cat, will help in identifying sources of the neutrino events. In order to search for the most possible candidates we apply cuts on X-ray and gamma-ray fluxes of the sources in those catalogs, and then we study the statistical significance of correlation by using invariant statistics and Monte Carlo simulations for different classes of sources.

1. Introduction

IceCube Neutrino Observatory, the world's largest neutrino detector, has detected 54 neutrino events within 1347 days with energy between 20 TeV and 2.3 PeV [1, 2]. Shower events, most likely due to ν_e or ν_τ charge current νN interactions and also due to neutral current νN interactions of all flavors, dominate the event list (39 including 3 events with 1–2 PeV energy) while track events, most likely due to ν_μ charge current interactions, constitute the rest. Among a total of 54 events about 21 could be due to atmospheric neutrino ($9.0^{+8.0}_{-2.2}$) and muon (12.6 ± 5.1) backgrounds. A background-only origin of all 54 events has been rejected at a 6.5-σ level [2]. Therefore a cosmic origin of a number of neutrino events is robust. The track events have on average $\sim 1^\circ$ angular resolution, but the dominant, shower events have much poorer angular resolution, $\sim 15^\circ$ on average [2]. Searching for sources of these events is now one of the major challenges in astrophysics. Pinpointing the astrophysical sources of these neutrinos is difficult, due to a large uncertainty in their arrival directions.

High energy cosmic rays (CRs) can interact with low energy photons and/or low energy protons to produce neutrinos and high energy gamma rays inside the source or while propagating to the Earth. So a multi-messenger study of neutrinos, Cosmic Rays (CRs) and gamma rays can identify the possible astrophysical sources. In our first attempt to search for sources we tried to see a correlation with Ultra-High Energy (UHE) CRs with the earlier 37 cosmic neutrino events detected by IceCube [3].

Here we study correlation of IceCube neutrino events with TeVCat [4], *Swift*-BAT 70 month X-ray source catalog [5] and 3LAC source catalog [6]. To do specific correlation study we use

different cuts on the observed energy flux of these sources, and also different sets of source types, and showed the results of this study.

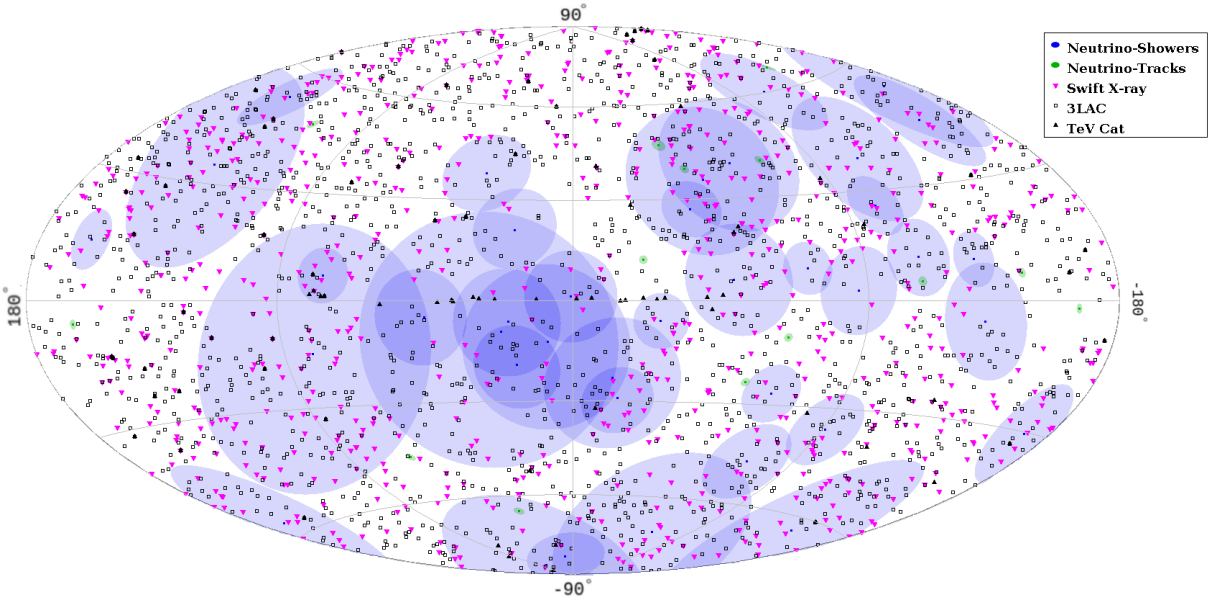


Figure 1. Sky map of the 52 IceCube cosmic neutrino events with error circles, and sources from different catalogs in Galactic coordinate system.

2. IceCube neutrino events and Source catalogs

For our analysis we have considered all 52 IceCube detected neutrino events. Two track events (event numbers 28 and 32) are coincident hits in the IceTop surface array and are almost certainly a pair of atmospheric muon background events [1]. Therefore we excluded them from our analysis. Fig. 1 shows the sky map of these 52 events in Galactic coordinates, with reported angular errors.

We have used 3 different source catalogs to do the correlation study. *Swift*-BAT 70 month X-ray source catalog [5], *Fermi* Third Catalog of Active Galactic Nuclei (3LAC) [6], TeVCat [4]. The sky map in Fig. 1 shows the extragalactic sources from these catalogs.

The *Swift*-BAT 70 month X-ray source catalog includes 1210 objects, from which only 785 are extragalactic sources, shown in Fig. 1. In our previous study [3] we found 18 sources from this catalog that are correlated simultaneously with UHECRs and IceCube neutrino events. The Pierre Auger observatory Collaboration has also found an anisotropy at $\sim 98.6\%$ CL in UHECRs with energy ≥ 58 EeV and within $\sim 18^\circ$ circles around the AGNs in the *Swift*-BAT catalog at a distance ≤ 130 Mpc [7] . These 18 sources mostly have an X-ray energy flux $\geq 10^{-11}$ erg cm^{-2} sec^{-1}. So, in the present analysis we use all the sources from this catalog which have a flux $\geq 10^{-11}$ erg cm^{-2} sec^{-1}. This condition decreased the number of sources to 687.

The TeVCat contains sources that are detected with very high energy (VHE) gamma rays with energy ≥ 50 GeV. It includes 161 sources, out of which 22 are unidentified sources. This is the highest energy source catalog, particularly interesting for ν production. Sky map in Fig 1 contains TeVCat extragalactic sources.

The *Third Catalog of Active Galactic Nuclei (AGNs)* detected by Fermi LAT (3LAC) [6] is a subset of the *Fermi* LAT *Third Source Catalog (3FGL)* [8]. The 3FGL catalog includes 3033 sources detected above a 4σ significance (test statistic > 25) on the whole sky, during the

first 4 years of the Fermi mission (2008-2012). The original 3LAC sample includes 1591 AGNs from 3FGL, though 28 are duplicate associations. An additional cut had also been performed to exclude the Galactic plane region ($|b| \leq 10^{\circ}$) where the incompleteness of the counterpart catalogs significantly hinders AGN association. However, in this paper, we chose to study what we call the "extended 3LAC" sample of 1773 sources, that includes sources of the Galactic plane, and that could be associated to several neutrino events. In the extended 3LAC sample, 491 sources are flat spectrum radio quasars (FSRQs), 662 are BL Lacs, 585 are blazars of unknown type (BCU), and 35 are non-blazar AGNs.

3. Statistical method for Correlation study

To study correlation between cosmic neutrinos and sources from different catalogs separately, we map the Right Ascension and Declination (RA, Dec) of the event directions and sources into unit vectors on a sphere as

$$\hat{x} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)^T,$$

where $\phi = RA$ and $\theta = \pi/2 - Dec$. The scalar product of the neutrino and source vectors ($\hat{x}_{\rm neutrino} \cdot \hat{x}_{\rm source}$) therefore is independent of the coordinate system. The angle between the two vectors

$$\gamma = \cos^{-1}(\hat{x}_{\rm neutrino} \cdot \hat{x}_{\rm source}), \tag{1}$$

is an invariant measure of the angular correlation between the neutrino event and source directions [3, 9]. Following ref. [9] we use a statistic made from invariant γ for each neutrino direction $\hat{x}_i$ and source direction $\hat{x}_j$ pair as

$$\delta\chi_i^2 = \min_j(\gamma_{ij}^2/\delta\gamma_i^2), \tag{2}$$

which is minimized for all j. Here $\delta\gamma_i$ is the 1-σ angular resolution of the neutrino events. We use the exact resolutions reported by the IceCube collaboration for each event [1].

A value $\delta\chi_i^2 \leq 1$ is considered a "good match" between the i-th neutrino and a source directions. We exploit distributions of all $\delta\chi_i^2$ statistics to study angular correlation between IceCube neutrino events and sources in catalog. The distribution with observed data giving a number of "hits" or $N_{\rm hits}$ with $\delta\chi^2 \leq 1$ therefore forms a basis to claim correlation. Note that in case more than one source direction from the catalog are within the error circle of a neutrino event, the $\delta\chi^2$ value for UHECR closest to the neutrino direction is chosen in this method.

We estimate the significance of any correlation in data by comparing $N_{\rm hits}$ with corresponding number from null distributions. We construct null distributions by randomizing only the RA of the sources, keeping their Dec the same as their direction in the catalog. This *semi-isotropic null* is a quick-way to check significance. We perform 100,000 realizations of drawing random numbers to assign new RA and Dec values for each event to construct $\delta\chi^2$ distributions in the same way we did with real data.

We calculate statistical significance of correlation in real data or p-value (chance probability) using frequentists' approach. We count the number of times we get a random data set that gives equal or more hits than the $N_{\rm hits}$ in real data within $\delta\chi^2 \leq 1$ bin. Dividing this number with the total number of random data sets generated (100,000) gives us the p-value. We cross-check this p-value by calculating the Poisson probability of obtaining $N_{\rm hits}$ within the $\delta\chi^2 \leq 1$ bin, given the corresponding average hits expected from the null distribution. We found the $N_{\rm hits}$ distribution in $\delta\chi^2 \leq 1$ does not follow the Poisson distribution.

4. Results and Discussions

We used the 45 HBL (high-frequency peaked BL Lacs) sources listed in TeVCat for our first correlation study with the neutrino events. A similar correlation study was carried out in [10],

using HBL sources with the 3 year IceCube neutrino events. Our study showed a p-value 0.58 with frequentists method. There are 16 neutrino events that correlated within 1σ error reported by IceCube with different HBLs, which is almost the same as for the null distribution. The Poisson distribution p-value, defined as $\lambda^k e^{-\lambda}/k!$, where k and λ are the $N_{\rm hits}$=16 for real data and null distribution respectively, is 0.1 for $\delta\chi^2 \leq 1$. The total distribution is shown in Fig. 2, where the blue histograms represent the number of neutrino events correlated with the HBLs while the red line represents the number in the case of the null distribution for different $\delta\chi^2$.

The *Swift*-BAT 70 month X-ray source catalog includes 657 sources with observed energy flux more than 10^{-11}erg cm^{-2} sec^{-1}. The study of correlation with neutrino events showed a p-value 0.825 with 39 $N_{\rm hits}$ for the neutrino data and nearly 40 for the null distribution. The detail study of the correlation is shown in Fig. 3.

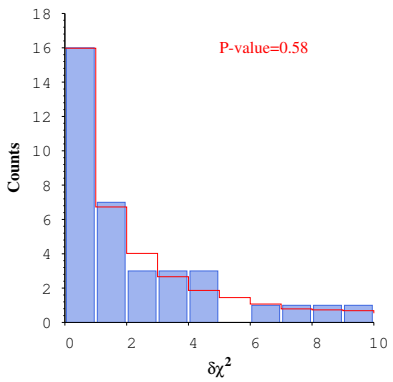


Figure 2. Correlation study of the 45 HBL sources from TeVCat with the IceCube neutrino events. The blue histogram shows the number of times the neutrino events are correlated with the HBLs within different $\delta\chi^2$, while the red line is for the null distribution.

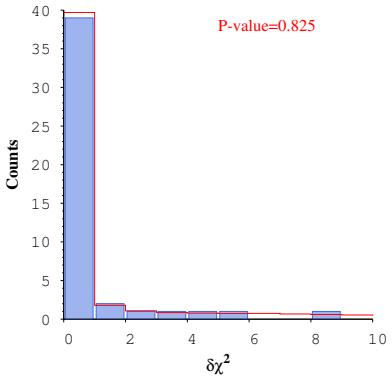


Figure 3. Correlation study of the *Swift* BAT X-ray catalog sources with energy flux more than 10^{-11}erg cm^{-2} sec^{-1} with the IceCube neutrino events. The blue bars shows the number of sources correlated with the neutrino events and the red line is for the null distribution of the sources.

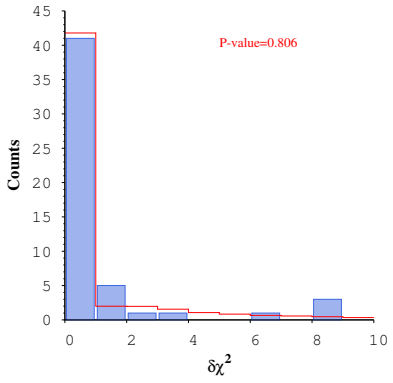


Figure 4. Correlation study of 1773 sources from the extended 3LAC source catalog with the IceCube neutrino events. The blue bars shows the number of sources correlated with the neutrino events and the red line is for the null distribution of the sources.

Our correlation study of all the 1773 sources in the extended 3LAC catalog gives a p-value 0.806 having 41 $N_{\rm hits}$ for the 52 neutrino events, as shown in Fig. 4. The distribution of the extended 3LAC sources according to their observed energy flux is shown in Fig. 5. It can be seen that most of the sources are populated in the region of energy flux 10^{-11}erg cm^{-2} sec^{-1}, and

the population decreases abruptly at higher flux. So, we took a set of sources with energy flux $\geq 10^{-11}$erg cm^{-2} sec^{-1}. It so decreased the number of sources in the set to 652. The correlation study of the neutrino events with these 652 sources show a p-value 0.763, having 39 $N_{\rm hits}$ in $\delta\chi^2 \leq 1$, shown in Fig. 6.

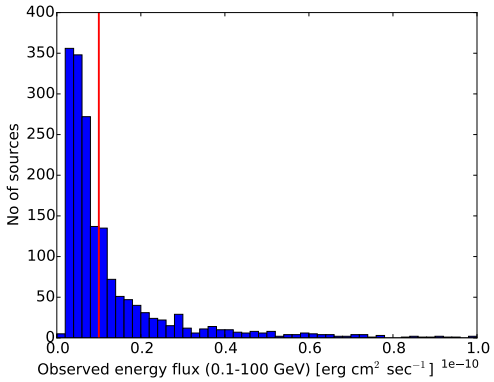


Figure 5. Distribution of energy flux of the extended 3LAC sources. The red vertical line indicate the energy flux 10^{-11}erg cm^{-2} sec^{-1}

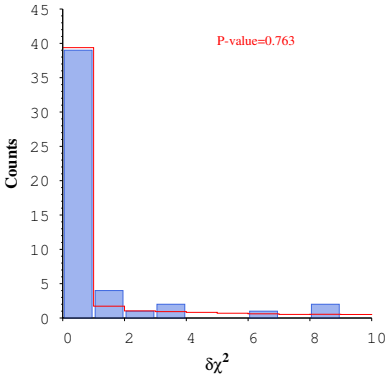


Figure 6. Correlation Study for sources from the extended 3LAC catalog with energy flux $\geq 10^{-11}$erg cm^{-2} sec^{-1}.

For further correlation studies we have used the 662 BL Lac source set from the extended 3LAC catalog. The correlation p-value for these sources is 0.764 as shown in Fig. 7. Similarly for the 491 FSRQ sources from the extended 3LAC catalog the p-value is 0.784, shown in Fig. 8. For BL Lac and FSRQ sources we found 39 and 38 $N_{\rm hits}$ respectively.

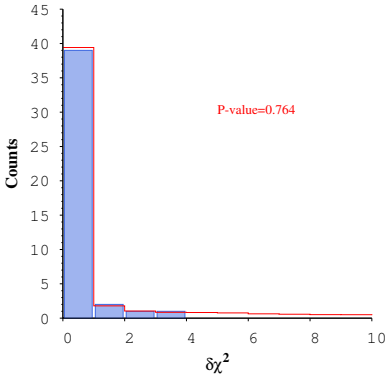


Figure 7. Correlation study of IceCube neutrino events with the BL LAC sources from the extended 3LAC catalog.

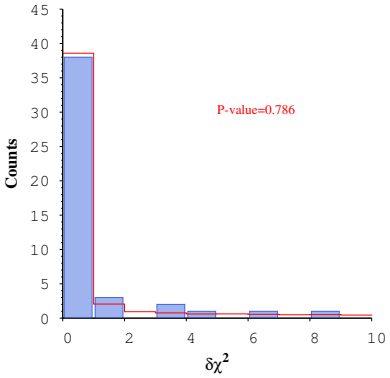


Figure 8. Correlation study of IceCube neutrino events with the FSRQ sources from the extended 3LAC catalog.

We have performed the correlation study of the IceCube neutrino events with different types of sources, as TeVCat HBL, 3LAC BL Lacs and FSRQs, but we have not found any statistically significant result for these sets. We have also taken subsets of the 3LAC catalog and the sources observed by the *Swift* during 70 months, by putting constraints on the observed energy flux, but the results remained similar. However these studies help us to discarding different types of astrophysical sources as the origin of the IceCube neutrino events.

5. Summary

The IceCube neutrino observatory has detected at least 54 neutrino events within the 30TeV–2PeV energy range. The Origin of these events is still a puzzle for both particle physics and astrophysics. In our project we have tried to find correlations of the arrival direction of these events with the positional direction of sources from the TeVCat, *Swift* and 3LAC catalogs. In order to test correlations we have used the minimum $\delta\chi^2$ invariant statistics, as in [3, 9]. Out of the 52 neutrino events, 16 were correlated with HBLs from TeVCat, but the statistical significance of this correlation is given by a $p-$value 0.58. Similarly we study correlations of neutrino events with the sources from the *Swift* and 3LAC catalogs, having observed an energy flux $\geq 10^{-11}\mathrm{erg\,cm^{-2}\,sec^{-1}}$. We also found a poor statistic significance. The FSRQs and BL Lacs from the 3LAC catalog also showed a low significant statistics in this correlation study.

References

[1] Aartsen M *et al.* (IceCube) 2014 *Phys.Rev.Lett.* **113** 101101 (*Preprint* 1405.5303)
[2] Halzen F 08-13 July,2015 25th International Workshop on Weak Interactions and Neutrinos.
[3] Moharana R and Razzaque S 2015 (*Preprint* 1501.05158)
[4] Wakely S P and Horan D 2008 *International Cosmic Ray Conference* **3** 1341–1344
[5] Baumgartner W H, Tueller J, Markwardt C B, Skinner G K, Barthelmy S, Mushotzky R F, Evans P A and Gehrels N 2013 *APJS* **207** 19 (*Preprint* 1212.3336)
[6] Ackermann M *et al.* (Fermi-LAT) 2015 (*Preprint* 1501.06054)
[7] Aab A *et al.* (Pierre Auger) 2015 *Astrophys.J.* **804** 15 (*Preprint* 1411.6111)
[8] Acero F e 2015 *APJS* **218** 23 (*Preprint* 1501.02003)
[9] Virmani A, Bhattacharya S, Jain P, Razzaque S, Ralston J P *et al.* 2002 *Astropart.Phys.* **17** 489–495 (*Preprint* astro-ph/0010235)
[10] Sahu S and Miranda L S 2015 *Eur.Phys.J.* **C75** 273 (*Preprint* 1408.3664)

A comparative study of the three empirical solar models in North West Province, South Africa.

T.S. Mulaudzi[1] , N.E. Maluta[1] , V Sankaran[1] and F Nemangwele[1]
[1]University of Venda, Department of physics, P/Bag X 5050, Thohoyandou, 0950
Author to whom correspondence should be addressed
corresponding author's: sophie.mulaudzi@univen.ac.za

Abstract. Energy crisis in South Africa (SA) is causing a lot of problems and has a negative impact on the growth of our economy. There is a dire need to implement the 2020 strategies to harness renewable energy and evidently the knowledge of the amount of solar irradiance received at different sites in SA is . With this knowledge the renewable energy systems can be meaningfully developed so as to sustain the outdoor conditions. The use of pyrheliometers and pyranometers to measure direct and global solar radiation respectively, also play an important role. However these instruments cannot be installed in many areas as they are expensive. In such situation it is economically viable to resort to a suitable theoretical for the estimation of global solar radiation. It is of a vital importance that a model to be selected should give reasonable estimates. This paper gives a comparative study of three modified empirical solar radiation models (Angstrom; Hargreaves & Samani and Glower & McCulloch) on a horizontal surface from sunshine hours and temperatures of different stations in North West Province, SA ($-25.8080^{\circ}; 25.5430^{\circ}$). A five year meteorological data from the four Agricultural research Council (ARC) stations were used to estimate the global solar radiation for this region. The estimated monthly solar irradiance data was compared with observed data using the statistical parameters such as, the mean bias error (MBE) and root mean square (RMSE). The Angstrom and temperature based models give better estimations for North West province.

1. Introduction

Global solar radiation on a horizontal surface passes through several layers of the earth whereby some portion of this radiation is reflected; scattered and or absorbed. Pyranometers can be used to measure the amount of global solar radiation at a particular area or satellite. Due to lack of funds, our country cannot afford to install pyranometers everywhere. The ARC and South African Weather Services (SAWS) have systematic measurements of some meteorological data in some areas at North West Province. The ARC stations are situated far much apart, so it is necessary to predict the global solar radiation at areas where there is no measuring instrumentation. Estimation of solar radiation is possible provided there are suitable empirical models. The knowledge of local global solar radiation is of vital importance for suitable design of energy systems, building photovoltaic (PV) stations, and a good evaluation of the thermal environment within buildings.

In this paper, the Angstrom regression coefficients are determined using Zabara model [1-2] that uses sunshine hours. Three empirical models, Angstrom-Prescott linear model [2 - 5]; Glower and McCulloch model [6] and Hargreaves – Samani equation [7-8] estimated the global solar radiation for the three ARC stations, that is, Koster; Ventersdorp and Delareyville at North West Province in South Africa (SA).

2. Methodology

The main aim of the present work is to compare the three empirical models and select a suitable model for estimating global solar radiation in North West Province, SA. The following models are used :

2.1 Angstrom – Prescott model

The global solar radiation is related to the amount of sunshine hours received in a particular area is given by the relation [2 – 4]:

$$\frac{H}{H_0} = a + b * \frac{S_a}{S_0} \quad (1)$$

Where H is the average daily global solar radiation, H_0 is the average extra-terrestrial solar radiation in MJ.m^-2 day^-1, a and b are regression coefficients, S_a is the daily observed sunshine duration and S_o is the possible sunshine duration or the maximum daylengths in hours. The ratio, $\frac{H}{H_0}$, represents the clearness index – the percentage deflection by the sky of the incident global solar radiation. This indicates the availability of solar radiation and changes in atmospheric conditions. $\frac{S_a}{S_0}$ is the relative sunshine hours – the measure of cloud cover. $\boldsymbol{a}$ and $\boldsymbol{b}$ are the regression coefficients that depend on the latitude and seasonal variation. $\boldsymbol{a}$ is the measure of the overall atmospheric transmission for totally cloudy conditions while $\boldsymbol{b}$ is the rate of increase of the clearness index with relative sunshine hour.

2.2 Glower and McCulloch model

Global solar radiation depends on the latitude and sunshine duration of the particular area under study and it is valid only for latitudes that are less than 60° [6].

$$\frac{H}{H_0} = a * \cos Lat + b * \frac{S_a}{S_0} \quad (2)$$

2.3 Hargreaves and Samani model

The equation given below shows the relationship between the clearness indices and the average daily maximum and minimum air temperatures [7].

$$\frac{H}{H_0} = k_r \sqrt{(T_{max} - T_{min})} \quad (3)$$

where k_r is an empirical constant = 0.16 for inland.

A five year daily data of sunshine hours, maximum and minimum air temperatures and global solar radiation data on the horizontal surface are provided by ARC and SAWS. The average daily extra-terrestrial solar radiation, H_0 and maximum possible sunshine hours, S_0 are computed using the relations given by Muzathik, Mulaudzi,et al [8-10] . The average daily extra-terrestrial solar radiation depends on the declination angle, δ, sunset hour angle, ω_s and the latitude [11-14].
The regression coefficients are computed monthly for the period of 5 years (2007 – 2011) using the following relations [12, 14]:

$$a = 0.395 - 1.247 * \left(\frac{S_a}{S_0}\right) + 2.680 * \left(\frac{S_a}{S_0}\right)^2 - 1.674 * \left(\frac{S_a}{S_0}\right)^3 \quad 4(a)$$

$$b = 0.395 + 1.384 * \left(\frac{S_a}{S_0}\right) - 3.429 * \left(\frac{S_a}{S_0}\right)^2 + 2.055 * \left(\frac{S_a}{S_0}\right)^3 \quad 4(b)$$

To compute the extra-terrestrial solar radiation and the maximum possible sunshine hours, we have used Mat-Lab software. We have also written the programme in order to determine the daily regression coefficients of each station under study using equations 4(a) and (b).

3. Results and discussion

The graphs below show the comparison of the estimated global solar radiation (H) determined from the three selected models with the ground measured data received from ARC. The study has been carried out for five years and some samples of the curves are presented below. The three empirical models mentioned in section 2 were used to calculate the estimated global solar radiation of each of the three selected stations in North West Province. Due to seasonal variations, the annual regression coefficients per station are as follows: Ventersdorp (a = 0.2153, b = 0.53480), Koster (a= 0.2156, b = 0.5254) and Delareyville (a = 0.2155, b = 0.5267). The average regression coefficients of North West Province are $a = 0.216$ and $b = 0.525$. The observed minimum and maximum temperature data requested from ARC were the input in equation 3 to estimate the daily global solar radiation Comparisons of the estimated and the measured global solar radiation are represented in Figures 1-6.

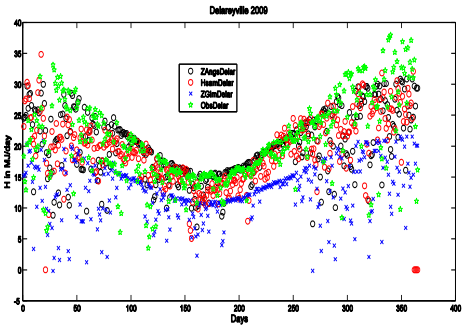


Figure 1: H Delareyville 2009 versus days of the year.

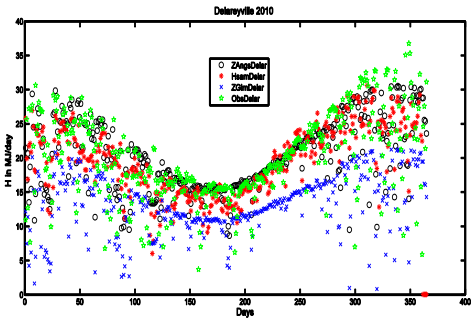


Figure 2: H Delareyville 2010 versus days of the year.

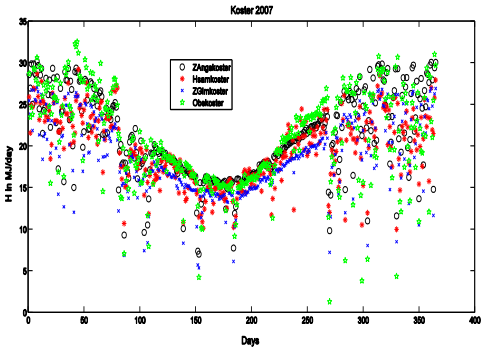


Figure 3: H Koster 2007 versus days of the year.

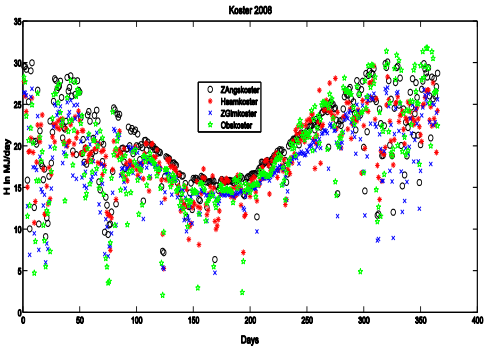


Figure 4: H Koster 2008 versus days of the year.

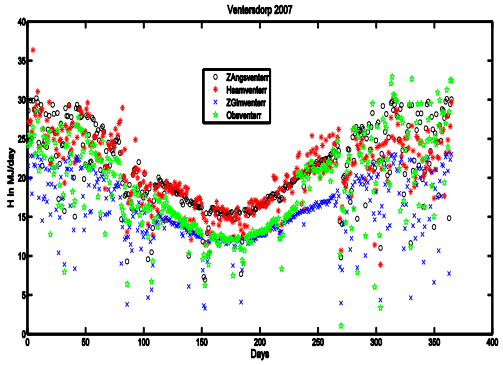


Figure 5: H Ventersdorp 2007 versus days of the year.

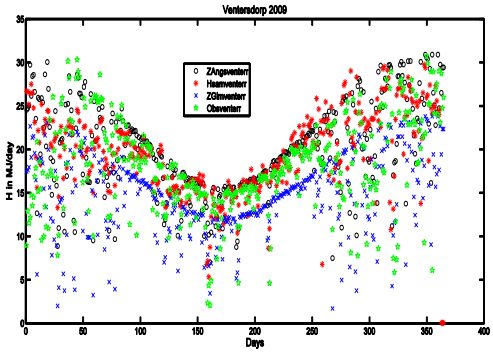


Figure 6: H Ventersdorp 2009 versus days of the year.

It is observed from the graphs that Angstrom and Hargreaves – Samani models give better estimation data for the stations almost every year. Variation of estimated global solar radiation data is compared with the ground measured data and we found that these two models have more accuracy than Glower and McCulloch model. Glower and McCulloch model underestimates the global solar radiation at Ventersdorp and Delareyville. All these models are suitable for Koster station. So, it is advisable to employ Glower & McCulloh model for the estimation of global solar radiation at the neighbourhood of Koster station. Figures 1-6 show that the global solar radiation data varies from (5.00 – 30.0) MJ.m^-2 day^-1 in summer. This variation is caused by the change in weather conditions. It is due to wind and water droplets since the rain fall during summer.

The statistical analysis for the Angstrom model was performed using the following equations [15]:

$$\text{Mean bias Error (MBE)} = \frac{1}{n}\sum_{i=1}^{n}(H_{est} - H_{obs}) \quad (5)$$

$$\text{Root Mean Square error (RMSE)} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(H_{est} - H_{obs})^2} \quad (6)$$

where H_{est} and H_{obs} are the estimated and observed global solar radiation respectively.

The average annual MBE and RMSE data for each station at North West are tabulated below:

Table 1: The MBE and RMSE of Ventersdorp; Koster and Delareyville stations.

	Ventersdorp (RMSE)	Ventersdorp (MBE)	Koster (RMSE)	Koster (MBE)	Delareyville (RMSE)	Delareyville (MBE)
2007	0.0103	-0.0065	0.0083	-0.0011	0.0115	0.0001
2008	0.0221	-0.0134	0.0109	-0.0036	0.0108	-0.0038
2009	0.0150	-0.0056	0.0166	-0.0101	0.0138	0.0000
2010	0.0227	-0.0107	0.0175	-0.0158	0.0101	0.0011
2011	0.0114	-0.0026	0.0179	-0.0146	0.0115	0.0000

4. Conclusion

The average daily global solar radiation of the observed and the predicted data were compared graphically. The modified Angstrom and Hargreaves – Samani models are more suitable for the estimation of global solar radiation at North West Province of South Africa the observed and the predicted daily global solar radiation values are comparable.

5. Acknowledgement

We would like to acknowledge the ARC and SAWS for the meteorological data they have supplied to us. Furthermore, we thank the University for the Research Support.

References

[1] Zabara K 1986 Estimation of the global solar radiation in Greece. *Solar and wind technology* **3**: 267 – 272.

[2] Rashid M A, Mamun R A, et al. 2012 Evaluating the solar radiation system under the climatic condition of Dhaka, Bangladesh and computing the Angstrom coefficients. *International Journal of Natural Sciences* **2** (1): 38-42.

[3] Duffie J A, Beckam W A, 1991 Solar engineering of thermal process. New York: Wiley.

[4] Angstrom A, 1924.Solar and terrestrial radiation. *Quarterly Journal of the Royal Meteorological Society* **50**: 121 – 125.

[5] Prescott J A, 1940 Evaporation from water surface in relation to solar radiation. *Transactions of the Royal Society of South Australia* **64** (3): 114 – 118.

[6] Glower J, McCulloch J S G, 1958 The empirical relation between solar radiation and hours of sunshine. *Quarterly Journal of the Royal Meteorological Society* **84**: 172.

[7] Hargreaves, G.H. and Z.A. Samani, "Reference crop evapotranspiration from temperature," Transaction of ASAE, vol. 1(2), pp. 96-99. 1985.

[8] Maluta E N, Mulaudzi T S and Sankaran V, 2013 Estimation of Global Solar Radiation on Horizontal Surface from Temperature Data for Vhembe District in Limpopo Province, South Africa. *International Journal of Green Energy* **11**: 1-11.

[9] Muzathik A M, Nik W B W, et al.; 2011 Daily global solar radiation estimation based on sunshine hours. *International Journal of mechanical and materials Engineering* **6** (1):75-80.

[10] Mulaudzi T S, Sankaran V and Lysko M, 2013. Solar radiation analysis and regression coefficients for the Vhembe Region, Limpopo province, South Africa. *Journal of Energy in South Africa* **24** (3):1-7.

[11] Ertekin C, Yaldiz O, 2000 Comparison of some existing models for estimating global solar radiation for Antayla. *Energy Conversion and Management* **41**: 311-330.

[12] Jamil Ahmad M. and Tiwari G.N., 2010. Solar radiation models – review. *International Journal of Energy and Environmental* **1**(3): 513- 532.

[13] Kumar R. and Umanand L., 2005. Estimation of global radiation using the clearness index model for sizing photovoltaic system. *Renewable Energy* **30**: 2221 – 2233.

[14] Medugu D.W. and Yakubu D., 2011.Estimation of mean monthly global solar radiation in Yola – Nigeria using angstrom model. *Advances in Applied Science Research* **2**(2): 414 – 421.

[15] Rahimi I., Bakhtiari B.,etc., 2012.Calibration of Angstrom equation for estimating solar radiation using meta-Heuristic harmony Search Algorithm Case Study: Mashhad-East of Iran). *Energy Procedia* **18**: 644-651.

A quasi-periodicity in the optical polarization of the blazar PKS 2155-304

N.W. Pekeur[1], R. Taylor[1, 2], S.B. Potter[3] and R.C. Kraan-Korteweg[1]

[1] Astrophysics, Cosmology and Gravity Centre (ACGC), Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, SA

[2] The Astrophysics Group, University of the Western Cape, Modderdam Road, Private Bag X17, Belville 7530, SA

[3] South African Astronomical Observatory, PO Box 9, Observatory 7935, Cape Town, SA

E-mail: nikki.pekeur@ast.uct.ac.za

Abstract. We report the detection of a quasi-periodic oscillation (QPO) in the polarized flux of the BL Lac PKS 2155-304 during a state of increased gamma-ray activity. Measurements were obtained with the HIgh Speed Photo-POlarimeter (HIPPO) of the South African Astronomical Observatory (SAAO) using a temporal resolution of 5 minutes. On 24 July 2009, the polarization degree appeared to be modulated by a periodic component. Simultaneous gamma-ray measurements with the High Energy Stereoscopic System (H.E.S.S.) showed that the quasi-periodicity coincided with an increase in the gamma-ray flux of the source. A periodogram of the polarized flux reveals the presence of a prominent peak at a frequency corresponding to a period of ∼30 minutes. This is the first detection of QPO activity in the optical polarization of an active galactic nucleus (AGN), potentially opening up a new method of studying the AGN phenomenon. Mechanisms capable of producing such a short timescale periodicity are briefly discussed.

1. Introduction

Active galactic nuclei (AGN) and Galactic black hole binaries (BHBs) are believed to be mass-scaled analogies, both powered by accretion of matter onto supermassive black holes (SMBH, with masses of $10^6 - 10^9 M_\odot$) and massive black holes ($\sim 10 M_\odot$ for BHBs). Observations of quasi-periodic oscillations (QPOs) are fairly common for BHBs and serve as an important diagnostic tool, e.g. providing accurate estimates of the black hole mass [1, 2]. On the contrary, very few claims of QPOs have been made for AGN, e.g. [3, 4]. PKS 2155–304 is one of a small number of AGN for which convincing evidence of quasi-periodic brightness variations have been found. The source is classified as a blazar, a subclass of radio-loud AGN for which the relativistic jet is oriented close to the line-of-sight of the observer [6]. PKS 2155–304 exhibits significant variability on timescales ranging from a few minutes [7] to years [8, 9]. Most recently, a ∼4.6 hour periodic component was detected in the $X-$ray light curve of the source [5]. Here, we report the possible discovery of a QPO in the polarized flux of PKS 2155–304. The optical polarization of the source was monitored for roughly three days in July 2009, during which amplitude fluctuations on timescales ranging from minutes to hours were detected. Gamma-ray observations were obtained independently over the same period, with indications of two $\gamma-$ray flares. In this paper we discuss the implications of the coincidence of these multi-wavelength

measurements. A brief description of the polarization and very high energy (VHE, photons with energies above GeV) $\gamma-$ray observations is given in section 2, followed by the data analysis in section 3. The results are discussed in section 4 and a summary of our findings are presented in section 5.

2. Observations

A description of the polarization and gamma-ray variability of PKS 2155–304 in July 2009 follows.

2.1. Optical polarization

The optical polarization of PKS 2155–304 was measured with the HIgh-sPeed Photo-POlarimeter (HIPPO [10]), a rapidly rotating dual-channel photo-polarimeter operated by the South African Astronomical Observatory (SAAO). The instrument was mounted on the 1.9 m Radcliffe telescope of the SAAO. The intra-day variability (IDV) of the source was monitored from 25 – 27 July 2009 for a total of 9.23 hours, using a time-resolution of 5 minutes. The linear and circular polarization of the source was recorded in the $I-$ and $B-$band. Observations commenced one day after a VHE $\gamma-$ray flare was recorded with the High Energy Stereoscopic System (H.E.S.S.). A summary of the average polarization degree (or fractional polarization) and polarization angle[1] are presented in Table 1 which lists the modified Julian date (MJD), length of observation (T_{obs}) and the mean daily polarization degree and polarization angle (calculated by taking the average of the $I-$ and $B-$band measurements). The table demonstrates that the polarization degree increased from roughly 3% – 9% over a three day period (MJD = 55037 to 55038) and that the orientation of the polarization angle changed rapidly from 90° to 68° in the span of one day.

Table 1. The mean daily optical polarization of PKS 2155–304 in July 2009.

MJD	T_{obs} (min)	p (%)	θ (°)
55037	251	3.7 ± 0.3	88 ± 2.5
55038	105	7.0 ± 0.3	67 ± 1.0
55039	197	8.3 ± 0.7	68 ± 0.5

2.2. Gamma-rays

Our observations overlapped with an increase in the $\gamma-$ray activity of the blazar. The results from the long-term monitoring campaign of PKS 2155–304 with the High Energy Stereoscopic System [11] between 19 and 27 July shows two prominent increases in the VHE flux I above the threshold energy of 300 GeV, as illustrated in Fig. 1. The polarization degree is indicated by the square markers and is superimposed on the VHE light curve. The $\gamma-$ray flares (marked as Flare 1 and Flare 2 on Fig. 1) are detected above a median flux of 1.9×10^{-11} ph cm^{-2} s^{-1}, consistent with the baseline ($\sim 10^{-11}$ ph cm^{-2} s^{-1}) reflected by the long-term measurements of the source [11]. The first $\gamma-$ray flare lasts approximately four days from MJD = 55033 to 55037, peaking at $I = (8.0 \pm 0.4) \times 10^{-11}$ ph cm^{-2} s^{-1} on MJD = 55035 (or 23 July 2009). Immediately following this, another increase in the $\gamma-$ray brightness is seen (Flare 2). The VHE flux rises from $(1.9 \pm 0.4) \times 10^{-11}$ ph cm^{-2} s^{-1} to $(5.6 \pm 0.4) \times 10^{-11}$ ph cm^{-2} s^{-1} over three days (from MJD = 55037 to 55039), which could be indicative of a developing flare. However, we are unable to verify this since the 2009 H.E.S.S. campaign concluded on MJD = 55039.

[1] The polarization angle is measured from the North Celestial Pole towards the Celestial East.

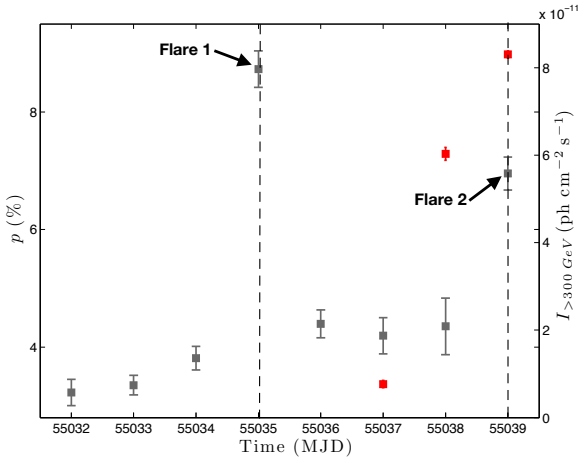


Figure 1. The $\gamma-$ray flux I of PKS 2155–304 during July 2009 for photons with energies > 300 GeV. Two prominent increases in the $\gamma-$ray flux are detected, marked as "Flare 1" and "Flare 2" respectively. For Flare 2 we have simultaneous polarization measurements. The $I-$band polarization degree p is represented by the red symbols and are superposed on the VHE light curve (indicated by the grey symbols). Note that the increase in the fractional polarization precedes the increase in $\gamma-$ray flux of Flare 2. The left scale shows p (%), while the scale on the right shows I (ph cm^{-2} s^{-1}).

3. Data Analysis

The daily polarization degree and $\gamma-$ray flux of PKS 2155–304, as well as amplitude fluctuations in the intra-day polarization degree, are examined below.

3.1. Simultaneous polarization and $\gamma-$ray measurements

The first simultaneous VHE $\gamma-$ray and optical polarization observations of PKS 2155–304 during an active state were recorded for Flare 2. A comparison between the fractional polarization and VHE flux shows that the increase in polarization degree precedes the rise in $\gamma-$ray flux by approximately one day. At radio to optical wavelengths the primary emission mechanism in blazars is synchrotron radiation, which produces polarized emission. VHE photons are produced by up-scattering the lower-energy synchrotron photons via the inverse Compton mechanism in a process known as synchrotron self-Compton (SSC) emission [12, 13, 14]. Since both processes are seeded by the same particle population correlations between fluctuations in the synchrotron and $\gamma-$ray emission are expected. It is therefore likely that the increase in the polarization degree and the gamma-ray flux are correlated and that the same evolving emission region is responsible for both.

3.2. A quasi-periodic oscillation in the intra-day polarization?

Inspection of the IDV on MJD = 55037, at the onset of the second $\gamma-$ray flare, suggests that the amplitude of the polarization degree is modulated on a timescale of ~ 30 min, as illustrated in Fig. 2. These short-term amplitude modulations are superposed on a slowly decreasing baseline over the 290 minutes of continuous observation. It is noteworthy that such short-term modulations are not seen on MJD = 55038 and 55039. Since the most robust measurements were recorded in the $I-$band (higher signal-to-noise ratio compared to the $B-$band) we focus our analysis on the $I-$band polarization, which is represented by the Stokes Q and U fluxes.

To test for the presence of periodic components in the optical polarization we computed the Lomb-Scargle periodogram [15], which measures the amount of variability power as a function of temporal frequency (ν) relative to the variance (σ^2) of the observations. The periodogram of the de-trended Stokes Q and U polarized fluxes is displayed in Fig. 3. The Q periodogram reveals

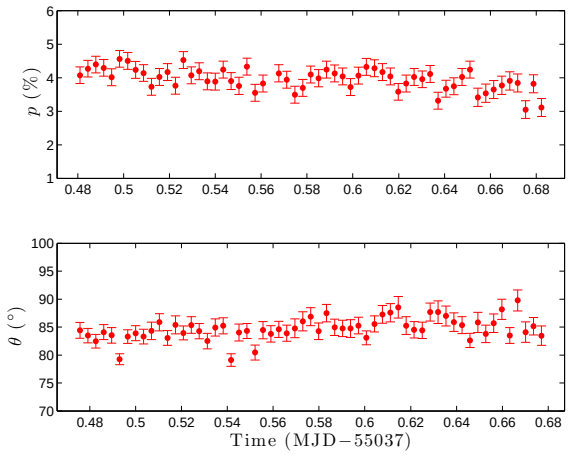


Figure 2. The intra-day variability of the I–band polarization of PKS 2155–304 as seen on 25 July 2009 (MJD = 55037). The top panel displays the polarization degree p (%), while the position angle of the electric vector θ (°) is shown in the bottom panel. The polarization degree appears to modulated by a periodic component at ∼30 minute intervals.

the presence of two dominant frequency components at $\nu = 563 \pm 10$ μHz and $\nu = 1293 \pm 9$ μHz. Each peak signal is well-described by a Gaussian function for which the errors are estimated from the full-width half-maximum and the peak height [16]. The amplitude of the Stokes Q flux therefore appears to be modulated by two periodic components, while no dominant peaks are detected in for U. The lack of detection in U is consistent with a variation in polarized intensity (as opposed to angle), as the observed polarization position angle $\theta \approx 90°$ (see Table 1 and Fig. 2), places virtually all the polarized emission in Q.

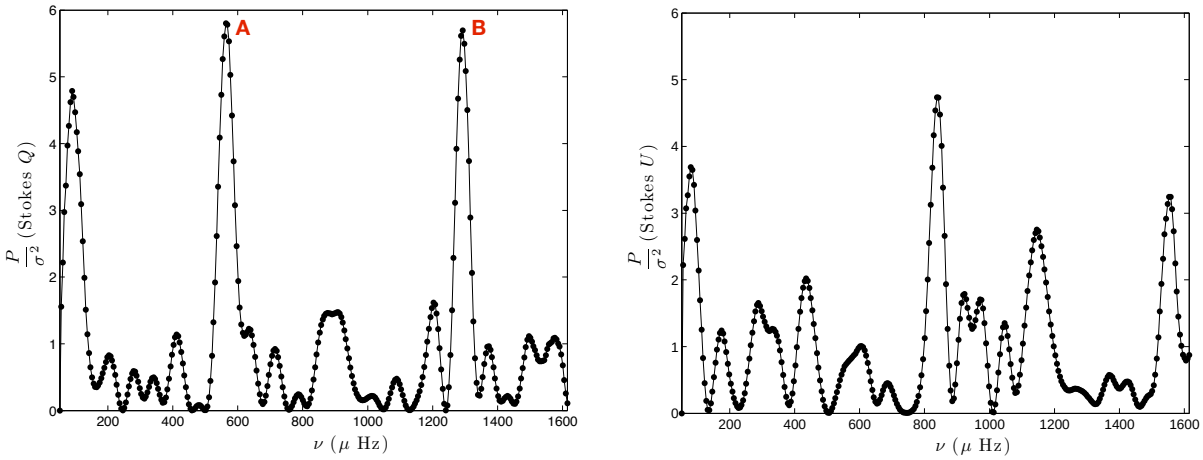


Figure 3. The periodogram of the Stokes Q and U polarized flux, shown in the left and right panel, respectively. Two prominent peaks are detected for Q at $\nu = 5.94 \times 10^{-4}$ Hz (corresponding to a period $T = 30$ min) and 12.9×10^{-4} Hz ($T = 13$ min), labelled as **A** and **B**, respectively. No dominant peaks are detected for U.

The phase-folded light curve of the Stokes Q flux is displayed in Fig. 4. Although the amplitude modulation is small, a cyclic trend appears to be present for both periods. A non-linear least-squares fit to the data yields a best-fit sinusoid that agrees well with the phase-folded diagram.

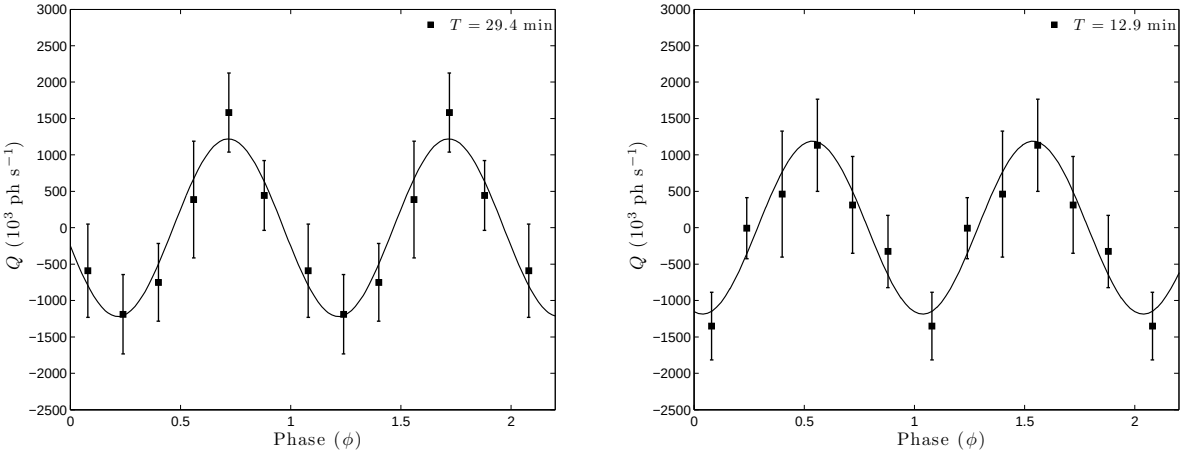


Figure 4. The phase-folded light curve of the Stokes Qflux, averaged in six phase bins, for $T = 30$ min and $T = 13$ min. Two cycles are plotted for clarity. A cyclic trend appears to be present for both periods. The solid line represents the best-fit sinusoid for each case.

4. Discussion

Since blazar emission is dominated by radiation from the relativistic jet, any change in the direction of the emitted radiation with respect to the observer, even for small changes in the viewing angle, will give rise to brightness fluctuations due to Doppler boosting. Fast variations are typically attributed to shocks propagating down the jet [17, 18]. Quasi-periodic fluctuations can arise if the shock encounters helical structures in the jet, for example the magnetic field, electron density and or the jet flow. Then, successive peaks in the emitted flux will be observed every time the shock completes a turn. Turbulence behind the shock can also explain the presence of short-lived QPOs. For such turbulent flows, the largest eddies will dominate the variability and the period will be determined by their turnover times. Finally, it might also be possible to detect emission from orbital motions of rotating hot-spots near the accretion disk or corona of the central engine if the blazar is in a quiescent state. However, since optical polarization in blazars are the result of synchrotron radiation from the jet, the QPO observed in the polarized flux of PKS 2155–304 most likely originates in the jet.

5. Summary and Conclusions

Monitoring of the intra-day optical polarization of PKS 2155–304 revealed the first detection of QPOs in the polarized flux of an AGN. Two modulating components with periods of 30 min and 13 min were identified in the Stokes Q polarized flux. Overlapping gamma-ray observations of the source showed that it was in a high state of activity, experiencing two $\gamma-$ray flares. The first simultaneous optical polarization measurements of the source during a high-state was recorded for the latter flare. Although the physical cause of this QPO is unclear, comparison with the VHE light curve showed that the source experienced a $\gamma-$ray flare two days before the appearance of the QPOs and another peak two days later. The oscillations could therefore be related either to the late phase of post flare activity for the first flare or to the onset of the latter flare, which could suggest that the QPOs are part of a longer-lived phenomenon within the blazar jet.

References

[1] Remillard, R A & McClintock, J E 2006 *Annual Rev. of A&A* **44** 49
[2] McClintock, J E & Remillard, R A & Rupen, M P & Torres, M A P *et al.* 2009 *ApJ* **698** 1398

[3] Gierliński, M & Middleton, M & Ward, M & Done, C 2008 *Nature* **455** 369
[4] Rani, B & Gupta, A C & Joshi, U C & Ganesh, S 2010 *ApJ Lett.* **719** L153
[5] Lachowicz, P & Gupta, A C & Gaur, H & Wiita, P J 2009 *A&A* **506** L17
[6] Urry, C M & Padovani, P 1995 *PASP* **107** 803
[7] Aharonian, F *et al.* 2007 *ApJ Lett.* **664** L71
[8] Fan, J H & Lin, R G 2000 *A&A* **355** 880
[9] Kastendieck, M A & Ashley, M C B & Horns, D 2011 *A&A* A **123** 531
[10] Potter, S B & Buckley, D A H & O'Donoghue, D & Romero-Colmenero, E *et al.* 2010 *MNRAS* **402** 1161
[11] Abramowski, A *et al.* 2014 *A&A* **571** A39
[12] Maraschi, L & Ghisellini, G & Celotti, A 1992 *ApJ Lett.* **397** L5
[13] Aharonian, F *et al.* 2009 *ApJ Lett.* **696** L150
[14] Abdo, A A *et al.* 2010 *ApJ* **716** 30
[15] Scargle, J D 1982 *ApJ* **263** 835
[16] Condon, J J 1997 *PASP* **109** 166
[17] Marscher, A P & Gear, W K 1985 *ApJ* **298** 114
[18] Marscher, A P *et al.* 2008 *Nature* **452** 966
[19] Bl&ford, R D & Königl, A 1979 *ApJ* **232** 34
[20] Rani, B & Wiita, P J & Gupta, A C 2009 *ApJ* **696** 2170
[21] Gupta, A C & Srivastava, A K & Wiita, P J 2009 *ApJ* **690** 216

Optical Observations of the Be/X-ray Binary A0538-66

A F Rajoelimanana[1], P A Charles[2,3], P J Meintjes[1] and L J Townsend[3]

[1]Department of Physics, University of the Free State, PO Box 339, Bloemfontein 9300, SA
[2]School of Physics and Astronomy, Southampton University, Southampton SO17 1BJ, UK
[3]ACGC, Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, SA

E-mail: rajoelimananaa@ufs.ac.za

Abstract.

We present recent results on the study of long-term photometric and spectroscopic properties of the Be/X-ray binary system A0538-66, using photometric data from the MACHO and OGLE-IV archives and spectroscopic observations from the Southern African Large Telescope (SALT). The previously reported 421 d superorbital modulation in the early MACHO data (1992–1996) appears to have shortened in the second half of the data (1996–2000). The OGLE-IV data (2010–present) shows a long-term variation but on a shorter timescale (around 250 d) and smaller amplitude than observed in the MACHO data. The spectroscopic observations from SALT enable us to probe both the evolution of the Be star envelope which can be traced by the Hα profiles and the details of its interaction with the neutron star in its presumed highly eccentric (e$\sim$0.7) orbit. The high-resolution spectra allow us to derive a better spectral classification. In addition, we derive a refined orbital period and ephemeris from the recent OGLE-IV light curves.

1. Introduction

Be/X-ray binaries (BeX) represent the largest sub-class of high-mass X-ray binary, which consists of a neutron star (usually an X-ray pulsar) orbiting an early-type non-supergiant Be star in a wide ($P_{orb} \sim 10 - 300$ days) and eccentric ($0.1 < e < 0.9$) orbit. The neutron star can interact and accrete from the Be equatorial disk during periastron passage, giving rise to periodic outbursts that may be observed over a wide range of wavelengths (e.g. X-ray, optical, IR).

The periodic recurrent transient X-ray source, A0538-66, was discovered in 1977 when two X-ray outbursts, separated by $\sim$ 17 d, were observed with the *Ariel V* satellite [1]. Further outbursts were observed with the *HEAO-1* modulator collimator which was found to recur every 16.66 d [2]. An improved *HEAO-1* position led to the identification of its optical counterpart as a bright (B$\sim$15) Be star in the Large Magellanic Cloud (LMC)[3, 4]. This implies that its peak X-ray luminosity was $\sim 10^{39}$erg s^{-1}, substantially super-Eddington for a 1.4 $M_\odot$ neutron star, and still the highest in its class. The discovery of the 69 ms X-ray pulsation in one of the X-ray outbursts observed with the Einstein Observatory indicates that the compact object is a neutron star [5] .

The optical counterpart was classified as a B2 III-V star and was found to reach 12$^{\rm th}$ mag during X-ray maximum [6]. Furthermore, this optical brightness increases are accompanied by the appearance of He II λ4686 emission [6, 7]. These huge outbursts had diminished by late 1983, and none have been observed since. However, much smaller optical outbursts (few tenths of a magnitude) were seen in both optical and X-ray observations and appeared to recur every 16.65 d which was interpreted as the orbital period of the system [8, 9].

The existence of dedicated telescopes to regularly observe the LMC, such as the MACHO project, has helped in understanding the long-term behaviour of the source. The early ($\sim$5 years) MACHO light curve displays large amplitude variability on a timescale of 421 d [10]. Within this 421 d modulation, the much shorter timescale, 16.6515 d outbursts are clearly visible, but remarkably they are only seen at certain phases of the long-term 421 d cycle. They occur during optical minimum, but not when the star reaches its maximum (quiescent) brightness, which led to the suggestion that these long-term superorbital variations were a result of the formation and destruction (or evaporation) of the circumstellar disc around the Be star [10, 11, 12].

2. Optical light curves from the MACHO and OGLE projects

The MACHO project has regularly monitored photometrically the LMC from 1992 to 2000. The data were taken simultaneously in two wide passbands, a blue band (4500 – 6300 Å) and a red band (6300 – 7600 Å). The complete MACHO light curve (1993-2000) is publicly available at the MACHO website[1] and the data are given in instrumental magnitudes. The field of A0538-66 also lies within the field regularly monitored by the OGLE-IV survey in which the source is identified as the OGLE object LMC 518.26.21298. Its real-time OGLE-IV light curve is available at the X-ray variables OGLE monitoring (XROM) website[2] [13]. Unfortunately, A0538-66 is not covered by the OGLE-II and OGLE-III projects.

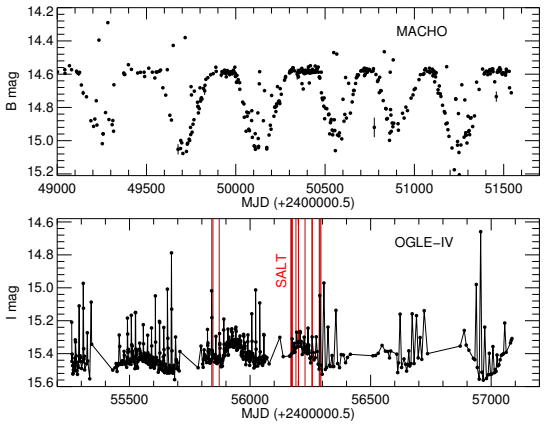


Figure 1. MACHO (top) and OGLE-IV (bottom) light curves of A0538-66. Vertical lines denote times of SALT/RSS observations

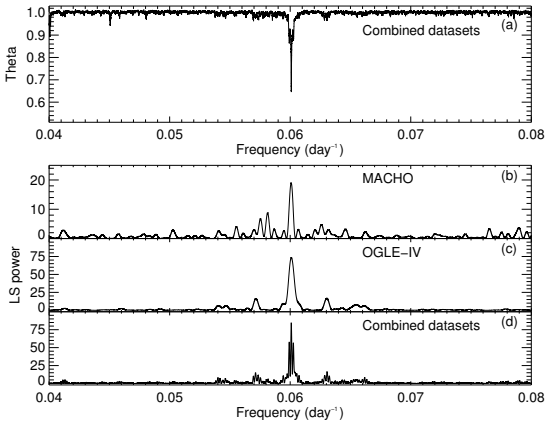


Figure 2. Power spectrum (PDM and LS) of the MACHO (b), OGLE-IV (c) and combined (a) and (d) datasets, showing a peak at 16.6409 d.

The complete MACHO light curve of A0538-66 displays large amplitude long-term variations of $\sim$0.5 mag as reported by [10]. From the complete light curve it appears that the period has slightly shortened in the second half of the data (see Figure 1). We performed Lomb-Scargle periodogram analysis of the two halves of the MACHO data and found the peak period to have

[1] http://wwwmacho.anu.edu.au
[2] http://ogle.astrouw.edu.pl/ogle4/xrom/xrom.html

changed from 439.75 d to 350.5 d (with uncertainties of ∼1 d). The power spectrum of the complete MACHO data shows a large peak at a period of 391.38 d which is slightly shorter than the previously reported 421 d superorbital period. The OGLE-IV light curve also shows a long-term variation but of a smaller amplitude and shorter timescale (∼ 250 days) than that seen in the MACHO light curve (see Figure 1).

In the MACHO data, the amplitude of the regular outbursts is very much dependent on the brightness level, they do not occur at optical maxima and are very strong (∼0.7 mag) at optical minima. This tells us that at optical maxima the equatorial disk is completely dissipated and we only see the naked B star. However, in the OGLE-IV data the 16.65 d regular orbital outbursts are always present, even though the source is at its superorbital maximum. The amplitude of these outbursts clearly varies through the long-term cycle. They are very weak at optical maxima, as the Be star has a smaller size disk during that time, and becomes stronger as the circumstellar disc grows (optical fades).

We have removed the long-term variations from the light curves by splitting them into several segments, the periodic outbursts events were first removed by sigma-clipping the bright points for each segment and then we fitted and subtracted a low-order polynomial from the data. We have run Lomb-Scargle (L-S) periodogram and phase dispersion minimization (PDM) on both individual and combined datasets over the period range ∼14–20 d with a resolution of 10^{-6} cycle d^{-1}. The power spectrum of the detrended MACHO, OGLE-IV and combined light curves (see Figure 2) shows significant peaks at 16.6434±0.0024, 16.6398±0.0026 d and 16.6409± 0.0003 d, respectively, which are slightly shorter than the previously reported values by [10]. These three periods are all consistent within 1-sigma (68% confidence) level. We note that the stated uncertainties on the periods are 1-sigma uncertainties and was calculated by using a Monte-Carlo simulation with 10 000 iterations.

Figure 3 represents the light curves folded on the period of 16.6409 d, showing not only one but two peaks in each orbital cycle. This might be caused by a misalignment between the orbital plane of the neutron star and the spin axis of the Be circumstellar disc. If we have a misaligned system, the neutron star interacts twice with the Be circumstellar disc every orbital cycle. From the folded light curves, we estimate the ephemeris for maximum light to be:

$$\mathrm{HJD_{outburst}} = 2455674.48 \pm 0.03 + \mathrm{n} * 16.6409 \pm 0.0003$$

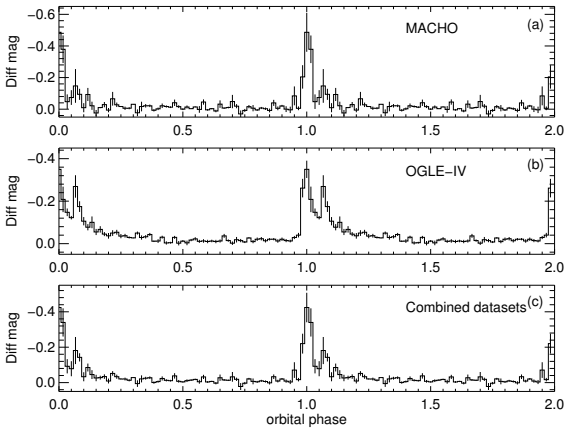


Figure 3. MACHO (a), OGLE-IV (b) and combined (c) light curves folded on the orbital period of 16.6409 d.

3. Optical spectroscopy with SALT

3.1. Broad-band spectra and Hα line profiles evolution

We have also obtained long-slit spectroscopic observations of A0538-66 with the Robert Stobie Spectrograph (RSS) [14, 15] on the Southern African Large Telescope (SALT) [16, 17] at SAAO, Sutherland. Broad-band spectra with a wavelength range of approximately $\lambda\lambda$3100–9000 and a dispersion of 1.0 Å pixel^{-1} were obtained by using two settings of the 900 lines mm^{-1} grating (PG0900). An example of broad-band spectra taken near optical maxima (smaller or no disk) and during the Be phase (with a disk present) are shown in figure 4.

Near the optical maximum (27–28 Aug 2012), the optical spectrum is dominated by the photospheric spectrum of the underlying B star in which all Balmer lines are in absorption. This is a signature of a smaller size or fully dissipated disk as the Balmer emission lines are formed in the circumstellar disk around the Be star. However, as the disk forms the Balmer (and sometimes He I) emission lines start to reappear. Figure 5 shows the time evolution of the Hα line profiles of A0538-66 which are highly variable both in terms of shape and strength. The profiles have various shapes such as single-peaked (07 Oct 2011), double-peaked (e.g. 01 Jan 2013), pure absorption (e.g. 27 Oct 2012) and shell profiles (e.g. 30 Aug 2012). Furthermore, it clearly shows V/R variations which are long-term changes in the relative strength of the violet (V) and red (R) components of the observed Hα emission profiles.

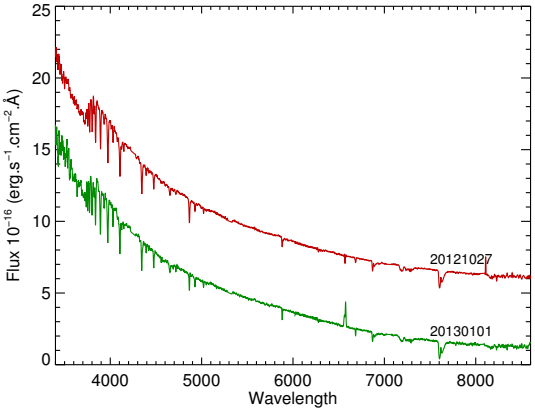


Figure 4. Broad-band spectra of A0538-66 taken near the optical maxima (20121027) and when the star is in a Be phase (20130101).

3.2. Spectral classifications

Figure 6 shows the average high-resolution blue spectrum ($\lambda\lambda$4000–4700) of A0538-66 obtained with the SALT/RSS in which all main spectral features have been identified. The spectra are mostly dominated by Balmer and neutral helium (He I) lines. The absence of the HeII 4686 line indicates a spectral type later than B0.5. The relative weakness of the Mg II 4481 line suggests a spectral type earlier than B2. The strength of the absorption lines of the C III + O II blends at 4070 Å and 4650 Å also indicates a B1 spectral-type source.

For luminosity classification, the ratio between the Si III 4552 and the He I 4387 was used which suggests a luminosity III (giant) star. The strength of the OII lines at 4070(blend), 4415 and 4640(blend) also indicates a luminosity class of III. The strength of these OII lines is rising with increasing luminosity. We conclude, therefore, that the optical counterpart of A0538-66 is a B1e III star which is slightly earlier spectral type than the previously reported (B2 III-V). As a comparison the spectrum of a standard B1 III (HD 147165) and B1 V (HD 144470) from the Walborn & Fitzpatrick spectral atlas [18] were plotted in Figure 7.

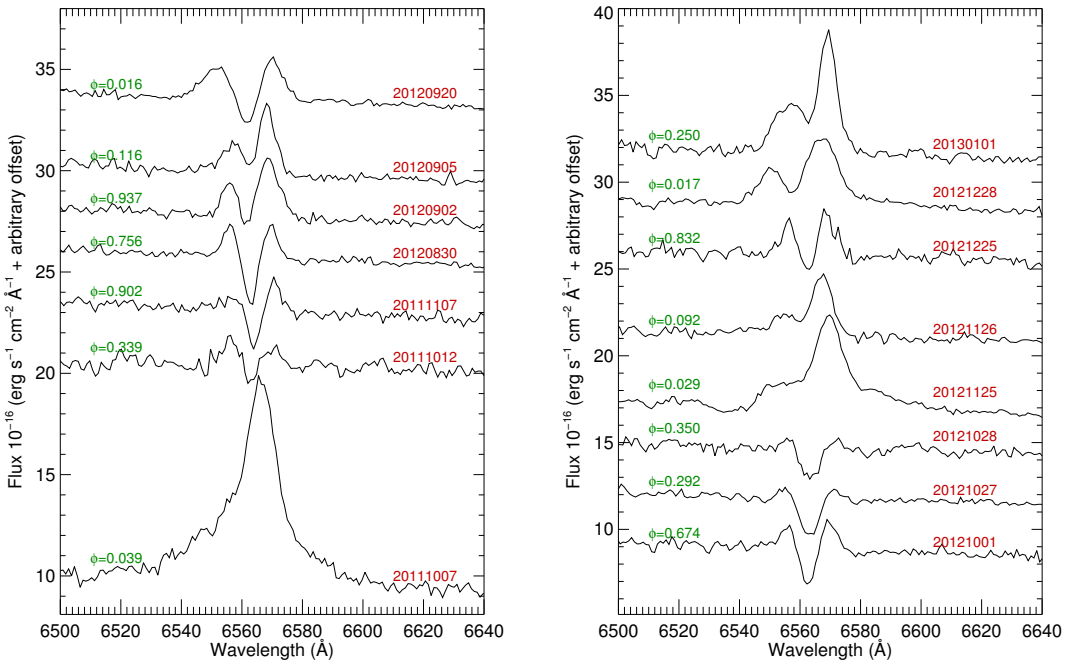


Figure 5. SALT/RSS spectra of the Be/X-ray system A0538-66 showing the evolution of its Hα emission line profiles.

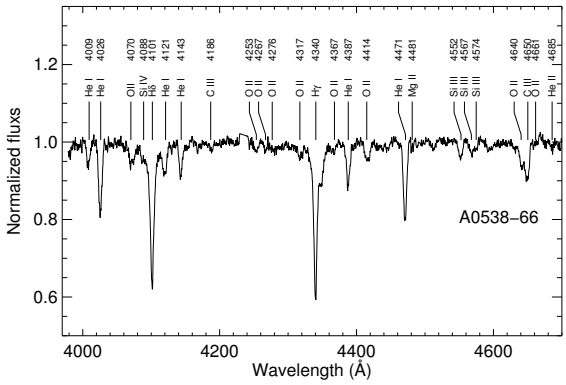


Figure 6. High-resolution spectra of A0538-66 in which all identified lines are also marked.

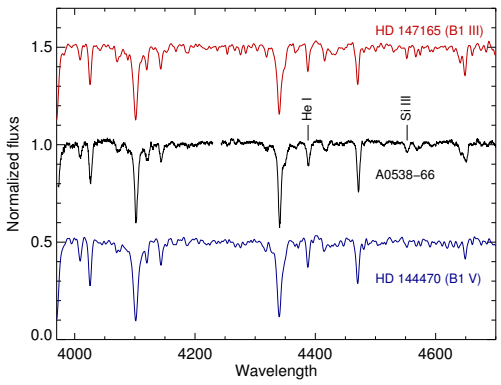


Figure 7. Blue spectra of A0538-66. The spectra of the standard B1 III (HD 147165) and B1 V (HD 144470) are also plotted

4. Conclusion

We have presented the long-term photometric and spectroscopic behavior of the LMC BeX source A0538-66. The previously reported 421 d superorbital modulation has evolved to 350 day in the second half of the MACHO observations and has shortened to $\sim$ 250 d in the OGLE-IV data. A refined orbital period of 16.6409$\pm$ 0.0003 d was derived from data including the most recent OGLE-IV light curve. The folded light curve on this period has a double-peaked profile which we suggest as arising from the misalignment between the Be disk and the orbital plane of the binary system. We have also revised its outburst ephemeris to $2455674.48 \pm 0.03 + n * 16.6409 \pm 0.0003$. The evolution of the H$\alpha$ line profiles in our spectra imply that the size of the equatorial disk

around the Be star varies over the years. We refined the spectral classification of its optical counterpart to B1e III which is earlier than the spectral type found in the literature, namely B2 III-V. Further optical and X-ray monitoring are planned with finer timing around periastron.

References

[1] White N E and Carpenter G F 1978 *MNRAS* **183** 11–5
[2] Johnston M D, Bradt H V, Doxsey R E, Griffiths R E, Schwartz D A and Schwarz J 1979 *ApJ* **230** 11–4
[3] Johnston M D, Griffiths R E and Ward M J 1980 *Natur* **285** 26
[4] Pakull M and Parmar A 1981 *A&A* **102** L1–L4
[5] Skinner G K, Bedford D K, Elsner R F, Leahy D, Weisskopf M C and Grindlay J 1982 *Natur* **297** 568–570
[6] Charles P A, Booth L, Densham R H, et al 1983 *MNRAS* **202** 657–82
[7] Densham R H, Charles P A, Menzies J W, van der Klis M and van Paradijs J 1983 *MNRAS* **205** 1117–33
[8] Skinner G K 1980 *Nature* **288** 141–143
[9] Skinner G K 1981 *SSRv* **30** 441–446
[10] Alcock C, Allsman R A, Alves D R, et al 2001 *MNRAS* **321** 678–84
[11] McGowan K E and Charles P A 2003 *MNRAS* **339** 748–756
[12] Rajoelimanana A F, Charles P A and Udalski A 2011 *MNRAS* **413** 1600–1622
[13] Udalski A 2008 *AcA* **58** 187–192
[14] Burgh E B, Nordsieck K H, Kobulnicky H A, Williams T B, O'Donoghue D, Smith M P, Percival J W 2003 *SPIE* **4841** 1463–1471
[15] Kobulnicky H A, Nordsieck K H, Burgh E B, Smith M P, Percival J W, Williams T B, O'Donoghue D 2003 *SPIE* **4841** 1634–1644
[16] Buckley D A H, Swart G P, Meiring J G 2006 *SPIE* **6267** 62670Z
[17] O'Donoghue D, Buckley D A H, Balona L A, et al 2006 *MNRAS* **372** 151–162
[18] Walborn N R and Fitzpatrick E L 1990 *PASP* **102** 379

Rotation Curves and Bars: Accounting for Non-circular Motions in Barred Spiral Galaxies

Toky H. Randriamampandry[1], Claude Carignan[1], Francoise Combes[2] and Nathan Deg[1]

[1]Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa
[2]LERMA, Observatoire de Paris, College de France, PSL, CNRS, Sorbonne Univ., UPMC, F-75014, Paris, France

E-mail: tokyr@ast.uct.ac.za

Abstract. The rotational velocities derived from the gas are often used to study the distribution of the luminous and the dark matter content of disk galaxies by assuming circular gas motions and that disk galaxies are axisymmetric. Unfortunately, this is not the case for barred galaxies since the observed velocities are the combination of an azimuthally averaged circular velocity and the non-circular flows induced by the gas streaming along the bar. The tilted-ring method implemented in ROTCUR is commonly adopted when deriving rotation curves of galaxies, but ROTCUR overestimates/underestimates the rotation velocities when the bar is perpendicular/parallel to the major axis of the disk. The DiskFit algorithm is specifically designed for barred galaxies. Unfortunately, when the bar angle ϕ_{bar} is close to 0 or 90 degrees DiskFit fails due to the degeneracy of the velocity components. We propose a method of correcting rotations based on numerical simulations of galaxies. We correct the rotation curve derived from the tilted-ring method based on a numerical simulation of a galaxy with similar properties and projections as the observed galaxy. NGC 3319 has a bar which is almost aligned with the major axis and it is used as a case study. The correction is obtained from simulations and applied to the observed rotation curve of this particular galaxy. The mass model analysis shows that the two rotation curves give different results for models with either a dark matter halo or using MOND. This shows the importance of correcting for non-circular motions when performing mass model analysis.

Keywords : Cosmology: dark matter – galaxies: kinematics and dynamics – structure – spiral

1. Introduction

The distribution of mass in an axisymmetric galaxy is well constrained by the rotational velocities derived from the gas using the following assumptions: that the gas is moving in circular orbit and the galaxy is axisymmetric. This is largely the case for spiral galaxies, given the very low velocity dispersion in the gas. However, for barred galaxies, part of the observed velocities are due to non-circular motions induced by the gas streaming along the bar (gas moving along the bar).

The tilted ring method [15] is commonly used to derive the rotation curve of galaxies. This method divides galaxies into rings and estimates the velocities and other kinematical parameters for each ring. It has been implemented in the GIPSY [1] task ROTCUR [3]. Accurate modeling has shown that the mass estimates of barred spiral galaxies may be wrong by up to 20% depending on the orientation of the bar relative to the major axis of the galaxy [6] if the tilted ring method is used to derive the rotation.

The publicly available code DiskFit [16] is specifically designed for barred galaxies. It applies a bi-symmetric flow model with fixed phase in the disk plane to the data. DiskFit does not work when the bar is almost parallel to the major or minor axis because of the degeneracy of the velocity components [14].

We propose a method of correcting rotation curves using results from numerical simulations of galaxies. The corrections are obtained from simulation with similar properties and projections as the observed galaxy. In this paper, we use a snapshot with similar properties to NGC 3319. The expected rotation curve (due to the different components:gas, stars and dark matter in the simulation) is calculated from the gravitational potential. The expected circular velocity is given by:

$$< V^2_{expected} >= r < \frac{\partial \Phi}{\partial r} >= r < F_r > \quad (1)$$

where F_r is the radial force from the particles calculated azimuthally on a grid and Φ is the gravitational potential.

The expected rotation curve $V_{expected}$ is then compared with the rotation curve obtained from mock velocities maps downloaded from the GalMer database using ROTCUR. The correction is the difference between the expected rotation curve and the ROTCUR rotation curve in the bar region. We refer the reader to [12] for more details.

NGC 3319 is chosen as a case study. NGC 3319 has a bar which is almost parallel with the major axis (Φ_{bar} = 2 degrees). Mass models using the observed rotation curve and the corrected rotation curve of NGC 3319 are performed to show the effect of the non-circular motion on the results.

The simulation is explained in section 2, and the method and a case study in section 3. Finally the mass model results are discussed in section 4.

2. The simulation

Simulations from the GalMer project [4] are used in this work. The purpose of the GalMer project was to create a database for merging and isolated galaxies using high resolution N-body/hydrodynamic simulations. The GalMer database consists of a large set of Tree-SPH simulation of galaxies spanning a wide range of morphological types and sizes. The snapshot of the giant Sb model (gSb) at the epoch T=500 Myrs is used in this case study. The gSb model was chosen because it has similar bar properties and buge-to-disk (B/D) ratio with NGC 3319. The GalMer database uses a go-on-the-fly algorithm which enables the user to obtain the masses, the velocity maps and other properties for every snapshot. The radial velocities maps are obtained by projecting the velocity components of each particle into a line-of sight (LOS) radial velocity. The pixel size depends on the zoom that is chosen by the user on the online tool (see [4] for details). The LOS radial velocity is

$$V_r = V_X cos\phi cos\theta + V_r sin\phi sin\theta + V_r cos\theta \,, \quad (2)$$

where V_X, V_Y and V_Z are the velocity components for each particle, ϕ and θ are the azimuthal and polar angles respectively The GalMer database does not provide the observed uncertainties for the mock velocity fields. The velocity maps and other properties are publicly available from the GalMer database (http://galmer.obspm.fr).

3. Case study: NGC 3319

NGC 3319 is a barred spiral galaxy located at a distance of 14.3 Mpc. The 21 cm neutral atomic hydrogen (HI) observations were made with the C and D configurations of the Very Large Array (VLA). The data was retrieved from the archive and reduced using the standard AIPS package. More details about the observations and data reductions can also be seen in [9]. Previous studies (e.g [6]) have shown that the inner part of the rotation curve could be underestimated if the bar is aligned with the major axis. Therefore, the observed rotation curve derived from ROTCUR need to be corrected for the non-circular motions before being used for mass model analysis.

The correction was estimated from the simulation by calculating the difference between the expected rotation curve from the gravitational potential and the rotation curve obtained using the tilted ring method applied to a mock velocity field of the simulation downloaded from the GalMer database

(http://galmer.obspm.fr). The snapshot with bulge-to-disk ratio (B/D) and bar properties similar to NGC 3319 was selected for the correction. A B/D=0.24 is found for NGC 3319 from the 3.6 μm surface luminosity profile which is comparable with B/D=0.25 (total mass ratio) for the gSb model. The obtained correction was then scaled in terms of velocity using the V_{max} ratio and in terms of radius using the ratio between the disk scale lengths of the model and NGC 3319 (for more details see [12]) before it is added to the observed rotation curve of NGC 3319.

3.1. Mass model

The mass model is done by comparing the observed rotation curves with the quadratic sum of the luminous and dark matter components using a chi-squared minimization technique. The mass models were carried out using the GIPSY tasks ROTMAS and ROTMOD.

For a mass model with dark matter component, the rotation curve is

$$\mathrm{V}^2_{\mathrm{rot}} = \mathrm{V}^2_{\mathrm{gas}} + \mathrm{V}^2_{*,\mathrm{disk}} + \mathrm{V}^2_{*,\mathrm{bulge}} + \mathrm{V}^2_{\mathrm{halo}} \,, \tag{3}$$

where V_{gas} is the gas contribution, $V_{*,disk}$ the stellar disk contribution, $V_{*,bulge}$ the stellar bulge contribution and V_{halo} is the contribution of the dark matter component.

3.1.1. Gas and stellar contributions The gas contribution was derived from the gas intensity map and the stellar contribution from the IRAC1 3.6 micron surface brightness profile. The method described in [11] was used to convert the luminosity profile into density profile assuming a constant mass-to-light ratio. The result from [5] obtained using chemo-spectrophotometric galactic evolution (CSPE) models are used for the disk and bulge mass-to-light ratios.

A bulge-disk decomposition was performed to obtain the contribution from the bulge to the rotation curve.

3.1.2. Dark matter contributions The dark matter component is calculated from a theoretical density profile. Here, we only use the pseudo-isothermal sphere (ISO) dark matter profile. The ISO density profile is given as

$$\rho_{\mathrm{ISO}}(\mathrm{R}) = \frac{\rho_0}{1 + \left(\frac{\mathrm{R}}{\mathrm{R_c}}\right)^2} \,, \tag{4}$$

where ρ_0 is the central density and R_c the core radius of the halo. The velocity contribution of a ISO halo is given by

$$\mathrm{V}_{\mathrm{ISO}}(\mathrm{R}) = \sqrt{4\pi\, G\, \rho_0\, R_c^2 \left(1 - \frac{R}{R_c}\, atan\left(\frac{R}{R_c}\right)\right)} \,, \tag{5}$$

3.1.3. Mass model with no dark matter component The Modified Newtonian Dynamic or MOND is also considered. Milgrom stipulated that there is no need for DM when fitting galaxies rotation curve if the Newtonian gravity is modified below a critical acceleration a_0 [8]. In the MOND framework, the gravitational acceleration of a test particle is given by

$$\mu(x = g/a_0)g = g_N \,, \tag{6}$$

where g is the gravitational acceleration, μ(x) is the standard MOND interpolating function and g_N the Newtonian acceleration (see [8] for more details). For the MOND mass model, the rotation curve is

$$\mathrm{V}^2_{\mathrm{rot}} = \mathrm{V}^2_{\mathrm{gas}} + \mathrm{V}^2_{*,\mathrm{disk}} + \mathrm{V}^2_{*,\mathrm{bulge}} \,. \tag{7}$$

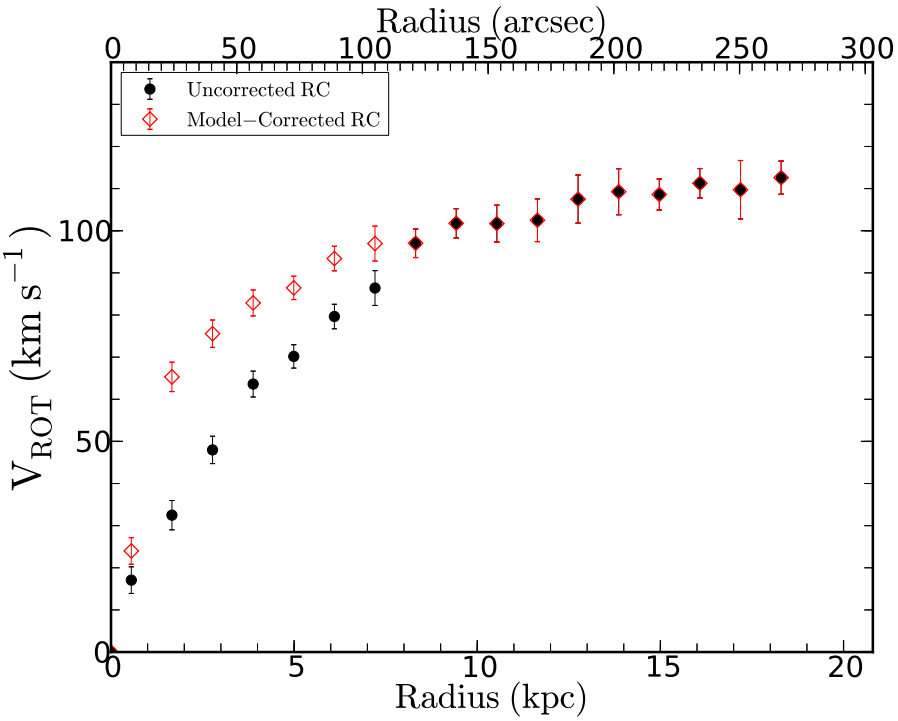


Figure 1. Comparison between the model-corrected and uncorrected rotation curves.

Table 1. Results for the ISO dark matter halo and MOND models using the uncorrected and corrected rotation curves (RCs).

Halo Model	Parameters	Uncorrected RC	Corrected RC
ISO	ρ_0 $^{(a)}$	5.5±0.9	74.1±15.1
	R_c $^{(b)}$	6.9±1.1	1.3±0.1
	Υ_d $^{(c)}$	0.2	0.2
	Υ_b $^{(d)}$	0.4	0.4
	χ^2_r	1.3	0.7
MOND	a_0 $^{(e)}$	0.5±0.1	0.7±0.1
	Υ_d $^{(c)}$	0.2	0.2
	Υ_b $^{(d)}$	0.4	0.4
	χ^2_r	3.2	5.3

Notes: the central dark matter density $^{(a)}\rho_0$ is given in units of 10^{-3} $M_\odot$ pc^{-3}, $^{(b)}$ R_c is core radius in kpc , $^{(c)}\Upsilon_d$ is the mass-to-light ratio for the stellar disk, $^{(d)}\Upsilon_b$ for the bulge and $^{(e)}a_0$ in 10^{-8} cm s^{-2}.

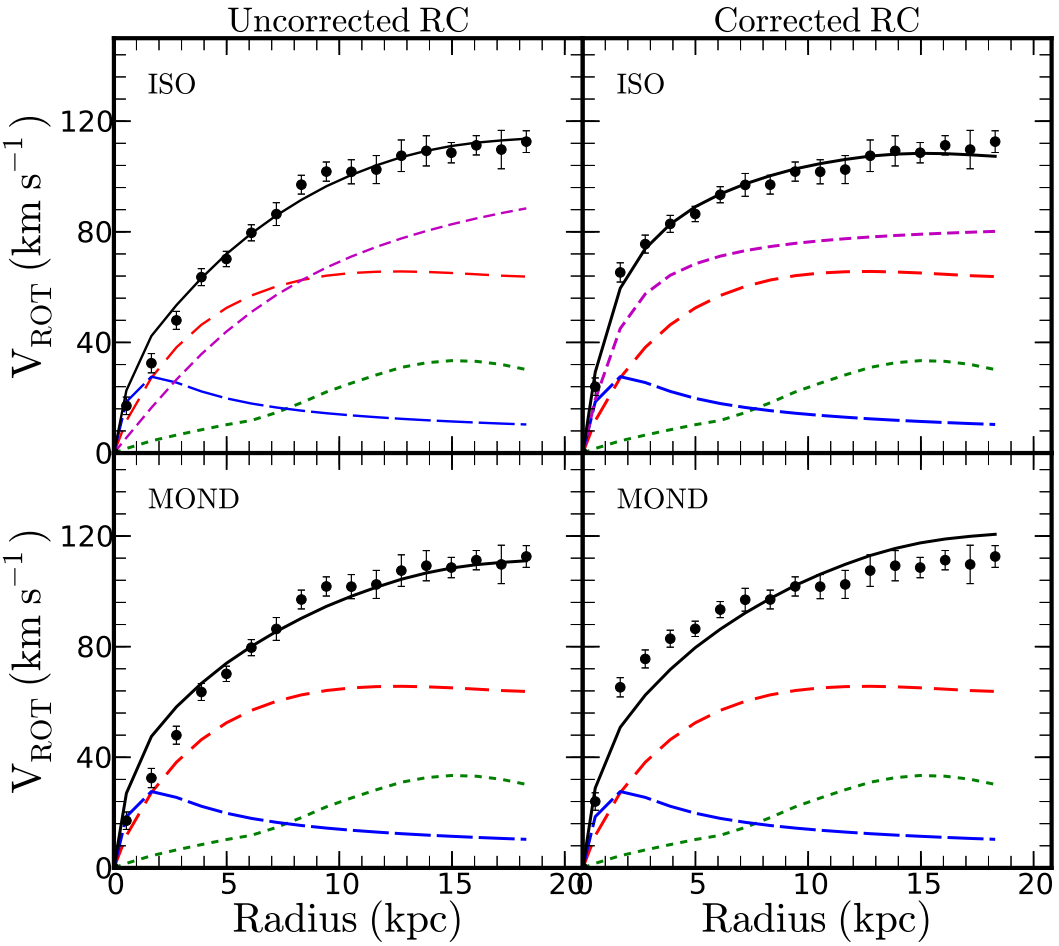


Figure 2. Mass models using ISO DM halo model (top) and MOND model (bottom) for the uncorrected and model-corrected rotation curves. The observed rotation curve is shown as filled black circles with error bars, the gas contribution as dashed green lines, the stellar disk contribution as dashed red lines, the stellar bulge contribution as long-dashed blue lines and the contribution from the dark matter halo as dashed magenta lines. The continuous lines are the best fit model to the rotation curve.

4. Results and Discussions

A comparison between the uncorrected and corrected rotation curves is given in figure 1, where the black circles are the rotation curve obtained from the tilted ring method and the red diamonds are the rotation curve after it has been corrected for non-circular motions.

The mass model results for the uncorrected and corrected rotation curves are presented in figure 2 for the ISO and MOND models and the results are summarized in table 1. For the ISO models, the dark matter halo is much more centrally concentrated when the corrected rotation curve is used, with a $R_C \sim 1$ kpc compared to ~ 5 kpc for the uncorrected rotation curve and a larger ρ_0. It is clear that the corrected and uncorrected rotation curve give very different results for the dark matter parameters. This shows the importance of the non-circular motions on the observed rotation curve and the inferred dark matter properties. The inner part of the rotation curve is under-estimated when the bar is parallel to the major axis, therefore a portion of the mass will be missed if the rotation curve is not corrected and different

conclusions could be drawn from the mass model.

MOND (with constant M/L's) cannot reproduce either the corrected or uncorrected rotation curves. However, even in this case, there is a difference between the inferred a_0's. It is important to note that we are not comparing the goodness-of-fit of the MOND and ISO models. Rather, we are comparing the inferred model parameters using either the corrected or uncorrected rotation curves. The results are significantly different in both cases. This illustrates both the necessity of properly corrected rotation curves for non-circular motions and the promise of using simulations to obtain these corrections for bar-induced motions.

We will be running new simulations of NGC 3319, which will allow us investigated the effect of the orientation of the bar on the observed rotation curve of this galaxy.

Acknowledgments

We would like to thanks the anonymous referees for their constructive comments which have greatly improved the quality of this paper. CC's work is based upon research supported by the South African Research Chairs Initiative (SARChI) of the Department of Science and Technology (DST), the SKA SA and the National Research Foundation (NRF); ND and TR's work are supported by a SARChI South African SKA Fellowship.

References

[1] Allen, R. J., Ekers, R. D., Terlouw, J. P., & Vogelaar, M. G. R. 2011, Astrophysics Source Code Library, 9018
[2] Barazza, F. D., Jogee, S., & Marinova, I. 2008, Formation and Evolution of Galaxy Disks, 396, 351
[3] Begeman K. G., 1989, A&A, 223, 47
[4] Chilingarian, I. V., Di Matteo, P., Combes, F., Melchior, A.-L., & Semelin, B. 2010, A&A, 518, A61
[5] de Denus-Baillargeon, M.-M., Hernandez, O., Boissier, S., Amram, P., & Carignan, C. 2013, ApJ, 773, 173
[6] Dicaire, I., Carignan, C., Amram, P., et al. 2008, MNRAS, 385, 553
[7] Kuzio de Naray, R., Arsenault, C. A., Spekkens, K., et al. 2012, MNRAS, 427, 2523
[8] Milgrom M., 1983b, ApJ, 270, 365
[9] Moore, E. M., & Gottesman, S. T. 1998, MNRAS, 294, 353
[10] Navarro J. F., Frenk C. S., & White S. D. M. 1996, ApJ, 462, 563
[11] Oh, S.-H., de Blok, W. J. G., Walter, F., Brinks, E., & Kennicutt, R. C., Jr. 2008, AJ, 136, 2761
[12] Randriamampandry, T. H., Combes, F., Carignan, C., & Deg, N. 2015, MNRAS, 454, 3743
[13] Schoenmakers, R. H. M., Franx, M., & de Zeeuw, P. T. 1997, MNRAS, 292, 349
[14] Sellwood, J. A., & Sánchez, R. Z. 2010, MNRAS, 404, 1733
[15] Rogstad, D. H., Lockhart, I. A., & Wright, M. C. H. 1974, ApJ, 193, 309
[16] Spekkens, K., & Sellwood, J. A. 2007, ApJ, 664, 204

Studying stellar populations of luminous red galaxies to probe the Hubble parameter $H(z)$

A. L. Ratsimbazafy[1], C. M. Cress[2,3], S. M. Crawford[4], M. Smith[5]

[1]Centre for Space Research, North-West University, Potchefstroom 2520, SA
[2]Centre of High Performance Computing, CSIR, 15 Lower Hope St, Rosebank, Cape Town 7700, SA
[3]Physics Department, University of the Western Cape, Private Bag X17, Cape Town 7535, SA
[4]South African Astronomical Observatory, PO Box 9 Observatory, Cape Town 7935, SA
[5]School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK

E-mail: Ando.Ratsimbazafy@nwu.ac.za

Abstract. There have been a number of attempts to measure the expansion rate of the Universe using age–dating of Luminous Red Galaxies (LRGs). Assuming that stars in LRGs form at the same time, age–dating of two populations of LRGs at different redshifts can provide an estimate of the time difference associated with the corresponding redshift interval (dz/dt). This gives a direct estimate of the Hubble parameter $H(z)$ at the average redshift of the two populations. We explore the validity of this method by using two different sets of data. Firstly, we select a homogeneous sample of passively evolving galaxies over $0.10 < z < 0.40$ from the Sloan Digital Sky Survey Data Release Seven (SDSS-DR7) catalogue by applying a refined criteria, which is based on absolute magnitude. Secondly, we carry out a series of observations on the Southern African Large Telescope (SALT) to obtain spectra of LRGs at two narrow redshift ranges $z \simeq 0.40$ and $z \simeq 0.55$ in order to calculate the Hubble parameter $H(z)$ at $z \simeq 0.47$. We utilise different methods of age-dating including full spectral fitting on high signal-to-noise galaxy spectra from our sample. We present the $H(z)$ estimates and their cosmological constraints using the two different data sets.

1. Introduction

The expansion rate of the Universe can potentially be measured by age–dating Luminous Red Galaxies (LRGs). This technique is well known as *Cosmic Chronometers* (CC), and was originally proposed by [1]. Most cosmological probes only measure the expansion rate integrated along a line-of-sight. However, the CC method allows a measurement at a specific redshift, and this can provide tighter constraints on cosmological parameters. The Hubble parameter is given by

$$H(z) = -\frac{1}{(1+z)}\frac{dz}{dt} \tag{1}$$

where dz/dt can be approximated by determining the time interval Δt corresponding to a given Δz, where Δz is centered at redshift z. The overall method is valid when assuming that if most stars in LRGs formed near the beginning of the Universe at a similar time, then measuring the age difference between ensembles of LRGs at two different redshifts provides the differential quantity $\Delta z/\Delta t$ required to estimate the Hubble parameter $H(z)$. Using a differential quantity

significantly reduces systematics associated with the age determination. Some systematic errors in the absolute age measurements have been a subject of discussion but such errors could be cancelled by using the relative ages of these galaxies, for example in Ref. [1, 2, 3]. Furthermore, only the galaxy evolution that takes place between the redshifts, where the difference is taken, is the most important while using this method.

A number of scientists have attempted to use this method to track the evolution of $H(z)$ up to redshift $z \sim 2$, and these results have been used to put constraints on cosmological parameters, for example in Ref. [2, 3, 4, 5, 6, 7, 8]. In each measurement, the authors assumed that LRGs are massive, passively–evolving elliptical galaxies that are homogeneous populations forming their stars at high redshift, and they fit single–burst equivalent ages to the galaxies.

Here, we present two different data sets including the existing archival data from the Sloan Digital Sky Survey Data Release Seven (SDSS–DR7) and data observed with the Southern African Large Telescope (SALT), $H(z)$ estimates and their cosmological constraints.

2. SDSS–LRGs

We selected a homogeneous sample of passively evolving elliptical galaxies from the SDSS catalogue by using different selection criteria which are based on the rest–frame magnitude and colour. The original sample was in the redshift range $0.10 < z < 0.40$. Further selection was applied to it to create a quiescent sample free from any emission lines and a sample which contains massive galaxies. We divided the redshift range into 15 redshift bins with a step of $\delta z = 0.02$. The full spectrum fitting was chosen to obtain Simple Stellar Population (SSP) equivalent ages and other parameters. We used the ULySS[1] (University of Lyon Spectroscopic analysis Software) package for the fits. ULySS [9] fits the entire spectrum with a model in the form of a linear combination of non–linear components, corrected for the kinematics and multiplied by a polynomial at the same time. The use of a multiplicative polynomial makes this method insensitive to the effects of the flux calibration uncertainties and the Galactic extinction.

In order to improve the fits, all galaxy spectra within a redshift bin were combined to obtain a high signal–to–noise (S/N) ratio spectrum. We used the SSP models called GALAXEV [10] (BC03). These models were generated using the STELIB library with a resolution of about 3Å FWHM across the whole spectral range 3200 – 9500 Å, the Chabrier Initial Mass function (IMF) with a mass of 0.1 to 100 $M_{\odot}$ and a slope of -1.35 and Padova 1994 isochrones. These models cover ages from 0.1 to 20 Gyr and [Fe/H] from -2.3 to 0.4 dex, consisting of 696 SSPs. The fitting was performed over the whole wavelength coverage of the co–added spectra, and its reliability was checked (with Monte Carlo simulations, χ^2 and convergence maps) before validating the SSP results.

Due to the lack of the number of galaxies to be combined to form the high S/N ratio spectra, only ages at high redshift $z > 0.20$ were used to measure $H(z)$. From establishing the age–redshift relation, we adopted the linear fitting procedure to fit ages in a specific redshift bin $t(z_i)$. The slope of the fit was therefore the value of dz/dt which is related to the Hubble parameter at an effective redshift $z_{eff} = (z_{max} + z_{min})/2$, according to the Equation 1. Precisely,

$$H(z_{eff}) = -\frac{1}{(1 + z_{eff})} \left(\frac{dt}{dz}\right)^{-1} \tag{2}$$

The redshift difference Δz should not be too wide or too narrow in order to preserve the same number of galaxies in each redshift bin, to maximise the number of $H(z)$ measurements and to minimise the relative error bars. After testing different Δz (from lowest 0.8 to highest 0.16), the optimised Δz was set to be 0.14 with which all ages were used and the number of ages were the same. Therefore, 8 age data points were used in each subsample, leading to 3 values of $H(z)$ at

[1] Available at http://ulyss.univ-lyon1.fr/

the following effective redshifts $z \simeq 0.32$, $z \simeq 0.30$ and $z \simeq 0.28$. The corresponding error on the $H(z)$ was obtained by using the error propagation technique. The stacking technique provides us less than 10% uncertainties in age estimates which leads to better $H(z)$ estimates compared with the values found in the literature.

The $H(z)$ results match the ΛCDM cosmology, see Figure 1. When testing different SSP models, we found some model dependence on the derived ages, hence different $H(z)$ measurements. The sensitivity and the different ingredients of models led to some changes on $H(z)$ estimates. This is one of the sources of systematic errors when adopting the CC technique. A further investigation on this is in progress. The possibility of the contribution of the young stellar populations in the relative age determination was also tested. We fitted two stellar components; one young (from 0.1 Gyr to 2 Gyr) and one old (from 2 Gyr – age of the Universe) and found that the fraction of the young population is negligible.

The comparison plot between our measurements and the other observational Hubble parameters $H(z)$ is shown in Figure 1. Data sets from Ref. [3, 4, 5, 6, 7, 11] are included to check the $H(z)$ evolution up to $z \sim 2$. Note that all $H(z)$ shown in this plot are obtained by applying the same technique of CC, except the results from [11] which are from the radial Baryonic Acoustic Oscillations (BAO) method. These data points are included in our analysis in order to compare our results on cosmological constraints (see section 4) with the previous works.

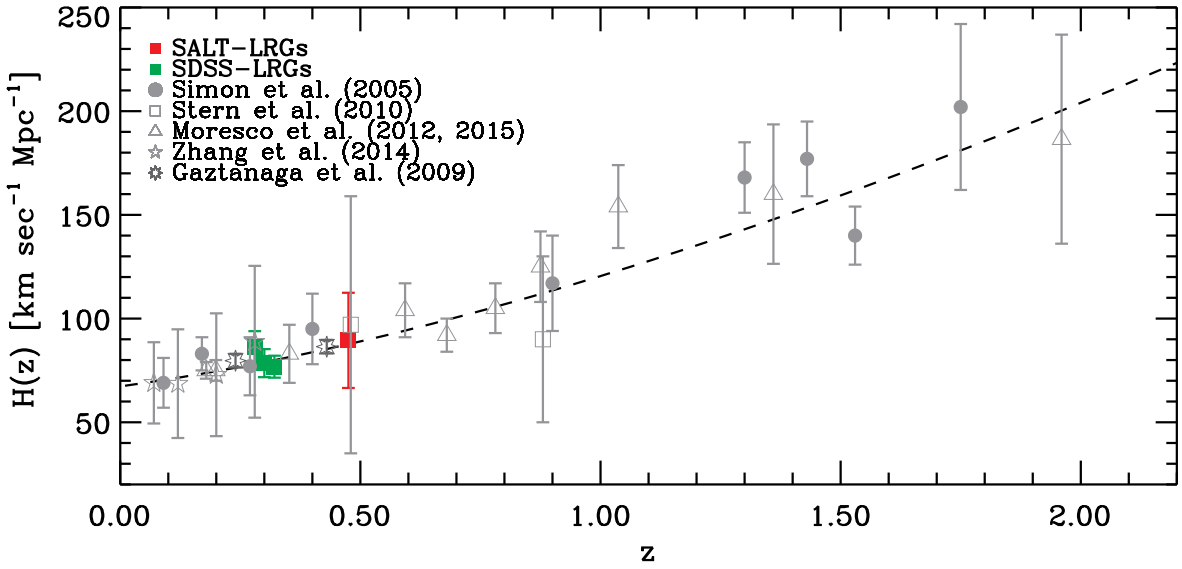


Figure 1. All available observational $H(z)$ data points. The green points are our measurements using SDSS–LRGs and the red point for using SALT–LRGs. The dashed line is the theoretical $H(z)$ of a flat ΛCDM cosmology model by *PLANCK* assuming $H_0 = 67.3$ km s^{-1} Mpc^{-1} and $\Omega_m = 0.315$. For the SDSS–LRGs : $H(z \simeq 0.32) = 86.3 \pm 7.6$ km s^{-1} Mpc^{-1}, $H(z \simeq 0.30) = 78.5 \pm 6.8$ km s^{-1} Mpc^{-1} and $H(z \simeq 0.28) = 76.8 \pm 5.3$ km s^{-1} Mpc^{-1} were estimated when fitting a straight line on the ages over $0.24 < z < 0.40$, $0.22 < z < 0.38$, and $0.20 < z < 0.36$, respectively. For the SALT–LRGs: $H(z \simeq 0.47) = 89 \pm 23$ km s^{-1} Mpc^{-1} was measured using the differential ages.

3. SALT–LRGs

We [12] found that a 3% measurement would be viable from a large redshift program targeting LRGs, and explored optimal observational set–ups for the experiment using the SALT telescope. We then obtained, reduced and analysed 16 long–slit spectra of LRGs recently observed with the SALT telescope. These galaxies were selected from the 2dF–SDSS LRG and QSO (2SLAQ) and MegaZ–LRGs catalogues at redshift $z \simeq 0.40$ and $z \simeq 0.55$. Our selection was based on stellar mass, brightness and emission lines of a galaxy in order to have a sample of old and massive passively–evolving galaxies. For all observations, the basic reductions were already performed by the semi–automated code PySALT[2] [13]. Further reduction was performed by following the standard long–slit data reduction technique with IRAF[3]. All spectra were corrected for foreground Galactic extinction.

Although in our previous work [12] we did find that LRGs may be better described by slightly extended star formation histories, we only fitted the LRGs to SSPs in order to compare this work to previous studies. We performed the full spectrum fitting after matching the resolution of both model and observed spectra. The reliability of the fits were also checked carefully. Despite the number of galaxies at $z \simeq 0.55$ (only 6 as compared to 10 at $z \simeq 0.40$), we found an age–redshift relation, that is, the mean age at $z \simeq 0.40$ is older (3.88 ± 0.20 Gyr) than that at $z \simeq 0.55$ (2.80 ± 0.18 Gyr). We found younger ages than the total sample of red galaxies in [4] at both redshifts but we note that we have a small number of galaxies in our sample. The error on the mean age was obtained by applying the standard error propagation technique.

The mean age of the sample at each redshift was used to measure the differential ages. Applying Equation 1 to the intermediate redshift between $z \simeq 0.40$ and $z \simeq 0.55$, we obtained a new observational Hubble parameter $H(z \simeq 0.47) = 89 \pm 23$ km s^{-1} Mpc^{-1} using SSP equivalent ages (see Figure 1). The most comparable measurement at the same redshift is by [4], who measured a value of $H(z) = 97 \pm 60$ km s^{-1} Mpc^{-1} at $z = 0.48$. Our value is consistent with the standard cosmology model with the parameters $\Omega_m = 0.315$, and $H_0 = 67.3$ km s^{-1} Mpc^{-1}.

We have addressed the estimated number of galaxies required to obtain the desired precision in [12]. In reality, we could not reach the estimated accuracy because of the observation constraints such as the quality of the observed spectra in terms of S/N ratio, or the fact that the number of observed objects could not complete our original estimation. A larger and better quality sample would help us to reach our main goal for the future. The relatively small sample presented here was part of an initial pilot study and further observations will be required to improve on the measurement.

4. Cosmological Constraints

We made use of our $H(z)$ measurements combined with the recent and available $H(z)$ in the literature (all data points plotted in Figure 1) to investigate their constraining power on the determination of cosmological parameters. We also combined those $H(z)$ data with other cosmological measurements: BAO, Wilkinson Microwave Anisotropy Probe (WMAP) and Hubble Space Telescope (HST) to better determine the constraints on the parameter calculations. The first calculation was done with our own code, the second one was performed by using the publicly available code `cosmoMC`. Both calculations were based on the Markov Chain Monte Carlo (MCMC) techniques to determine the cosmological parameters in the standard ΛCDM models, such as the density of matter parameter today Ω_m, the cosmological constant Ω_Λ, the spatial curvature parameter today Ω_k and the Hubble constant today H_0.

[2] pysalt.salt.ac.za

[3] Image Reduction and Analysis Facility, a software system distributed by the National Optical Astronomy Observatories (NOAO). Available at http://iraf.noao.edu/

Our new $H(z)$ measurements impact current constraints on cosmological parameters even though only 4 data points were added to the old ones (27 data points). The 1D marginalised probability distribution of using 31 data points shows more serried distribution than using the old data sets, while they tighten each parameter contour of the confidence level in the 2D marginalisation plots. The best fit parameters are shown in Table 1 for a flat ΛCDM model.

We further combined the above observational $H(z)$ measurements with the other major data sets such as WMAP nine year data (WMAP9), BAO and H_0 HST to investigate their constraints on the determination of the cosmological parameters. The latest results of the Cosmic Microwave Background (CMB) temperature and polarization power spectra from WMAP9 [14] were used as well as the BAO measurement from:

- 6dF Galaxy Redshift Survey (6dFGRS) at z =0.106 [15]
- SDSS–LRGs DR7 sample at $z = 0.35$ [16]
- SDSS–LRGs DR9 sample at $z = 0.57$ [17]
- WiggleZ survey at $z = 0.44$, $z = 0.60$ and $z = 0.73$ [18]

The last data set used was HST measurement of the Hubble parameter today $H_0 = 73.8\pm$ 2.3 km s^{-1}Mpc^{-1} [19]. This value was obtained from the magnitude–redshift relation of 235 Supernovae type Ia (SNIa) observed with HST.

For a flat ΛCDM, we obtained the constraints shown in Table 2. It shows mainly the effect on cosmological parameters of adding four basic data combinations: WMAP9, WMAP9+H_0 HST, WMAP9+BAO and WMAP9+BAO+H_0 HST. Through these results, we can see that the overall joint constraints are consistent with the standard cosmological model. We find that adding the $H(z)$ data sets and H_0 HST shows noticeable constraints on H_0, Ω_m parameters. This is also due to the better constraint on the current Hubble constant H_0. WMAP9+Hz joint analysis does not provide any improvement over the WMAP9 alone since the WMAP9 constraint is much more precise than the $H(z)$ constraint. The WMAP9+BAO data alone only give very weak constraints on Ω_m. However, a combination of WMAP9+BAO with the measurements of Hubble parameter $H(z)$ significantly improves the 1-σ error on both parameters. The parameter values of $H_0 = 69.4 \pm 1.3$ km s^{-1}Mpc^{-1} and $\Omega_m = 0.295^{+0.011}_{-0.012}$ are consistent with results of [8], where only 25 observational $H(z)$ data points were used. More external $H(z)$ data sets (for example from the most recent BAO measurements, *PLANCK* data, latest SNIa data sets) are needed to further investigate on cosmological constraints, and determine different cosmological parameters.

Table 1. Constraints on cosmological parameters from LRG measurements of $H(z)$. Marginalised constraints on H_0, Ω_m, and Ω_Λ at 1-σ for a flat ΛCDM model

Model Parameter	27 data points	31 data points
H_0	$68.9^{+2.7}_{-2.5}$	68.6±2.3
Ω_m	$0.320^{+0.047}_{-0.064}$	$0.321^{+0.048}_{-0.060}$
Ω_Λ	$0.679^{+0.047}_{-0.050}$	$0.675^{+0.060}_{-0.048}$

5. Conclusions

The differential age or *Cosmic Chronometers* technique is an interesting way to measure the expansion rate of the Universe, since it gives a direct measurement between two distinct redshifts. It uses massive and passive ellipticals. We have explored this technique by analysing both existing archival SDSS–DR7 data and new spectra obtained with the SALT telescope. The full spectrum fitting with BC03 models has been applied to extract the SSP equivalent ages.

Table 2. Combining LRG results from $H(z)$ with other measurements. Marginalised constraints at 1-σ on H_0, Ω_m parameters obtained for a flat ΛCDM model. Hz means including $H(z)$ data sets.

	H_0	Ω_m
WMAP9+Hz	$68.7^{+2.6}_{-2.3}$	$0.319^{+0.048}_{-0.062}$
WMAP9+H_0 HST+Hz	$71.3^{+1.8}_{-1.6}$	$0.270^{+0.034}_{-0.043}$
WMAP9+BAO	$68.1^{+3.8}_{-4.7}$	$0.299^{+0.016}_{-0.022}$
WMAP9+BAO+Hz	69.4±1.3	$0.295^{+0.011}_{-0.012}$
WMAP9+BAO+H_0 HST	$72.5^{+0.7}_{-0.6}$	0.284±0.003
WMAP9+BAO+H_0 HST+Hz	70.4±1.1	0.290 ±0.010

Using SDSS–LRGs, three measurements of $H(z)$ were obtained at three different redshifts: $H(z \simeq 0.32) = 86.3 \pm 7.6$ km s^{-1} Mpc^{-1}, $H(z \simeq 0.30) = 78.5 \pm 6.8$ km s^{-1} Mpc^{-1} and $H(z \simeq 0.28) = 76.8 \pm 5.3$ km s^{-1} Mpc^{-1}. One additional $H(z)$ measurement was estimated using SALT–LRGs, which has a value of $H(z \simeq 0.47) = 89 \pm 23$ km s^{-1} Mpc^{-1}. The four $H(z)$ values combined with external data sets have been employed to constrain the cosmological parameters. We have obtained noticeable constraints and parameter values of $H_0 = 69.4 \pm 1.3$ km s^{-1}Mpc^{-1} and $\Omega_m = 0.295^{+0.011}_{-0.012}$, which are consistent with the standard cosmological model and very consistent with previous results.

Further investigations on cosmological constraints should be done by adding more external $H(z)$ data sets, for example from the new *PLANCK* data, latest SNIa data sets, recent BAO measurements. The systematic errors between SSP models remain a subject of discussion. Further analysis on this will be performed.

Aknowledgments

ALR thanks funding from National Research Foundation. Some of the observations reported in this paper were obtained with the Southern African Large Telescope (SALT) under program 2011-3-RSA_OTH-026 and 2012-1-RSA_OTH-013 (PI: Ando Ratsimbazafy).

References

[1] Jimenez R & Loeb A 2002 *ApJ* **573s** 37
[2] Jimenez R, Verde L, Treu T, Stern D 2003 *ApJ* **593** 622
[3] Simon J, Verde L, Jimenez R 2005 *PhRvD* **71** 123001
[4] Stern D, Jimenez R, Verde L, Kamionkowski M, Stanford S A 2010 *JCAP* **2** 8
[5] Moresco M, et al. 2012 *JCAP* **8** 6
[6] Moresco M 2015, *MNRAS* **450** 16
[7] Zhang C, et al. 2012, *RAA* **14** 1221
[8] Zheng W, et al. 2014 *International Journal of Modern Physics D* **23** 1450051
[9] Koleva M, Prugniel P, Bouchard A, Wu Y 2009 *A&A* **501** 1269
[10] Bruzual G, & Charlot S 2003 *MNRAS* **344** 1000
[11] Gaztañaga E, Cabré A, & Hui L 2009 *MNRAS* **399** 1663
[12] Crawford S M, Ratsimbazafy A L, Cress C M, et al. 2010 *MNRAS* **406** 2569
[13] Crawford S M, Still M, Schellart P, et al. 2010 *PROCSPIE* **7737** 773725
[14] Hinshaw G, Larson D, Komatsu E, et al. 2013 *APJS* **208** 19
[15] Beutler F, Blake C, Colless M, et al. 2011 *MNRAS* **416** 3017
[16] Padmanabhan N, Xu X, Eisenstein D J, et al. 2012 *MNRAS* **427** 2132
[17] Anderson L, Aubourg É, Bailey S, et al. 2014 *MNRAS* **441** 24
[18] Blake C, Brough S, Colless M, et al. 2012 *MNRAS* **425** 405
[19] Riess A. G., et al. 2011 *ApJ* **730** 119

Implementation of a goodness-of-fit test for finding optimal concurrent radio and γ-ray pulsar light curves

A S Seyffert[1], C Venter[1], A K Harding[2], J Allison[3] and W D Schutte[3]

[1] Centre for Space Research, North-West University, 11 Hoffman Street, Potchefstroom, 2531, South Africa

[2] Astrophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

[3] School for Computer, Statistical, and Mathematical Sciences, North-West University, 11 Hoffman Street, Potchefstroom, 2531, South Africa

E-mail: 20126999@nwu.ac.za

Abstract. Since the launch of the *Fermi* Large Area Telescope in 2008 the number of known γ-ray pulsars has increased immensely to over 200, many of which are also visible in the radio and X-ray bands. Seyffert et al. (2011) demonstrated how constraints on the viewing geometries of some of these pulsars could be obtained by comparing their observed radio and γ-ray light curves by eye to light curves from geometric models. While these constraints compare reasonably well with those yielded by more rigorous single-wavelength approaches, they are still a somewhat subjective representation of how well the models reproduce the observed radio and γ-ray light curves. Constructing a more rigorous approach is, however, made difficult by the large uncertainties associated with the γ-ray light curves as compared to those associated with the radio light curves. Naively applying a χ^2-like goodness-of-fit test to both bands invariably results in constraints dictated by the radio light curves. A number of approaches have been proposed to address this issue. In this paper we investigate these approaches and evaluate the results they yield. Based on what we learn, we implement our own version of a goodness-of-fit test, which we then use to investigate the behaviour of the geometric models in multi-dimensional phase space.

1. Introduction

The *Fermi* Large Area Telescope (LAT) is the primary instrument on the *Fermi Gamma-ray Space Telescope* mission which was launched on 11 June 2008. To date *Fermi* LAT has detected more than 200 new γ-ray pulsars, a dramatic increase over the 6 or 7 known at the time of *Fermi* LAT's launch [1]. This sudden abundance of pulsars visible in both the radio and γ-ray bands has opened the door to multiwavelength investigations that were not possible before.

These studies have, however, been made difficult by the comparatively low γ-ray photon flux (typically), as compared to the radio photon flux for any particular pulsar. This disparity manifests itself in the relative uncertainties on the observed intensities in the two wavebands, making traditional goodness-of-fit techniques very ineffective when trying to fit the radio and γ-ray light curves (LCs) simultaneously. This has led some investigators to adopt so-called by-eye methods to find optimal simultaneous fits (e.g., [2, 3]). A number of attempts have, however,

been made to address this issue directly, and this paper presents our latest such attempt. The most notable attempts to date ([4, 5]) artificially inflate the uncertainties accompanying the radio LCs in an attempt to grant the radio and γ-ray data equal weight in the determination of the optimal fit. As we will show, neither of their approaches accomplish this fully. We will then describe our proposed solution using these two approaches as reference.

2. The geometric pulsar models

We use an idealized picture of the pulsar system, wherein the magnetic field has a retarded dipole structure [6]. The γ-ray emission originates in regions of the magnetosphere (referred to as 'gaps') where the local charge density is sufficiently lower than the Goldreich-Julian charge density [7]. These gaps facilitate particle acceleration and consequently radiation. We assume that there are constant-emissivity emission layers embedded within the gaps, in the pulsar's corotating frame. The location and geometry of these emission layers determine the shape of the γ-ray LCs, and there exist multiple models describing their geometry. In this paper only the two-pole caustic (TPC, [8]) γ-ray model will be examined. For the radio emission we use a simple empirical hollow cone model [9]. For a more detailed description of these models, see our previous work [3].

Fig. 1 (left) shows an example of a phaseplot produced using the TPC model. It is an equirectangular projection skymap of the emission of the pulsar, with rotational phase ϕ on the horizontal axis, and the observer's line of sight ζ on the vertical axis. For each point in (ϕ, ζ) space it gives the relative intensity per solid angle of the emission that would be observed if the observer's line of sight would pass through that point as the pulsar rotates. A phaseplot thus contains the projected all-sky γ-ray emission of a pulsar over one complete rotation.

The equator of the neutron star is at $\zeta = 90°$, and its magnetic poles are at $\zeta = \alpha$ and $\zeta = 180° - \alpha$. Thus, each value of α corresponds to a unique phaseplot. Fig. 1 shows an example of how a predicted LC can be obtained by cutting a phaseplot at a constant value for ζ (e.g., for a pulsar with $\alpha = 50°$). The predicted LCs are independently normalised so their respective maxima are 1.

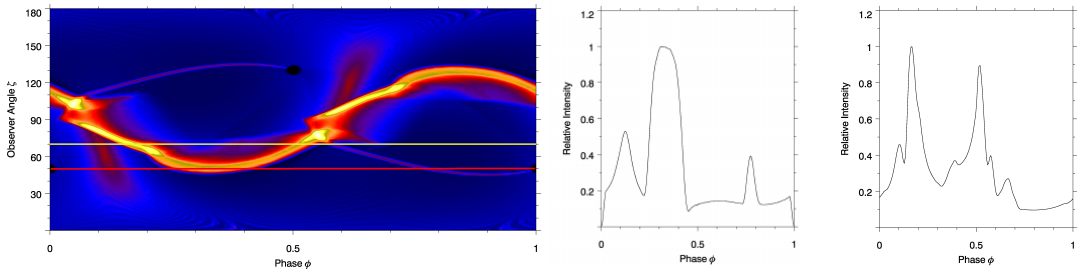


Figure 1. An example of constant-ζ cuts through a phaseplot for $\alpha = 50°$. The cuts correspond to values of $\zeta = 50°$ (red) and $\zeta = 70°$ (yellow). We see that the caustics on the phaseplot present as peaks in the LC.

3. The fitting methods

3.1. The by-eye method

The by-eye method considers the LCs associated with the entire domain of α and ζ by generating so-called atlases. Each atlas consists of a set of predicted γ-ray and radio LCs generated at each point on a grid of (α, ζ) values covering the (α, ζ) domain at a chosen resolution. The data LCs are then compared to the model LCs at each point in this grid, and a subjective decision is made regarding the quality of each fit. Once the entire space has been scanned, a best-fit

contour emerges, and constraints on the viewing geometry (α, ζ) of the pulsar can be derived. An important aspect to note is that not only are the shapes of the LCs reproduced, but also the radio-to-γ phase lag. For an example of this method's application, see our previous study [3].

The by-eye method is not very rigorous, and would possibly yield different results when applied by different people. The results, however, do seem to be a good first attempt at finding concurrent fits when considered alongside other fits obtained independently ([10], [3]). Furthermore, this method allows us to discern which features of the LC most likely contribute to a specific LC's rejection, which in turn gives us insight into what effects are responsible for the shape of the final solution contours. Since the statistical methods aim to fit the same models (in our case at least) to the same data, we can use the knowledge garnered from the by-eye contours as a sanity check when considering the contours produced by the statistical methods.

3.2. The statistical methods

The statistical methods developed to date implement a modified version of Pearson's χ^2 goodness-of-fit test, with the test statistic defined as

$$X = \sum_{i=1}^{n_{\text{bins}}} \left(\frac{E_i - O_i}{U_i} \right)^2, \tag{1}$$

with E_i, O_i, and U_i corresponding to the model (expected) intensity, observed intensity, and error (uncertainty) on the observed intensity for the ith bin (of n_{bins} bins) [11]. This test statistic is χ^2 distributed with $N = n_{\text{bins}} - n_{\text{parameters}} - 1$ degrees of freedom. Ideally the minimum value obtained for X (as a function of the tested model's parameters) will be roughly N, indicating that there is some set of parameters for which the model fits the data well. Since the current models are still only approximate representations of the actual phenomena, this unfortunately rarely occurs, with the best fits having values for $X_{\text{min}} >> N$. To compensate for this, and to allow constraints to be derived from our imperfect models, the value of X is normalised such that its minimum is N.

This method has been applied successfully for LCs from a single waveband, but has proven to be less useful when trying to fit observed LCs from multiple wavebands concurrently. Calculating the values of X for the LCs in each waveband and naively combining them (with equal weighting) for each set of parameter values, yields constraints that tend to be dominated by the waveband that has lower relative uncertainties. In our case, the radio goodness-of-fit test statistic dominates the overall fit.

This problem has been addressed in the methods developed thus far, but with limited success. One approach is to somehow artificially inflate the radio uncertainties (e.g., [4, 5]). How this is done differs depending on the method, but the results are invariably an improvement on a naive goodness-of-fit method. Fig. 2 shows how these artificially inflated uncertainties look (specifically the method used by [4]), and how the resulting values of X change.

The motivation behind these statistical methods is that for us to be able to add the χ^2 profiles for the radio and γ-ray goodness-of-fit tests, the values for X associated with the best radio and best γ-ray fit should be of roughly comparable size. This is best accomplished by adjusting the uncertainties of the radio LC. These methods do not work effortlessly for all pulsars, though. For some pulsars, one needs to increase the degree to which the radio uncertainties are inflated, while for others a decrease is required [4]. This seems to indicate that there might still be room for improvement.

4. Our proposed improvement

The method we propose takes a more direct route toward obtaining comparable radio and γ-ray values for X. In considering the statistical methods above we realised that ensuring comparable

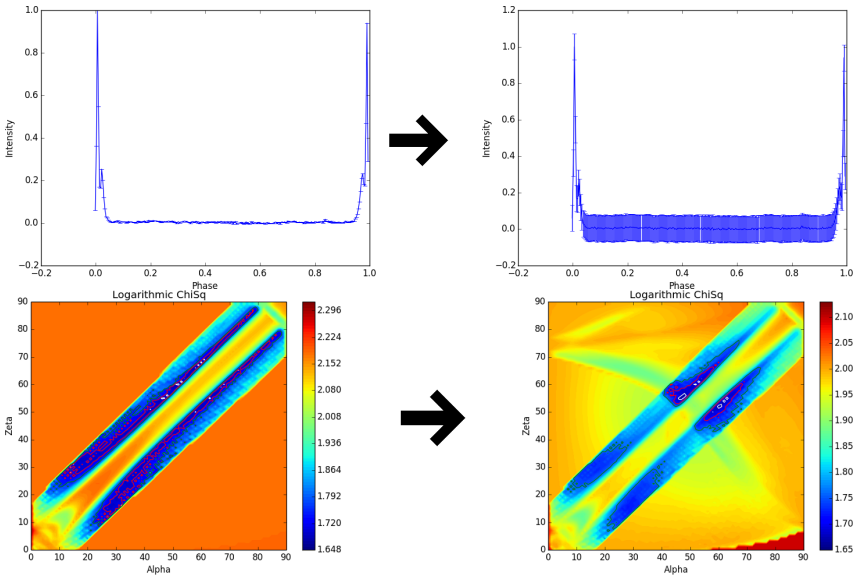


Figure 2. The top set of figures shows an example of radio uncertainties artificially inflated in a manner similar to that proposed by [4]. The bottom set of figures shows how this influences the resulting values for X for the combined radio and γ-ray fits. The left-most of the χ^2 maps finds its best fits along the diagonal, which means that the radio fit dominates. In the right-most χ^2 map we see the cross-diagonal features of the γ-ray fit emerging. The resulting contours become more constraining as both models start contributing more equally to the fitting process.

minima for the χ^2 maps is not sufficient as the minima will generally not be colocated in (α, ζ) space. Stated more clearly, we realised that the minimum of the γ-ray test statistic should be compatible with all the values of the radio test statistic, not just its minimum, and vice versa. This ensures that the minimum γ-ray test statistic value is not averaged with a (non-minimum) radio test statistic value that is orders of magnitude larger, when combining the radio and γ-ray χ^2 maps, since the radio will still dominate in that case.

To accomplish this compatibility we scale the radio test statistic values such that their dynamic range (the difference between their maximum and minimum) is equal to that of the γ-ray test statistic values. In effect we force the radio and γ-ray test statistic values to carry equal weight in the determination of the best concurrent fit. This scaling is done globally over the space defined by the parameters for which we are deriving constraints.

The procedure employed proceeds as follows:

(i) **Calculate the unscaled χ^2 maps, using Eq. (1), for the radio and γ-ray fits independently.** In our case each model LC has four parameters associated with it, which means these maps will have four dimensions. The LC parameters are α, ζ, ϕ_0 (zero point in phase), and A (the LC amplitude).

(ii) **Minimise each map over all parameters that are independent for the two wavebands.** For the concurrent fits in this study only A can be minimised over in this step. Since we assume that the radio and γ-ray emission originates from different regions in the same magnetic field, the value of ϕ_0 cannot be varied independently, and can therefore not be minimised over in this step. The resulting maps in our case have three dimensions.

(iii) **Scale the γ-ray map to have a minimum equal to the appropriate number of degrees of freedom [5].** For simplicity, we assume here that this scaled test statistic is χ^2 distributed.

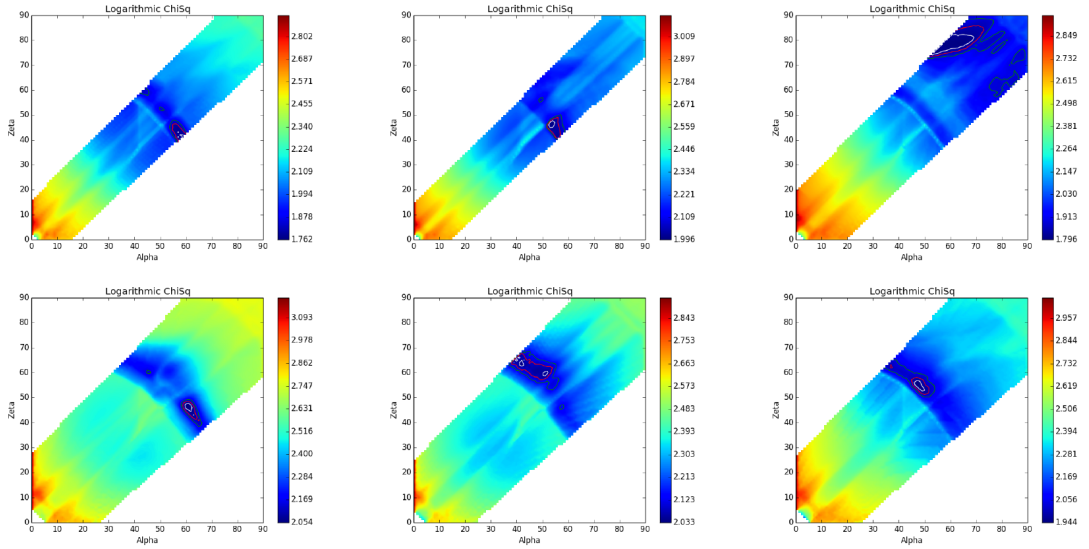


Figure 3. The resulting combined χ^2 maps for the six pulsars studied in our previous work [3]. The white, red, and green contour lines represent 1σ-, 2σ-, and 3σ-like contours respectively, but should be considered cautiously. The white regions are where the radio model predicts no LC.

(iv) **Scale the radio map such that it has the same dynamic range as the freshly-scaled γ-ray map.** The scaled radio test statistic is also assumed to be χ^2 distributed, in accordance with the assumption made for the scaled γ-ray test statistic.

(v) **Sum the two maps.** The resulting combined map will have as many dimensions as the models have coupled parameters (parameters that cannot be varied independently). In our case the coupled parameters are α, ζ, and ϕ_0. The number of degrees of freedom for this combined test statistic is the sum of the minima of the constituent maps [12]. This follows from the assumption made in *(iii)* and *(iv)* for the two constituent maps.

(vi) **Minimise the combined map over all coupled parameters except α and ζ.** For our study, and more generally, these two parameters are the ones for which constraints are to be derived. The resulting map therefore has two dimensions.

(vii) **Obtain the confidence contours using the 1σ, 2σ, and 3σ values of a χ^2 distribution with 2 degrees of freedom.** Since the lowest value the minimum of the combined map can have is the sum of the minima of the constituent maps (in the case of colocated minima), we can derive the contours by adding the above-mentioned 1σ, 2σ, and 3σ values to this lowest minimum [12].

5. Results

Implementing this approach for the six pulsars we previously studied yields constraints that compare very well to those obtained by eye [3]. The inclusion contours shown were derived as discussed in step *(vii)* in the previous section. They should, however, only serve as an estimation of how the actual contours may look since it is not yet clear how the actual contours should be computed.

6. Conclusion and Future Work

The inclusion contours obtained using our proposed method compare favourably with those for the by-eye method, which in turn compare well with independently derived constraints [3].

Although the the aim of these statistical methods is not to reproduce the results obtained using the by-eye method, the results of the by-eye method serve as a sanity check which the statistical methods should pass.

The χ^2 maps produced using our method have a number of properties which may prove useful in future studies. Chief among these properties is that the contours are calculated relative to a value other than the minimum of the combined map. This allows our method to reject an unsuccessful model combination even though some normalisation is done in obtaining the combined map. Another useful property is that the combined map remains essentially unchanged if step *(iii)* is omitted, effectively only being scaled. This allows for the possibility of comparing the goodness of fit of different model combinations, and therefore identifying the most successful model combination.

The interplay between our confidence in the geometric models and the uncertainties associated with the observed LCs also seems to be an important one to understand. It is useful to differentiate between evaluating which predicted LCs fit *best*, and whether or not the observed LCs are fit *well*. For a single model, determining the goodness of fit for each possible solution not only tells us how *well* the data is fit, but also allows us to determine which of those solutions is the *best* fit. This correlation does not hold when we generalise our method to include multiple models and multiple datasets. The breakdown occurs because we calculate the combined test statistic as the sum of the two constituent test statistics, making it sensitive to changes in the scaling of either of the single-model test statistics. Since our confidence in each model could be expressed as a scaling of the test statistic, effectively weighting their contribution to the combined test statistic, it is difficult to ensure that the single-model test statistics are properly scaled before summation. Our approach to the scaling of the individual test statistics is an attempt to address this problem directly.

The most important open issue in this study is how the confidence contours are to be defined. The specification given in Sec. 4, based on the assumptions regarding the distributions associated with the radio and γ-ray test statistics, still needs to be tested. We are currently developing a Monte Carlo simulation strategy aimed at doing this. The method derived here will also be used to compare a variety of multi-model combinations based on concurrent fitting of pulsar LCs.

Acknowledgments

This work is based on research supported by the National Research Foundation (NRF) of South Africa (Grant Numbers 90822 and 93278). Any opinions, findings, and conclusions or recommendations expressed are that of the authors, and the NRF accepts no liability whatsoever in this regard. A.K.H. acknowledges support from the NASA Astrophysics Theory Program. C.V. and A.K.H. acknowledge support from the *Fermi* Guest Investigator Program.

References

[1] Thompson D J 2001 *AIP Conf. Series* **558** 103–14
[2] Venter C, Harding A K and Guillemot L 2009 *Astrophys. J.* **707** 800
[3] Seyffert A S 2011, Venter C, Harding A K, Johnson T J *Proc. of The 56th Ann. SAIP Conf.* 531–36
[4] Johnson T J et al. 2014 *Astrophys. J. Suppl. S.* **213** 1–54
[5] Pierbattista M et al. 2015 *Astron. Astrophys.* **575** 1–88
[6] Deutsch A J 1955 *Ann. d'Astrophys.* **18** 1–10
[7] Goldreich P and Julian W H 1969 *Astrophys. J.* **157** 869–80
[8] Dyks J and Rudak B 2003 *Astrophys. J.* **598** 1201–6
[9] Story S A, Gonthier P L and Harding A K 2007 *Astrophys. J.* **671** 713–726
[10] Weltevrede P et al. 2010 *Astrophys. J.* **708** 1426–41
[11] Avni Y 1976 *Astrophys. J.* **210** 642–646
[12] Lampton M, Margon B and Bowyer S 1976 *Astrophys. J.* **208** 177–90

Numerical modelling of hydrodynamical astrophysical outflows: application using the PLUTO code

I P van der Westhuizen[1], B van Soelen[1], P J Meintjes[1], S J P K Riekert[2] and J H Beall[3]

[1]Department of Physics, University of the Free State, Bloemfontein, 9301, SA
[2]High Performance Cluster, University of the Free State, Bloemfontein, 9301, SA
[3] St. John's Collage, Annapolis, MD, 21401, USA

E-mail: VanDerWesthuizenIP@ufs.ac.za

Abstract. In order to gain a better understanding of how structures form and evolve in astrophysical outflows it is necessary to build a numerical model to simulate the motion of material in these environments. Due to the large length scales of astrophysical outflows compared to the Debye radius of the ejected particles, a fluid dynamical approach provides a suitable model for the large scale motion of the outflow. In this study a 3 dimensional relativistic fluid dynamical model was set up on a structured mesh with dimensions of 64x64x64 units. A uniform ambient medium was defined on the mesh grid and a nozzle of radius 1 was created on the initial z boundary to inject a jet with a Lorentz factor of 10. The opensource magnetohydrodynamical code PLUTO was used to evolve the numerical model with time. The PLUTO code uses high resolution shock capturing schemes to evolve the time dependent partial differential conservation equations on the structured mesh. This model led to the formation of a collimated relativistic beam surrounded by a cocoon of backflow material. Asymmetric turbulence was found, which caused instabilities in the central beam and led to the formation of shocks in the jet. These results are consistent with those seen in previous studies, which were used in order to validate the simulation.

1. Introduction

Observational studies have shown that many astrophysical sources produce collimated outflows of material. Some of these objects include radio loud Active Galactic Nuclei (AGN), accreting binary systems and Young Stellar Objects (YSO) [1]. High resolution imaging of sources such as radio loud AGN have revealed the presence of complex small scale structures in the outflows. These structures include radio blobs which are time dependent and propagate through the outflows at superluminal velocities [2]. The formation and propagation of these time dependent structures may be a cause of variability in sources containing relativistic outflows. The processes which cause the formation of such structures within the relativistic outflows are not well understood and they are often too complex for analytical analysis. In such cases numerical simulations provide a powerful analysis tool to investigate the formation and evolution of these outflows. In this study we focus on AGN jets and their properties, since their structures have been observed over a wide range of the electromagnetic spectrum.

AGN are some of the most energetic objects in the universe and emit large amounts of radiation over a wide range of the electromagnetic spectrum. The non-thermal emission received from radio loud AGN is dominated by the jet component emerging from the central engine [3]. The biggest challenge with creating a numerical model for such sources is the large range of length scales which have to be taken into consideration. This includes the effects of ion interactions in the plasma on atomic length scales as well as the large scale hydrodynamic motion of the plasma, which in the case of AGN jets, can stretch over kiloparsec scales. We can, however, simplify the problem by considering a relativistic jet that has to conserve charge neutrality, which causes the Debye radius of the electrons and protons in the jet to be negligible when compared with the jet radius. This allows us to create an appropriate model using purely relativistic magnetohydrodynamics [4].

In this paper a preliminary relativistic outflow model was created and evolved using the PLUTO opensource code. The initial model was developed to test the computational intensiveness of the simulations on the University of the Free State High Performance Cluster and to validate the base code before incorporating more complex effects into the model. In this paper section 2 will focus on the numerical methods used to create the relativistic outflow model as well as the setup of the simulation, section 3 will give the preliminary results and compare our simulation to previous studies. Finally section 4 contains a short conclusion.

2. Numerical method

The simulation of a relativistic hydrodynamical outflow is created by numerically solving the fluid dynamical conservation equations on a structured mesh grid. We implemented this using the PLUTO *ver.* 4.0 opensource code [5]. The basis of computational fluid dynamics, the PLUTO code and our simulation setup is discussed in the sections below.

2.1. Computational fluid dynamics

The motion of a relativistic fluid can be described by a set of partial differential equations, which are called the hydrodynamical conservation equations and has the following general form

$$\frac{\partial \vec{U}}{\partial t} + \nabla \cdot \vec{\vec{T}}(U) = \vec{\vec{S}}(U), \tag{1}$$

where $\vec{U}$ is a column vector consisting of the conserved variables, $\vec{\vec{T}}(U)$ is a tensor containing the flux vectors as a function of the conserved variables in every direction and $\vec{\vec{S}}(U)$ is a tensor containing the source terms, which are additional terms introduced by effects such as viscosity and gravitational forces [6].

For an ideal relativistic fluid with no magnetic field the variables in equation 1 are described by:

$$\vec{U} = \begin{bmatrix} \rho\Gamma \\ \rho\Gamma^2 h\mathbf{v} \\ \rho\Gamma^2 h - p \end{bmatrix}, \quad \vec{\vec{T}}(U) = \begin{bmatrix} \rho\Gamma\mathbf{v} \\ \rho\Gamma^2 h\mathbf{v}\mathbf{v} + p\mathbf{I} \\ \rho\Gamma^2 h\mathbf{v} \end{bmatrix}, \quad \vec{\vec{S}}(U) = 0, \tag{2}$$

where ρ is the density, p is the pressure, h is the entropy, Γ is the Lorentz factor, $\mathbf{I}$ is a 3x3 unit matrix and $\mathbf{v}$ is the velocity vector.

In order to completely describe a fluid we need an equation of state (EoS), which relates different quantities in the fluid to each other. The EoS may vary from one substance to another, therefore the choice of equation plays an important role in the simulation. For example, the caloric EoS which relates the internal energy to the pressure and density of an ideal fluid is given by

$$e = \frac{p}{\rho(\gamma_{ad} - 1)}, \tag{3}$$

where e is the internal energy of the fluid and $\gamma_{ad} = C_p/C_v$ is the adiabatic index [7]. A complete introduction to computational fluid dynamics is given in [6].

Based on previous studies we chose to characterize the relativistic hydrodynamical simulation based on 5 variables namely, the Lorentz factor Γ, the Mach number M_b, the jet to ambient density ratio η and the adiabatic index γ_{ad}. Using these variables we can recover quantities such as the pressure, the energy and the velocity of the fluid [7].

For example by combining equation (3) with the definition of the sound speed in a fluid ($C_s = \frac{v}{M_b}$) we can solve for the pressure as,

$$p = \frac{(\gamma_{ad} - 1)\rho C_s^2}{\gamma_{ad}(\gamma_{ad} - 1 - C_s^2)}. \tag{4}$$

Therefore, using the proper EoS, all of the properties of the fluid can be calculated.

2.2. PLUTO relativistic magnetohydrodynamics code

To evolve the fluid dynamical simulation the opensource modular relativistic magnetohydrodynamical code PLUTO *ver.* 4.0 was used.[1] The code was specifically designed for supersonic time-dependent flows containing discontinuities, which makes it ideal for the simulation of relativistic astrophysical outflows. PLUTO uses High Resolution Shock Capturing (HRSC) algorithms to solve the fluid dynamic conservation equations and evolve them with time. It contains different modules for hydrodynamic (HD), relativistic hydrodynamic (RHD), magnetohydrodynamic (MHD) and relativistic magnetohydrodynamic (RMHD) models, which allows the code to incorporate different effects based on the conditions of the simulation. The modular nature of the code also allows the addition of effects such as gravity, viscosity and radiative cooling to be incorporated in the calculations [5].

The PLUTO code integrates and evolves equation (1) using three computational steps. The first step is to use interpolation to construct boundary values for each cell on the structured mesh. The interpolation is done based on the centre averaged values that are assigned for each cell. It then solves a Riemann problem to determine the flux vectors across each cell boundary. Finally the code evolves equation (1) with time. The PLUTO code follows these three steps regardless of which physics module is used. The effects of the different physics modules are incorporated by a conversion in the variables of equation (1) before the computational steps [5].

2.3. Simulation setup

For testing and validation purposes a 3 dimensional numerical model of a relativistic hydrodynamical jet was constructed. For this model we considered a Cartesian static mesh grid of $64 \times 64 \times 64$ units. In this simulation arbitrary units were assigned to all variables to avoid truncation errors when working with the extremely large or small values that accompany *cgs* units. In this simulation 1 unit length corresponds to the radius of the jet nozzle at the initial inlet. An initial rest background medium was assigned to the mesh grid at $t = 0$ with a uniform density and pressure. A nozzle containing the jet material was set up on the $z = 0$ boundary. Only RHD was used in this simulation and therefore no magnetic field was assigned. A pressure matched model was set up with the pressure of the material recovered using equation (4). The density of the jet medium was normalized in arbitrary units such that $\Gamma\rho_{jet} = 1$. To validate the simulation we used parameters similar to those of previous studies such as [7] and [9]. These parameters are listed in table 1.

The boundary condition for the $z = 0$ boundary was set to reflective to simulate the production of two symmetric jets on either side of the central engine. All other boundary

[1] The code was implemented by [5] and is available at http://plutocode.ph.unito.it/.

Table 1. Variables used in the set up of the initial conditions for the preliminary RHD jet simulation.

Parameter		Value (arbitrary units)
Lorentz factor	Γ	10
Density ratio	η	10^{-3}
Jet density	ρ_b	0.1
Ambient density	ρ_{am}	100
Mach number	M_b	3.0
Adiabatic index	γ_{ad}	5/3

conditions were set to outflow so that the matter could escape freely. The simulation was run at a resolution of 8 points per unit length. A constant inflow of jet material with $\Gamma = 10$ in the positive z direction was used. The piecewise parabolic interpolation method was used to determine the boundary values of each cell. The Hartman Lax van Leer Riemann solver with contact discontinuity (HLLC) was used to solve the Riemann problem for this model, while the time stepping was done using characteristic tracing [8].

3. Results and Discussion

The proper density, pressure and proper velocity distributions produced by the simulation were visualized as two dimensional slices through the y-axis. Figure 1 shows these slices at three different time steps. From this visualization we can see that the structure of the jet consists mainly of 6 regions. The outer region consists of the uniform unperturbed ambient medium. A terminal bow shock is formed around the body of the jet, produced by the inflow of material that compresses the surrounding ambient medium. As the ambient medium passes through the bow wave its pressure increases forming a shocked region. The high pressure and density of this region causes a discontinuity to form at the surface between the shocked medium and the jet material. This is called the working surface. The speed at which the working surface propagates through the medium determines the propagation speed of the jet.

The centre of the jet consists of the high Lorentz factor ($\Gamma = 10$) material that is flowing through the nozzle into the computational domain. This region will be referred to as the relativistic beam of the jet. As the jet material propagates through the beam of the jet it will reach the working surface at the head of the jet. At this boundary the interaction between the jet and the high pressure ambient material causes the jet material to decelerate, converting the kinetic energy into internal energy. The collision between the lower density jet material and the higher density ambient medium also causes a backflow of material surrounding the beam of the jet. This backflow forms a region of low density material between the jet and the ambient medium which prevents the interaction of jet material with the high pressure ambient medium. This region (called the cocoon) assists in keeping the jet collimated over the propagation distance [9].

The turbulent interaction of material at the head of the jet causes vortices to form in the cocoon. These vortices in the cocoon are further amplified by the formation of Kelvin-Helmholtz instabilities at the surface between the cocoon and the higher pressure ambient medium. Periodic shock waves are also formed in the central beam of the jet. These shocks are generated by perturbations in the boundary between the beam of the jet and the surrounding cocoon, and are the result of a pressure difference between material in the beam and the cocoon along with

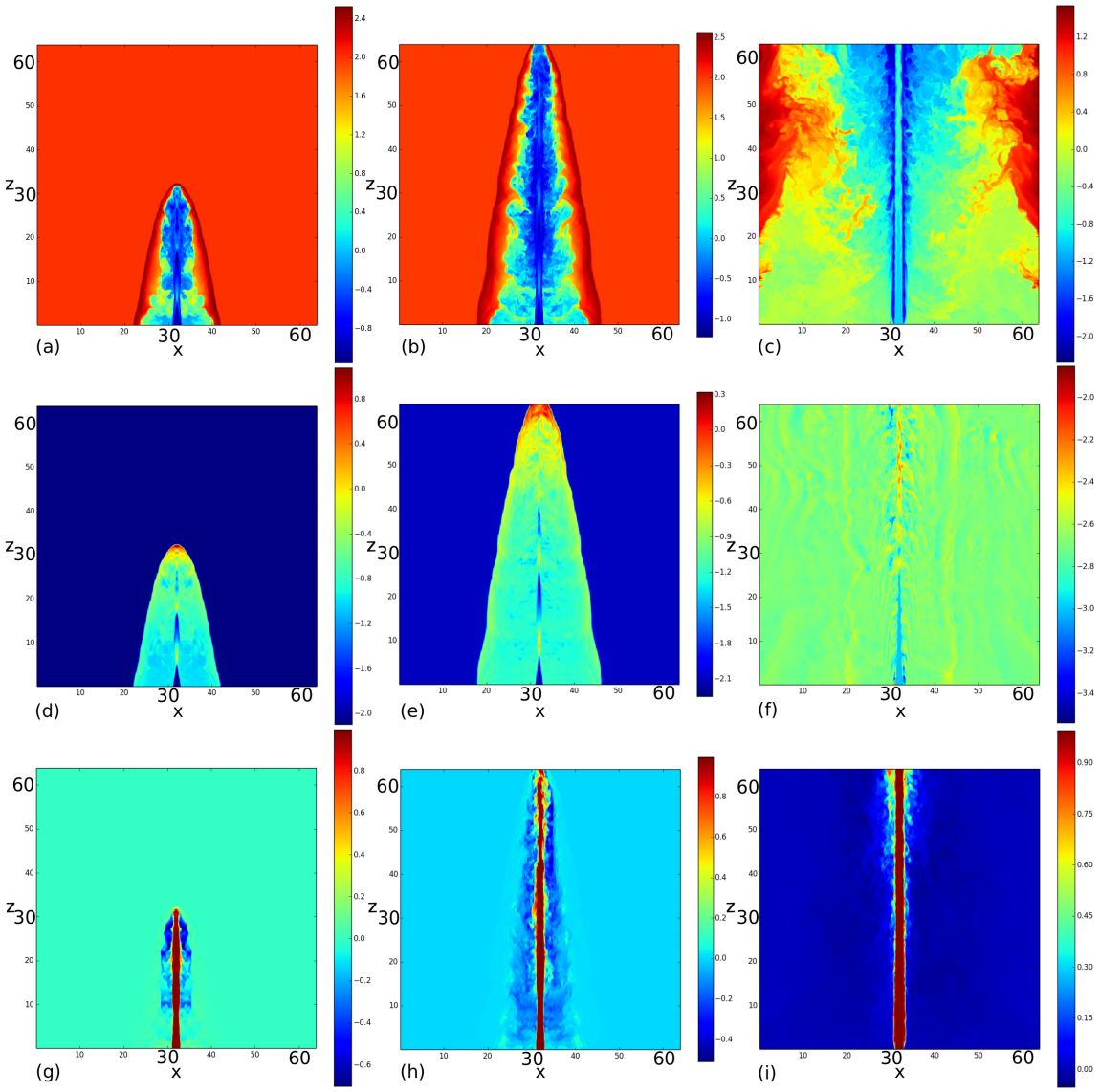


Figure 1. Two dimensional visualizations of the preliminary simulation through the y-axis of the jet, showing the proper density at (a) $t = 100$, (b) $t = 200$, (c) $t = 4300$; the pressure at (d) $t = 100$, (e) $t = 200$, (f) $t = 4300$, and the proper velocity component in the z direction at (g) $t = 100$, (h) $t = 200$, (I) $t = 4300$. Logarithmic scales are shown for the density and pressure plots in arbitrary units, while the velocity plots have a linear scaling in units of c.

the turbulence generated in the cocoon [7].

As the simulation evolves with time the cocoon of the jet expands outwards into the surrounding medium. The turbulence on the surface of the beam grows with time due to the interaction between the cocoon and jet material. This amplifies perturbations in the beam of the jet to form wavelike structures in the flow.

The simulation was run on the UFS HPC on 270 cores with a CPU time of 106098.27 h. This corresponded to 7384.7 units of simulation time. As seen in figure 1 the head of the jet leaves the computational domain at a time of 200 units, which allows for enough time for the formation of a stable jet. The real time necessary to complete the simulation amounted to 405.65 h. The

simulation used a total of 112.8 Gb of memory. Higher resolution of the simulation was sacrificed for a decrease in computational resources. Additional computing power will be necessary when more complex effects such as magnetic fields are included in the simulation.

4. Conclusion

A 3D simulation of a relativistic outflow was created and evolved over time. The simulation shows a collimated central beam with little deceleration surrounded by an outer cocoon. Asymmetric structures formed in the jet, which is in accordance with observations of AGN sources and previous studies. The results shown are comparable to those in [7] and [9]. Small scale differences in the asymmetric turbulence are present in our results when compared to [7]. We attribute these differences in the jet morphology to a difference in the chosen parameters, a difference in grid structure and the fact that the model in that study was based on an axis symmetric two dimensional jet.

More complex effects such as viscosity, magnetic fields and a variable flow can now be incorporated into our model. The next step in our study will be to incorporate variable injection of jet material to investigate the formation of structures corresponding to a non-uniform flow inside the beam of the jet as well as how these structures propagate with time. This fluid dynamical simulation is not limited to the AGN case but can also be adopted for a variety of astrophysical objects such as microquasars, YSO, and X-ray binaries.

Acknowledgments

Acknowledgement is given to the HPC unit at the University of the Free State for their assistance in the installation and set up of software used to run simulations on the UFS HPC.
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 are not necessarily to be attributed to the NRF.
This work is based on the research supported in part by the National Research Foundation of South Africa for the grant 87919. 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] De Gouveia Dal Pino E M 2005 Astrophysical jets and outflows *AdSpR* Issue 5 **35** 908-924
[2] Ghisellini G, Padovani P, Celotti A and Maraschi L 1993 Relativistic bulk motion in active galactic nuclei *ApJ* **407** 65-82
[3] Böttcher M 2011 Modeling the spectral energy distributions and variability of blazars *2011 Fermi & Jansky: Our Evolving Understanding of AGN* (*preprint arXiv:1205.0539*)
[4] Begelman M C, Blandford R D and Rees M J 1984 Theory of extragalactic radio sources *Rev. Mod. Phys.* **56** 255-351
[5] Mignone A, Bodo G, Massaglia S, Matsakos T, Tesileanu O, Zanni C and Ferrari A 2007 PLUTO: a numerical code for computational astrophysics *ApJ* **170** 228-42
[6] Toro E F 2009 *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction Third Edition* (Berlin: Springer) Chapter 1 pp 1-40
[7] Marti J M, Mller E, Font J A, Ibez J M and Marquina 1997 Morphology and dynamics of relativistic jets *ApJ* **479** 151-163
[8] Mignone A and Bodo G 2006 An HLLC Riemann solver for relativistic flows - II. Magnetohydrodynamics *MNRAS* **368** 1040-54
[9] Rossi P, Mignone A, Bodo G, Massaglia S and Ferrari A 2008 Formation of dynamical structures in relativistic jets: the FRI case *A&A* **488** 795-806

Optical spectroscopy of PSR B1259-63 around the 2014 periastron passage

B van Soelen[1], P Väisänen[2], I Sushch[3], L Klindt[1], PJ Meintjes[1], A. Odendaal[1] and R. Armstrong[4]

[1]Department of Physics, University of the Free State, Bloemfontein, 9301, South Africa
[2]South African Astronomical Observatory, PO Box 9, Observatory 7935, Cape Town, South Africa [3]Centre for Space Research, North-West University, Potchefstroom, 2520, South Africa
[4] SKA South Africa

E-mail: vansoelenb@ufs.ac.za

Abstract. PSR B1259-63 is a gamma-ray binary system which consists of a 48 ms pulsar in a 3.4 year orbit around a Be star. Of the five known gamma-ray binary systems, it is the only one where the nature of the compact object is conclusively known. This makes it an extremely important target for multi-wavelength observations. Near periastron, the interaction between the pulsar and stellar winds creates a shock front, producing non-thermal emission from radio to TeV energies. The 2010 periastron passage was the first that was observed with the *Fermi*-LAT telescope and, approximately 30 days after the 2010 periastron passage, an unexpected GeV flare was observed at a time when the other non-thermal emission was already decreasing. A repeat flare event (though on a lower scale) was observed after the 2014 periastron passage. We report on spectroscopic observations undertaken with the Southern African Large Telescope (SALT) from approximately 33 days before until 78 days after the 2014 periastron passage. These observations confirm the variability of the Hα and He-I lines during this period, as was previously reported. Combined with multi-wavelength results, this suggests that the circumstellar disc is disrupted during periastron. The reported multi-wavelength results are also briefly discussed.

1. Introduction

Gamma-ray binaries present a distinct class of binary systems that are characterized by a Spectral Energy Distribution (SED) that peaks in the gamma-ray regime (see [1] for a detailed review). All these systems consist of an early main-sequence star and a compact object, but it is only in PSR B1259-63/LS 2883 that the compact object is clearly detected as a pulsar. The 48 ms pulsar is in a $\sim$ 3.4 year orbit around a Be star [2] which has been classified as a late O-type (O 8-9V) type star [3].

Close to periastron the pulsar passes behind the Be star's circumstellar disc, as indicated by the eclipse of the pulsar [4]. A shock forms due to the interaction of the pulsar and stellar winds, resulting in particle acceleration and subsequent non-thermal emission from radio to TeV gamma rays. The flux of the non-thermal emission is, in general, highest around the period of periastron, with an indication of maxima around the time of the disc crossings (see [5] and references therein for discussion of the multi-wavelength observations around the 2010 periastron passage). However, the GeV gamma-ray emission (as observed with *Fermi*-LAT during 2010) showed only a slight detection around periastron, but showed a large flux increase

around 30 days after periastron, at a period when the other multi-wavelength emission was already decreasing [6]. This emission peaked at an implied luminosity of ~ 100 per cent of the spin-down luminosity of the pulsar. A number of theoretical models have been investigated, but there is no clear consensus as to the cause of this emission [7, 8, 9, 10, 11, 12].

During 2014, PSR B1259-63/LS 2883 was monitored by a number of multi-wavelength observatories around the periastron passage that occurred on 2014 May 4. The GeV gamma-ray emission was re-detected at a similar orbital phase, though there was no detection at periastron and the detected flux was lower [13, 14].

Optical spectroscopic observations around the 2010 periastron passage showed a variation in the equivalent width (EW) of the Hα and He I λ6678 emission lines, as well as a change in the symmetry of the lines [5]. However, there were no observations around the period of the *Fermi* flare. In order to monitor the changes in the circumstellar disc around periastron we used the Robert Stobie Spectrograph (RSS) to undertake long-term monitoring of the binary system from 33 d before, until 78 d after, periastron ($t_p - 33$ d and $t_p + 78$ d, respectively, where t_p is the epoch of periastron).

2. Observations

Observations were undertaken with the RSS configured for a wavelength coverage of 6176.6 – 6983.0 Å with a resolution of $R = 11\,021$ at the central wavelength. The system was successfully observed 25 times, with each observation consisting of 3-4 exposures, with an integration time of between ~ 476 to 500 s. The exposures were averaged and reduced following the standard IRAF reduction processes.[1] Spectral shape correction was performed using observations of the spectroscopic standard LTT4364.

The EW width of the Hα emission line was measured by integrating over the line, following the standard IRAF procedures.[2] The He I λ6678 line lies in a larger, stellar absorption feature and since the emission line is always present a correction for this absorption feature would require the underlying intrinsic stellar absorption spectrum to be accurately modelled. Therefore, the EW of the double peak He I λ6678 line has been measured by fitting two Gaussian profiles (relative to the continuum level) and integrating over the area of the fitted lines.[3] While this ignores the underlying absorption feature, it allows for a better comparison to be made to the data presented in [5].

3. Results

The results for the Hα and He I λ6678 lines are discussed below.

3.1. Hα line

The Hα line is in emission throughout all observations. The average EW is -55.2 ± 0.8 Å before $t_p - 20$ d, which is consistent with previous measurements away from periastron [3, 16], and rises until the line strength peaks at -72.7 ± 0.7 Å around $t_p + 13$ d, before decreasing again to the pre-periastron value (figure 1).

While the line remains single peak, it is asymmetric and is better fitted by multiple components. This is believed to be due to the blending of the underlying double (multiple) emission peaks arising from the circumstellar disc. ‘

[1] The data reduction was performed using the NOAO/CCDRED and NOAO/TWODSPEC packages.
[2] The analysis was performed using the tools available within the NOAO/ONEDSPEC package.
[3] The line profile fits were performed using the FITYK software package [15].

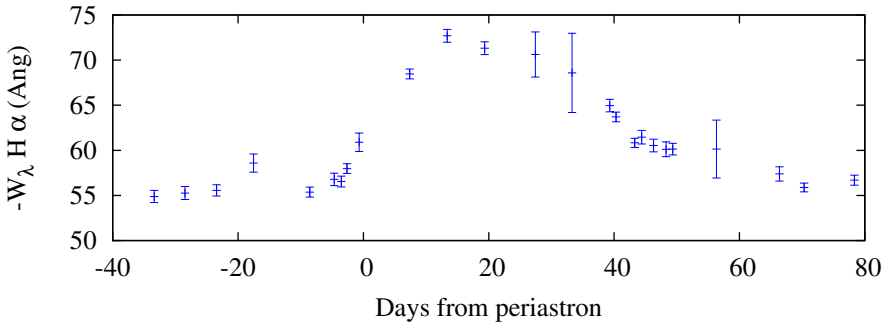


Figure 1. Variation in the Hα EW around periastron.

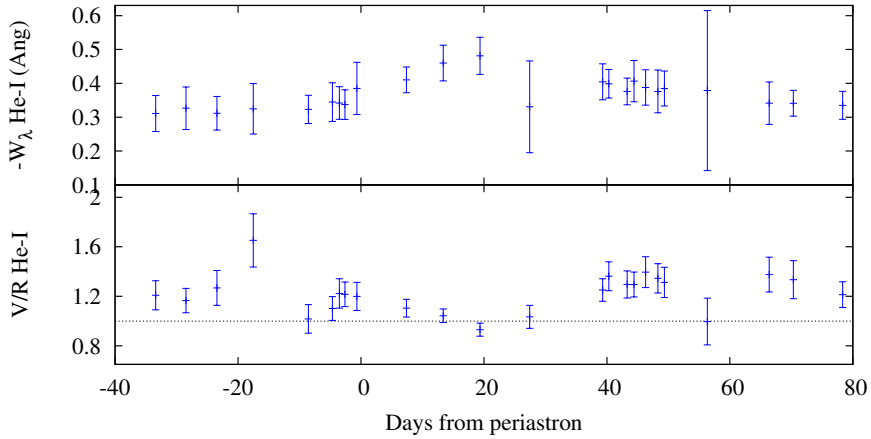


Figure 2. Variation in the He I λ6678 EW (top) and V/R (bottom) around periastron.

3.2. He I λ6678 line

The variation in the EW of the He I λ6678 line follows the same general trend as the Hα line, peaking at -0.48 ± 0.05 Å around t_p+19 d (figure 2-top). The line remained double peaked throughout all observations and exhibits a change in the relative height of the violet (V) and red (R) peaks, as shown in the V/R variation (figure 2 bottom).

While the variation in V/R is relatively smooth, there is a marked change at t_p-17.5 d where the violet line becomes stronger. This occurs around the period of the first disc crossing and corresponds to the marked increase in the EW of the Hα line. This may indicate a significant disruption of the circumstellar disc around this period or the introduction of instabilities. See for example the Smooth Particle Hydrodynamical (SPH) simulation undertaken by [17].

The change in the V/R around t_p+56 d is believed to be due to the lower signal to noise of the observation, as can be seen in the larger errors associated with the Hα and He I λ6678 EWs. The point is plotted for completeness.

4. Discussion

The variation in the Hα emission line around the 2014 periastron passage is consistent with what was previously observed around the 2010 periastron passage [5]. However, the absolute EW is lower around periastron in 2014 than it was in 2010/2011. After t_p+40 d, the EWs are comparable. We attribute this difference to the intrinsic variability of the circumstellar disc associated with Be stars.

The variation in the EW and V/R can be understood as the possible truncation of a circumstellar disc around periastron, as well as the indication of density structures in the disc around periastron. The EW of the Hα line was previous used by [5] to model the mass of the circumstellar disc in LS 2883 using the earlier model of [18]. The authors suggested a rapid decrease in the mass of the circumstellar disc after periastron that may be associated with the *Fermi*-LAT flare. A similar effect may be evident during the 2014 periastron passage [19].

Since the circumstellar disc can influence the shape of the shock front, detailed studies of the disc variability around the period of the *Fermi* flare may provide more clues to the underlying emission processes. Unfortunately, the observations reported here were hampered by weather and technical difficulties with RSS around $\sim t_p$+30 d, as indicated by the larger errors.

Unlike in the 2010 periastron, it has been suggested that there may be a correlation between the X-ray and GeV emission, with an increase in the X-ray emission detected with *Swift*-XRT around t_p+30 d [20] (see also [13] for corresponding modelling).

5. Conclusion

The optical observations with the RSS around the 2014 periastron passage present a similar trend as to previous observations. The newer observations provide a better resolution and show orbital variation in the emission line as well as shorter time scale changes, possibly associated with the pulsar disc crossing. A comparison to multi-wavelength observations may help develop a more detailed model for the interaction in this binary system.

Acknowledgments

All observations reported in this paper were obtained with the Southern African Large Telescope (SALT).

References

[1] Dubus G 2013 *A&ARv* **21** 64
[2] Johnston S Manchester RN, Lyne AG, Nicastro L and Spyromilio J 1994 *MNRAS* **268** 430
[3] Negueruela I, Ribó M, Herrero A, Lorenzo J, Khangulyan D and Aharonian FA 2011 *ApJ* **732** L11
[4] Johnston S, Manchester RN, Lyne AG, D'Amico N, Bailes M, Gaensler BM and Nicastro L 1996 *MNRAS* **279** 1026-36
[5] Chernyakova M *et al.* 2014 *MNRAS* **439** 432
[6] Abdo AA *et al.* 2011 *ApJ* **736** L11
[7] Dubus G and Cerutti B 2013 *A&A* **557** A127
[8] Khangulyan D, Aharonian FA, Bogovalov SV and Ribo M 2012 *ApJ* **752** L17
[9] Kong SW, Cheng KS and Huang YF 2012 *ApJ* **753** 127
[10] Takata J, Okazaki AT, Nagataki S, Naito T, Kawachi A, Lee SH, Mori M, Hayasaki K, Yamaguchi MS and Owocki SP, 2012 *ApJ* **750** 70
[11] Mochol I and Kirk JG 2013 *ApJ* **776** 40
[12] Sushch I and Böttcher M 2014 *Jour. High Energy Astrophysics* **3** 18
[13] Tam PHT, Li KL, Takata J, Okazaki AT, Hui CY and Kong AKH *ApJL* **798** L26
[14] Caliandro GA, Cheung CC, Li J, Torres DF and Wood K 2015 astroph:1504.06343
[15] Wojdyr M 2010 *J. Appl. Crystallogr.* **43** 1126
[16] Van Soelen B, Meintjes PJ, Odendaal A and Townsend LJ 2012 *MNRAS* **426** 3135
[17] Okazaki AT, Nagataki S, Naito T, Kawachi A Hayasaki K, Owocki SP and Takata J 2011 *PASJ* **63** 893
[18] Grundstrom ED and Gies DR 2006 *ApJ* **651** L53
[19] Chernyakova M *et al.* 2015 *MNRAS* **454** 1358
[20] Bordas P, Zabalza V, Romoli C, Khangulyan D and Puehlhofer G 2014 *ATel* **6248** 1

The unusually strong coronal emission lines of SDSS J1055+5637

Hartmut Winkler

Department of Physics, University of Johannesburg, South Africa

E-mail: hwinkler@uj.ac.za

Abstract. Many Seyfert galaxies display weak 'coronal' emission features corresponding to [Fe VII], [Fe XI] and [Fe XIV] in their optical spectra, whereas elsewhere these lines seem to be entirely absent. These lines appear to highlight zones in the nucleus irradiated by high-energy photons. The presence of these zones and the conditions therein as determined by the relative line strengths and profiles impose important constraints on the physical models of active galactic nuclei, and Seyferts in particular. In 2009 the discovery was announced of the highly unusual spectrum of SDSS J0952+2143, where the coronal lines are exceptionally strong. This paper presents a second object with abnormally strong coronal features, SDSS J1055+5637. The spectrum, line ratios and related parameters are compared to those of SDSS J0952+2143, three AGN with moderate coronal lines and one where the coronal lines are missing altogether. Possible mechanisms are discussed that may account for the stronger than usual coronal features.

1. Introduction

Emission lines associated with highly ionised iron have been noticed in the spectra of some Seyfert galaxies since some of the early investigations of this class of active galactic nuclei (AGN). These have been dubbed 'coronal lines' in view of the presence of similar emission in the spectrum of the solar corona [1]. Spectral lines identified in Seyferts include [Fe VII] $\lambda\lambda$ 5721, 6087 Å, [Fe X] λ 6374 Å, [Fe XI] λ 7892 Å and [Fe XIV] λ 5303 Å [2]. Studies of the highly ionised iron spectral line profiles have revealed subtle differences between these features and the prominent narrow lines associated with [O I], [O II], [O III], [N II] and [S II] [3]. In particular, the iron lines have been found to have a greater width and a Doppler shift compared to the remainder of the narrow line spectrum, and the magnitude of this discrepancy correlates with the degree of ionisation [4]. This has traditionally been interpreted as evidence of stratification of the narrow line region, with the iron emitting zone, particularly where the ionisation is greatest, presumed to be the innermost layer most exposed to photoionisation by high energy radiation generated in the vicinity of the AGN accretion disk surrounding a central black hole [5,6,7].

The Sloan Digital Sky Survey (SDSS) [8] has greatly increased the number of known AGN, and has unearthed a variety of unusual objects. One of these is SDSS J0952+2143, which, in addition to a peculiar multiple Balmer line profile, exhibited coronal lines with relative strengths a multiple of what has been hitherto observed [9,10]. The SDSS spectrum of this object recorded in December 2005 displays a [Fe VII] λ 6087 Å line strength of as much as 67% of the [O III] λ 5007 Å flux (even amongst AGN recognised as having strong coronal emission, this ratio is typically only 4% [7]). Follow-up spectroscopic observations in 2008 showed that the relative strength of the coronal lines in

SDSS J0952+2143 had by then decreased by about 50% for [Fe VII] and 80%-90% for higher levels of ionisation [10]. This unusual coronal line brightening event illustrates that processes other than photoionisation of narrow line region gas from a luminous central source may be responsible for the coronal emission, including exceptional supernova eruptions or tidal disruptions due to galaxy interaction [10]. This raises the question whether other data exists displaying similar hyper-strong coronal line emission.

This papers presents SDSS J1055+5637, another AGN with an SDSS spectrum showing coronal lines way stronger than normal [11], though not quite on the scale seen for SDSS J0952+2143.

2. Spectral analysis

This paper investigates archival wavelength and flux calibrated spectra acquired online from the SDSS database [8]. Figure 1 displays the SDSS spectrum of SDSS J1055+5637, where the abnormally strong coronal lines are immediately evident.

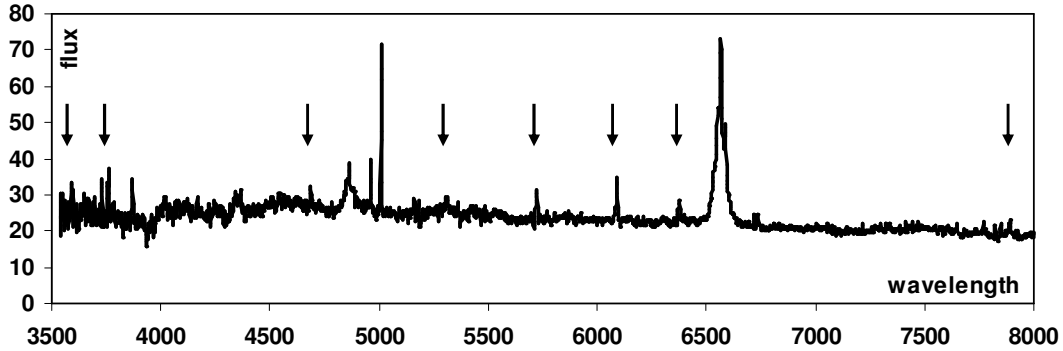
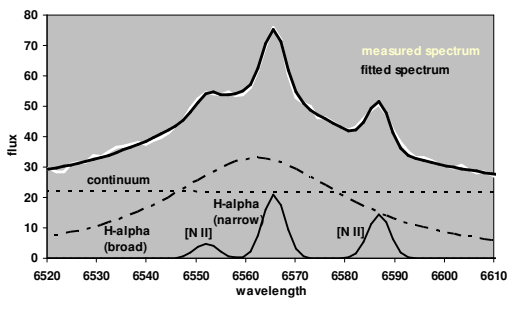


Figure 1. SDSS spectrum of SDSS J1055+5637. The flux is given in units of $\times 10^{-17}$ erg s $^{-1}$ m^{-2} Å^{-1} and the wavelength in Å. The arrows pinpoint the locations of the unusually strong coronal lines.

Continuum levels $C(\lambda)$ in the spectral neighbourhood of an emission line were estimated through the fitting of a power law of the form $C(\lambda) = C_0\lambda^{-\alpha}$ between two continuum points identified 'by eye' on either side of the line. The line flux F was hence determined to be the difference between the total and continuum fluxes in the wavelength range flanked by these points. A Gaussian or, where line blending was evident, several Gaussians were then fitted to the line profile (after subtraction of the continuum). Balmer broad line components were modelled by Lorentzian profiles rather than Gaussians. Figure 2 illustrates the outcome of the fitting of the spectral region around the hydrogen alpha line (λ 6563 Å). This fit was also performed 'by eye', constraining the three narrow lines to all be of the same width (in terms of Doppler velocities) and fixing the ratio between the two [N II] lines to be 3:1, which is a close approximation of the theoretical value [12]. This approach has been validated by the good match between the measured and modelled profiles.

Figure 2. Section of the spectrum in the vicinity of Hα for SDSS J1055+5637. The white line represents the measured spectrum, while the thick solid black line represents the fitted model. The components of this model are also shown in the figure as broken lines for the continuum and broad component and thin solid lines for the three narrow lines.



Emission line widths were always determined to be greater than the instrumental resolution of ~150 km s^{-1} characteristic of SDSS spectra. In particular, the coronal line *FWHM* (full width at half maximum) typically exceeded the corresponding value for emission lines such as [O III] by about 100 km s^{-1}. The coronal lines are furthermore blue-shifted by on average 75 km s^{-1} relative to [O III].

The flux ratios between specific emission lines are indicative of the physical conditions in the gas they originate in. In particular, specific [O III] and [S II] line ratios enable the estimation of the gas density and temperature [13]. Where photoionisation is the dominant emission mechanism, the ratios between iron emission lines from different degrees of ionisation provide clues about the system geometry and physical processes [5]. Finally, this work also includes a determination of the asymmetry parameter *A20* [14] for the [O III].

In addition to SDSS J1055+5637, for comparison purposes, this study also examines the spectra of three AGN with distinct coronal line emission, and one Seyfert with no iron emission lines detectable. Furthermore, the line ratios and some line parameters of SDSS 0952+2143 have been redetermined to contrast it with the primary object of this study. Table 1 lists the coordinates, redshift, reddening $E(B–V)$ at those coordinated by the interstellar medium (ISM) in our Galaxy (obtained from Schlegel et al [15]) and the date the original spectrum was recorded. Note that the redshifts of all six objects in Table 1 are similar, meaning that the size of the region captured within a slit is similar in all cases. The spectra of two of these additional objects are displayed in Figure 3 (the remaining ones may be inspected on the SDSS website or in the online ZORROASTER catalogue [16]).

Table 1. Details of Seyfert galaxies described in this paper.

Object	R.A. (2000)	Dec. (2000)	z (SDSS)	$E(B–V)$ (ISM)	Date of SDSS observation
SDSS J1055+5637	10h55m26.4s	+56°37′13″	0.0748	0.008	9 Apr 2002
SDSS J0952+2143	09h52m09.6s	+21°43′13″	0.0790	0.028	30 Dec 2005
SDSS J1355+6440	13h55m42.8s	+64°40′45″	0.0754	0.013	16 Mar 2001
SDSS J1429+4518	14h29m25.1s	+45°18′32″	0.0749	0.012	30 Mar 2003
SDSS J2356–1016	23h56m54.3s	–10°16′05″	0.0741	0.032	22 Aug 2001
SDSS J1314+2606	13h14m47.1s	+26°06′24″	0.0720	0.015	28 Feb 2006

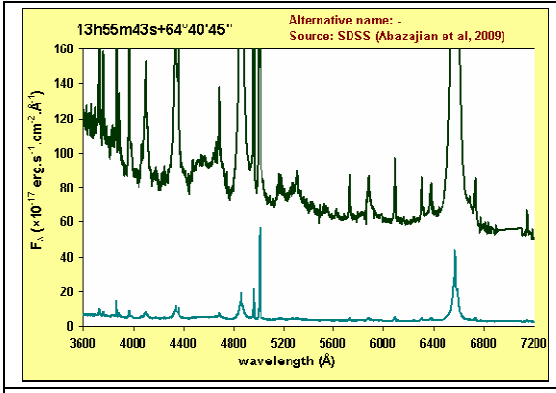


Figure 3(a). Spectrum of SDSS J1355+6440. The vertical scale refers to the top spectrum. The lower spectrum is a scaled down version of the top one.

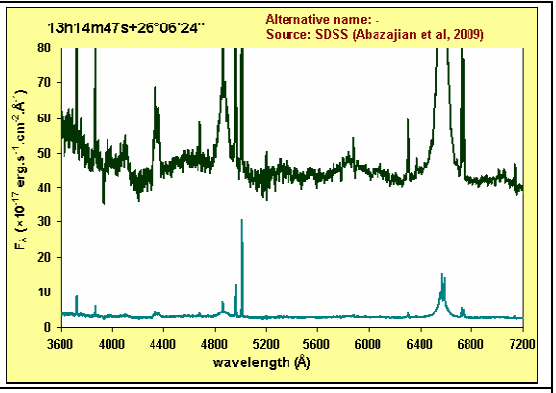


Figure 3(b). Spectrum of SDSS J1314+2606. As in Figure 2(a), the lower spectrum is a scaled down version of the upper one.

Table 2 lists the line fluxes of all emission features of the entire AGN sample studied here. Based on the line strength uncertainties available at SDSS for each spectrum evaluated here, flux errors will

be of the order of 5% for strong lines down to about 20% for the weaker emission features. Any corrections due to underlying absorption will only affect higher order Balmer lines, and will even then only be minor in view of the virtual invisibility of most features associated with the galaxy background component (e.g. Na I).

Table 3 then presents the results of the determination of a variety of parameters obtained from the line profile fitting and the analysis of the line flux ratios. The full width at half maximum of the hydrogen beta line (*FWHM-β*) indicates the broad line region gas velocities (projected towards the observer). The narrow line widths (*FWHM-n*) and asymmetry parameters quoted in the table are those of [O III] λ 5007 Å, which is strong in all spectra. The reddening parameter $E(B–V)$ was determined from the flux ratio of Hα to Hβ. This was done separately from the broad and narrow components. Column 7 records the [Fe X] λ 6374 Å to [Fe VII] λ 6087 Å line flux ratio, which best illustrates how different ionisation levels are represented in each of our AGN sample. The last columns list the free electron density n_e deduced from the [S II] $\lambda\lambda$ 6716,6731 Å doublet and the free electron temperature T_e calculated from the relative strength of [O III] λ 4363 Å compared to the other [O III] lines.

Table 2. Line fluxes of the objects investigated in this study (in units of 10^{-15} erg s^{-1} m^{-2}). An (↑) indicates the presence of the line, but that it could not be measured with accuracy due to it being blended with the line above.

λ (Å)	Identification	SDSS 1055+5637	SDSS 0952+2606	SDSS 1355+6440	SDSS 1429+4518	SDSS 2356–1016	SDSS 1314+2606
3587.2	[Fe VII]	0.50	0.61	1.60	-	4.15	-
3727	[O II]	0.57	0.81	6.18	3.48	18.2	5.38
3759.9	[Fe VII]	0.92	1.31	3.40	0.58	6.29	-
3868.74	[Ne III]	0.64	0.56	7.86	2.28	9.33	2.41
4101.73	Hδ	-	-	19.0	7.05	10.3	3.32
4340.46	Hγ	2.14	0.87	46.8	12.2	22.8	7.87
4363.21	[O III]	↑	↑	4.18	-	4.04	0.72
4550	Fe II band	weak	-	50.6	14.8	22.6	-
4685.68	He II	0.50	1.18	4.06	0.94	4.96	0.76
4861.33	Hβ broad	5.08	2.02	92.7	27.0	66.2	20.8
4861.33	Hβ narrow	0.30	0.43	5.54	1.78	7.68	1.57
4958.91	[O III]	0.85	0.52	26.2	8.07	31.5	8.85
5006.84	[O III]	2.40	1.61	76.5	25.6	89.4	26.1
5250	Fe II band	weak	-	26.0	-	13.2	-
5302.86	[Fe XIV]	0.75	1.05	1.00	-	-	-
5721.11	[Fe VII]	0.73	1.04	2.27	-	3.79	-
5875.62	He I	-	-	7.23	5.74	7.38	0.92
6086.92	[Fe VII]	1.23	1.44	3.79	1.24	6.27	-
6300.30	[O I]	-	-	2.61	1.26	2.83	1.38
6374.51	[Fe X]	0.75	2.46	2.83	0.47	4.87	-
6562.80	Hα broad	19.5	16.5	302	126	243	82.4
6562.80	Hα narrow	1.26	1.75	↑	10.1	24.0	7.51
6583.45	[N II]	↑	↑	↑	10.6	22.2	8.56
6716.44	[S II]	0.22	0.49	1.58	2.32	7.43	3.02
6730.82	[S II]	0.25	0.43	2.03	1.79	6.42	2.66
7135.8	[Ar III]	-	-	1.39	0.59	2.17	0.65
7325	[O II]	-	-	2.39	-	2.10	-
7891.94	[Fe XI]	0.72	1.76	no data	-	3.13	-

Table 3. Parameters determined from the spectral properties.

Object	*FWHM-β* [km s^{-1}]	*FWHM-n* [km s^{-1}]	*A20* [O III]	*E*(*B*–*V*) (broad)	*E*(*B*–*V*) (narrow)	*F*(6374)/ *F*(6087)	n_e (cm^{-3})	T_e (K)
SDSS J1055+5637	2170	229	0.060	0.21	0.30	0.61	700	-
SDSS J0952+2143	1700	351	0.173	0.95	0.25	1.70	300	-
SDSS J1355+6440	2350	284	0.241	0.38	-	0.76	1000	30000
SDSS J1429+4518	1850	172	0.095	0.60	0.40	0.38	100	-
SDSS J2356–1016	2350	337	0.148	0.14	0.00	0.77	300	25000
SDSS J1314+2606	2960	256	0.147	0.23	0.43	-	300	18000

3. Discussion

3.1. SDSS J1055+5637

The measurements in this work confirm that the coronal emission lines in SDSS J1055+5637 are indeed exceptionally strong (e.g. *F*(6087)/*F*(5007) = 0.51). The ratios of coronal line fluxes to those of the strongest [O III], [O II] and [O I] emission features in this object exceed those determined in other AGN with coronal lines [7, 11] (with the exception of SDSS J0952+2143).

If one does not consider the coronal lines, SDSS J1055+5637 has a very typical Seyfert 1.5 spectrum. The widths, line profiles and gas characteristics are normal for the class. The broad line component is blue-shifted by 150 km s^{-1} compared to the corresponding narrow lines, highlighting a systemic motion of the broad line region gas.

The most plausible explanation for the unusual strength of the coronal lines is that this effect is due to an earlier very bright phase of the inner nuclear zone consisting of the accretion disk and its immediate surroundings. The possible sequence of events thereafter is represented in Figure 4. The light echo from this outburst would pass through the broad line region within a few days, causing this zone to brighten (a). Thereafter the light burst reaches the innermost parts of the narrow line region, where the coronal lines are believed to originate, and which would then result in a brightening of that zone as seen in the spectrum investigated (b). Some time after that the emission from that zone again fades as the pulse reaches further parts of the narrow line region (c).

It is probable that spectra of this object from a different epoch would, if they were available, display rather different coronal line emission features. Future spectroscopy of SDSS J1055+5637 is therefore strongly encouraged.

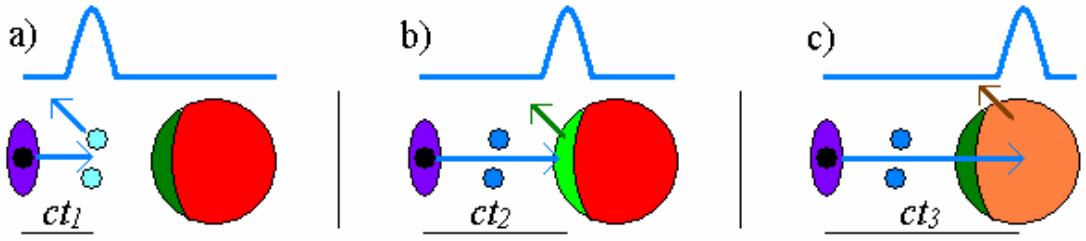


Figure 4. Illustration of the reverberation of a light pulse generated in the accretion disk (at the left of each of the diagrams). The diagrams are not drawn to scale. The pulse reaches the fast-moving broad line gas clouds in (a), the coronal line emitting zone at the inner part of the narrow line region in (b), and the other parts of the narrow line region in (c).

3.2. Other objects

SDSS J0952+2143 has been comprehensively studied in the past [9,10], and will not be discussed in depth here. Note that the high *F*(6374)/*F*(6087) ratio was only a temporary phenomenon presumably

associated with the immediate aftermath of an eruptive event, and that for subsequent spectra this ratio was much closer to the other values in Table 3 [10].

SDSS J1355+6440 (also referred to as VII Zw 533), SDSS J1429+4518 and SDSS J2356–1016 are more conventional broad-line Seyferts with coronal line spectra of varying prominence. The first mentioned of these displays a relatively large narrow line asymmetry and a greater electron density.

SDSS J1314+2606 (also known as Arp 60) is characterised by relatively wide broad lines, with the width of the He I line in Figure 3(b) apparently exceeding that of the Balmer lines.

4. Conclusion

SDSS J1055+5637 has been confirmed to be a Seyfert galaxy with the second-strongest set of coronal emission lines identified to date. It is postulated that these unusually strong features are due to a light echo transiting the coronal line formation zone following a powerful brightening of the inner nucleus. A later epoch spectrum would be required to confirm this hypothesis.

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] Oke J B, Sargent W L W 1968 *Astrophys. J.* **151** 807
[2] Penston M V, Fosbury R A E, Boksenberg A, et al. 1984 *Mon. Notices Roy. Astron. Soc.* **208** 347
[3] Appenzeller I, Östreicher R 1988 *Astron. J.* **95**, 45
[4] Grandi S 1978 *Astrophys. J.* **221** 501
[5] Korista K T, Ferland G J 1989 *Astrophys. J.* **343** 678
[6] Rodriguez-Ardila A, Prieto M A, Viegas S, Gruenwald R 2006 *Astrophys. J.* **653** 1098
[7] Gelbord J M, Mullaney J R, Ward M J 2009 *Mon. Notices Roy. Astron. Soc.* **397** 172
[8] Ahn C P, Alexandroff R, Allende Prieto C, et al. 2014 *Astrophys. J. Suppl.* **211** A17. SDSS data accessible at http://skyserver.sdss.org/
[9] Komossa S, Zhou H, Wang T, et al. 2008 *Astrophys. J.* **678** L13
[10] Komossa S, Zhou H, Rau A, et al. 2009 *Astrophys. J.* **701** 105
[11] Nagao T, Taniguchi Y, Murayama T 2000 *Astron. J.* **119** 2605
[12] Storey P J, Zeippen C J 2000 *Mon. Notices Roy. Astron. Soc.* **312** 813
[13] Osterbrock D E, Ferland G J 2005, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*, 2nd Ed., ISBN 978-1-891389-34-4, University Science Books
[14] Whittle M 1985 *Mon. Notices Roy. Astron. Soc.* **213** 1
[15] Schlegel D J, Finkelbeiner D P, Davis M 1998 *Astrophys. J.* **500** 525
[16] Winkler H 2013 *ZORROASTER online catalogue*, http://www.uj.ac.za/EN/Faculties/science/departments/physics/Pages/ZORROASTER.aspx

Division D2 – Space Science

Modelling ground conductivity for computing the electric field associated with geomagnetically induced currents using the Finite Element Method. (A mid latitude case study)

E Matandirotya [1,2], P J Cilliers [1,2] , and R R Van Zyl [1]

[1]French South African Institute of Technology, Department of Electrical, Electronics and Computer science, Cape Peninsula University of Technology, Cape Town, South Africa
[2]South African National Space Agency, Space Science Directorate, Hermanus, South Africa

E-mail:electdom@gmail.com

Abstract: The study of geomagnetically induced currents (GIC) in technological systems connected to the Earth such as power lines and pipelines during adverse space weather conditions requires the computation of the electric field induced in the Earth. These computations can be achieved through solving Maxwell's equations with appropriate boundary conditions. COMSOL Multiphysics, a finite element method (FEM) simulation package is used to compute the electric field resultant from measured geomagnetic field variation under the assumption of a horizontally layered Earth model. The study is based on the calculation of the GIC in a transformer grounding point at an electrical substation in South Africa for which a previous study determined the conductivity profile. The investigation aims at studying the effects of varying the number, thickness and conductivity of layers when computing the electric field associated with GIC, thus, enhancing the understanding of the distribution of the induced horizontal electric field within the various layers of the Earth during a geomagnetic storm. The measured GIC and the modelled GIC are compared to determine the best representation of the Earth. This kind of study is important in understanding the layers which matter most in the effective modelling of GIC for this particular substation. The results based on this case study indicates that for a layered Earth model where the top layers have a low conductivity compared to the underlying layers the deeper high conductivity layers have a significant influence on the accuracy of the modelled GIC.

1. Introduction

Geomagnetically induced currents (GIC) are enhanced during geomagnetic storms. A geomagnetic storm arises when the magnetosphere is highly disturbed. This occurs when the interplanetary magnetic field turns southward for a significantly long period [1]. The variations in the magnetosphere-ionosphere currents have a great influence on the geomagnetic perturbations on the surface of the Earth. As such, the magnetosphere-ionosphere currents are considered to be the primary source of surface electric field [2]. The variations of the surface magnetic field determine the characteristics of the surface electric field. The extent to which space weather affects grounded technological systems is an expanding research arena. Such knowledge aids in understanding the vulnerability of several technological systems to GIC during adverse space weather.

An investigation of the variations of the induced electric field in the various layers of the Earth is presented. The conductivity profile used in the model was derived by Ngwira et al., [3] for the Earth below a South African 400 kV substation at Grassridge (33.7°S, 25.6°E).

For the computation of the induced field, finite element method (FEM) software package, COMSOL Multiphysics is used. This is a rarely explored technique in the computation of electric fields connected with GIC computation. A comparison is made of GIC derived from an electric field

modelled using the FEM with measured GIC at Grassridge during the geomagnetic storm of October 29, 2003. The case study uses geomagnetic field data measured at Hermanus (34.4°S, 19.2°E) as model input.

2. GIC modelling in mid- and low- latitude

The computation of GIC in any technological system requires firstly the determination of the surface electric field where the system is grounded. The second step requires determination of the system parameters which are unique and dependent on system configuration [4]. The electric field associated with GIC can be computed or modelled in several ways. In reality, the computations are complex due to the inhomogeneity of the ground conductivity and the non-uniformity of the electric field. Hence some reasonable assumptions are made in order to simplify the computations.

For middle- and low- latitude regions, the plane wave model is often deemed adequate as the regions are not subjected to auroral electrojets. The model assumes that the magnetic field propagates as a plane wave directed vertically down to the Earth resulting in components of the field being parallel with the surface of the Earth. The GIC computations require magnetic field time series data which is typically only measured at magnetic observatories. The coupling between the magnetic field due to the telluric currents and those due to the space currents are neglected to further simplify the computations [5].

In cases where the surface electric field is computed over a layered Earth (consisting of horizontal layers of different conductivities as shown in Figure 1, the apparent conductivity of each layer is considered [2]. The horizontal components of the electric field are given by equation 1 and 2

$$E_y = -\mu_0^{-1} Z B_x \quad (1)$$

$$E_x = \mu_0^{-1} Z B_y \quad (2)$$

where E_y is the East-directed component of the electric field, E_x is the North-directed component of the electric field and B_x and B_y are the corresponding orthogonal horizontal magnetic field components, Z is the surface impedance and μ_0 is the permeability of free space.

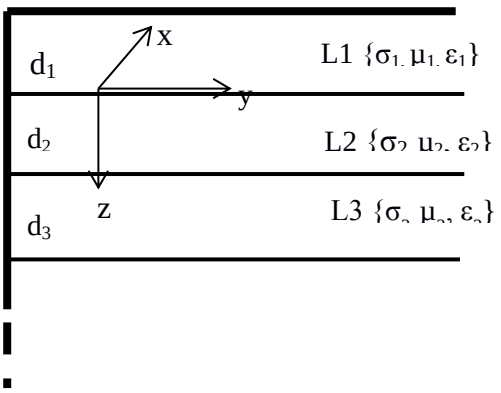


Figure 1. A 2D representation of the Earth's structure consisting of 3 layers, L1, L2 and L3 of depth d_1, d_2 and d_3. Each layer has a different conductivity (σ), permeability (μ) and permittivity (ε). In reality the number of layers can vary. The layers are all assumed to have infinite extent in the x-direction and the bottom layer is assumed to have infinite extent also in the z-direction. The dimension of the layers in the y-axis is finite to allow localisation of the computations

3. Computing the electric field by means of the finite element method

Techniques such as the complex image method (CIM) [6], method of auxiliary sources (MAS) [7] have been applied in the computation of magnetic field and electric field associated with GIC modelling. Recent studies are considering the use of Finite element method [8] to compute the electric field associated with GIC. Only the horizontal components (E_x and E_y) of the electric field are computed and equation 3 is used for GIC computation. E_z is always zero if there is no horizontal variation in the Earth's conductivity, thus the Earth's electric field is always horizontal [5]. The coefficients (*a* and *b*) are computed separately using appropriate methods [9].

$$\mathrm{GIC(t)} = a\mathrm{E_x}(t) + b\mathrm{E_y}(t) \quad (3)$$

where *a* and *b* are the network coefficients which can be derived from the network topology and impedances.

FEM is a technique that divides the problem domain into non-overlapping simple elements to reduce the complexity of the structural geometry. The partial differential equation (PDE) governing the field variable under investigation is solved for each element. This reduces the complexities by finding a solution of the PDE for each. The element solutions are integrated to give the overall solution for the whole structure.

Within the COMSOL Multiphysics package, the *ACDC module* was used as it is designed to solve Maxwell's equation [10]. In FEM procedures the general steps that are followed are, defining the geometry, allocating material properties, imposing appropriate boundary conditions and selecting the type and size of the elements (mesh). Computations are handled by the FEM software and the results are analysed by means of graphic displays.

3.1 The model

The construction of the layered Earth model is done in a 2D domain in the y-z plane. The interest is therefore in the variation of the induced electric field $\mathrm{E_{(x,y)}}$ along the z- axis. The most important thing at this stage is to make sure that all layers are properly joined to each other for electrical continuity. This is achieved through the built-in "*union*" function of the COMSOL software.

3.2 Magnetic field data

The measured magnetic field is introduced in the model as a piecewise cubic interpolated function applied to the surface of the Earth. One magnetic field component is loaded at a time, that is either the $\mathrm{B_x}$ or $\mathrm{B_y}$ component. In the current case, magnetic field data measured at one-minute intervals are used.

3.3 Boundary Conditions

To keep the magnetic field uniform within the model, a "perfect magnetic conductor" boundary condition was set on the left and right edges of the flat Earth model. This boundary condition, implying **n** x **H** = 0, prevents tangential components of the magnetic field at the edges of the model. **n** is the unit normal to the surface. The "magnetic insulation" boundary condition was set at the bottom of the model to ensure that the induced currents can only flow in the y- direction.

3.4 *The system equation*

The layers of the Earth are considered to be sub-domains. For each sub-domain, Ampere's law, (equation 4) is solved.

$$\sigma\frac{\mathrm{d}\mathbf{A}}{\mathrm{dt}} + \nabla\times\left(\mathbf{B}(\mu_0\mu_\mathrm{r})^{-1}\right) - \sigma\mathbf{v}\times\mathbf{B} = \mathbf{J} \quad (4)$$

where **A** is the magnetic potential vector, **B** is the applied magnetic flux density, μ_r is the relative permeability **J** is the current density, **v** is the velocity of the conductor which is 0 in the current case as there is no moving media in the model, μ_0 is the permeability of free space and σ is the conductivity of the Earth . The electric field (**E**) is then inferred from the expression in equation 5:

$$\mathbf{E} = \mathbf{J}\sigma^{-1}. \quad (5)$$

3.5 The Mesh

An *extremely fine* mesh size of triangular elements was used to provide a reasonable compromise between spatial resolution and computation time. This yielded a total number of 11137 elements and 23096 degrees of freedom. The size of the mesh has a significant bearing on the computation time.

While fewer elements may take a shorter time to converge, the accuracy of the solution may be compromised.

3.6 Computation

The final step of the process is to allow the computation of the desired variables. A *time dependent* study was used for the computation of the electric field within the several layers of the domain. The electric field at the interface between Earth and air is then used to compute the GIC that is likely to be driven into any grounded technological system in the vicinity.

4. Modelled Cases

The study considers 3 cases. These layer combinations were selected so that characteristics of the electric field could be deduced from these three situations which are close representation of the ground conductivity structure. For case 1, the purpose was to evaluate if the poorly are sufficient for calculating the electric field associated with the GIC. Case 2 evaluate the effects of a less resistive layer underlying a high resistivity layer on the calculated electric field. Case 3 evaluate the effects of the terminating half-space on the surface electric field.

Case 1: 7 layers (high resistivity layers only). Resistivity in Ωm = [440,420,400,380,330,300,260] and the corresponding depth in km [1.1, 3.9, 7, 14, 24, 50, 100]

Case 2: 2 layers, (layer 1 has a resistivity which is the average of the high resistive layers on top while layer 2 has a low resistivity). Resistivity in Ωm = [361, 10] and depth in km = [200, 190]

Case 3: 3 layers (Same layers as for Case 2, plus a deepest layer of infinite extent and medium resistivity). Resistivity in Ωm = [361, 10, 16] and depth in km = [200, 190,∞]

In all cases, the depth of the layers as originally derived are preserved. The model inputs are the resistivity and magnetic field.

5. Results discussion and conclusion

Dong et al, [8] evaluated the accuracy of 2D FEM electric field modelling by comparing FEM simulated electric field with analytical values. This current study has demonstrated the applicability of FEM in electric field modelling through the good correlation between measured GIC and GIC derived from the electric field obtained by FEM. The use of FEM comes with an advantage of handling complex geometry related to the geological structure. The aim of this study is identify the suitable Earth conductivity profiles to be used for GIC modelling for the Grassridge substation.

The results presented focuses on the period with high magnetic perturbations on October 29, 2003 i.e. during 06:00-08:00 hours UT. The network coefficients used for the GIC modelling are a = -80 $AkmV^{-1}$ and b = 15 A kmV^{-1}, which were derived by Koen [11] for the Grassridge substation. Figure 2, 3 and 4 show the electric field variation and GIC modelled for the selected period. For evaluation of the modelled GIC, the maximum modelled GIC, correlation coefficient and the root mean square of the difference between the modelled and measured results for the three cases and are calculated. The maximum measured GIC for this particular storm was 12.56 A.

Figure 2 shows that for Case 1 the magnitude of the induced electric field reduces with depth. For these 7 consecutive layers, the conductivity is actually increasing with depth. The GIC modelled with the surface electric field has 91% correlation with the measured GIC. The maximum value of the modelled GIC during this storm was 12.0 A and the RMSE of 0.98 A.

For the second case which considers an average of the 7 layers (low conducting) above a highly conducting layer, the modelled GIC has a correlation coefficient of 89 %, RMSE of 1.00 A and maximum modelled GIC of 12.35 A . Figure 3 shows how the variation of the electric field on the surface compares with the maximum field induced in the two layers. What is interesting to note is that the electric field induced on the surface is the same as that induced on the first layer at some point. For example, during the period 06:30-07:00, the surface electric field (E_y) is the same as that in the first layer, but this is not true of E_x for the same period. However, the same can be said for E_x

during the period 07:00-08:00. The reasons for this behavior in the electric field are yet to be investigated.

The third case which considers the deepest layer to have infinite extent results in a GIC with a correlation coefficient of 88 %, RMSE of 1.15 A and a maximum modelled GIC of 10.78 A.

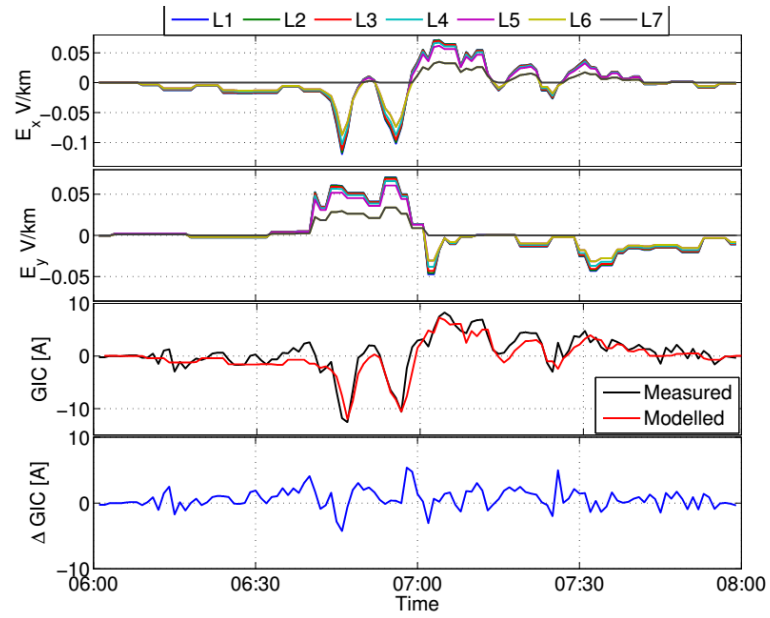


Figure 2. The estimated horizontal field components for each layer, the modelled GIC and the difference between the measured and modelled GIC for Case 1.

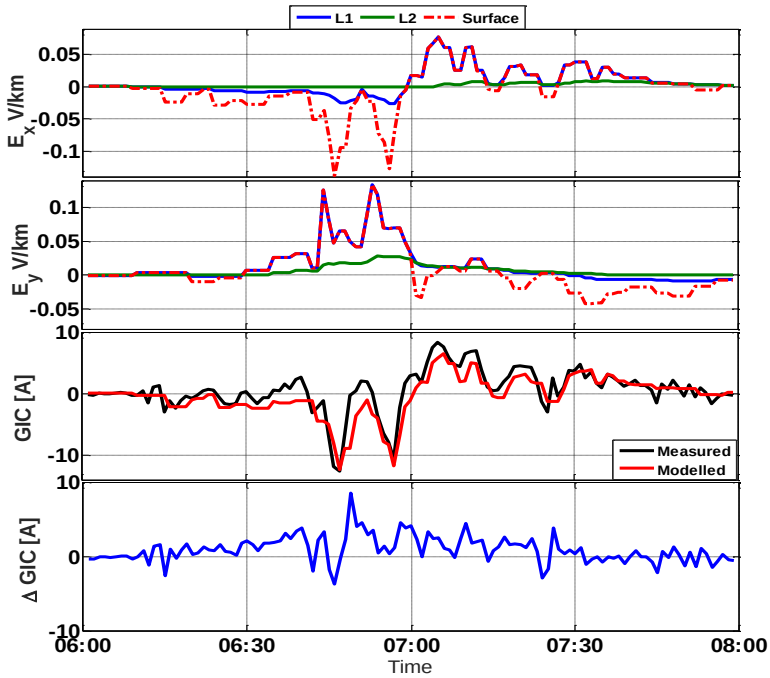


Figure 3. The estimated horizontal electric field components for each layer, the modelled GIC and the difference between the measured and modelled GIC for Case 2.

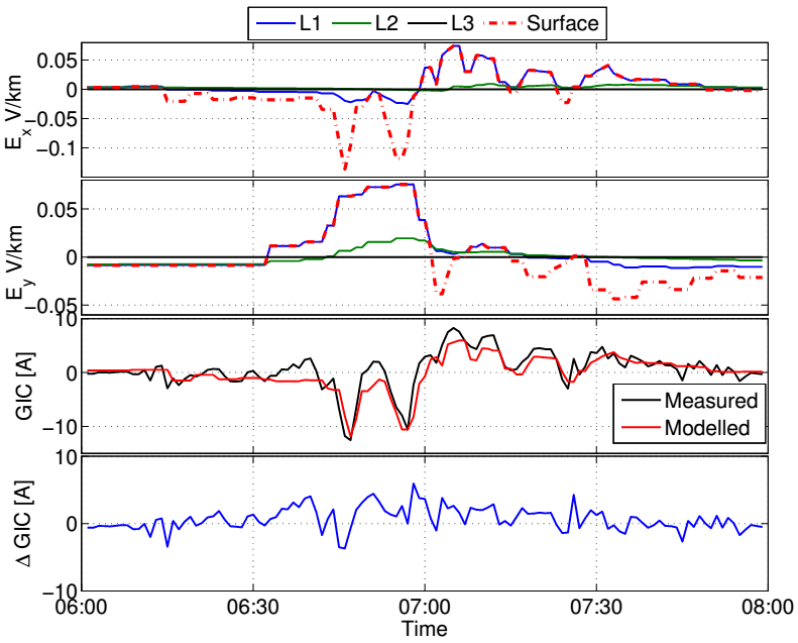


Figure 4. The estimated horizontal field components for each layer, the modelled GIC and the difference between the measured and modelled GIC for Case 3.

It was noted that the FEM software introduces an arbitrarily phase shift observed in the modelled and measured GIC, (time lag in figure 2 and lead Figures 3 and 4). The lag/lead of the modelled GIC was adjusted for maximum correlation with the measured GIC.

From the current results it is indicated that the configuration of layers considered for a GIC model can have a significant impact on the maximum modelled GIC, however we note that the correlation coefficient and the root mean square error obtained from the for the three cases are very close to each other. These results therefore indicate that quality GIC estimation could be achieved with simplified Earth representation instead of a 7-layer structure such as Case1. Another significant conclusion that the deeper layers have a significant contribution to the surface electric field induced during a geomagnetic storm.

Acknowledgements

We appreciate the financial support provided by South African National Space Agency (SANSA), Space Science Directorate for conducting the research and the French South African Institute of Technology (FSATI) for covering the conference travelling costs.

References

[1] Tsurutani B T, Gonzalez W D, Tang F, and Lee T 1992 *Geophysical research letters.* **19** 73-76
[2] Pirjola R 2002 Surveys *in Geophysics.* **23** 71–90
[3] Ngwira C M, Pulkkinen A, McKinnell L-A ,and Cilliers P J 2008 *Space Weather*. **4**
[4] Boteler D H, Pirjola R, and Nevanlinna H 1998 *Advances in Space Research.* **22** 17–27.
[5] Pulkkinen A 2003 Geomagnetic induction during highly disturbed space weather conditions: studies of ground effects, Ph.D. thesis. University of Helsinki, Helsinki, Finland.
[6] Boteler D H and Pirjola R J 1998 *Geophysical Journal International*. **132** 31–40.
[7] Shepherd G S and Shubitidze F 2003 *Journal of atmospheric and solar-terrestrial physics.* **65** 1151–1160.
[8] Dong B, Danskin D W, Pirjola R J, Boteler D H and Wang Z Z 2013 *Annales Geophysicae.* **31**, 1689-1698.
[9] Boteler D 2014 *J. Space Weather Space Clim.***4** A21-p1 A21-p11
[10] COMSOL.Inc 2012 Introduction to the AC/DC Module. http://www.comsol.com.
[11] Koen J 2002 Geomagnetically Induced Currents in the Southern African Transmission Network, Ph.D. thesis, University of Cape Town, Cape Town, South Africa.

An integrated software based analytical model for the signal path efficiency of the HartRAO lunar laser ranger optical system

S C Ndlovu[1, 2]**, L Combrinck**[1, 2, 3]**, N P Nkosi**[1, 2] **and R C Botha**[1]
[1] Space Geodesy Programme, Hartebeesthoek Radio Astronomy Observatory, Krugersdorp, South Africa,
[2] Programme of Land Surveying (Geomatics), School of Engineering, University of KwaZulu-Natal, Durban, South Africa
[3] Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa

E-mail: sphume@hartrao.ac.za

Abstract. The Lunar Laser Ranger (LLR) system under development at HartRAO (25.8900° S, 27.6853° E) will accurately measure the Earth-Moon distance through the use of laser pulses. The complete signal chain needs to be optimally configured to ensure that the transmitted pulses reaches the retroreflectors located on the Moon and a signal is returned to the Earth-fixed receiving telescope. We discuss the hardware components and software used in HartRAO's LLR system to achieve optimal signal path efficiency. This includes thorough descriptions of the laser source, optical components used throughout the coudé path, atmospheric transmission efficiency and retroreflector's optimal reflectance value. The use of the link budget equation in this work estimates the number of photons that are expected to be received; this result has a direct relationship with total system efficiency. An integrated model for HartRAO's LLR is an essential tool to enable optimal signal path (optical and electrical) efficiency and is useful in estimating the expected number of photon returns for given observational parameters.

1. Introduction

The Lunar Laser Ranger (LLR) system under development at HartRAO (25.8900° S, 27.6853° E) will accurately measure the Earth-Moon distance by the use of laser pulses [1-3]. It will accurately measure the time of flight (TOF) for pulses transmitted to the Moon's five arrays of corner cube reflectors and returned back to the Earth fixed receiving telescope. The measured TOF is basically multiplied by the speed of light and divided by two in order to obtain the separation between the surfaces of Earth and Moon. Several other corrections need to be included for high accuracy, these include general relativistic range delay and LLR station position changes due to solid Earth tide corrections. The success of this technique depends highly on a number of parameters which include the quantum efficiency of the detector, characteristics of the transmit and receive optics, transmission properties of the atmosphere and cirrus cloud cover as well as the reflector's effective cross section.

The obtained number of photons resulting from the returned laser pulses is a key technique in lunar science and the General Theory of Relativity (GTR). The LLR data analysis is the most effective

technique to study the interior of the Moon and dynamics of the Earth Moon system [4]. It has also contributed to studies of the lunar core, detection of lunar free libration, evaluation of the strong principle of equivalence for massive bodies and time variability of the universal gravitational constant [5-7]. The LLR technique has also provided measurements of the Moon's tidal acceleration, geodetic precision of the lunar orbit and its rotation where it contributes to the determination of the Earth orientation parameters, such as nutation, precision and polar motion [8-10].

A good site for laser ranging is determined by the magnitude of atmospheric transmittance which depends on the behaviour and amount of different particles in the atmosphere (e.g. aerosols, dust, water vapour, thermal and density variations and air mass). The fluctuations in the atmosphere and the presence of cloud cover can cause a propagating laser beam to diverge and thus, adversely affect lunar laser ranging data quantity and quality. Nickola (2012) investigated these fluctuations by determining the astronomical seeing conditions for the HartRAO LLR site selection and characterisation. Once the HartRAO LLR is completed and fully operational, it will be moved to Matjiesfontein in the Karoo (a semi-desert region in South Africa) where the astronomical seeing conditions will be better and less cloud cover will occur for precise laser ranging.

The aim of this work is to discuss the vital components (hardware and software) that will ensure success for HartRAO lunar laser ranging. An integrated software based analytical model will provide estimates of signal path efficiency for the HartRAO lunar laser ranger optical system. After installation at Matjiesfontein the system could be supplemented with collocated space geodesy systems (GNSS, SLR, VGOS and DORIS) to form a node of the IAG Global Geodetic Observing System [1].

2. HartRAO LLR System Description

The HartRAO LLR system consists of two TME_{00} laser sources, one for Lunar Laser Ranging and the other for Satellite Laser Ranging (SLR). Both these laser sources have a wavelength of 532 nm and produce laser pulses of different repetition rates (see Table 1). The standard deviation for the sync out pulse jitter for these lasers is less than 0.1 ns. The output beam diameters of the transmitted lasers are; 1 m for LLR and 0.2 m for SLR. For LLR, the laser beam is transmitted through the transceiver 1 m optical telescope and directed to the retroreflector mirrors mounted on the Moon surface. The laser transmitter for SLR is mounted and properly aligned parallel with the 1 m receiving telescope to ensure that the transmitted laser signals are received through the 1 m aperture.

Table 1: The specifications of the laser installed at HartRAO for Satellite/Lunar Laser Ranging system.

Laser Specifications	SLR	LLR
Output Energy, mJ	0.5	120
Repetition rate, Hz	1000	20
Beam Diameter, mm	~3	~12
Pointing stability, µrad	<30	<50
Beam Diameter, m	0.2	1.0

Each laser pulse is transmitted via the same coudé path which forms the laser ranging optical train, starting from the optical table and endings at the transmitting telescope. The optical train includes the optical components that will play a key role in ensuring that the transmitted laser is directed to the distant retroreflector mirrors and laser signals are reflected to the Earth “fixed” receiving telescope. It consists of an average transmit and receive optics efficiency of greater than 90% which maximises the success of laser ranging. The complete optical layout (see Figure 1) consists of a laser sources (SLR/LLR) mounted inside the two boxes, respectively, to prevent dust contamination and reduce thermal variations.

The 1 metre cassegrain optical telescope mount configuration is of Azimuth-Elevation configuration and consists of transceiver (transmitting and receiving) optical mirrors. A hyperbolic 0.3 m secondary mirror mounted on a metal spider structure to the front of the telescope tube directs the returned signal to the detector mounted at the back of the telescope. The telescope is equipped with servo drives and a high accuracy steering and pointing software that maximises high precision tracking capabilities [11]. This optical system will be integrated with a solid state photon detector, event timer, precipitation and visibility sensors, range gate generator, start diode and programmable pulse repetition frequency (PRF) system.

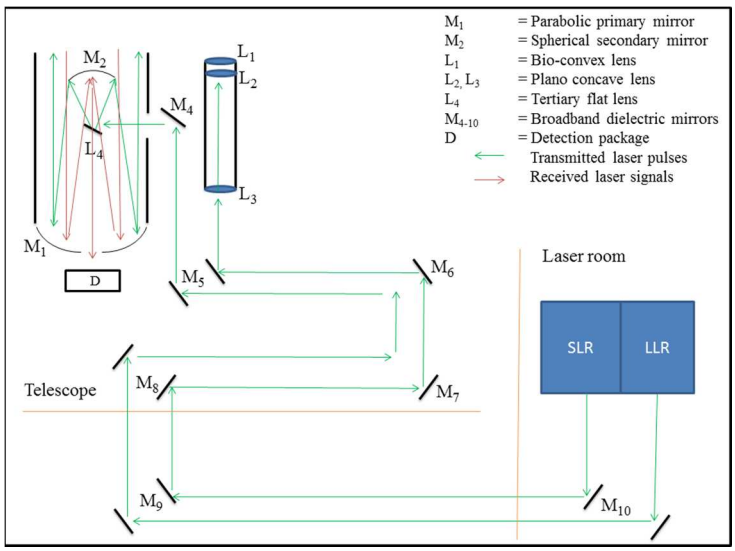


Figure 1: The schematic diagram of the HartRAO LLR system. Laser pulses for both SLR and LLR propagate through the same coudé path. Each laser has a 532 nm wavelength.

3. Laser signal path optimisation

There are many detrimental factors that affect propagating laser beams [12-16]. Thermal and density fluctuations in the atmosphere are some of the limiting factors whenever operating laser ranging. Analysis of these atmospheric fluctuations and development of an integrated model and system that enables optimal efficiency of a lunar laser ranger signal path will improve the chances of effectively targeting the cube retroreflectors mounted on the lunar surface.

One of the crucial considerations that can enable optical signal path efficiency is to develop a sensitive laser pulse detection system which operates with Single Photon Avalanche Diodes (SPADs). The SPADs help to register the arrival of a single photon and the detection of this photon is time-tagged. Using a single photon detection system is ideal for the superior performance of detectors at the single photon level [17-19].

4. Simulations, results and analysis

The estimated number of returned photons per minute for HartRAO LLR was obtained by developing a mathematical tool which estimates the relationship between the returned number of photons and the varying and fixed link budget equation parameters. The mathematical tool is used to indicate “worse and optimal” parameter values which influence the return signal, presented as an estimate of expected number of returned photons for the HartRAO station. The link budget equation is expressed as, [20]

$$n_p = \eta_q \left(E_T \frac{\lambda}{hc} \right) \eta_t G_t \sigma \left(\frac{1}{4\pi R^2} \right)^2 A_r \eta_r T_a^2 T_c^2 \qquad (1.1)$$

where η_q is the quantum detector efficiency, E_T the total laser pulse energy, η_t the transmit optics efficiency, G_t the transmitter gain, σ the satellite/Moon reflector optical cross-sectional area, R the slant range, A_r the effective area of the receiving telescope aperture, η_r the receive optics efficiency, T_a is the one-way atmospheric transmission and T_c is the cirrus cloud cover.

An LLR photon estimator computer program, written in C++, is under development at HartRAO. It uses known and estimated parameters to determine the expected number of photon returns under various scenarios (see Table 2). This is all done to achieve signal path optimal efficiency to yield an improvement in the return-energy of the laser for accurate ranges to the corner cube retroreflectors.

Table 2: The relationship between varying parameters and the number of photons reflected from Apollo 11 Laser Ranging Retroreflectors.

Parameter	Worst value	Optimal value
Transmit optics efficiency	0.4	0.9
Slant range (km)	399929	347929
Detector quantum efficiency	0.4	0.7
Receive optics efficiency	0.4	0.9
Atmospheric transmission	0.02	0.81
Cirrus transmission (Cloud cover)	0.1	1
Returned photons/minute	0.003	12

The influence of the beam divergence on a transmitted laser beam reduces the number of photons that illuminate the corner cube reflectors and reflected back to the Earth 'fixed' receiving telescope. There is a two-way (uplink and downlink) spread on the laser beam as it propagates through the atmosphere under various scenarios. The variations on the returned number of photons depend on the initial beam diameter transmitted through the telescope, slant range, atmospheric transmission, cirrus cloud cover transmittance and the effectiveness of the lunar reflectors (see Figure 2 and Figure 3). The optimal values of atmospheric and cirrus cloud cover transmissions (see Table 2) are referred to, in this article, as favourable weather conditions (see Figure 3). These values, together with the other optimal link budget equation parameter values, yield a maximum number of returned photons.

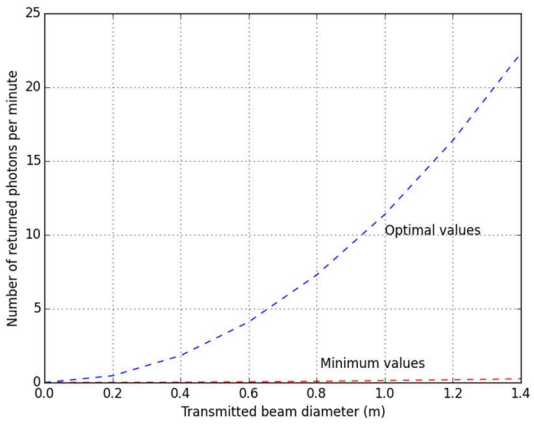


Figure 2: The dependence of the number of returned photons per minute on the output beam diameter and variable link budget equation parameters. The blue dotted line represents optimal values while the red dotted line represents minimum values.

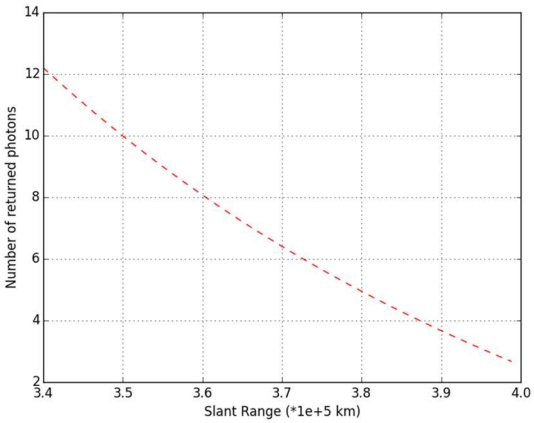


Figure 3: The expected number of returned photons per minute as calculated for the HartRAO LLR system under favourable weather conditions.

The number of returned photons linearly varies with the lunar reflector cross section. A study on the effective area of the corner cube has revealed that, at arbitrary incident angle, θ_{inc}, the area is reduced by the factor, [21]

$$\eta(\theta_{inc}) = \frac{2}{\pi}\left(\sin^{-1}\mu - \sqrt{2}\tan\theta_{ref}\right)\cos\theta_{inc}, \tag{1.2}$$

where the quantity $\mu = \sqrt{1-\tan^2\theta_{ref}}$, θ_{ref} is the internal refracted angle as determined by Snell's law. In this case, the refractive index for retroreflector's optical material type is $n = 1.45$ for fused silica. Thus, the peak optical cross-section in the centre of the reflected lobe decreases as the incident angle increases (see Figure 4) [21], i.e.

$$\sigma_{eff}(\theta_{inc}) = \eta^2(\theta_{inc})\frac{\pi^3\rho D^4}{4\lambda^2}, \tag{1.3}$$

where ρ is the reflectivity of the retroreflector which is typically equal to 0.78 for aluminium-coated back faces and 0.93 for uncoated Total Internal Reflection (TIR) surfaces, D is the retroreflector diameter and λ is the wavelength.

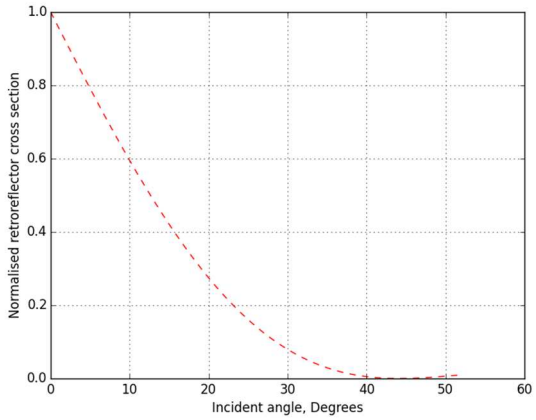


Figure 4: The normalised cross-section as a function of incident angle for fused silica retroreflectors.

5. Conclusion

We have successfully developed an integrated software based model that will enable optimal signal path efficiency for the HartRAO's LLR system [22, 23]. An integrated software based analytical model for the signal path efficiency of the HartRAO lunar laser ranger optical system is an essential

tool that can improve and ensure photon returns when ranging to the Moon. Our estimated signal return rate is a true reflection of the LLR photons returns.

Acknowledgements
The authors wish to thank the Department of Science and Technology (South Africa), Observatoire de la Côte d'Azur (France) and NASA-GSFC (United States of America) for their support.

References

[1] Combrinck L 2007 http://geodesy.hartrao.ac.za/pastevents/workshop2
[2] Combrinck L 2010 *Sciences of Geodesy-I: Satellite Laser Ranging* (Berlin, Springer)
[3] Combrinck L 2011 62nd IAC
[4] Kopeikin SM 2008 16th International Workshop on Laser Ranging
[5] Williams JG, Boggs DH, Turyshev SG and Ratcli JT 14th International Laser Ranging Workshop 155-161
[6] Williams JG and Boggs D H 2008 16th International Workshop on Laser Ranging
[7] Williams JG, Turyshev S G and Boggs D H 2009 Int. J. Mod. Phys. D **18** 1129-1175
[8] Anderson JD, Gross M, Nordtvedt K L and Turyshev SG 1996 Astrophys. J. **459** 365
[9] Müller J, Nordtvedt K, Schneider M and Vokrouhlicky D 2006 http://cddis.gsfc.nasa.gov/
[10] Merkowitz SM 2010 Living Rev. Relativity **13** 7
[11] Combrinck L 2014 19th International Workshop on Laser Ranging
[12] Cook RJ 1975 J. Opt. Soc. Am. A **65** 942-948
[13] Roddier F 1981 Prog. Opt. **19** 281-376
[14] Andrews LC and Phillips RL (Bellingham, SPIE)
[15] Wei HY and Wu ZS 2008 J. Elec. Wav. and App. **22**(5-6) 787-802
[16] Ndlovu SC and Chetty N 2014 Cent. European J. Phys. **12**(7) 466-472
[17] Degnan J 2003 J. Geodyn. **34** 551-594
[18] Prochazka I, Hamal K and Sopko B 2004 J. Mod. Opt. **51** 1289-1313
[19] Dirkx D, Noomen R, Prochazka I, Bauer S and Vermeersen L L A 2014 19th International Workshop on Laser Ranging
[20] Degnan JJ 1993 A Review, American Geophysical Union
[21] Degnan J 2012 17th international Workshop on Laser Ranging
[22] Ndlovu SC, Combrinck L, Exertier P, Akombelwa M and Chetty N 2014 19th International Workshop on Laser Ranging
[23] Ndlovu SC, Combrinck L, Exertier P, Akombelwa M and Chetty N 2015 South African J. Geol., **119**(1) 77-82

Division E – Education

The light bulb effect: University students' problem solving cognitive processes in a physics problem solving skills test

C Albers, D Clerk and D Naidoo

School of Physics, University of the Witwatersrand, Private Bag 3, WITS 2050

E-mail: claudia.albers@wits.ac.za

Abstract. A diagrammatic cognitive process representation, called the 'Light Bulb effect', with the ultimate intended purpose of teaching students problem solving skills through empowering them metacognitively, is used here for probing the question of what makes a problem difficult. The 'light bulb effect' is used, in this study, as a way to analyse students' problem solving skills on entering a first year Physics major course, at the University of the Witwatersrand. Analysis of student written answers to a test designed to probe some essential problem solving skills revealed that students had difficulty with the algebraic manipulation of equations, understanding displacement and taking logical steps that are unfamiliar or not immediately obvious. In this representation, these are called 'light bulbs'. The more linked 'light bulbs' are required in reaching a solution to a problem, the more difficulty the students have with solving that problem. This diagrammatical representation thus suggests a way of determining the level of difficulty of any physics problem, through the determination of the number of cognitive skill links and the number of not obvious logical steps, required to solve a specific problem.

1. Introduction

A problem is a situation facing an individual for which there is no obvious path to the solution so the problem solver has to create a path [1]. Thus problem solving is both a part of everyday life and an important aspect of learning physics. In physics, it is used as a learning strategy through which students get to practice the application of principles and concepts they are learning [2]. However, teaching students problem solving has met with very little success as this is usually done by a teacher, to whom the problem is not new, who presents the problem and the solution in a very clear and logical manner. The student who learns the solution will then apply it when he recognizes a similar problem. But because most teachers do not mention that a new or real problem usually needs many attempts at solving it [3], the student gives up when he faces a problem that he does not recognize [1]. However, successfully solving a problem that is new or unrecognized is dependent on metacognitive monitoring and students can improve their problem solving performance through metacognitive training [4, 5].

Metacognition is both knowledge of one's cognitive processes and regulation of those processes [6]. The specific skills used in a problem solving process are the essence of metacognition [7]. In order to be able to help students gain metacognitive tools and thus help them with problem solving it is important to first identify the different cognitive processes involved in problem solving and then determine some way of putting them together so that their interaction is clear. It is the aim of this study to identify the cognitive processes involved in physics problem solving and to determine which of these processes make a problem more or less difficult for students. Since 'one picture is worth one

thousand words', in this study, a diagrammatic cognitive process representation is designed and used to analyse problem solving cognitive processes.

There are several cognitive skills that play a role in solving physics problems. These skills can be divided into several basic categories:

1. Language skills of several forms needed to understand the question or problem description. These include the use of normal English words, terms specific to physics, pictorial representations and mathematical expressions.
2. Mathematical manipulations involving for example: Algebra, Arithmetic, Calculus, or Geometry.
3. Knowledge of physical principles or laws, for example: Newton's second law of motion.
4. Writing answers to the problem in some mathematical form or through the use of a definition, or a relationship.

A large number of problems encountered by both students and physicists at the forefront of physics require only the combination of several skills from the four above categories, in some logical sequence, for the successful invention of a path to the final solution. But there are also problems where a logical step or the 'figuring out of something' that is not immediately obvious is also required. This is when a problem solver looks at the several 'strings' of given information and at the various 'strings' of manipulations he can do, and suddenly realizes how to put it together in order to arrive at the answer. This is the *realization moment*. This is the 'light bulb' moment. This is the moment when it may seem as if a light bulb goes off inside his mind. It will be demonstrated in this study that students find these types of problems, where a 'light bulb' moment is required, much more difficult to deal with than the problems where only the linking of several skill categories leads to the answer.

2. Method and Analysis

A test was set to probe students' problem solving skills and in the process of setting and writing solutions for the test, the different cognitive processes in use in each of the questions was identified. The different processes were then represented as either a string or a bulb and linked in a diagram that attempts to represent the complete problem solving process with a diagram containing various diagrammatic elements. These elements are either a string, represented by a straight or a curved line, or a bulb, represented by a circle. The two bulb elements and an example of a string element are shown in Figure 1. Each of the string elements are accompanied by a code comprising of a letter and a subscript. All the string cognitive processes required in the test are shown in Table 1.

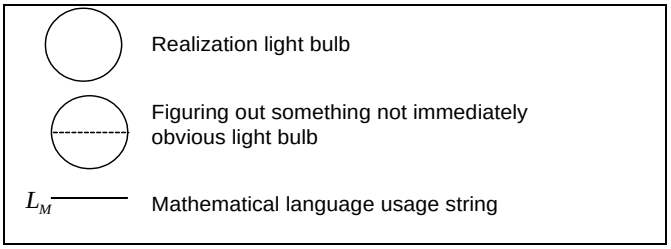


Figure 1: Three diagrammatic problem solving cognitive process representation elements

The number of students who wrote the test was 310 and a sample of 100 student test answers were analysed. The initial diagrammatic representation was modified slightly if the analysed answers showed that there were some possible additional links. The bulbs represent a cognitive process that requires greater mental effort and so are viewed in the analysis as representing a higher level process than the processes that involve only strings. Thus levels were assigned to the diagrams according to the number of bulbs in them. If a problem solving process did not contain any bulbs it was assigned a problem solving process level of 0 (L0). If there was one bulb it was assigned level 1 (L1), if it contained five bulbs, as the last question in the test, it was assigned level 5 (L5).

Table 1: String cognitive process elements involved in the problem solving cognitive processes relevant to the test given to the first year Physics major students, at the beginning of 2015

String code	Cognitive process description
L_E	Reading and understanding English language such as the language used in the wording of a question and usually includes terms specific to physics
L_M	Usage of mathematical language, usually used in reading and understanding given mathematical expressions
L_P	Reading and understanding information in a pictorial or graphical format
M_{Ar}	Use of arithmetic
M_{Al}	Algebraic mathematical manipulation
M_{Cal}	Mathematical manipulation using calculus
M_G	Mathematical process using Geometry
R_P	Remembered physical concept, law or principle
A_{DL_E}	Answer in the form of a definition using English words and terms specific to physics
A_{RL_E}	Answer in the form of a relationship using English words and terms specific to physics
A_M	Answer in mathematical language, may be a number and may contain variables. Needs to be simplified as far as possible

The first question in the Problem Solving Skills Test is shown in Figure 2. The diagrammatic representation, of the cognitive processes, required in solving this problem is shown next to the question. This question requires the use of the English language and physics terms (L_E), mathematical language (L_M) to read and understand the given mathematical expression, and the use of algebra (M_{Al}). These three strings need to be linked in order to arrive at the answer, which is in mathematical format (A_M). No realization or 'light bulb' process is required for this question so it is a level 0 (L0) process diagram. Notice that the progression of time is represented in going from left to right, with the answer appearing on the right or at the end of the time progression. But since the string processes can be combined in any order, the initial time point is the point where the three strings link.

The focal length, *f*, of a lens is given by the following equation:

$$\frac{1}{2a}+\frac{3}{5a}=\frac{1}{f},$$

where *a* is length and is measured in centimetres (cm). Solve for *f* in terms of *a*.

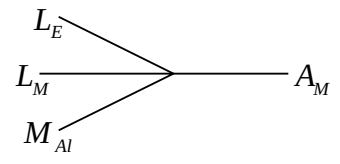


Figure 2: Question 1 in the Problem Solving skills test with the corresponding cognitive process diagram appearing on the right. This is a 4 string L0 diagram.

Question 2, shown in Figure 3, has five parts, so each is treated as a separate problem and a separate diagram is drawn for each part. There are two possible diagrams for parts (b) and (c) of question 2 because it was discovered after examining the students' test answers that some students knew and used calculus instead of going through the 'light bulb' process or realization, they would use if calculus was unknown to them. The realization required was that the given expression matched the given equations so that the initial speed and acceleration could be identified. The R_P string was added to the diagram, for part (e), after it was determined, from looking at the answers the students gave, that quite a few

students were not familiar with the term: displacement. This process was initially included in the L_E string as a normal physics term but it became obvious that it would be more appropriate to include it as a separate string process, namely as a 'remembered physical concept'. Parts (b), (c) and (f) contain L1 diagrams. The realization required in part (f) was that a straight line would be obtained from the given information only if *x/t* or *v* was plotted along the *y*-axis.

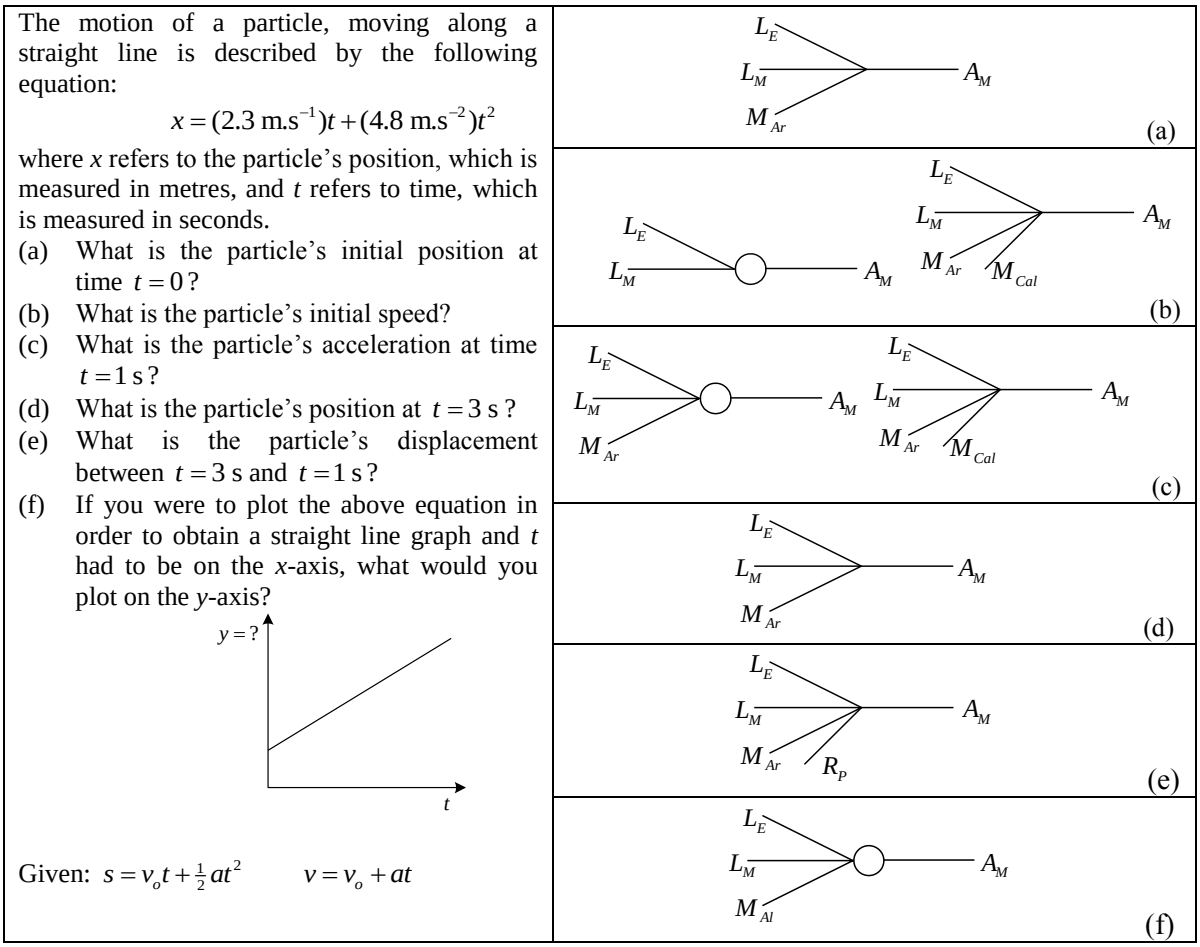
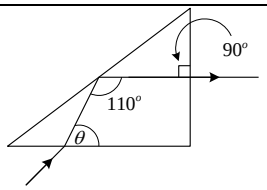


Figure 3: Question 2 in the Problem Solving skills test with the corresponding cognitive process diagrams, for the different parts of the question, appearing on the right.

Question 3, shown in Figure 4, requires a realization that the diagram in the question, shows a quadrilateral or a straight line intersecting two parallel lines. This realization is then combined with mathematical geometric and arithmetic cognitive processes, after which the solution emerges. The other processes required the understanding of English and physics terms (L_E) , and the understanding of pictorial language (L_P). The relevant cognitive process diagram, on the right in Figure 4, shows all these processes at a glance. It is L1 diagram with 5 string processes.

Figure 5 shows the final question, question 4, in the Problem Solving Skills Test. This is the only question where the 'figuring out' bulbs appear. Also it is the first time that bulbs are used in the answer part of the process. The strings going through the three middle bulbs are viewed as a single string each, so the diagram has 5 bulbs and 7 strings. The diagram suggests that this is the most complicated problem solving process the students face in the test.

A ray of light enters a right-angled prism through one of its faces, undergoes internal reflection on another face of the prism and exits through the third face, as shown in the figure below. Find the angle θ. Give reasons for your answer.



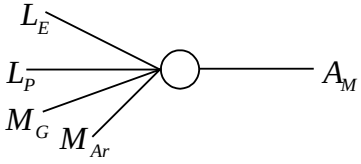


Figure 4: Question 3 in the Problem Solving skills Test with the corresponding cognitive process diagram appearing on the right. The diagram represents a L1 process

Write down an explanation of Newton's second law of motion that you would give to your little 6 year old sister who has not yet done any Physics at school. Make sure you define all the physics terms you use in a way that she can understand.

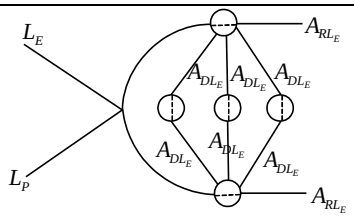


Figure 5: Question 4 in the Problem Solving skills Test with the corresponding cognitive process diagram appearing on the right. Lines without an assigned code are just linking different parts of the process.

3. Results

The results obtained from the analysis of the Problem Solving Test, given to first year Physics major students at the beginning of the academic year, in 2015, appears in Table 2. The first row in the table shows the question number. The second row shows the level of the question according to the 'light bulb' process representation, and the third row shows the maximum possible marks assigned to each question. The total number of marks in the paper was 25 and the number of student test answers analysed was 100. The number of students who obtained a perfect score for each question is shown in the fourth row. The fifth and sixth rows show the main reason why students did not obtain full marks for the question.

Table 2: Student performance in the Problem Solving Test at the beginning of 2015.

Question number	**1**	**2a**	**2b**	**2c**	**2d**	**2e**	**2f**	**3**	**4**
Question level	L0	L0	L1	L1	L0	L0	L1	L1	L5
Marks assigned	5	1	2	2	1	2	3	4	5
Number with perfect score	**72**	**99**	**69**	**21**	**73**	**65**	**36**	**63**	**8**
Main difficulty	M_{Al}	A_M	L1	L1	M_{Ar}	R_P	L1	L1	L1
No. with this difficulty	23	1	28	76	27	21	56	20	35

The main difficulty encountered by the students in answering question 1 was dealing with the required algebraic manipulation (M_{Al}). In question 2a, only 1 student had difficulty, and it was because he/she did not understand the expected way of writing the answer (A_M). He wrote in answer: '*At the start*'. In comparing the diagrams for questions 1 and 2a we see that both diagrams contain 4 strings but the diagram for 2a contains an Algebra string instead of an Arithmetic string. Algebra is a more abstract and therefore a more complex process making the question more difficult for the students. The main difficulty with question 2d was arithmetic (M_{Ar}) and in question 2e, it was remembering and understanding the displacement concept (R_P). In all the questions where a light bulb or realization cognitive process was required, the greatest difficulty was achieving realization (L1). In question 4, where 5 light bulbs, were required, 35 students demonstrated they could not achieve a single bulb process and 18 were able to mentally construct one of the required bulbs. One of these 18 students

wrote: '*When you push something it will move in the direction you are pushing*'. This student was able to define force, in words a six year old child could understand, by using the word 'push' and therefore achieved one bulb cognitive process.

Student performance in the four L0 questions, namely 1, 2a, 2d and 2e, was better than in the L1 questions, namely 2b, 2c, 2f and 3. The best score was obtained in question 2a, where students only needed to use very simple arithmetic (M_{Ar}) in combination with understanding the wording (L_E) and the given expression (L_M), in the question, as seen from the top process diagram in Figure 3. The L1 question with the lowest number of students obtaining a perfect score is 2c, where the process seems at first similar to that of 2b but a possible realization is using the units ($m.s^{-1}$ in 2b) rather than comparing equations. This realization works in 2b but not in 2c. There was also possibility of answering using calculus, which changes the cognitive process from a L1 to a L0, with the addition of the M_{Cal} string. Very few (5) students used calculus. Other L1 questions, 2f and 3, also demonstrate that students find L1 questions more challenging then L0 questions as the number of students obtaining a perfect score was 36 and 63, respectively, which is below the numbers obtained for L0 questions. The question with the lowest number of perfect scores was question 4, a L5 question, where only 8 students obtained a perfect score.

4. Discussion and Conclusion

The analysis of the Problem Solving Test at the beginning of 2015 summarized in Table 2, shows that students perform the worst on questions that require a higher level cognitive process as determined by the Light Bulb diagrammatic representation. This representation gives, at a glance, information on the cognitive processes required by different questions so that it is easy to highlight differences between questions. For example: When a comparison is made between the process diagrams for questions 1 and 2d, we see that the difference is that only arithmetic is required in 2d but question 1 requires some algebraic manipulation (M_{Al}). Also, the very low student performance in question 4 is not surprising, when one draws a process diagram for it, which requires 5 bulb processes, and therefore a corresponding greater cognitive effort from the students. Thus, the 'light bulb' effect representation can be used to determine why students find one question more difficult than another, and one of its more powerful applications may turn out to be in teaching metacognitive skills which will hopefully lead to better problem solving performance.

References

[1] D. Gil-Perez, A. Dumas-Carre, M. Caillot & J. Martinez-Torregrosa, "Paper and pencil problem solving in the physical sciences as a research activity," Studies in Science Education, **18**, 137-151 (1990).

[2] R. M. Garrett, "Problem-solving in science education", Studies in Science Education, **13**, 70-95 (1986).

[3] Martinez, M. (1998). What is problem solving? Phi Delta Kappan, 79, 132. Retrieved from http://www.questia.com/PM.qst?a=o&se=gglsc&d=5001331555.

[4] K. Teong, "The effect of metacognitive training on mathematical word-problem solving", Journal of Computer Assisted Learning, **19**, 46-55 (2003).

[5] T. Gok, "The general assessment of problem solving processes and metacognition in physics education", Eurasian Journal of Physics and Chemistry Education, **2**, 110-122 (2010).

[6] J. Flavell, Metacognitive aspects of problem solving. In L.B. Resnick (Eds.), *The nature of intelligence*, pp.231-236 (New Jersey: Lawrence Erlbaum Associates, 1976).

[7] M. Goos, P. Galbraith, P. Renshaw, "Socially mediated metacognition: Creating collaborative zones of proximal development in small group problem solving", Educational Studies in Mathematics, **49**, 193-223 (2002).

Effect of guided inquiry laboratory activities on first-year physics students' views on the nature of science

V M Baloyi[1], W E Meyer[1] and E Gaigher[2]
[1] Department of Physics, University of Pretoria, Private bag X20, Hatfield, Pretoria, South Africa, 0028
[2] 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 the effect of 'explicit reflective guided inquiry' (ERGI) versus traditional recipe-based laboratory practical activities on first-year university students´ views on the nature of science (NOS). The ERGI laboratory activities adopted the 'Physics by Inquiry model'. Ninety seven first-year Bachelor of Science physics students participated in this study at a well-established South African university, which compared guided inquiry and traditional recipe-based laboratory approaches. The students were divided randomly in a control group that did traditional recipe-based practical activities and an experimental group, that did guided inquiry-based practical activities. Both groups had the same reflective questions on an aspect of the NOS at the end of the practical activities. At the end of the practical course, data were collected using the VNOS-Form C questionnaire and follow-up focus group interviews were conducted. Additionally, there was also a practical test that consisted of both a hands-on and a written section. The results showed that the experimental group students developed better understanding of three aspects of NOS: tentative, empirical and difference between observations and inferences. However, students' conceptions on the difference between theory and law, role of imagination and creativity, influence of social and cultural values and notion of using universal scientific method in the development of knowledge were similar for the two groups. This study demonstrated that there is significant effect of guided inquiry-based laboratory practical activities on the students' understanding of NOS in first-year physics course compared to traditional laboratory approach.

1. Introduction

This study is informed by the goals of the Introductory Physics Laboratory of the American Association of Physics Teachers and the National Curriculum Assessment Policy Statement (1- 4) which encourage Inquiry-Based Science Education (IBSE) which is believed to enhance the deep understanding of nature of science (NOS), scientific inquiry (SI) process skills and science literacy [5]. NOS and SI are different constructs but they are intertwined with regard to the goals of current science education [6]. Teaching students how to conduct SI engages students in the practices of science including conducting scientific investigations and performing laboratory practical activities like scientists to address questions and formulate explanations using creative and critical thinking [3]. When students are engaged in scientific investigations and practical activities, they use observations and inferences to formulate conclusions and empirically based explanations [7]. Understanding of the difference between observations and inferences, informed conceptions of the tentativeness,

subjectivity, difference between theory and law, and role of social and cultural values associated with the development of scientific knowledge are aspects of NOS [8] which are also linked with conceptions about SI [6].

The current study was undertaken at a South African university where the laboratory component of an undergraduate physics course has been redesigned from a traditional to an inquiry-based format. The rationale behind this transformation is that the current practice of doing the practical laboratory activities is procedural [9] which, according to experience, does not enhance students' thinking skills [10]. Conversely, the ERGI laboratory activities may bring about changes in students' thinking and problem-solving skills [11]. It is believed that asking intellectually guiding questions in inquiry-based activities directs students' thinking and helps a teacher to understand their thinking [12]. Here the ERGI laboratory activities served as a context to the NOS questions [13].

Socrates used questioning in teaching to assist a student to think, analyse and seek for new information [14]. Such questioning may assist the development of understanding NOS, which is considered as knowledge that needs to be understood rather than to be known [15]. This study has explored the degree to which explicit-reflective questions integrated with the ERGI laboratory activities, as suggested by Clough [15], influence the students' views on NOS. Explicit-reflective instruction includes: focusing students' attention to targeted NOS aspects through discussion and written work, encouraging students to both think about how their investigations are linked to the NOS features and how their ERGI laboratory activities resemble the work of scientists [16]. We investigate if students may develop a better understanding of targeted NOS aspects when they are combined with physics content to be learnt in the ERGI laboratory activities [6]. These correlations have not yet been experimentally studied in the South African context, where students are traditionally taught via a recipe-based method. The results of this study may enhance the understanding of current science laboratory practices, learning processes and the potential effects of inquiry-based instruction at university level.

Inquiry has a wide range of overlapping meanings among different scholars. Inquiry is broadly defined as scientific investigations that encourage classroom practices such as posing questions which focus at knowledge attainment and development [17]. IBSE originated from the deep conceptualisation of the scientific practice and processes [18]. The fundamental principles of IBSE include discovering natural laws, relating information into real-life contexts, developing critical thinking, and encouraging positive attitudes towards science learning [19]. The four levels of inquiry include: confirmation, structured, open, and guided inquiry [20], which were formulated to guide teachers in scaffolding students' inquiry activities. This study is informed by Physics by Inquiry model (PbI) in which students' learning is facilitated by guiding questions and the teaching approach utilised is guided inquiry [21]. In PbI, students are guided through step-by-step questions and activities and make observations which they may use to formulate their models in physical sciences. The purpose of questioning encourages the development of thinking and reasoning skills in students [21]. Our research question for this study was: *To what extent does explicitly reflective guided inquiry-based instruction in practical laboratory activities influence first-year Physics students' views on the nature of science?*

2. Literature review

A previous study examined metacognition as means to increase the effectiveness of inquiry-based science education [22]. The results demonstrated that metacognitive reflection integrated with inquiry had an effect in the teaching of scientific process and scientific thought and resulted in preparing students to be better critical thinkers and more scientifically literate. Another study explored the effect of reflective discussions following inquiry-based laboratory activities on students' views of NOS [23]. Results indicated that the inquiry-based laboratory activities enhanced students' views of the targeted NOS aspects more than implicit inquiry-based instruction.

The explicit–reflective approach has been used in professional development programs and has enriched practicing elementary teachers' views of NOS [24]. It was demonstrated that after using

explicit reflective instruction in inquiry-based context, elementary students developed informed views of NOS [25]. Vhurumuku [26] examined the influence of a short explicit-reflective NOS course on undergraduate science students. The participants' views changed from naïve to reasonably informed. Another study explored the effects of a semester-long NOS course on two teachers' conceptions of NOS using the discursive approach which drew from History, Philosophy and the Sociology of Science [27]. The results showed that teachers' views of NOS changed from inadequate to informed. Abd-El-Khalick's [28] exploratory and interpretive study examined: (a) the influence of an explicit reflective activity-based approach implemented in the context of a science content course, (b) the ability of participants to apply the acquired NOS. The results showed that participants were able to translate NOS understandings in the context of atomic structure covered in the course to that in dinosaur extinction.

Research on gender differences shows that boys' experiences in science and mathematics are different from those of girls throughout their life, which causes confusion and low confidence among girls [29]. Additionally, from an early age, girls show a lower level of interest in physics as compared to their male counterparts. One study tracked the performance differences across gender differences of learners from grade 8 [30]. The results showed an increase in performance differences between males and females. Usually the average early yearly increase of males for academic performance is always greater than that of females. In the tenth grade female learners' attitudes toward science learning start to increasingly become negative [31]. Research shows that by the end of the fifth grade, girls usually show less interest in many areas of physics than boys [32] due to other factors such as teacher behaviour and traditional gender stereotypes [33].

Another study on gender differences regarding academic performance shows that females have better learning vocabulary, grammar, and writing skills than males [34]. It is believed that stereotypical beliefs relating to sex roles and social expectations may have encouraged females to have more positive attitudes towards languages and reading [35]. Conversely, males outperform females in standardised tests in physics because they have better spatial skills. These results have been attributed to higher numerical and spatial abilities of boys [36], coupled with social expectations for boys to have positive attitudes toward science and mathematics [35].

3. Methodology

The current study followed an experimental design using a mixed methods approach [37], located within the positivism paradigm. Positivism allowed us to explore the cause and effect of ERGI laboratory activities on students' views on the NOS [25].Convenience sampling was used as the study intended to examine the effect of the ERGI laboratory activities at a specific university on students' views on the NOS [37]. Participation was voluntary and students could withdraw at any time. The population was comprised of 220 first year BSc students. Of these, only, 97 students consented to participate in the study and were randomly assigned to either experimental or control groups [37]. The experimental group did the ERGI laboratory activities designed using PbI [21]. The control group did the traditional recipe-based laboratory activities. Students in both groups worked in small groups and the intervention programme comprised eight weekly practical sessions.

The VNOS-Form C questionnaire [9] was used to evaluate students' views on NOS at the end of the practical course. Collected data was analysed using the guidelines in Lederman, Abd-El-Khalick, Bell and Schwartz [9]. The description frame for NOS aspects by Bell, Lederman and Abd-El-Khalick [38] was used as analysis criteria. A scoring scale used to evaluate students' responses: A score of 1 represented naïve views, 2 represented mixed views and 3 represented informed views of NOS. Focus group interviews were used to validate findings from the questionnaire [9]. VNOS data were sorted randomly and analysed by three researchers to establish reliability [9]. Qualitative data was translated into scores that were analysed quantitatively using descriptive statistics [39]. Numerical test scores of open-ended VNOS questionnaire of both the experimental and control groups were added to obtain an average score per individual. Scores were also statistically analysed and compared.

4. Results and discussions

Figure 1 shows the students' scores per NOS aspect, for the control and experimental groups. The students in the experimental group performed slightly better than students in the control group in all the aspects of NOS. However, in no category was the difference between the two groups statistically significant, but the overall (average) score did show a significant difference (p = 0.006). We attribute the difference in VNOS score to the ERGI laboratory activities which guided students in the discovery of science knowledge. For example, at the end of the alternating current practical activity, students were asked the following explicit reflective question: *In this experiment, you were guided by the practical notes. Do you think that all scientists use one scientific method when developing scientific knowledge?* One student (V26) gave the following informed view: *No they use different methods according to their different ways of thinking and analysing their practical notes or other scientific practical notes.* This finding agrees with previous research that revealed that explicit-reflective instruction is more effective than implicit instruction in enhancing learners' views of the NOS [13].

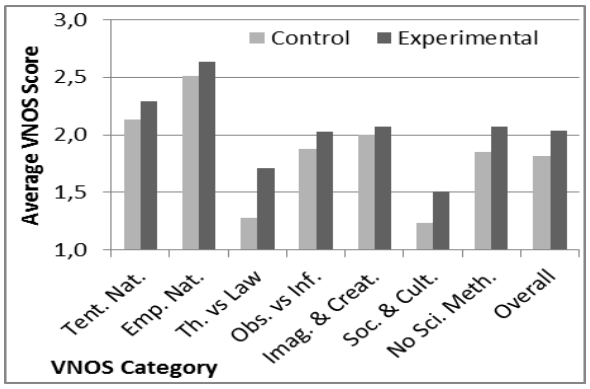


Figure 1: Students' VNOS score averages on seven aspects of NOS.

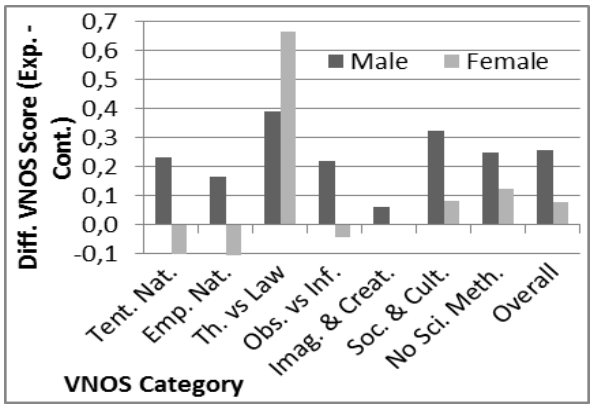


Figure 2: Difference between experimental and control groups' VNOS scores, shown separately for males and females

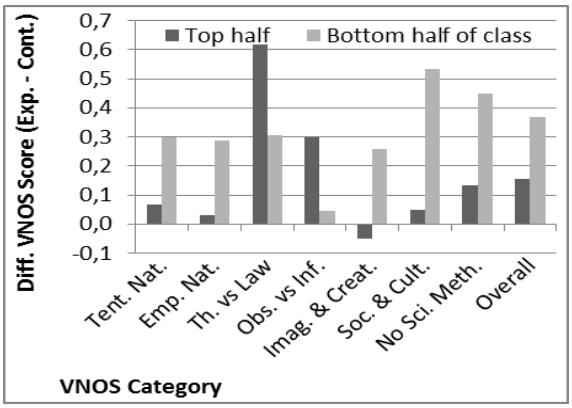


Figure 3: Difference between experimental and control groups' VNOS scores, shown separately for the top and bottom halves of the sample.

Figure 2 shows the male and female students' scores per NOS aspect, expressed as the difference between scores of the experimental and the control groups. Generally, males seemed to have benefited most from the ERGI laboratory activities, with overall scores significantly better for the experimental group (p = 0.006) while the benefit in females was not significant (p = 0.3). In all categories, with the exception of understanding the difference between theory versus law, the males benefited more from the ERGI laboratory activities. This agrees with a previous study [40], where inquiry-based laboratory activities were also found to be differentially effective for males and females. Gender differences in academic performance have been reported to be linked with science skills and other factors associated with learning experiences [41]. Research shows that males are more interested in fixing objects, building models, and searching for action-oriented activities while females prefer life science activities

comprising the development and caring of plants and animals [42]. Additionally, research has shown that gender differences with regard to academic performance start to increase throughout the secondary school [43]. We ascribe the significant enhancement in the understanding of the difference between theories and laws by the females to the above-mentioned finding that females are generally better with language skills, and would therefore better understand the difference in semantics.

Figures 3 compares the academic top and bottom half of the sample across the NOS aspects. The vertical axis indicates the difference between the scores of the experimental and the control group. Overall, the top half did not show a significant effect (p = 0.33) while the bottom half did show a strong effect (p = 0.01). In all categories, except for the difference between theory and law, the bottom half of class showed a much stronger effect than the top half of the class. The bottom half (low performing students) showed a significant effect on the understanding of social and cultural values (p = 0.03) and in the understanding of scientific methods (p = 0.02). We believe that this difference in performance might be ascribed to the guided inquiry questions being aimed at the academically poorer performers. This was done to ensure that most students would be able to perform the practical activities. High performers could foresee the answers, and therefore benefitted less from the guiding of their thought processes. The exception to the rule was the understanding of the difference between theories and laws, which is, as mentioned earlier, a more complex concept, and therefore it was grasped better by the higher performers.

The findings in this study show that many students had incomplete understanding of NOS. Many students did not understand the difference between theory and law and the role of social and cultural values, but understood the tentative and empirical nature of science. Many students showed an informed view that scientists are not certain of the atomic model and that it might change in the light of further discoveries. For example, during the focus group interviews student (V8) expressed the following informed view: *Scientists have refined their model of the atom over many years. At the moment scientists are quite sure that the model explains the known data. The alpha back scattering experiment was used to find that the atom consisted of a positively charged nucleus and negative particles far away from the nucleus.* The ERGI laboratory activities showed a positive effect on all NOS aspects, especially the two weakest points, i.e. understanding the difference between theory and law and the social and cultural embeddedness. Generally, the males seem to have benefited more than females from the ERGI laboratory activities, with the exception of understanding the difference between theory and law, where female students did much better. Academically weaker students seem to have benefited more from the ERGI laboratory activities than high performing students with the exception of theory and law, and observation and inferences. These results are consistent with the findings in other studies: [13, 22, 23, 26, and 28].

The results of this study should not be generalized to all first year undergraduate Physics courses as the entrance requirements and also the type of Physics courses may have different essence and motivation at different institutions. However, results could be generalized within the population of first-year physics students in the university from which the sample was drawn. It may be concluded that the experimental group showed better understanding in all NOS aspects and statistically significant difference overall than the control group under conditions where reflection was part of practical content. Evidence-based insight into first-year physics students' views on the NOS may help in transforming undergraduate science courses and may contribute to faculty and researchers' understanding of the perceptions of science held by undergraduates, assisting university lecturers to improve scientific literacy in future scientists and diverse university graduates.

Acknowledgements
Funding was provided by the Sasol Inzalo Foundation.

References

[1] Department of Education 2012 Curriculum and Assessment Policy Statement: Physical Sciences Grade 10 – 12 (Pretoria: Government Printer)

[2] Feynman R 1998 *Am. J. Phys* **66** 483
[3] National Research Council 2012 A *Framework for K-12 Science Education Practices, Crosscutting Concepts, and Core Ideas* (Washington, DC: National Academy Press)
[4] Next Generation Science Standards Consortium of Lead States 2013 *Next generation science standards by states, for states* (Washington, DC: National Academies Press)
[5] Laugksch R C 2000 *Sci. Educ.* **84** 71
[6] Schwartz R S, Lederman N G and Crawford B A 2004 *Sci. Educ.* **88** 610
[7] American Association for the Advancement of Science 1989 *Science for all Americans* (New York: Oxford University Press)
[8] Lederman N G, Abd-El-Khalick F, Bell R L and Schwartz R S 2002 *J. Res. Sci. Teach.* **39** 497
[9] Banchi H and Bell R 2008 *Sci. Child.* **2** 26
[10] McDermott L C 1993 *Am. J. Phys.* **61** 295
[11] McDermott L C, Heron P R L, Shaffer P S and Stetzer M R 2006 *Am. J. Phys.* **74** 763
[12] Clough M P 2007 *Iowa. Sci. Teach. J.* **34** 2
[13] Khishfe R and Abd-El-Khalick F 2002 *J. Res. Sci. Teach.* **39** 551
[14] Crockett C 2004 *Educ. Leadership.* **61** 34
[15] Clough M P and Olson J K 2008 *Sci. Educ*.**17** 143
[16] Abd-El-Khalick F 2005 Modelling science classrooms after scientific laboratories *NSF Inquiry Conf. Proc.*
[17] Blanchard M R, Southerland S A and Granger E M 2008 *Sci. Teach. Educ.* **93** 322
[18] Narode R, Heiman M, Lochhead J, and Slomianko J 1987 Teaching thinking skills *Science* (Washington, DC: National Education Association)
[19] Kyle W C Jr, Bonnstetter R J, McCloskey J and Fults B A 1985 *Sci. Child.* **23** 39
[20] Bell R L, Smetana L and Binns I 2005 *Sci.Teach*.**72** 30
[21] McDermott L C and the Physics Education Group at the University of Washington 1996 *Physics by Inquiry, Vols. I and II* (New York, NY: John Wiley & Sons Inc)
[22] Seraphin K D, Philippoff J, Kaupp L and Vallin L M 2012 *Sci. Educ. Int.* **23** 366
[23] Yacoubian H A and BouJaoude S 2010 *J. Res. Sci. Teach.* **47** 1229
[24] Akerson V L and Abd-El-Khalick F S 2003 *J. Res. Sci. Teach.* **40** 1025
[25] Akerson V L and Donnelly L A 2010 *Int. J. Sci. Educ.* **32** 97
[26] Vhurumuku E 2010 *Afr. J. Math. Sci. Tech. Educ.* **14** 99
[27] Ogunniyi M B 2006 *Afr. J. Math. Sci. Tech. Educ.* **10** 93
[28] Abd-El-Khalick F 2001 *J. Sci. Teach. Educ.* **12** 215
[29] Linn M C 1980 *Sci. Educ.* **64** 237
[30] Bacharach V R, Baumeister A A and Furr R M 2003 *J. Genet. Psychol.* **164** 115
[31] Alexakos K and Antoine W 2003 *Sci. Teach.* **70** 30
[32] Hoffmann L 2002 *Learn. Instr*. **12** 447
[33] Haussler P and Hoffmann L 2002 *J. Res. Sci. Teach.* **39** 870
[34] Yarborough B H and Johnson R A 1980 *J. Psychol.* **106** 55
[35] Meece J, Bower Glienke B and Burg S 2006 *J. School. Psychol.* **44** 351
[36] Buzhigeeva M I 2004 *Russ. Educ. Soc.* **46** 76
[37] Creswell J W 2009 Research Design *Qualitative, Quantitative and Mixed Methods Approaches* (3rd Ed.) (Thousand Oaks, CA: Sage)
[38] Bell R L, Lederman N G and Abd-El-Khalick F 2000 *J. Res. Sci. Teach.* **37** 563
[39] Caracelli V J and Greene J C 1993 *Educ. Eval. Policy. An.* **15** 195
[40] Fraser B J and Wolf S J 2008 *Res. Sci. Educ.* **38** 321
[41] Linn M C and Pulos S 1983 *J. Educ. Psychol.* **75** 86
[42] Farenga S J and Joyce B A 1997 *Educ.* **117** 563
[43] National Center for Education Statistics (NCES) 2000 *The NAEP Science Scale* Retrieved June 29 2005 from http://nces.ed.gov/nationsreportcard/science/scale.asp

Division F – Applied Physics

Performance Comparison between the Traditional Intensity Modulation Direct Detection and Coherent Detection in a High Speed Optical Fibre Communication System

T V Chabata[1], D Kiboi Boiyo[1], E K Rotich Kipnoo[1], R R G Gamatham[2], A W R Leitch[1], T B Gibbon[1]
[1]Physics Department, Nelson Mandela Metropolitan University, P.O. Box 77000, Port Elizabeth 6031, South Africa. Tel: +27 41-504-2141.
[2] The Square Kilometre Array (SKA) Project, South Africa.

Email: Tichakunda.Chabata@nmmu.ac.za

Abstract: - Passive optical networks (PONs) have become a dominant approach for the fibre-to–the-home (FTTH) network deployments. Reliable and spectral efficient technologies are a necessity for extending the unamplified transmission reach in the FTTH environment. An optical coherent detection scheme that supports higher modulation formats and increases the receiver sensitivity is implemented. A 10 Gb/s data stream, intensity modulates a 1550 nm distributed feedback (DFB) laser with a direct detection scheme. The same modulated signal is enhanced by mixing it with a continuous wave (local oscillator) placed at the receiver in a homodyne coherent detection scheme. The enhanced mixed signal is then demodulated to evaluate and to compare the link performance of the direct detection and coherent cases. A back to back and a transmission through 30 km single mode fibre were simulated for the two transmission modalities. The coherently detected scheme improved the receiver sensitivity by 12.1 dBm at an acceptable bit error ratio (BER) of 10^{-9} as compared to the traditional intensity modulation direct detection (IMDD) scheme.

1. Introduction

The recent injection of new communication applications coupled with an exponential increase in users has further congested the already deployed intensity modulation direct detection (IMDD) optical communication networks [1]. One of the important goals in optical fibre communication system is to transmit the highest data throughput over the longest distance without costly signal regeneration. Amplitude modulated coherent detection scheme that has advantages of increased receiver sensitivity, increased transmission reach and is more spectral efficient is considered an alternative to meet the current communication requirements in a passive optical network in the fibre-to-the-home (FTTH) technology [2]. The coherent detection scheme can be incorporated into the already deployed fibre communication network thereby reducing the cost by maintaining the technology and form of the existing communication links.

In this paper, a traditional intensity modulated direct detection (IMDD) scheme is compared to intensity modulated coherent detection scheme (IMCD). Coherent detection is a modification of the tradition IMDD by beating the modulated signal with a continuous wave signal of the same frequency at the receiver. In a homodyne coherent scheme, phase matching of the transmitted signal and local oscillator is of vital importance in enhancing the quality of the signal at the receiver. Bit error ratio (BER) measurement which is an effective and reliable quantitative evaluation technique was used to compare the performance of the two communication modalities [3]. The receiver sensitivity at an acceptable BER of 10^{-9} and the corresponding penalties after transmission through a G. 652 fibre were determined.

2. Theory

A photodetector in an optical communication link is responsible for demodulating the optical signal to electrical form. Photodetectors are square law detectors. They convert the incident optical power (P_{in}) into a current and transform the incident electric field by a square law into an electrical power. The electric field of the received modulated signal in a direct detection scheme is expressed as:-

$$E_S(t) = A_S Exp[-i\omega t + \theta_s] \quad (1)$$

where A_s, ω and θ_s represent the amplitude, frequency and phase of the received signal.

In a direct detection scheme, the detected photocurrent I_p is given by:

$$I_P = \Re . P_{in} = \Re . A_{eff} . \frac{1}{2}\sqrt{\frac{\varepsilon_0}{\mu\mu_0}} . E_0^2 = K . E_0^2 \quad (2)$$

The electrical power, P_{el} generated in the detector with the Ohmic resistor is then given by:-

$$P_{el} = R . I_P^2 = R . (\Re P_{in})^2 \quad (3)$$

where R is the resistance and $\Re$ is the receiver sensitivity.

Coherent demodulation involves beating in a photodiode the received optical signal with a continuous wave (CW) signal. The electric field for the signal from the LO can be expressed as [3]:-

$$E_{LO}(t) = A_{LO} Exp[-i\omega_{LO} t + \theta_{LO}] \quad (4)$$

where A_{LO}, ω_{LO} and θ_{LO} represent the amplitude, frequency and phase of the LO signal.

Considering identical polarization on both signals, the total optical power into the photodetector can be expresses as:-

$$P_{(t)} = P_S + P_{LO} + 2\sqrt{P\, P_{LO}} \cos(\omega_{IF} + \theta_S - \theta_{LO}) \quad (5)$$

where:- $P_s = KA^2_s$, $P_{LO} = KA^2_{LO}$ and $\omega_{IF} = \omega_s - \omega_{LO}$, P_s is the optical power of the received signal, P_{LO} is the optical power of the LO, K is a constant of proportionality and ω_{IF} is the intermediate frequency. The photocurrent is proportional to the incident power and can be expressed as:-

$$I_{(t)} = RP_s + RP_{LO} + 2\sqrt{P_S P_{LO}} \cos(\omega_{IF} + \theta_S - \theta_{LO}) \quad (6)$$

Considering only the AC-coupled output of the photodiode and assuming that $P_{LO} >> P_s$, the above equation simplifies to:-

$$I_{(t)} = 2R\sqrt{P_S P_{LO}} \cos(\omega_{IF} + \theta_S - \theta_{LO}) \quad (7)$$

From (7) it is clearly shown that signal information can be retrieved from the amplitude (P_s), frequency (ω_s) or phase (θ_s) of the optical signal. Depending on the frequency offset between the received signal and optical signal from the LO, coherent detection can be implemented in two forms known as homodyne and heterodyne schemes. In homodyne coherent detection, $\omega_{IF} = 0$, and in heterodyne coherent detection, $\omega_{IF} \neq 0$. *[3,4,5]*

An optimum Homodyne detection scheme is obtained when the photocurrent produced is a maximum. The maximum photocurrent is obtained by matching the frequencies ($\omega_{IF} = 0$) and the phase ($\theta_s - \theta_{LO} = 0$), of the two signal giving [6,7,8] :-

$$I_{max} = 2.R\sqrt{P_{LO} P_S} \quad (8)$$

The ratio of the coherently detected photocurrent (I_{CD}) to the directly detected (I_{DD}) is therefore:-

$$\frac{I_{CD}}{I_{DD}} = 2\sqrt{P_{LO} / P_S} \quad (9)$$

Equation (9) shows that the photocurrent in a coherently detected scheme is more than that in direct detection thereby improving the receiver sensitivity.

3. Research Design

The schematic diagram for the theoretical simulation using VPI Photonics software is shown in Figure 1 [4]. The scheme shows a complete optical transmission link comprising mainly of a transmitter, the transmission medium, local oscillator (LO) and the optical receiver. The intensity modulation direct detection scheme is performed when the LO is switched OFF and the coherent detection scheme is performed with the LO switched ON.

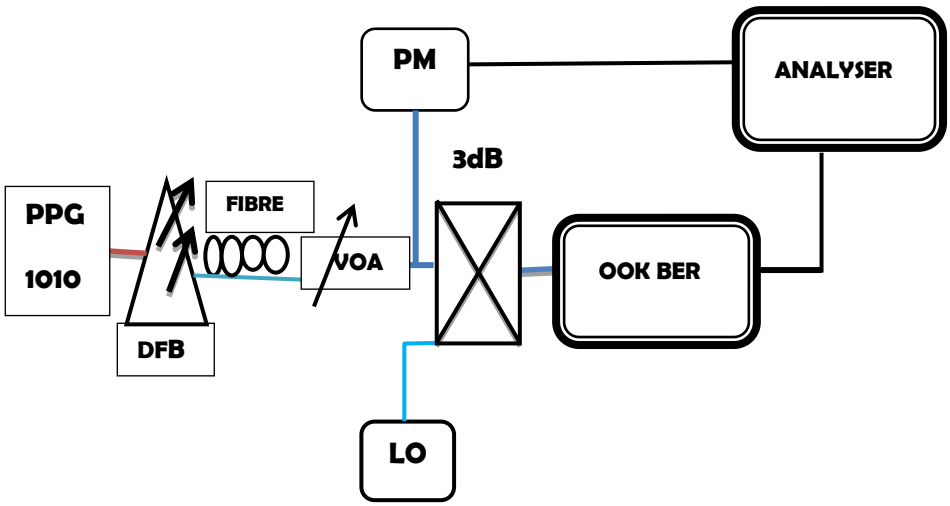


Figure 1: Simulation setup for an intensity modulation direct detection (IMDD) and intensity modulation coherent detection (IMCD) scheme

PPG: programmable pattern generator. DFB: distributed feedback laser. VOA: variable optical attenuator power PM: power meter. OOK BER On-Off keying bit error rate analyser.

The DFB laser signal at 1550 nm is directly modulated at a bit rate of 10 Gb/s by a Non-Return-to-Zero (NRZ) Pseudo-Random Binary Sequence (PRBS) signal and was launched into an optical fibre for transmission to the receiver. The variable optical attenuator (VOA) emulates the actual transmission link by varying the optical power that reaches the receiver and the signal power is measured by the power (PM). The direct intensity modulated signal is directly detected (DD) by the photodiode (PD) in the ON-OFF Keying BER (OOK BER) module. The OOK BER module computes the BER at different attenuation values and the result is plotted against the received optical power in the analyser module. For the coherent detection scheme the local oscillator is switched ON, the modulated signal is mixed with the continuous wave signal before detection at the receiver. The LO signal and the modulated signal are phase matched by first running a sweep at different initial phase angles of the LO thereby noting the best transmitted signal.

4. Results and Discussion

For both direct detection and coherent detection, a 3 mW signal is directly modulated with a 2^7-1, pseudo random bit sequence (PRBS). In a IMDD setup, the same signal is transmitted and is detected at the receiver. In a coherent detection scheme a 4 mW LO signal is mixed with the modulated signal before detection by a photodiode. The beating increases the optical signal power that the receiver gets resulting in improved receiver sensitivity.

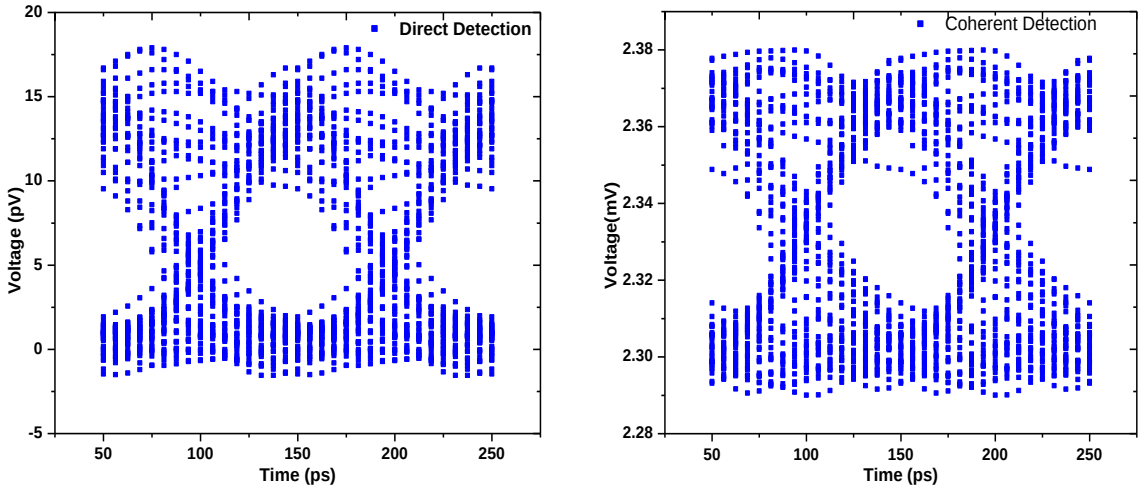


Figure 2: Eye Diagram for both direct detection and Coherent detection

Figure 2 shows eye diagram plots for back to back transmission for both IMDD and IMCD at an acceptable bit error ratio of 10^{-9} on a back to back transmission respectively. An eye diagram a qualitative and inituitive way to evaluate the signal quality. The more open the eye the better is the quality of the received signal .The IMDD transmission shows an open eye with a peak to peak voltage of 15 pV compared to a peak to peak voltage of 0.07 mV for the IMCD setup. The coherently detected signal is five fold greater than the traditional directly demodulated signal. This helps in enhancing the weak signal thereby improving the receiver sensitivity and increasing transmission reach.

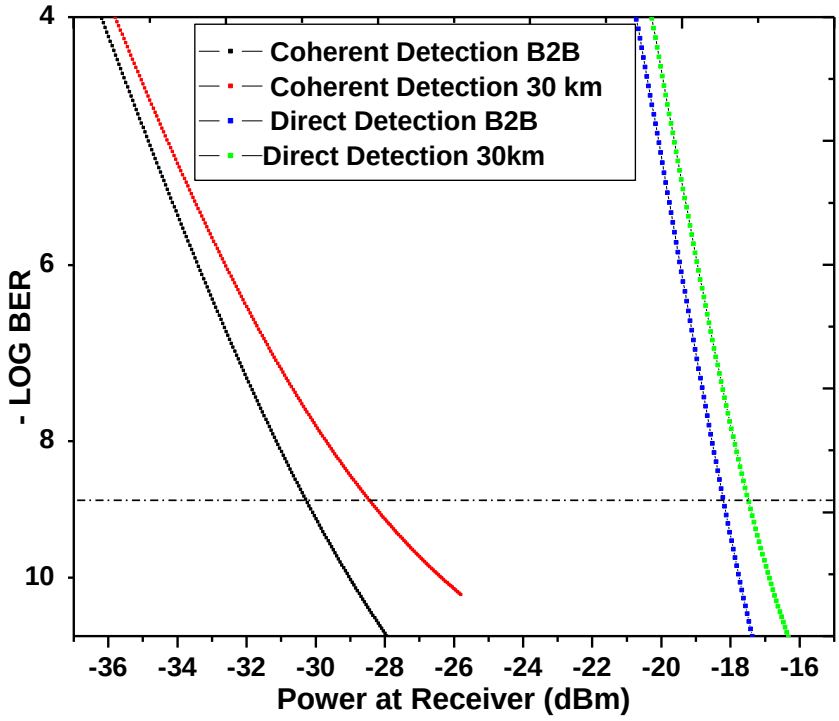


Figure 3: Bit error rate (BER) curves for back to back and after 30 km of G652 fibre

Figure 3 shows the BER curve for back to back and a transmission through a 30 km G. 652 fibre. A receiver sensitivity of -18.3 dBm and -30.4 dBm is obtained for IMDD and IMCD respectively. The IMCD has a 12.1 dB gain in receiver sensitivity compared to the IMDD. A penalty of 1.9 dB is obtained for the coherent detection scheme compared to 0.8 dB for direct detection scheme after transmission through 30 km of fibre. Power budget calculations indicate an unamplified transmission distance of 135 km for coherent detection scheme and 75 km for direct detection. In a PON the power

budget gives a splitting ratio of 1:126 and 1:1994 for IMDD and IMCD respectively. The higher receiver sensitivity, higher splitting ratio and a long unamplified transmission reach makes coherent detection scheme a technology of choice for application in FTTH technology.

5. Conclusion

In this paper a performance evaluation of the traditional IMDD scheme is compared to the IMCD. IMCD coupled with the advantages of increased receiver sensitivity and a longer unamplified transmission reach is considered an option to implement in the passive optical network in the FTTH technology. In addition coherent detection scheme can be used to demodulate more spectral efficient higher order modulation formats since information can be retrieved from the amplitude, phase and frequency of received signal [5,9].

Acknowledgement

We are grateful for Research Funding from: Telkom, Dartcom, Ingoma, CISCO, DST, NRF, CSIR, THRIP, and Scholarship Funding from ALC and NMMU.

References

[1] Rhodes R, Cheng N, Jensen J B, Monroy I T,"10 Gb/s Real-Time All- VCSELL Low Complexity Coherent Scheme for PONs". DTU Fotonik, Department of Engineering, Denmark, *OFC/NFOEC Technical Digest*. OSA .2012

[2] Jensen J B , Rodes R , Zibar D, Monroy I T. "Coherent Detection for 1550 nm, 5 Gbit/s VCSEL Based 40 km Bidirectional PON Transmission". DTU Fotonik, Technical University of Denmark, DK-2800 Denmark, *OFC/NFOEC Technical Digest. OSA.* 2011.

[3] Agrawal G P .2007 Non Linear Fibre Optics Academic Press 4th Edition. VPItransmissionMaker 9.3, © 2015 VPIPhotonics GmbH ELach, WIdler2011, "Modulation formats for 100G and beyond", Elsevier. Opt. Fib. tech. 17, pg 377-386

[4] Freude W, et al 2012. "Quality Metrics for Optical Signals: Eye Diagram, Q-factor, OSNR, EVM and BER", *Institute of Photonics and Quantum Germany, ICTON.*

[5] Abdullah M F L., Talib R .2012, "Multilevel signal analyzer tool for optical communication system". *Department of communication Engineering. University Tun Hussein Onn Malaysia. Institute of Advanced engineering and science*

[6] Krzysztof S , et al 2011,"30 Gbps 4-PAM transmission over 200m of MMF using an 850 nm VCSEL",*European Conference and Exposition on Optical Communications (ECOC), OSA.*

[7] Chabata T V, Gamatham R R G , Kourouma H Y S, 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.*

[8] T B Gibbon, et al "VCSEL transmission at 10 Gb/s for 20 km single mode fibre WDM-PONwithout dispersion compensation or injection locking", *Elsevier. Opt. Fib. tech.*vol **17**, pg 41-45 (2011)

[9] H FAN -9.0.2 Application note (04/08), "Optical signal to noise ratio and the Q-Factor in Fibre Optic Communication Systems"

Monte Carlo based estimation of the effect of different aerosol classes on solar irradiance in African atmospheric conditions

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. We present a study of the impact of aerosols on the ground level solar irradiance through their scattering and absorption of solar light. The degree of direct solar beam attenuation, as well as the angular and wavelength dependence of the diffuse (scattered sunlight) sky brightness strongly depends on the concentration, size distribution and nature of aerosol class. Aerosols common in the atmosphere in African conditions, such as biomass burning-generated smoke, wind-generated dust and urban-mixed (mixed particles from urban environment) all influence incoming sunlight in different ways. In this paper, a Monte Carlo approach is employed to track the movement of photons from the top of the atmosphere to the Earth's surface for a variety of atmospheric compositions characteristic of typical African localities. The results illustrate how variations in aerosol types change the amount of direct solar radiation reaching the ground, as well as the angular and spectral distribution of the diffuse light.

1. Introduction

The intensity and spectral distribution of solar radiation reaching the Earth's surface depends not only on solar position, but also on the local thickness (i.e. altitude) and composition of the atmosphere. The influence of atmospheric aerosols on solar spectral irradiance has been investigated in various studies and those studies show that aerosols have a big impact on incoming solar radiation attenuation due to their optical properties [1, 2]. A new method is applied to interpret this issue and it is presented in this the work.

An aerosol is defined in its simplest form as a solid or liquid particle suspended in a gas, and it includes a wide range of manifestations such as dust, smoke, fog, haze and smog [2]. Atmospheric aerosols range in size from a few nanometers to tens of micrometers in diameter [3].

Aerosols interact both directly and indirectly with solar light to affect the Earth's radiation budget and climate. They directly scatter sunlight to a new direction of propagation. This results in sky brightening or in radiation being returned back to space. As an indirect effect, aerosols in the lower atmosphere can modify the size of cloud particles, changing how the clouds reflect and absorb sunlight, thereby affecting the Earth's energy budget [1, 3].

The range and nature of the aerosol size distributions reveal the origin of their formation [4]. The scattering characteristics of atmospheric aerosols may change rapidly with time and meteorological

conditions due to evolution of the aerosol component. Aerosols have a range of sizes and reflective characteristics, meaning that they have different optical and physical properties such as the scattering phase function, the single scattering albedo, and the aerosol optical depth (AOD), all of which are wavelength dependent [5]. Those optical properties affect the impact of aerosols on atmospheric radiation, and as result play a major role in the solutions to the radiative transfer equation [4, 6, 7].

There are two fundamental approaches for measuring solar radiation reaching the ground. The first one is direct measurement which is based on ground based solar radiation readings, and it is the most accurate. This one requires a large number of measurements for a range of wavelengths over a representative time period and with adequate spatial coverage [5]. Such measurements are currently not available in the required quantity and quality for most locations, especially in Africa. The second approach is mathematical or computational modelling, which is usually the most viable approach to generate a complete set of local solar irradiance estimates. And it has to be validated by observed data.

Different studies have determined aerosol optical properties from microphysical quantities such as the particle complex refractive index and size distribution based on Mie scattering theory, which assumes that all particles are spherical and homogeneous [3, 5, 6].

In this paper, we investigate the spectral effect of the aerosol optical depth (AOD) on solar irradiance. We focus on the typical characteristics of the AOD for three types of aerosols commonly found in Africa: i) biomass burning generated smoke, ii) wind- generated dust and iii) urban –mixed [4, 6]. In this context we are looking at the distribution of scattering angle as function of scattered photons. The contribution of atmospheric aerosols effect on ground solar irradiance based on scattering angle is a concept which has not been adopted in previous studies. A Monte Carlo method is used here to simulate photons' walk and their new direction based on Henyey-Greenstein phase function. The number of photons scattered at given range of scattering angles determines how much solar light reaches the ground and how much is scattered back in atmosphere. This is the new approach which is studied here and it is well detailed in section3.

2. Mathematical formulation

We investigate computationally the influence of atmospheric aerosols on solar spectral irradiance. The incident flux from the top of the atmosphere passes through the atmosphere and gets attenuated by atmospheric constituents (including aerosols). The attenuated solar radiation is partly determined by the aerosol optical depth, which in turn depends on aerosol characteristics and total aerosol loading (i.e. aerosol optical depth).

Light is absorbed as it passes through the atmosphere and at the same time it is subject to scattering. Suppose that the spectral radiation intensity at wavelength λ outside the Earth's atmosphere is $I_0(\lambda)$ and the airmass is m. The amount of solar radiation transmitted through the atmosphere is calculated according to Lambert-Bouguer-Beer's law as follows:

$$I(\lambda, m, \tau) = I_0(\lambda)\exp(-m\tau) \quad (1)$$

In the equation above τ is a parameter describing atmospheric turbidity and is referred to as the optical depth. The exponential part corresponds to the transmission fraction, defined as that part of the solar beam unaffected by extinction in the atmosphere, and expressed as

$$T = \exp(-\tau m) \quad (2)$$

Another interpretation of this quantity is as the probability of a single solar photon passing through the atmosphere without deflection and T is therefore in the interval $0 \leq T \leq 1$.

The optical depth can be written as the sum of optical depth components $\tau = \sum \tau_i$. One of the mechanisms for light scattering in the atmosphere is Rayleigh scattering, which is caused by molecules in the atmosphere. Rayleigh scattering is particularly effective for short wavelength light (blue light) since its optical depth τ_R has a λ^{-4} dependence.

In addition to Rayleigh scattering, aerosols and dust particles are the major source of light attenuation in the visible range, and contribute to the scattering of solar light known as Mie scattering, which is the subject of this study. Light scattered by aerosols is characterized by a phase function $P(\xi,\lambda)$ which describes the angular dependence of light scattering and it is scattering angle (ξ) and wavelength (λ) dependent.

The aerosols' optical depth τ_a , which is a measure of the size and number of particles present in a column of air in the atmosphere, can be parameterized using the formula suggested by Ångstrom which defines this quantity in terms of an Ångstrom turbidity coefficient β and a wavelength exponent α [2].

$$\tau_a(\lambda) = \beta\lambda^{-\alpha} \tag{3}$$

The Ångstrom wavelength exponent characterizes the wavelength dependence of the AOD, and provides basic information on the aerosol size distribution. It is the slope of the logarithm of aerosol optical depth versus the logarithm of wavelength:

$$\alpha = \frac{d(\ln\tau_a)}{d(\ln\lambda)} \tag{4}$$

The single scattering albedo $\omega_0(\lambda)$ is a measure of the fraction of light intercepted and scattered by a single particle. It largely depends on complex part of refractive index, and particle size. For aerosol scattering, the Henyey-Greenstein phase function is used to determine the scattering angle distribution [7]

$$p_{aer}(\xi) = \frac{1-g_\lambda^2}{\left(1+g_\lambda^2-2g_\lambda\cos\xi\right)^{3/2}} \quad , \tag{5}$$

where g is referred to as the asymmetry factor. This lies in the range $-1 \le g \le 1$, which represents backscattering ($g=-1$) through isotropic scattering ($g=0$) to forward scattering ($g=1$). It indicates the portion of the light scattered forward and the portion scattered backward.

3. The modelling method and algorithm

We apply Monte Carlo simulations to estimate the solar irradiance in a horizontally homogeneous and plane parallel atmosphere with optical thickness $\tau(\lambda)$, illuminated by incident monochromatic flux $I_0(\lambda)$ at the top of the atmosphere. The procedure is based on the interpretation of photon-atmosphere interaction as a random process and the subsequent calculation of desired radiative characteristics. The fluxes measured at ground level as a function of wavelength are obtained as average values over a multiplicity of photon simulations.

The process of radiative transfer in the Monte-Carlo method is simulated as a multitude of photon movements. Coming from the top of atmosphere, the photon moves along a certain trajectory, which finishes either with a photon exiting the atmosphere, its absorption by the atmosphere, or its detection at the surface.

In our model we simulated the photon trajectories based on equation (2). A random number generator is used to provide random numbers r between 0 and 1. These are used to calculate whether the photon reaches the surface, the nature of the interaction (absorption or scattering) and the scattering angle (deflection angle ξ and azimuth ψ) [8]. And for each point new random number is generated.

In the first of these, if $r \le T$, the photon passes through the atmosphere it becomes transmitted. On the other hand, if $r > T$ the photon is scattered or absorbed. The new generated random number (r_1) is compared with single scattering albedo (ω_0) in order to simulate either scattering or absorption process. If $r_1 < \omega_0$, we simulate a scattered photon, otherwise absorbed. If the photon is scattered, it changes direction and the simulation continues from the scattering location. If the photon is absorbed, it is also annihilated, and therefore its trajectory ends [8].

Therefore, it is enough to simulate only three processes: the free path of a photon (i.e. the likelihood of interaction with the atmosphere), the interaction with the atmosphere (the absorption and scattering), and the scattering angle.

We computationally launched 10^4 photons from the top of the atmosphere at four specific wavelengths and tracked the progress of each one through the atmosphere. We applied this to different aerosols types investigated in this study.

The aerosol optical depth τ_a, single scattering albedo ω_0 and asymmetry parameter g used in this model are obtained from [6]. For τ_a we adopt typical values of α quoted for these aerosols by [6]. We investigate the scenario where $\tau_a(440\text{ nm}) = 1$, which corresponds to relatively high aerosol loading, and which therefore displays the effect of aerosols more clearly.

Table 1: Parameters associated with the aerosol types investigated

Wavelength (nm)	**Biomass- Burning** $\alpha = 2.0$			**Desert Dust** $\alpha = 0.5$			**Urban –Mixed** $\alpha = 1.8$		
	g	ω_0	τ_a	g	ω_0	τ_a	g	ω_0	τ_a
440	0.64	0.88	1.00	0.73	0.93	1.00	0.68	0.90	1.00
670	0.53	0.84	0.43	0.71	0.98	0.81	0.61	0.88	0.44
870	0.48	0.80	0.26	0.71	0.99	0.71	0.58	0.85	0.29
1020	0.47	0.78	0.18	0.71	0.99	0.66	0.57	0.83	0.21

In this paper, we only considered the case that the Sun is overhead. We furthermore only considered single scattering. Future work will incorporate other zenith angles and multiples scattering.

4. Results

We present the relationship between number of scattered photons and scattering angle from a simple Monte Carlo simulation. This is based on optical properties of biomass burning, desert dust and mixed urban aerosols. The dependence of scattered solar flux on wavelength is also presented. Scattering angle and scattering albedo together with AOD at different wavelength constitute a useful tool to determine the effect of aerosols on solar flux in our model.

4.1 The effect of biomass burning on solar flux

Figure 1 shows the photons scattered by biomass burning aerosols, at different wavelengths as function of scattering angle. This confirms that our simulation produces forward scattering [7], and that this effect becomes more pronounced at shorter wavelengths.

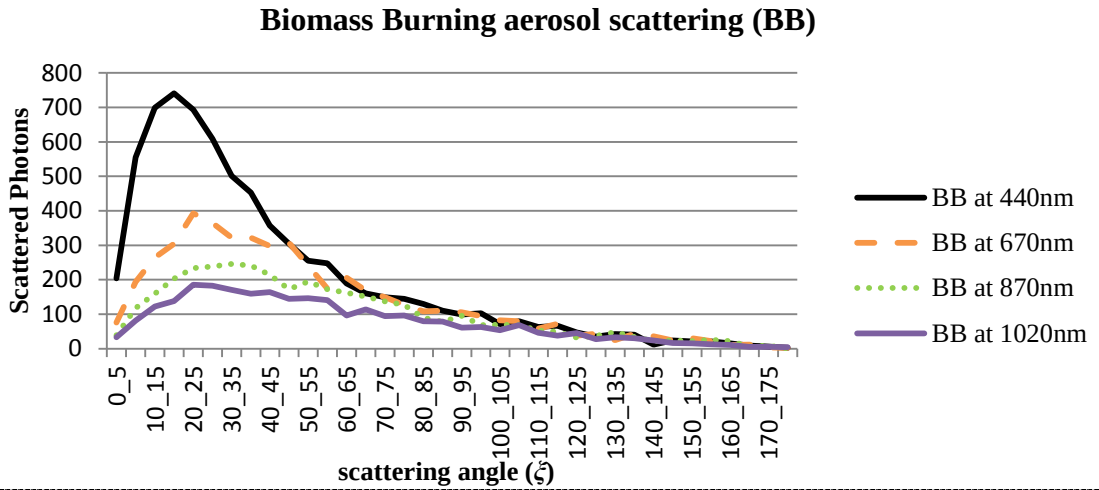


Figure 1. Distribution of scattered photons as a function of scattering angle for biomass burning aerosols.

4.2 The effect of desert dust on solar flux

Figure 2 shows the angular distribution of scattered photons at different wavelengths for desert dust aerosols. We note that our simulation correctly reproduces the strong forward scattering associated with this aerosol type.

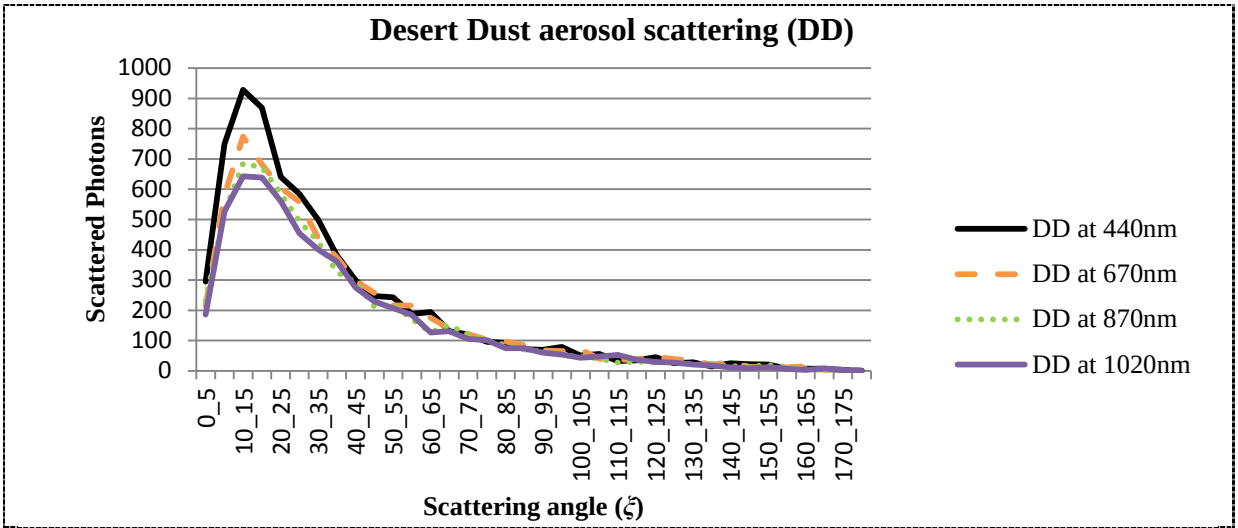


Figure 2. Distribution of scattered photons as a function of scattering angle for desert dust aerosols.

4.3 The effect of urban –mixed on solar flux

Figure 3 shows the angular distribution of scattered photons at different wavelengths for urban-mixed aerosols. As expected, the distribution here closely resembles the BB case.

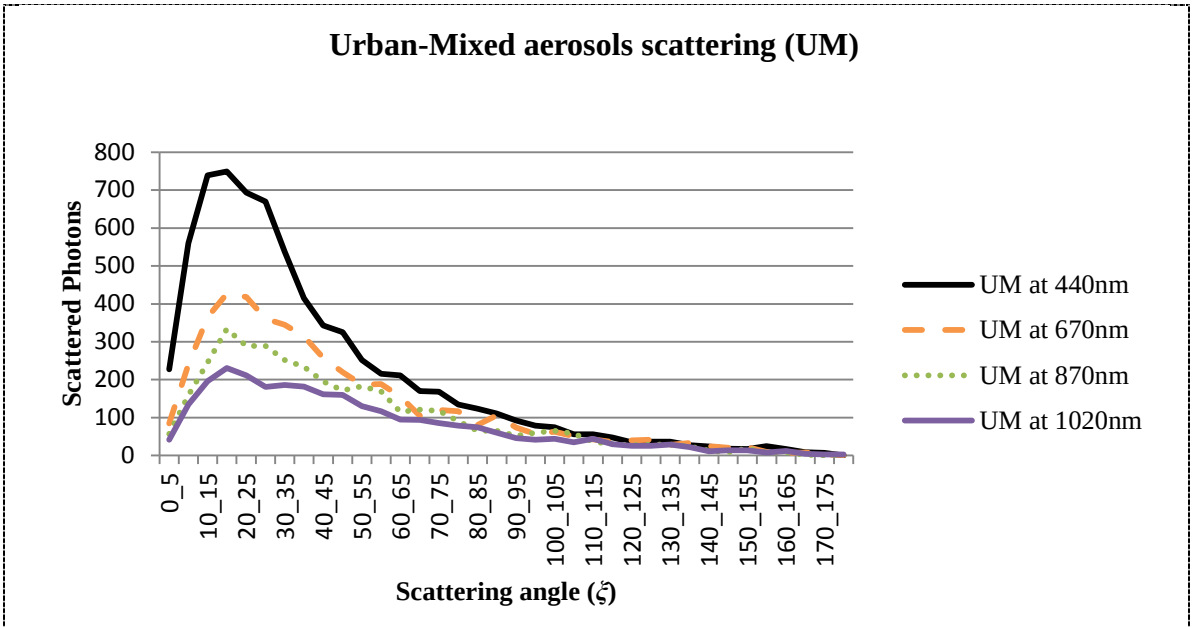


Figure 3. Distribution of scattered photons as a function of scattering angle for urban-mixed aerosols.

4.4 The wavelength dependence of aerosols optical depth

Figure 4 shows the wavelength dependence of the aerosol optical depth for aerosol types investigated here. In all cases the graph was calculated for an assumed value of $\tau_a(440\ \text{nm}) = 1$. Our simulations

therefore confirm the stronger wavelength dependence of the scattered portion for the BB and UM cases.

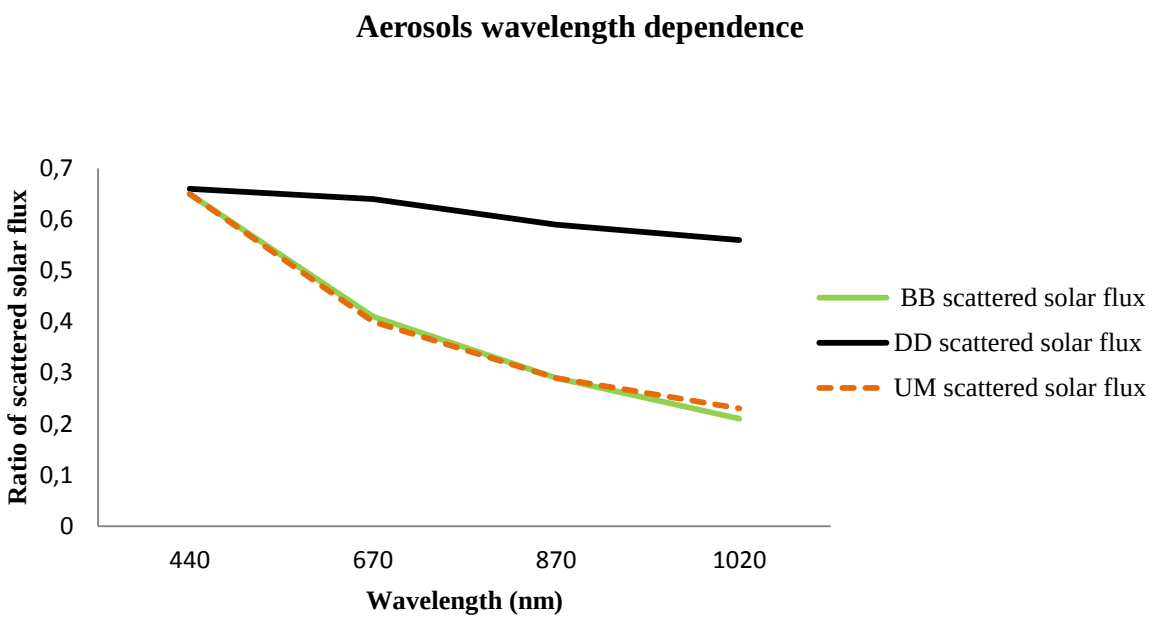


Figure 4. Relation between scattered solar flux and wavelength for $\tau_a(440\text{nm}) = 1$.

5. Conclusion and future work

We have presented the results of aerosol single scattering stochastic simulations on solar irradiance, and the scattered photon distribution as a function of scattering angle. The simulations were carried out at four wavelengths and for three aerosol types commonly present in the atmosphere over Africa. This distribution determines the aerosol contribution to the sky brightness and colour. Despite using the simplifying assumption that secondary scattering is small and the limited number of photons used in each simulation, we obtain scattering distributions that present a good reproduction of the true sky for the aerosol types tested, which can then be combined with the Rayleigh scattered components to construct the spectrum of the total skylight as a function of zenith angle. We are therefore able to confirm that the methodology used is suitable for further development to also model the sky colour for more complex arrangements such as off-zenith solar position and the consideration of multiple scattering events. The only minor scatter (i.e. distribution function plots) in our plots shows that good approximations can be achieved without prohibitively long computation time usage.

References

[1] Badarinath K V S, et al. 2006 *J. Atmospheric and Solar-terrestrial Physics* **69** 589-599
[2] Eck T E, et al. 1999 *Geophysical J.* **104** 333-349
[3] Papadimas C D, et al. 2012 *Atmospheric Chemistry and Physics* **12** 7185-7185
[4] Sun Z and Liu A, 2013 *Solar Energy* **98** 125-137
[5] Levoni C, Cervino M, Guzzi R and Torricella F 1997 *Applied Optics* **36(30)** 8031-8041
[6] Dubovik O, et al. 2001 *J. Atmospheric Sciences* **59** 590-608
[7] Kuznetsov A, Melnikova I, Pozdnyakov D, Seroukhova O and Vasilyev A 2012 *Remote Sensing of the Environment and Radiation Transfer. An Introductory Survey*, Springer Press, ISBN 978-3-642-14899-6
[8] Prahl S A, Keijzer M, Jacques S L and Welch A J 1989 *SPIE Institute Series* **5** 102-110

Characteristics and functions of the South African national measuring standard for force

S T S Dlamini
National Metrology Institute of South Africa, CSIR Campus, Meiring Naude Road, Brummeria, Lynnwood Ridge, Pretoria, South Africa.

E-mail: sdlamini@nmisa.org

Abstract. This paper describes the force comparator machine by which a national measuring standard for force is established. The design of the machine is described focusing on the mechanical structure of the force comparator. High-precision strain gauge force transducers are at the heart of the force comparator machine. Characteristics of these force measuring devices are discussed. The process of their calibration for industry using the comparator machine is explained and it is shown that it can be carried out to uncertainties as low as 0.03 %.

1. Introduction

The use of calibrated measuring equipment is very important in industries. In force measurements, there is a variety of force measuring instrumentation within South African industries such as space, mining, aircraft and automobile [1]. Measurement accuracy in these industries can only be determined by calibration against a force standard. In the past, South African industries had to calibrate their force measuring equipment against standards kept overseas. The national measuring standard for force described in this paper was established to support South African industries by providing accurate force calibration against a standard in a timely and cost effective way. This force standard consists of a set of highly accurate force transducers that are periodically calibrated overseas in an internationally accredited force laboratory which is able to perform calibrations with low uncertainties. Calibration is done in a form of a comparison between the standard force transducer and the reference force transducer used in industry.

2. Force comparator measuring system

The South African national measuring standard for force is currently a force comparator machine consisting of the following components: mechanical frame, hydraulic system, force transducer and electronic measuring system.

2.1. Mechanical structure of the force comparator.

The frame of the force comparator structure is made up of forged iron material. The stress permitted in the direction of load application, Y (see figure 1) is 310 MPa and perpendicular to direction of load application (X) is 210 MPa. The total permissible load is 7.8 MN with the breaking load of 14 MN. The extreme stiffness of the mechanical structure ensures that the structure remains rigid and stable up to the maximum operating load.

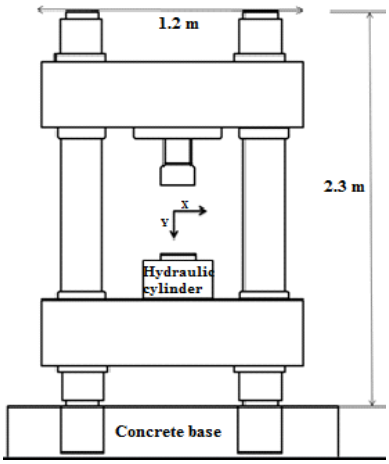


Figure 1. A schematic diagram of the frame of the force comparator machine.

Figure 2. The frame of the force comparator machine.

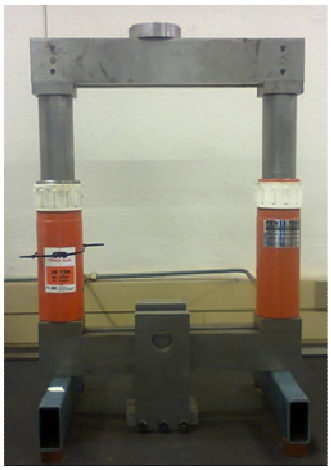


Figure 3. Comparator frame structure used for tensile force measurement.

The frame of the comparator machine shown in figure 2 is used for calibration of force transducer in compression in the range between 2 kN and 4500 kN. For calibration of force transducers in tension, the rig in figure 3 is used. Tensile measurement in the range of 2 kN to 250 kN can be performed (see section 3.1).

2.2. Hydraulic system.

The hydraulic system consists of hydraulic electric pump and a stepper motor coupled to the Budenberg pressure balance (see figure 4). These are used to drive the piston of the hydraulic cylinder and generate force within the comparator machine. The dual action hydraulic cylinder (figure 1) has a capacity to generate up to 4500 kN of force in compression and the double cylinders shown in figure 3 has capacity to generate 250 kN of tensile force. The hydraulic system is operated manually with a hand-pump for movement of the hydraulic cylinders piston up or down to generate force. The stepper motor coupled to the Budenberg pressure balance is used for sensitive and fine applications of force.

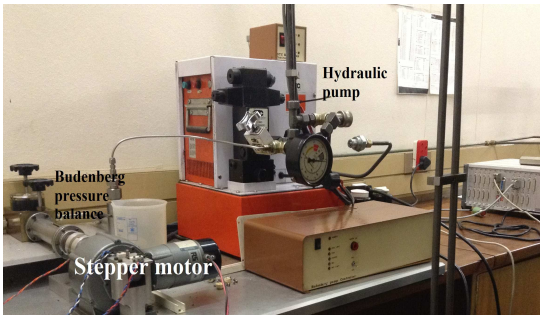


Figure 4. The hydraulic system consisting of the hydraulic pump and a stepper motor coupled to a Budenberg pressure balance.

2.3. Force measuring devices (force transducer).

A force transducer is a device that converts a force or load into a measureable output. There are several types of force transducer but strain gauge force transducer are by far the most common form of force transducer commercially available today [2]. The force transducers that are used with the

comparator machine as standards for force measurements are of accuracy class 0.5 and 00 as classified according to ISO 376 [3]. They are column design and are manufactured at HBM in Germany [4].

2.4. Force measuring devices (The digital measuring unit).
The basic output of strain gauge force transducer is a low level voltage signal (i.e., mV), therefore it must be gained by sophisticated methods. Through the use of signal conditioning and amplifiers, this signal can become a higher level voltage **[**5]. HBM signal amplifiers model DMP40S2 are used as read out units. Connected strain gauge force transducers are excited with a 225 Hz carrier frequency and the output signal of a selected force transducer is measured, amplified and displayed. The readings can be displayed in any desired unit such as mV/V, kN, kg, etc. [6] [7].

3. Calibration of force transducer using the comparator machine.

3.1. Calibration Setup
During calibration measurements of force transducer in compression (figure 5), the reference force transducer is attached to the upper coupling. The force transducer to be calibrated which is commonly known as the Unit Under Test (UUT) is coaxially aligned with the reference force transducer. This alignment band ensures that the force applied is in alignment with the UUT's principal load axis. In tension measurements (figure 6), the crossbeam is placed on the appropriate piston fittings. The reference force transducer is placed in the centre of the crossbeam and the loading button is placed between the tension to compression rig and the reference force transducer. The tension to compression top half is placed over the reference force transducer and bottom part fastens to the applicable screws. The UUT is fitted with suitable tension bars and attached to the tension rig.

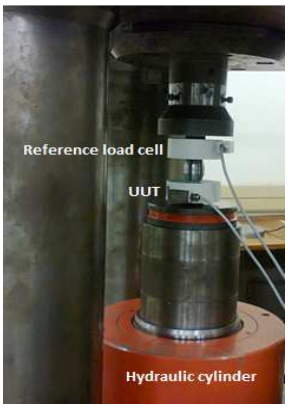


Figure 5. Calibration set up of force transducer in compression.

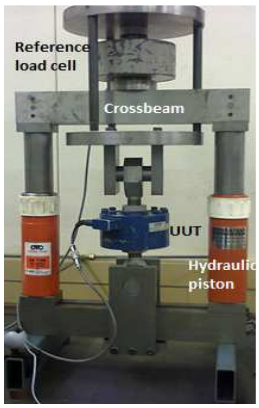


Figure 6. Calibration set up of force transducer in tension.

3.2. Uncertainty of the measurement
Calibration using the force comparator machine consists of applying precisely-known forces to the force transducer under calibration and recording the data from the indicator. At each calibration force, a combined standard uncertainty (u_c) expressed in units of force is calculated using the following equation [3] [8]:

$$u_c = \sqrt{\sum_{i=1}^{8} u_i^2} \qquad (1)$$

Where:
u_1 = standard uncertainty associated with the force comparator machine.

u_2 = standard uncertainty associated with reproducibility of calibration results.
u_3 = standard uncertainty associated with repeatability of calibration results.
u_4 = standard uncertainty associated with resolution of readout unit at zero.
u_5 = standard uncertainty associated with resolution of readout unit at measurement point.
u_6 = standard uncertainty associated with creep of the force transducer under calibration.
u_7 = standard uncertainty associated with drift in zero output
u_8 = standard uncertainty associated with temperature.

The equations for calculating each of the above standard uncertainties are derived and stated on the ISO 376 document which is an international standard for calibration of force transducer [3].

3.3. Calibration Procedure.

Force transducer are attached to their respective readout units and energised for about 30 minutes. Before the calibration forces are applied in a given mode (tension or compression), the maximum force is applied to UUT three times to exercise it. Two runs of increasing calibration forces are applied to the UUT and readings are recorded and used to calculate the repeatability uncertainty. The UUT is then rotated symmetrically on its axis to positions 120˚ and 240˚ and a run of increasing calibration forces is applied at each position and readings recorded to calculate the reproducibility uncertainty. The zero point is recorded before and after each run approximately 30 s after the force has been completely removed from the UUT. This is to calculate the uncertainty due to drift in zero for the force transducer. After the final zero reading is taken, the change in UUT output 300 s after removal of the calibration force is recorded to calculate creep. Laboratory temperature is recorded before and after the calibration is completed to calculate uncertainty of measurement due to change in temperature.

This procedure is followed during the calibration of force transducer to disseminate traceability in force measurement to South African industry. For the determination of the calibration measurement capability (CMC) of the comparator machine, reference standard force transducers of the force lab that have been calibrated by an internationally accredited NMI were calibrated against each other using the comparator machine following the above procedure. To determine the CMC of the machine in the range 2 kN to 20 kN, reference standard force transducer HBM Z4A (serial number: 184830136) of capacity 20 kN and class 00 was used to calibrate reference standard force transducer HBM C4 (serial number: 35178) of class 00 with capacity 20 kN in compression mode.

3.4. Calibration Results.

The calibration data obtained from HBM C4 (serial number: 35178) force transducer calibrated using HBM Z4A (serial number: 184830136) force transducer is shown in Table 1. The procedure as stated in section 3.3 was followed to perform the calibration. The temperature of the lab before and after the calibration was recorded and was found to be 18˚C and 19˚C respectively. The readout unit DMP40S2 connected to the HBM C4 force transducer indicated the measurements in mV/V units with resolution of 0.000001.

Using calibration data from Table 1, uncertainty contributors to the force measurements expressed in the units of force, kN can be calculated using equations as derived from the ISO 376 standard document. Table 2 shows values calculated of all uncertainty contributors. The machine uncertainty is taken as the combined standard uncertainty as quoted on the recent calibration certificate of the HMB Z4A force transducer issued by an internationally accredited NMI. The uncertainty due to reproducibility, repeatability, resolution of the readout unit for both at zero and at the point of measurement, creep and drift at zero are calculated using equations derived in the ISO 376 document. The uncertainty contribution of temperature to measurements is also calculated.

Table 1: Calibration data received from HBM C4 (serial number: 35178) load cell in compression mode.

Ref force	**Pre loads**			**Run 1**	**Run 2**	**Run 3**	**Run 4**
kN	**mV/V**	**mV/V**	**mV/V**	**mV/V**	**mV/V**	**mV/V**	**mV/V**
0,0	-0,000002	0,000030	0,000034	0,000002	0,000010	0,000011	0,000001
2,0				0,199938	0,199948	0,199927	0,199926
4,0				0,399857	0,399866	0,399855	0,399858
6,0				0,599760	0,599765	0,599774	0,599792
8,0				0,799647	0,799647	0,799684	0,799727
10,0				0,999519	0,999515	0,999582	0,999659
12,0				1,199376	1,199371	1,199471	1,199585
14,0				1,399217	1,399216	1,399348	1,399504
16,0				1,599045	1,599053	1,599213	1,599412
18,0				1,798858	1,798883	1,799067	1,799307
20,0	1,998662	1,998618	1,998628	1,998658	1,998709	1,998908	1,999187
Force indicated at zero 30 seconds after the removal of the maximum load.				0,000070	0,000060	0,000020	0,000040
Force indicated 300 seconds after the final maximum load has been removed (Creep effect)							-0.000030

Table 2: Calculated values of uncertainties contributors in force measurements by HBM C4 standard force transducer.

Machine Uncertainty (kN)	Reproducibility (kN)	Repeatability (kN)	Resolution Zero (kN)	Resolution @ Point (kN)	Creep (kN)	Drift @ Zero (kN)	Temperature (kN)
u_1	u_2	u_3	u_4	u_5	u_6	u_7	u_8
0,00000	0,00003829	0,00000000	0,000004	0,000004	0,0000	0,0002	0,0000000
0,00011	0,00003811	0,00005868	0,000004	0,000004	0,0000	0,0002	0,0001559
0,00020	0,00000768	0,00004920	0,000004	0,000004	0,0001	0,0002	0,0003118
0,00027	0,00009355	0,00002601	0,000004	0,000004	0,0001	0,0002	0,0004677
0,00032	0,00022987	0,00000117	0,000004	0,000004	0,0001	0,0002	0,0006235
0,00040	0,00040340	0,00002258	0,000004	0,000004	0,0002	0,0002	0,0007794
0,00048	0,00060558	0,00002851	0,000004	0,000004	0,0002	0,0002	0,0009353
0,00049	0,00082784	0,00000919	0,000004	0,000004	0,0003	0,0002	0,0010912
0,00056	0,00106166	0,00004510	0,000004	0,000004	0,0003	0,0002	0,0012471
0,00063	0,00129852	0,00014411	0,000004	0,000004	0,0003	0,0002	0,0014030
0,00070	0,00152997	0,00029758	0,000004	0,000004	0,0004	0,0002	0,0015588

At each reference force applied, u_c was calculated from all the uncertainty contributors as listed in section 3.2 and calculated and shown in table 2 using equation (1) and the results are shown in table 3. Expanded uncertainty (U) in measurements is derived by multiplying the u_c by the coverage factor of k = 2 and dividing by the magnitude of the reference force applied and reporting the results in percentages at the confidence level of approximately 95 % [8].

The CMC of the machine is then taken as the highest expanded uncertainty calculated in percentage rounded up to two decimal points. Therefore the CMC for the Comparator force machine in the range of 2 kN to 20 kN is 0.03 % when HBM C4 force transducer is used as reference standard in the machine to perform calibrations. The CMC for the machine in the other force ranges was also determined following the same method of calibrating two reference standard force transducer using the Comparator force machine. The CMC in the range of 20 kN to 200 kN was found to be 0.03 % in

both compression and tension. From 200 kN to 1000 kN in compression only, the machine has the CMC of 0.04 % and 0.08 % in 1000 kN to 4500 kN range.

Table 3: Determination of the expanded uncertainty (*U*) in measurements of HBM C4 standard force transducer

Reference force	Combined uncertainty	Coverage factor	Expanded uncertainty	Expanded uncertainty
F_R	u_c	k	U	U
kN	**kN**		**kN**	**%**
2,000	0,00027	2	0,001	0,027
4,000	0,00042	2	0,001	0,021
6,000	0,00059	2	0,001	0,020
8,000	0,00077	2	0,002	0,019
10,000	0,00100	2	0,002	0,020
12,000	0,00124	2	0,003	0,022
14,000	0,00149	2	0,003	0,023
16,000	0,00176	2	0,004	0,024
18,000	0,00205	2	0,005	0,025
20,000	0,00235	2	0,005	0,026

4. Conclusions

The calibration capability of the Comparator force machine depends primarily on the force transducer used as the reference standard to measure the force generated by the machine. Precision strain gauged force transducers of high accuracy connected to signal amplifier of high resolution are used as the main components of the Comparator force machine. Therefore the forces generated by the machine can be precisely measured and transferred during the process of calibration. The CMC expressed as the expanded uncertainty of the measurement was found to be 0.03 % in the range of 2 kN to 20 kN for the Comparator force machine.

5. References

[1] Schäfer A 2012 Examples and proposed solutions regarding the growing importance of calibration of high nominal forces *Proc. XX IMEKO world Congress* (Busan: Republic of Korea)

[2] Packard H 1981 *Practical strain gauge measurements* (USA**)**

[3] ISO 376: 2011, Metallic materials – Calibrations of force-proving instruments used for the verification of uniaxial testing machines.

[4] Bray A, Barbato G and Levi R 1990 Theory and Practice of force measurement (Academic Press) p 144.

[5] Schäfer A, Kitzing H 2012 DMP41-A new chapter of ultra-precision instrument for strain gauge transducers *Proc. XX IMEKO world Congress* (Busan: Republic of Korea)

[6] Kitzing H *A solid base for precision strain gauge measurements* (Hottinger Baldwin Messtechnik GmbH, Darmstadt, Germany)

[7] Kreuzer M *High-precision measuring technique for strain gage transducers* (Hottinger Baldwin Messtechnik GmbH, Darmstadt, Germany)

[8] BIPM IEC ISO IUPAP OIML 1993 Guide to the Expression of Uncertainty in Measurement International Organisation for Standardisation (Geneva: Switzerland)

Estimation of energy production decrease due to shading for the NamPower rooftop photovoltaic system

P Dobreva [*1,2]**, E E van Dyk**[1] **and F J Vorster**[1]
1 Nelson Mandela Metropolitan University, P.O. BOX 77000, Port Elizabeth- 6031, South Africa
2 University of Namibia, Private Bag 13301, Windhoek, Namibia

E-mail: pdobreva@unam.na

Abstract. Photovoltaic (PV) energy has become one of the most important renewable energy technologies and the installation of PV systems on rooftops in industrial, commercial and residential sectors has become common. The operation of these systems is not without challenges one of them being the effect of shade on the performance of the system. Potential shading of arrays or parts thereof needs to be carefully considered when designing a system. In this study an operational 63.45 kW rooftop PV system, on the roof of the NamPower building in Windhoek, was investigated. The focus of this paper is the effect of the partial shading that some of the module strings of the system experience during part of the year due to surrounding buildings. We estimate the loss in energy production due to shading using simulation software and compare this to actual performance data. In the analysis of the effects of shading the current-voltage (I-V) curves for module strings that experience varying levels and configurations of shading are compared to unshaded strings. The results highlight the importance of considering the effects of shading on system performance and illustrate potential negative impact of unsuitable string configurations.

1. Introduction

Solar energy in general and photovoltaic (PV) energy in particular have become one of the most important renewable energy options, recording strong levels of growth in both small-scale and large-scale projects. The improved cost effectiveness of the solar sector and the ever-increasing cost of electricity provided by national grids drives an increase in rooftop and larger scale PV installations [1]. The urban environment, where PV systems are often installed on rooftops, presents serious challenges to these systems. Rooftop PV installations almost inevitably experience shading due to roof topography, adjacent buildings and poles, trees, overhanging cables, etc.

Shading has direct and indirect effects on the energy production. The direct effect is the drop in energy output due to decrease in the radiation incident on the modules. The indirect effect is the drop in energy output due to electrical mismatch arising from shading part of a module or string of modules.

These effects are best illustrated by considering the changes caused by shading in the I-V curve of a string. Most modules incorporate bypass diodes that prevent the formation of hot-spots when a cell becomes reverse biased. A diode protects a group of cells that will be referred to as a sub-module or cell-string. The bypass diode starts conducting whenever a cell in its sub-module is shaded. The bypass

* To whom any correspondence should be addressed.

diode carries all the current generated by the unshaded parts of a module past the shaded part. Once a bypass diode turns on, the voltage of the module drops by an amount equal to the voltage of the sub-module protected by that diode. If a module has three bypass diodes and one of them starts conducting, then the voltage of the module drops by one-third (and, in addition, by the forward bias of the diode) [2]. This means that the sub-module whose bypass diode turns on is, effectively, removed from generation capacity and the output power of the module decreases by one-third. The number of sub-modules that are affected by shading can be read straight from the current-voltage (I-V) characteristic curve and the shift of the maximum-power-point (MPP) to lower voltage is easily discernable. It is well established that the decrease in energy output of PV cells, modules and systems due to partial shading is highly non-linear [3, 4], which means that the decrease in electrical performance is not proportional to the percentage of shaded area but is significantly higher.

Partial shading, unlike uniform shading, is particularly troublesome for PV plants. In uniform shading the whole module or string of modules experiences a uniform decrease in light intensity, which directly reduces the current but has little effect on the voltage, since the voltage only varies logarithmically with light intensity. In contrast, in the case of partial shading, part of a module/string of modules experiences shading while the rest is under full illumination. This results in switching on of bypass diodes so that the full current produced by the unshaded parts can be conducted but it causes significant reduction in the voltage. The decrease in the voltage is responsible for the non-linear decrease in power output of a module/string of modules under partial shading.

This study considers only the effects of shadings that arise with change in the sun's angle like shading due to nearby objects and inter-row shading (shading of the second row by the first and so on). The NamPower roof top grid-connected PV system experiences shading during part of the year yet no studies on the effects of partial shading on the output of this system have been conducted.

2. System Description

The PV system studied in this paper is installed on the rooftop of the national power utility of Namibia, NamPower in Windhoek (coordinates: 22.5^0 S, 17.1^0 E). Figure 1 shows a photo of the system taken from the North on 18 May 2015 at 15:29h local time. Initial shading due to an adjacent building is clearly visible. On 18 May the solar transit in Windhoek took place at 11:48h local time and the day was perfectly cloudless.

The system consists of 259 Solar World SW 245-poly modules and six SMA 12 000 Tripower inverters. The SW 245 modules consist of 60 cells connected in three sub-modules so that a bypass diode protects 20 cells. A sub-module is located along the length of the module.

Due to varying slope of the roof, 42 of the modules are installed at a tilt of 13° while the rest of the modules are installed at tilt of 7°. All modules are orientated at 5° W. The modules are distributed into seven strings of 22 modules and five strings of 21 modules. One string is connected to each inverter input.

The SMA Tripower inverters have two inputs, each with its own independent maximum-power-point-tracker (MPPT), with power split asymmetrically between the two (2/3 on Main and 1/3 on Secondary input). This allows strings of different lengths to be connected to the different inputs. The first 84 modules are installed at a module row pitch of 2.05 metres while the rest of the modules are installed at a pitch varying from 1.96 to 2.14 metres. All modules are arranged in portrait orientation. Inter-row shading occurs for a large part of the day during the winter months. The system experiences shading due to a nearby building and previously installed solar water heaters from March to September.

3. Modelling

The modelling of the system in this paper was done with PVsyst modelling software. PVsyst allows extensive input from the user on system orientation and configuration, de-rate factors and three-dimensional system modelling. The 3D-scene created in PVsyst allows for a detailed analysis of shading effects on a PV system at fifteen minute intervals on a string-by-string basis. Figure 2 shows a snapshot of the 3D-scene that was created to study the effects of partial shading on the NamPower system. The

image corresponds to 18 May, 15:30h. In this installation the strings run along the rows of modules (PV sheds) from North to South with String 1 starting on the first northernmost row. Inter-row shading is indicated by the yellow colour on the PV sheds while the grey colour indicates shading by nearby objects.

Figure 1. Photo of the NamPower system taken on 18 May 2015 at 15:29h from the North direction.

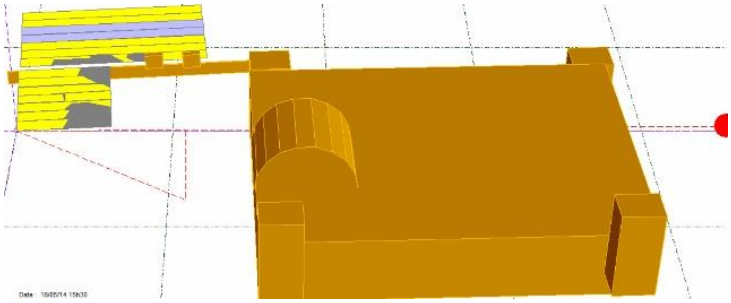


Figure 2. PVsyst 3D scene of shading on 18 May 15:30h as viewed from the North. The building in the foreground is casting shadow on the PV system in the background.

The electrical configuration was modelled according to the installation data. Two orientations (corresponding to the tilts 13^0 and 7^0) were selected and eight sub-arrays were defined. A sub-array connects to either the Main or to the Secondary inverter input. DC wiring loss was taken at 1% with respect to STC (standard-test-conditions) running conditions. The soiling de-rate factors were chosen to vary from 0% to 4% depending on the rainfall and the system owner's cleaning schedule. The light-induced-degradation (LID) factor was taken at default value and module efficiency loss was assumed to be -1.5% (indicating over-performance) as per module manufacturer data. The effect of wind was assumed to be the default value for semi-integrated modules with air-duct underneath and the unavailability was assumed to be one day per year [5].

4. Results

The effects of shading are analysed for a particular day – 18 May. For accurate modelling to take place, it is essential that an accurate 3D-scene is created. PVsyst uses the one-diode model to simulate the electrical behaviour of the modules. To assess the effects of partial shading, the electrical behaviour of each sub-module of each string connected to each MPP inverter input is computed.

On the 18th May the system experiences inter-row shading up to 8:45h in the morning and then, again, from 14:15h in the afternoon. The dramatic effects of inter-row shading on the power output of a string can be demonstrated by the behaviour of a particular string. Figure 3 shows the I-V curve of String 1 that extends over two rows, with 12 modules in the first row that is free from inter-row shading and 9 modules in the second row. At 14:15h the cells at the bottom of the modules in the second row become shaded while the others are fully illuminated. With the modules in portrait orientation, this means that all sub-modules (except for one) of the nine modules in the second row are affected, their bypass diodes start conducting and the voltage of the string decreases by an amount equal to the total voltage of all 26 sub-modules. This causes significant reduction in the power output of the string, as can be seen in Figure 3. The dashed curve indicates the I-V curve of the string if it were unshaded. The solid green curve is the I-V curve with shading on the string. The vertical solid blue line indicates the minimum voltage input of the inverter (which is 150 V in our case).

The voltage of the string at the operating current is just below 400 V and the MPP of the string has shifted to lower voltage. The total effect of the inter-row shading on String 1 at 14:15h on 18 May is a reduction of the power output of the string by 46.0% (from 3.37 kW to 1.81 kW) even though only 4.3%

of the string area is shaded. While the power decreases by only 17.5% due to drop in irradiance, it experiences a decrease of 28.5% due to electrical mismatch caused by the shading (giving a total loss of 46.0%).

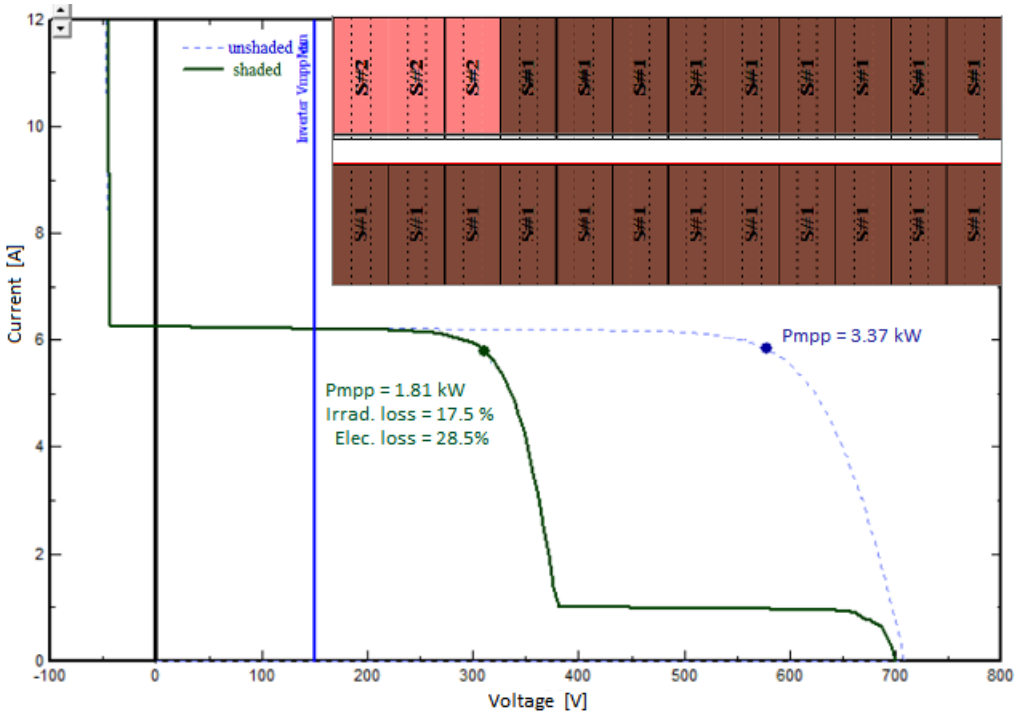


Figure 3. IV-curve of String 1 at 14:15h on 18 May. The inset shows initial inter-row shading on String 1 (S#1) and part of String 2 (S#2).

Any further shading of the modules in the second row will affect the output of the string only through the loss of the contribution of the single unshaded sub-module. Partial shading of modules in the first row will, however, further reduce the output of the string. The inset in Figure 4 shows String 1 at 15:30h when the adjacent building casts shadow onto it. Nineteen sub-modules in the front row experience shading due to the nearby building. The decrease in the light intensity causes a drop in current but the partial shading reduces the voltage of the string further by the voltage of those nineteen sub-modules. The IV-curve of String 1 at 15:30h is shown on Figure 4. The global maximum MPP of the string has shifted to a voltage below the inverter input threshold voltage of 150 V. The inverter then seeks the secondary maximum, which is 0.52 kW at slightly over 600 V, and which falls on the part of the I-V curve that is due to diffuse radiation only, resulting in loss of output of the fully illuminated part.

These examples and literature [6, 7] clearly illustrate the substantial loss in PV power caused by even a small amount of shading. The configuration of the shadow is significant only in the context of the number of sub-modules that it covers and the degree to which it shades them. It appears that the portrait orientation of the modules in the Nampower system is advantageous in the case of the shadow cast by the nearby building which strikes the modules along their length, affecting one sub-module after another with time and not all three at once.

The system was also modelled with the modules positioned in landscape orientation so that the inter-row shading would affect only a single sub-module at a time and not all three at once. The system would have produced 2.3% more energy on 18 May with modules in landscape orientation. The available roof area, however, doesn't permit the same capacity to be installed with modules in landscape orientation.

A comparison of the measured (on 18 May 2015) and modelled for that day energy output is shown in Figure 5.

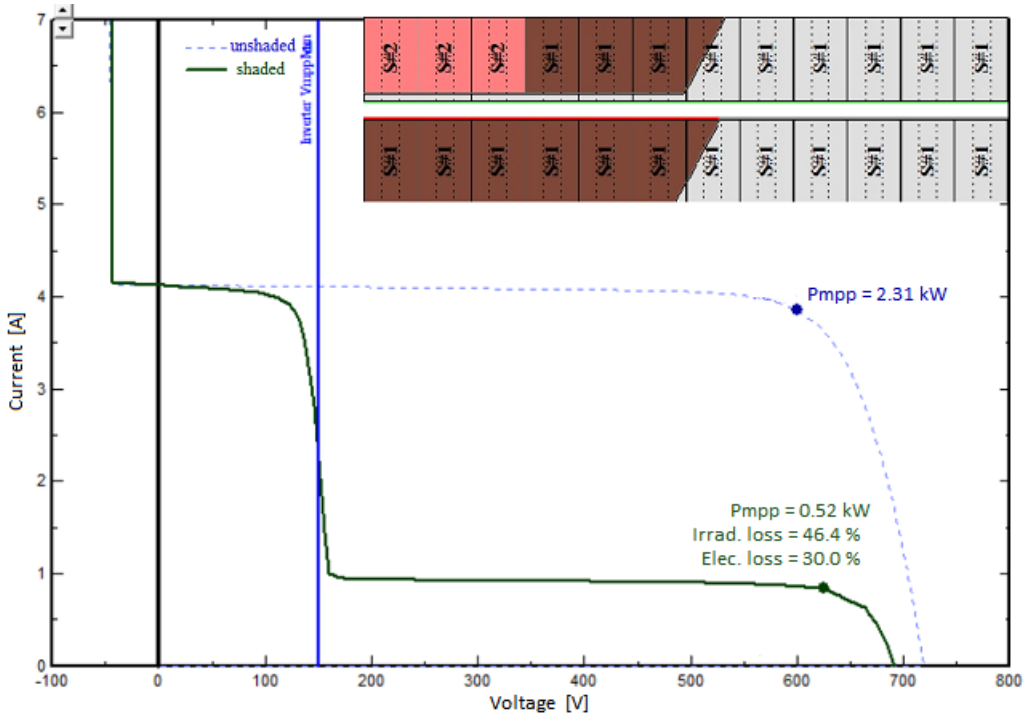


Figure 4. IV curve of String 1 at 15:30h on 18 May. The grey colour indicates shading

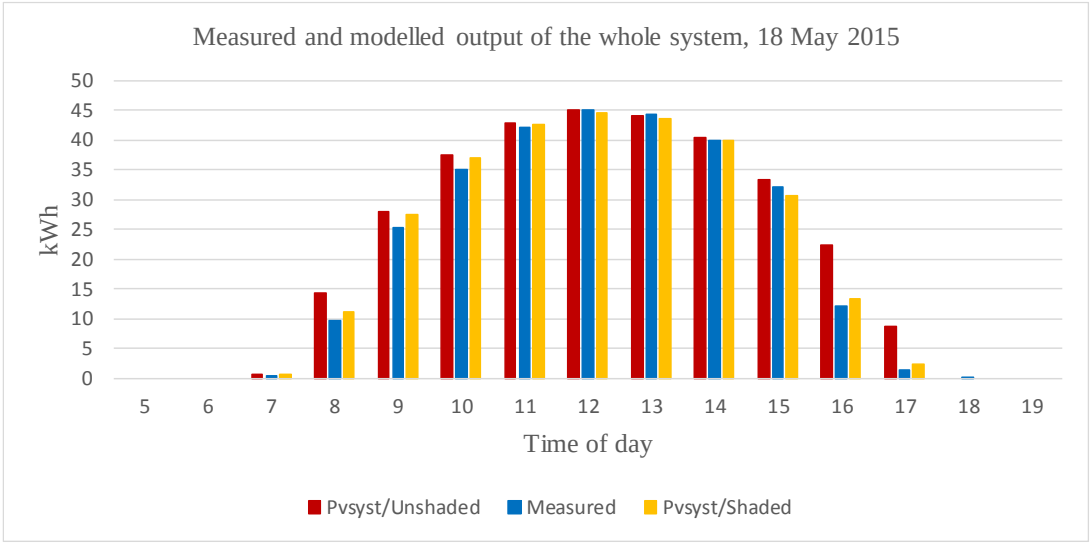


Figure 5. Measured and modelled hourly energy output for 18 May 2015.

The modelled energy output, with shading taken into account on 18 May, is 1.9% higher than the measured. Although the actual values are small in comparison with the overall uncertainty, estimated previously [8], there is a clear trend of good alignment between measured and predicted effects which become more pronounced with long term weather patterns and fluctuations. If the system were not experiencing partial shading, it would output 10% more energy on that day.

The interconnection of modules into strings also plays a role in the performance of a system under partial shading [9]. Further modelling was done by connecting together modules that experience similar

shading conditions. This requires numerous inter-row connections, thus increasing cable length and resistive losses.

5. Discussion and conclusion

Our model accurately predicts the behaviour of the NamPower PV system under partial shading and allows the energy losses due to shading to be estimated. In Figure 5 the model suggests that the system is completely shaded after 17:00h and has negligible output while the measured output at that time is not zero. Since sunset in Windhoek takes place at 17:17h on 18 May, it is possible that some light reflected from the surroundings is incident on the system before sunset, thus contributing to higher production at that time of the day; such effect, however, has not been considered in the model.

The effects of inter-row shading can be reduced if the modules are installed in a landscape orientation, if the roof area permits.

Modules should be connected horizontally along a row even if they are subject to dissimilar shading conditions in order to keep wiring length and resistive losses to a minimum.

Another factor that affects the performance of a PV system under partial shading is the number of strings connecting to an MPPT input of an inverter. If two or more strings that experience different levels of shading connect to the same MPPT input of an inverter, then the voltage mismatch arising between them as a result of the shading causes further losses. The NamPower system design has a single string per inverter input which reduces the negative effects of partial shading on energy output.

For the NamPower system the expected (modelled) electrical loss due to shading stands at 1.3% per annum with ohmic wiring loss of 0.3%. The loss due to shading is low since shading takes place in winter months and only outside hours of high insolation. The effects of unavoidable partial shadings can be mitigated at the design stage by avoiding, as far as possible, multi-string connection to same inverter MPPT input, using suitable module orientation and string configuration.

Acknowledgements

The authors would like to thank NamPower and Solar Age Namibia for the continuous support.

References

[1] Renewable Energy Policy Network for the 21st Century 2015 *Renewables 2015 Global Status Report* (Paris: REN21)

[2] Deline C 2009 *Partially shaded operation of a grid-tied PV system in Proc. 34th IEEE Photovoltaic Specialists Conf. (PVSC),* (*June 7-12 2009, Philadelphia, Pennsylvania)* pp 001268-001273

[3] Rauschenbach H S 1971 Electrical Output of Shadowed Solar Arrays *J IEEE Transactions on Electron Devices*. **18**, no. 8 pp 483-490

[4] Bishop J W, 1988 Computer Simulation of the Effects of Electrical Mismatches in Photovoltaic Cell interconnection Circuits *Solar Cells*. **25**, pp 77-89

[5] Luehl H, Personal communication, Solar Age Namibia, and T.Ndemutela, Nampower.

[6] Woyte A, Nijs J and Belmans R 2003 Partial Shadowing of Photovoltaic Arrays with Different System Configurations: Literature Review and Field Test Results *Solar Energy* **74** pp 217–233

[7] Quaschning V and Hanitsch R 1996 Numerical Simulaton of Current-Voltage Characteristics of Photovoltaic Systems with Shaded Solar Cells *Solar Energy*. **56**, no. 6, pp 513-520

[8] Dobreva P, van Dyk EE, Vorster FJ 2015 Energy yield modelling of PV systems operating in Namibian conditions *in Proc. 3rd Southern African Solar Energy Conference (SASEC), (11-13 May 2015, Skukuza, Kruger National Park, South Afric).* pp. 492-496

[9] Picault D, Raison B, Bacha S, de la Casa J and Aguilera J, 2010 Forecasting photovoltaic array power production subject to mismatch losses *Solar Energy*. **84**, pp 1301–09

Measuring the optical thermometry properties of La_2O_2S:Eu phosphor material

L J B Erasmus, H C Swart*, J J Terblans and R E Kroon

Department of Physics, University of the Free State, PO Box 339, Bloemfontein, 9300, South Africa

Email: *SwartHC@ufs.ac.za,

Abstract. This paper is focused on the investigation and measurement of optical thermometry properties of Lanthanum Oxysulphide doped with Europium (La_2O_2S:Eu) by utilising the Photoluminescence (PL) technique. After a literature study it was concluded that the optical thermometry properties of phosphor materials can be measured by several techniques. The technique used in this paper is the fluorescence intensity ratio technique. This technique consists of obtaining the fluorescence spectra of a phosphor material and monitoring the intensity ratio between two thermally coupled levels as a function of temperature. The PL system in the Physics department at the University of the Free State was modified to measure the fluorescence spectra of a phosphor material at temperatures range from room temperature to 700 K. The modified PL system was used for the measurements of the optical thermometry properties of La_2O_2S:Eu for this paper.

1. Introduction

It is generally known that the energy to light conversion efficiency of some inorganic phosphor materials is temperature dependent and therefore gives these phosphor materials their temperature sensing characteristics. Phosphor materials that exhibit this characteristic are also known as thermographic phosphors. [1] A generic phosphor thermometry system consists of an excitation source that is used to excite the phosphor material that is bonded to the surface of interest. [2] The emission of the phosphor material is then analysed and compared to pre-calibrated data to determine the temperature of the surface in question.

The system design in terms of the excitation source, phosphor material and detector will depend on the application in interest. The excitation source can vary between a flash lamp, Light Emitting Diode and a laser that can also be a continuous or a pulsed source depending on the technique that is used as will be explained in the theory section. [2] There are a large number of phosphor materials that have different responses in comparison to temperature range and sensitivity that can be matched to a variety of applications. A Photo Multiplier Tube (PMT) or a Charge Couple Device (CCD) can be used for detection of the emitted light from the phosphor material. [2]

Thermographic phosphors that are used in the thermometric technique are adaptable to the needs of a wide variety of situations. It provides a *non-contact* optical alternative for measuring temperature in contrast to other conventional techniques and can therefore also be employed in systems where other thermographic techniques are not suitable. [1]

2. Theory

There are several ways the temperature dependence of phosphor materials is manifested. The properties measured to characterise the temperature dependence of the phosphor materials are the intensity, rise and decay lifetime and the line shift and width of selected spectral features. [1] A summary of each of these techniques is given below.

2.1 Fluorescence Intensity and Fluorescence Intensity Ratio technique

When a light source excites a phosphor material, electrons are transferred to higher energy states which also return back to their ground states. This is known as electron transitions. Usually equilibrium between excitation and de-excitation is reached which results in a steady emission intensity that is emitted by the phosphor material. If the temperature increases a decrease in emission intensity is observed. This is due to an increase in the probability of non-radiative processes with an increase in temperature. Thus by calibrating the fluorescence intensity response as a function of temperature, temperature measurements can be made of the emission intensity. [2]

The fluorescence intensity based approach generally requires the simplest instrumentation and is therefore relatively cheap to implement. However difficulty arises with the maintaining of intensity calibration since the intensity of emission is also a function of other variables. These variables include non-homogenous illumination, light source instabilities, distance and angle of detector. [2] Thus a better approach that eliminates most of these variables is the Fluorescence Intensity Ratio (FIR) technique.

The FIR technique uses the ratio of two or more fluorescence emission lines at different wavelengths of a phosphor material to determine the temperature response. Phosphor materials that are good thermographic phosphors exhibit multiple emission response where some emission lines being less or more sensitive to change in temperature which makes the use of the FIR technique possible. [2-4]

2.2 Fluorescence Lifetime technique

When a phosphor is excited by a pulsed source, the resulting fluorescence can be observed as the fluorescence intensity rises and decays exponentially. The lifetime of the fluorescence is a function of temperature and is therefore a useful method of thermometry. The fluorescence lifetime approach does not suffer the same disadvantages as the intensity based approach because rise and decay rates can be measured in terms of time and therefore offers less uncertainty since this quantity can generally be determined with greater accuracy than optical intensities. However to measure the fluorescence intensity rise and decay one must use a fast analog-to-digital converter which could be costly. [1]

2.3 Fluorescence Line Shift and Line Width technique

Each emission line of a phosphor is characterised by a wavelength for which the intensity is a maximum. Line shifts may occur with a change in temperature. Each emission line also has a finite width, which is often designated as the full width half maximum which also changes with a change in temperature. Line shift and line width changes as a function of temperature are usually small and therefore not often used in fluorescence thermometry as compared to the intensity and lifetime techniques. [1]

In this study only the fluorescence intensity and FIR technique was used to characterise the temperature response of Lanthanum Oxysulphide doped with Europium (La_2O_2S:Eu). [5]

3. Experimental

Commercially available La_2O_2S doped with Eu, from Phosphor Technology, was selected for this study since it was reported in a previous study to shown a decrease in emission intensity for the temperature range 292 K to 333 K. [2] An X-ray diffraction (XRD) pattern of the La_2O_2S:Eu was measured with the D8 Advance Bruker XRD apparatus. The emission spectra containing the different emission peaks were monitored for the La_2O_2S:Eu from 200 nm to 580 nm by using a Cary Eclipse fluorescence spectrophotometer (Varian).

The Photoluminescence (PL) system in the Physics department at the University of the Free State is capable of measuring fluorescence spectra of a phosphor material at room temperature and thus it was necessary to modify the system for this study. Figure 1 shows a basic layout of the modified PL system. At the top left corner is the (a) Helium-Cadmium (Kimmon IK Series, model KR1801C) laser. The laser beam was filtered with a (b) band-pass filter (Newport 10LF10-325) to pass only a 325 nm laser beam that was used as the excitation source. The laser beam was directed (c) onto the La_2O_2S:Eu phosphor material (Phosphor Technology SKL63/F-R1) and the emitted light of the phosphor was directed towards (d) a fibre optic cable that was connected to (e) a spectroscope (Ocean Optics, Model USB2000+). The spectroscope uses a diffraction grating to divide the emitted light into its different wavelength components and a (f) CCD to generate an electrical signal from the light spectrum and the spectrum is then displayed and stored on (g) a computer.

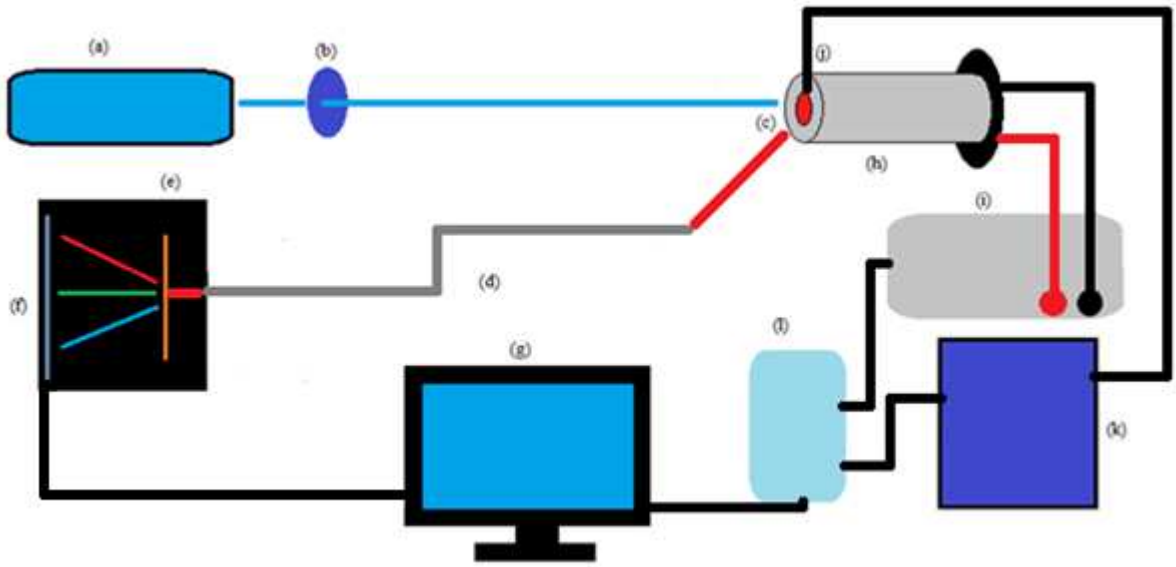


Figure 1. Experimental setup for analysing temperature dependent fluorescence intensity, (a) Helium-Cadmium laser, (b) filter, (c) La_2O_2S:Eu phosphor material, (d) fibre optic cable, (e) diffraction grating, (f) CCD, (g) computer, (h) heating element, (i) power supply, (j) thermocouple, (k) temperature display unit and (l) digital-to-analog and analog-to-digital converter.

The La_2O_2S:Eu phosphor material's temperature was changed by placing it in (h) a custom build heater and the temperature was controlled by a software program developed by the author. The temperature control software controlled (i) the power supply (Manson Switching Mode Power Supply, 1-30 VDC, 15 A, model HCS-3302) for (h) the heater. The temperature of the phosphor material was measured by a thermocouple (K-type) that was connected to (k) a unit that displayed the temperature as can be seen in figure 1. Both the temperature display unit and power supply was connected to (l) a digital-to-analog and analog-to-digital converter (uDAQ Lite Micro DAQ Data Acquisition, Serial D1000014395) that was also connected to a computer. The temperature control program uses a proportional-integral-derivative (PID) controller to control the temperature of the phosphor material.

The setup was used to excite the La_2O_2S:Eu phosphor material at a wavelength of 325 nm and to record the emission spectra at temperatures ranging from 300 K to 500 K in 10 K steps.

4. Results and Discussion

The XRD pattern of the La_2O_2S:Eu is shown in figure 2. As it can be seen in figure 2 there is a good correlation between the measured pattern and the known data for La_2O_2S (PDF-Number 27-0263). Therefore the La_2O_2S:Eu phosphor material can be recognised as well-crystallised particles in the hexagonal phase with cell parameters of a=4.051 Å and c=6.944 Å. The peak that is marked with a dash appears if it may be from Lanthanum Oxide (La_2O_3) (PDF-Number 05-0602) impurity.

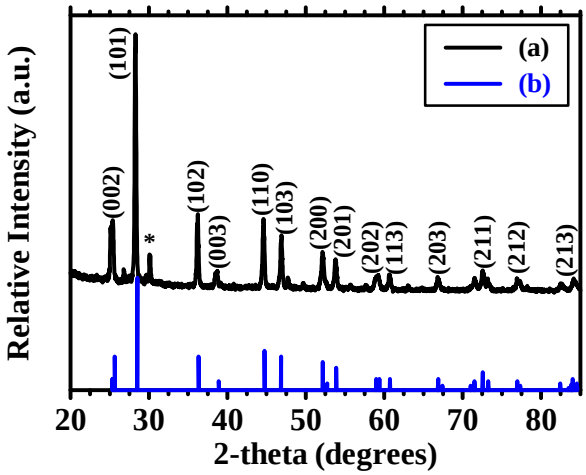


Figure 2. (a) Measured XRD pattern for the La_2O_2S:Eu phosphor material and (b) the standard data of the XRD pattern for the La_2O_2S:Eu phosphor material with peak indexing.

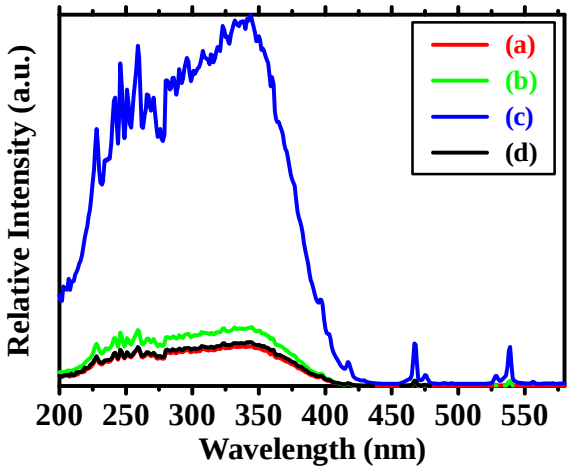


Figure 3. Excitation spectra of the (a) 594 nm, (b) 616 nm, (c) 624 nm and (d) 704 nm emission lines of La_2O_2S:Eu.

The excitation spectra (figure 3) of the 594 nm, 616 nm, 624 nm and 704 nm emission peaks were monitored for the La_2O_2S:Eu from 200 nm to 580 nm. It can be seen in figure 3 that the 325 nm He-Cd laser wavelength used in the PL system falls in the correct range for efficient excitation of the current phosphor.

The emission spectra of the La_2O_2S:Eu phosphor material was measured by the PL system at different temperatures as explained in the experimental section. Figure 4 shows the emission spectra for 5 different temperatures and it can be seen that the intensities of the emission lines decreased as the temperature increased. The integrated intensities were calculated for the six main peaks (that are numbered in figure 4) and plotted as a function of temperature in figure 5.

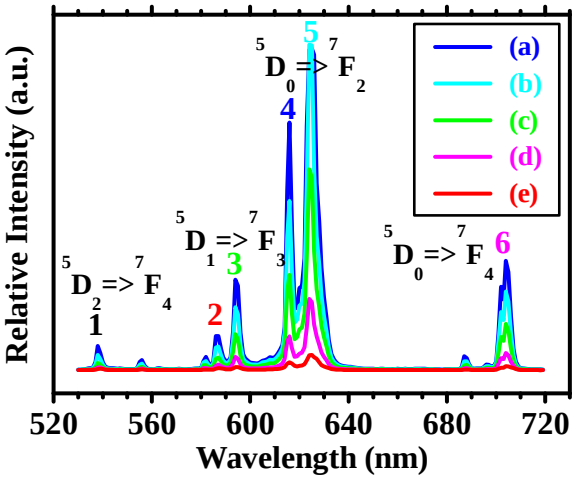


Figure 4. Emission spectra of La_2O_2S:Eu at 325 nm wavelength excitation at (a) 300 K, (b) 350 K, (c) 400 K, (d) 450 K and (e) 500 K with peak indexing corresponding to transitions of the Eu^{3+} ions. [6]

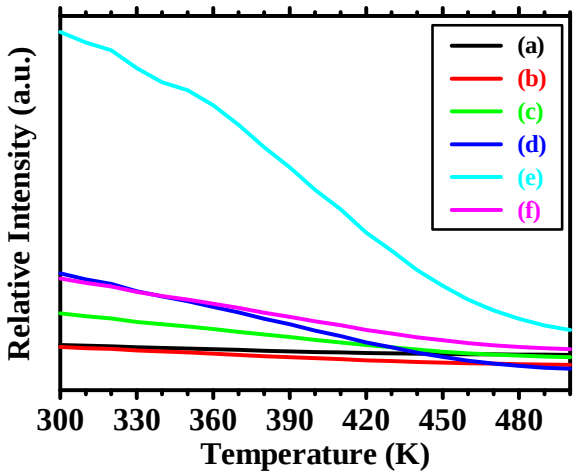


Figure 5. Emission intensity as a function of temperature for the (a) Peak 1 (535 nm - 544 nm), (b) Peak 2 (585 nm - 591 nm), (c) Peak 3 (591 nm - 599 nm), (d) Peak 4 (614 nm - 618 nm), (e) Peak 5 (618 nm – 630 nm) and (f) Peak 6 (699 nm – 709 nm) as numbered in figure 4.

To employ the FIR technique the intensity ratio of the different peaks were taken. The intensity ratio between some of the peaks as a function of temperature can be seen in figure 6. Figure 7 shows the Arrhenius plot for the same data as in Figure 6.

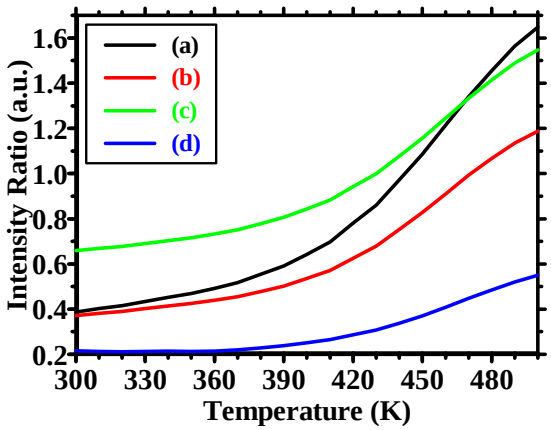


Figure 6: Intensity ratios between (a) Peak 1 / Peak 4, (b) Peak 2 / Peak 4, (c) Peak 3 / Peak 4 and (d) Peak 3 / Peak 5 as a function of temperature.

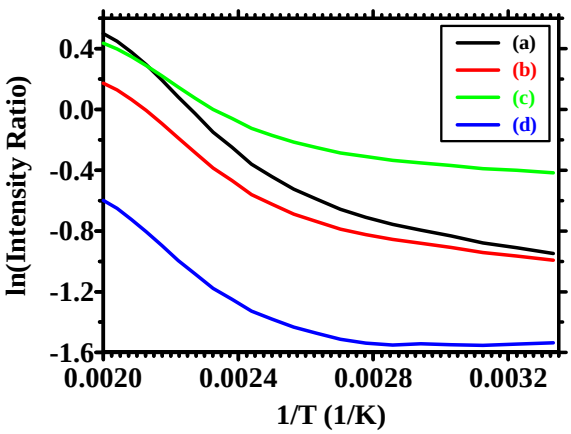


Figure 7: Arrhenius plot for intensity ratios between (a) Peak 1 / Peak 4, (b) Peak 2 / Peak 4, (c) Peak 3 / Peak 4 and (d) Peak 3 / Peak 5.

It can be seen in figure 6 that the ratios between different peaks intensities vary sufficiently that La_2O_2S:Eu could be used as a temperature sensor in the 300 K to 500 K temperature range. Thus by using the FIR technique the effect of the environmental variables could be eliminated and allow La_2O_2S:Eu to be used as a temperature sensor.

5. Conclusion

In this paper the measurement of optical thermometry properties of La_2O_2S:Eu by utilising the PL technique was successfully demonstrated. The fluorescence intensity technique where the fluorescence spectra of a phosphor material is obtained and monitored as a function of temperature was used to achieve this and by using the FIR it was shown that La_2O_2S:Eu could be used as a temperature sensor in the 300 K to 500 K temperature range. The PL system in the Physics department at the University of the Free State was also enhanced to allow the investigation of thermographic phosphors by using the fluorescence intensity technique. However the spectroscope must be replaced with a device that is capable of measuring higher resolution emission spectra to allow higher quality data.

Acknowledgements

The PL system used is supported by the Rental Pool Programme of the National Laser Centre (NLC). The authors would like to acknowledge S Cronjé for helping with the XRD measurements and the South African Research Chairs Initiative of the Department of Science and Technology (DST) and National Research Foundation (NRF).

References

[1] Allison S W and Gillies G T Remote thermometry with thermographic phosphors: instumentation and applications 1997 *Rev. Sci. Instrum.* 2615-28

[2] Khalid A H and Kontis K Thermographic phosphors for high temperature measurements: principles, current state of the art and recent applications 2008 *Open Access Sensors* **8** 5673-91

[3] Pandey A, Som S, Kumar Vijay, Kumar Vinod, Kumar K, Rai V K and Swart H C, Enhanced upconversion and temperature sensing study of Er^{3+}-Yb^{3+} codoped tungsten-tellurite glass 2014 *Sensors and Actuators B: Chemical* **202** 1305-1312.

[4] Pandey A, Rai V K, Kumar Vijay, Kumar Vinod, Swart H C, Upconversion based temperature sensing ability of Er^{3+}-Yb^{3+}codoped $SrWO_4$: An optical heating phosphor 2015 *Sensors and Actuators B: Chemical,* **209** 352-358.

[5] Binnemans K Interpretation of europium(III) spectra 2015 *Coordination Chemistry Reviews* **295** 6-15

[6] Bang J, Mostafa A, Abrams B, Williams L and Holloway P H Combustion synthesis of Eu-, Tb- and Tm- doped Ln_2O_2S (Ln=Y, La, Gd) phosphors 2004 *J. of Lumin.* **106** 177–185

Acceleration parameters for fluid physics with accelerating bodies

I.M.A. Gledhill[1], H. Roohani[2], A. Biobaku[2] and B. Skews[2]

[1]Aeronautical Systems Competency Area, DPSS, CSIR. PO Box 395, Pretoria 0001
[2]School of Mechanical, Industrial and Aeronautical Engineering, University of the Witwatersrand, Johannesburg

E-mail: igledhil@csir.co.za

Abstract. Theoretical work on transforming the Navier-Stokes equations into arbitrarily accelerating frames has included the continuity, momentum, and energy conservation equations. An analysis of the momentum equation in non-dimensional terms leads to an acceleration parameter that appears to be new in fluid physics, but is known in cosmology. A selection of cases for rectilinear acceleration has been chosen to illustrate the point that this parameter alone does not govern regimes of flow about significantly accelerating bodies, and reference must be made, above all, to the Mach number for transonic effects. Other parameters from the literature on impulsive start-up in wind tunnels are also shown to be useful in delimiting regimes of flow, such as the Freymuth start-up time. Two dominant effects in fluid dynamics with accelerating objects are shown to be flow history, a term being used to cover the difference between an instantaneous flow field with an accelerating body and the flow field about the same body at steady state, and the dependence of stagnation pressure on acceleration. The dependence of these effects on dimensionless parameters is explored.

1. Introduction

Increased agility is a topic of growing importance in aeronautics. In the design of aircraft, missiles, and Unmanned Aerial Systems, growing emphasis is being placed on the ability to out-manoeuvre another vehicle. Current missiles are launched under an acceleration of about 200 g, where g is the acceleration due to gravity, and may turn at 100 g. Earlier missiles have been ground-launched at 400 g. Accurate prediction of the transient loads using Computational Fluid Dynamics (CFD) is being explored, and in addition, methods of distinguishing regimes in which the aerodynamic effects of acceleration are significant are required in order to guide conceptual design.

Earlier reports and papers have provided theoretical support in the form of transformation of the Navier-Stokes equations between inertial and non-inertial frames [1,2,3,4] for arbitrary acceleration. A source term appears in the momentum equation in the non-inertial frame for linear acceleration. The magnitude of the linear acceleration term may be compared to that of the convection term, and provides a coefficient Q; a description is provided below. We term Q the acceleration parameter. In

addition, some guidance on low acceleration effects is available from experimental work on transient effects in the start-up of wind tunnels, and this work provides the Freymuth start-up time $t_F = \sqrt{L/a}$ [5]. These parameters are confined to a linear comparison of acceleration effects with convective phenomena.

In compressible flow with speed $v(t)$ and sound speed c, the Mach number, $M(t) = v(t)/c$, is of dominant importance. For the present applications, constant acceleration along the axis of the flare body, shown in figure 1, from $M = 0.1$ to $M = 1.0$ is modelled. A uniform acceleration of $\dot{M} = 3$ s^{-1} or $a_0 = \dot{v}(t) \sim 1041$ ms^{-2}, is modelled. Work in this paper is intended to contribute to understanding regimes of accelerating aerodynamics using Q, t_F and the Mach number.

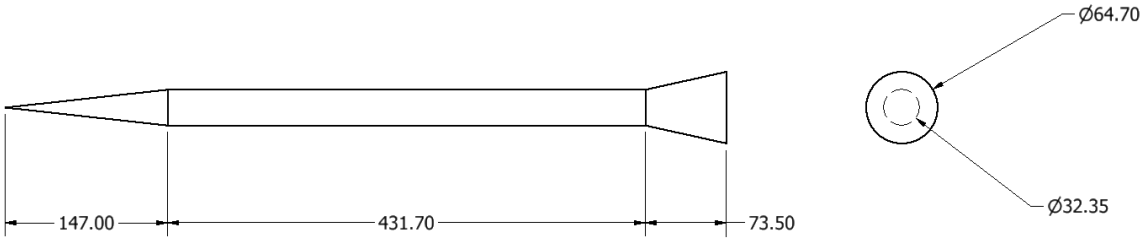


Figure 1. Flare geometry. Dimensions are in mm.

Dimensionless parameters only compare terms in the governing equations and do not describe the complicated non-linear behaviour of compressible fluids. The term flow history was introduced by Bassett [6] in connection with boundary layer development. Lilley showed flow history in compressible fluids using the Mach lines of a single particle [7]. After this paper, very little work was done analytically, and experiments have been very difficult to set up. For shapes of real interest it is necessary to explore the relevant flow fields with CFD. Roohani and Skews [8,9,10] have modelled objects of various shapes with CFD and provided evidence of flow history and inertial effects. Flow history can be visualised in terms of a lag between the current flow field and earlier flow fields. In other words, the current flow field reflects some of the unresolved features of earlier flow fields.

In this paper, the results of CFD calculation of drag forces on a slender body of rotation are discussed in order to explore regimes in which acceleration effects are easily categorised with flow parameters. The flare configuration has been explored by Forsberg *et al.* [3] and it was shown that for the specific case investigated, unsteady drag dropped in comparison with steady results through the transonic region. In work on aerofoils, shocks appear on curved surfaces in the transonic regime or as detached bow shocks at M > 1. The flare, by contrast, has a conical nose, on which attached oblique shocks should develop, and has expansion and compression corners at the nose and tail flare.

2. Theory

In an inertial frame Σ, the Navier-Stokes equations are the conservation laws for mass, momentum and energy. Transformation of mass and momentum conservation equations to an arbitrarily moving non-inertial frame Σ' is given by Forsberg [2]. For linear acceleration a_0 of the origin O′of Σ' relative to the origin O of Σ, an acceleration source term appears in frame Σ'. The equation is normalized to typical parameters, so that the coefficient of the convective term is 1. The coefficient of the time derivative term becomes the Strouhal number, $St = \frac{v_0 t_0}{L}$. The coefficient of the acceleration term is the acceleration parameter $Q = \frac{La_0}{{v_0}^2}$ where L is a typical length and $v_0$is a typical velocity; this parameter provides a linear estimate of the importance of acceleration-related effects in comparison to convective

flow. Inclusion of gravity $\boldsymbol{g}$ as an external force in Σ leads to an analogous term characterized by the Froude number, $Fr = \frac{v_0}{\sqrt{Lg}}$. The parameter $Q$is related to Fr^{-2}.

Typical length L is defined as the length of the flare and is shown with other parameters in Table 1. The pressure and temperature used were 101325 Pa and 300 K respectively. The Freymuth start-up time is an estimate of the time taken to traverse length L at acceleration a_0, starting from rest. It has been used as a normalising factor in examining the transient vorticity developing over aerofoils in wind-tunnel start-up studies, with a_0 referring to the flow acceleration. While this case is not equivalent to the acceleration of an object from rest, an object accelerating from rest in still air would take of the order of the Freymuth time to travel its own length L, and the factor may therefore be taken to be an indicator of the minimum time during which start-up transients may be expected. The Freymuth time indicates that start-up transients over this flare may be expected to last for 0.03 s, or when $M \sim 0.23$ in the present flare case.

Table 1. Typical parameters

Parameter	**Value**
Typical length L	652.3 mm
Sound speed c	347 ms^{-1}
Initial Mach number M_0	0.10
Typical Mach number M	1
Typical acceleration a_0	1041 ms^{-2}
Acceleration parameter $Q = \frac{La_0}{{v_0}^2}$	0.26
Freymuth start-up time $\sqrt{L/a_0}$	0.03 s

It is noted that at the slender conical nose with an attached oblique shock, stagnation properties are an approximation to surface pressures that are likely to be modelled.

3. Methodology

The model geometry is shown in figure 1. The primary interest of this model is in examining shocks, surface pressures and drag, and it is reasonable to use an axisymmetric approximation which will eliminate asymmetric vortex shedding in the wake.

The acceleration terms for the momentum and energy were provided as source terms in the non-inertial frame Σ' through user coding in a finite-volume cell-centred unstructured code ANSYS® Fluent. A second-order upwind scheme was used for shock capture. A two-dimensional hybrid quadrilateral-triangular unstructured grid, with adaptive mesh refinement, was used, and local Reynolds numbers y^+ were checked at the wall for compatibility with the Spalart-Allmaras turbulence scheme, which was chosen as a well-validated model in the transonic regime. The far field boundaries receive relative velocities that increase with the acceleration a, while the walls are no-slip walls with velocity zero relative to the body. Mesh independence was established to a level of 5% variation in loads between successive grids. A time step of 10^{-4} s was found to be small enough to capture the same loads and flow fields as time steps less than 10^{-4} s. A grid convergence study was performed on the global mesh size before checking results for pressure along the missile wall with adaptive gridding on the steady state solutions. Grid adaption was performed every 10 time steps in the transient solutions. Verification of the source terms was performed by inspecting the evolution of the pressure field across the domain.

4. Results

Drag is not normalized in acceleration cases, but presented in Newtons.

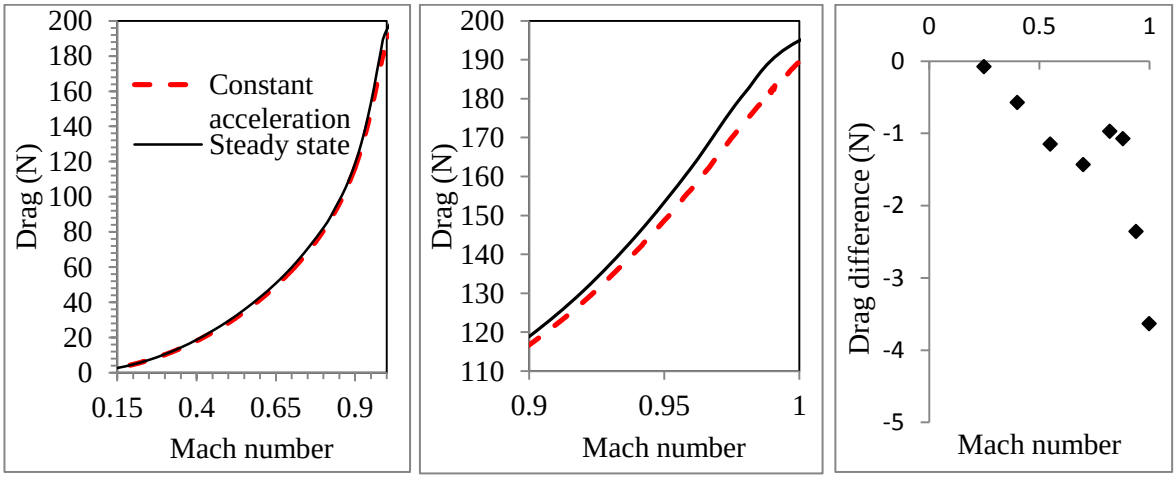


Figure 2. Drag and drag difference as a function of Mach number

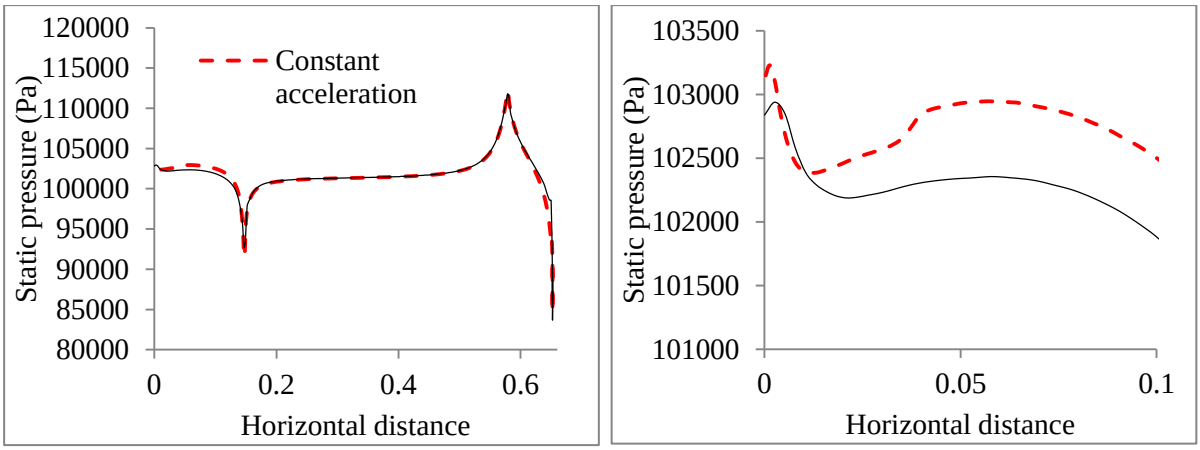


Figure 3. Pressure profiles for M=0.7 as a function of distance from the nose over (a) the length of the flare and (b) the stagnation region

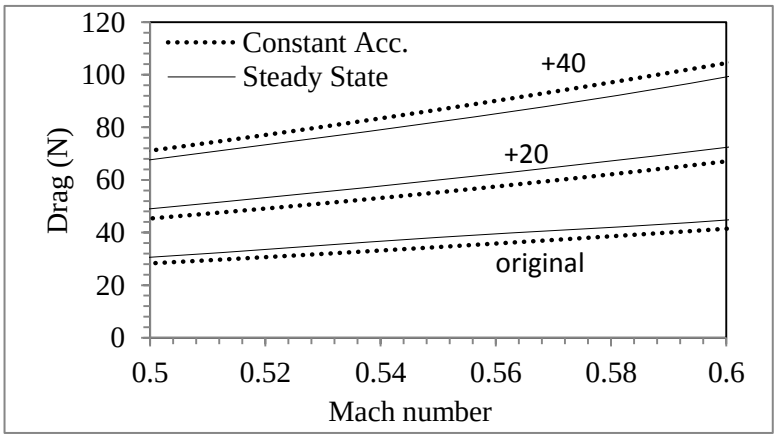


Figure 4. Drag in the subsonic region for increased calibre geometries

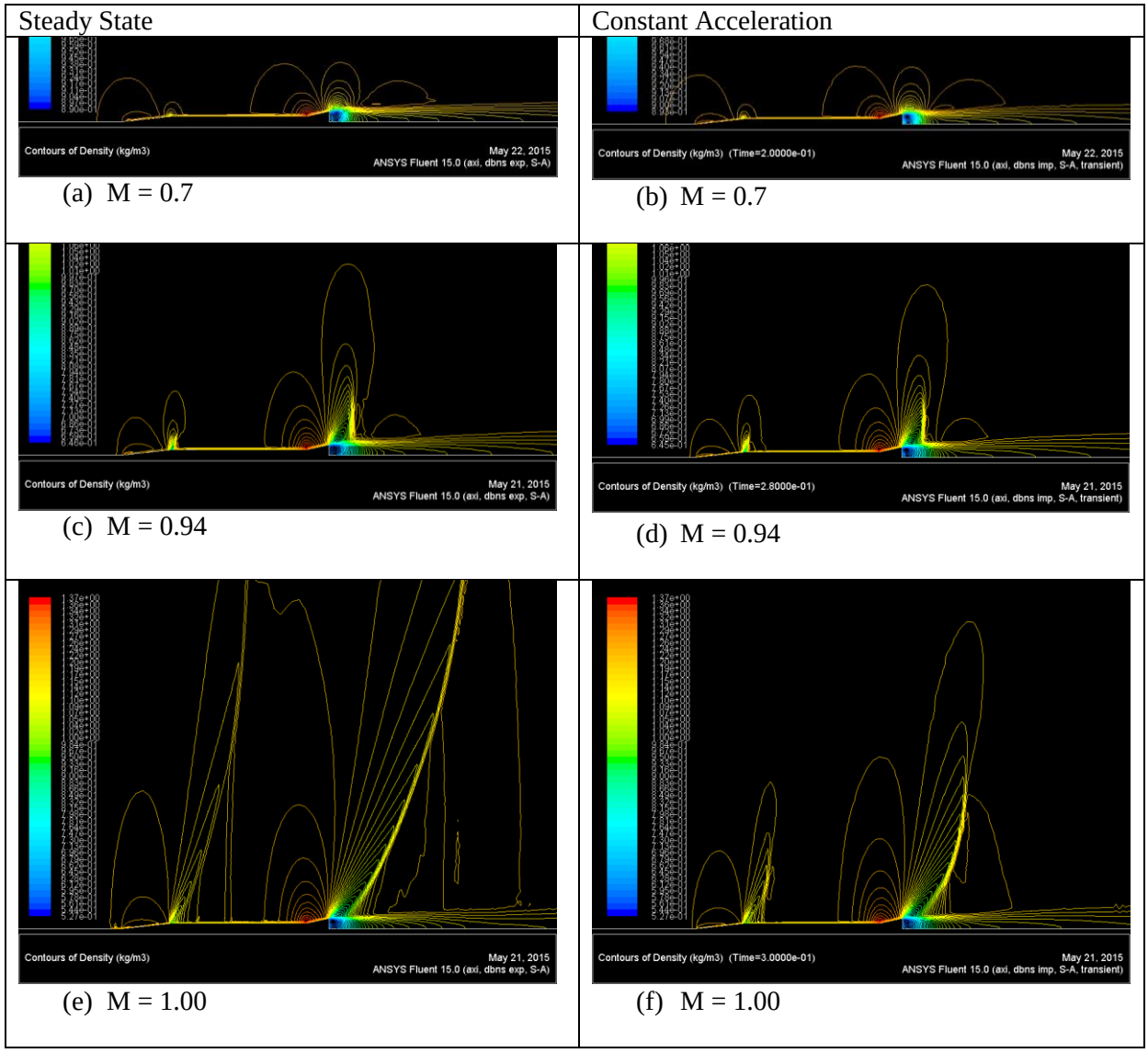


Figure 5. Instantaneous density flow fields. Left: steady state. Right: constant acceleration

5. Discussion of results

Figure 2 compares steady and unsteady drag at Mach numbers in the subsonic and transonic range. It is clear that for this particular geometry drag is usually lower in the unsteady case during acceleration. However, in previous work by Roohani and Skews [8,9,10] the opposite effect was observed in aerofoils and bluff bodies. This is because unsteady drag is not only a function of Mach number and acceleration, but it is also affected by the geometry of an object. In Figure 4 this phenomenon is illustrated by increasing the diameter of the flare by 20 mm and 40 mm and subjecting it to the same steady and unsteady conditions. It is shown that when the diameter is increased by 40 mm, acceleration results in an increase in drag compared to the steady state scenario. In figure 3 the pressure profile at Mach 0.7 as a function of distance from the nose shows that in the frontal area of the flare the pressure has increased due to acceleration. This is due to fluid inertia. At the tail flare the pressure is lower during acceleration, which can be attributed to flow history and boundary layer effects. At the flat base surface the minimum pressure during acceleration (84.5 kPa) is higher than the

minimum steady state pressure (83.4 kPa). Both these effects contribute to the lower drag observed during acceleration. Contours of density are shown in figure 5, where the steady state and the accelerating cases are compared at specific Mach numbers. At Mach 0.94 a recovery shock has developed in the wake of the flare in both cases. There appears to be a shock developing behind the expansion corner on the nose cone, which could be attributed to a separation bubble. This needs more detailed investigation as it appears to be a function of acceleration. In addition, the wake shock has strengthened in both cases, but extends a smaller distance from the centerline in the accelerating case. The decreased wave drag accounts for the sharp drop in drag for the accelerating case in the transonic region, shown in the data of figure 2(c).

The Freymuth time, which is a measure of the time in which start-up transients would be expected, is 0.03 s in this case, and the total trajectory is covered in 0.3 s. The start-up transients were not observed to be significant under these circumstances.

6. Conclusions

It has been shown that for a slender body of rotation, drag under acceleration of $1041 m/s^2$ from M = 0.1 is lower than drag in steady flow across the subsonic region, and that there is a significant decrease in unsteady drag near Mach 1. This is in contrast with previous work on aerofoils and bluff bodies. It is also shown that increasing the diameter of the slender body reverses the trend, so that drag under acceleration in the subsonic region is larger than steady drag. The decrease in drag is attributed to the effects of flow history over the rear of the flare body, including a decrease in wave drag as the wake shock is reduced in extent during acceleration. In these cases, flow history dominates flow inertia. However, as the diameter of the flare body is increased, flow inertia effects on the nose cone increase and dominate flow history effects, and the total drag rises above that of the steady state.

In this case, the Mach number is a clear indicator of changing compressible effects near Mach 1, where drag changes are very significant. The acceleration parameter Q is 0.26 based on the initial velocity, and very significant acceleration effects have been observed. This factor should be explored further in future work as an indicator of the relative significance of acceleration in understanding flow fields and aerodynamic loads.

References

[1] Löfgren P 1998 *Relative Motion in Fluid Mechanics* FFAP-B-066 (Stockholm: Flygtekniska försöksanstalten, The Aeronautical Research Institute of Sweden)

[2] Forsberg K 1999 *Treatment of a moving reference frame for discretized NS equations* (Stockholm: Flygtekniska försöksanstalten, The Aeronautical Research Institute of Sweden)

[3] Forsberg K, Gledhill I M A, Eliasson P and Nordström J 2003 Investigations of acceleration effects on missile aerodynamics using CFD 2003 (American Institute of Aeronautics and Astronautics) *AIAA-2003-4084*

[4] Gledhill I M A, Forsberg K, Eliasson P, Baloyi J and Nordström J 2009 Investigation of acceleration effects on missile aerodynamics using computational fluid dynamics *Aerospace Science and Technology* **13** 197–203

[5] Freymuth P, Bank W and Palmer M 1983 Visualization of accelerating flow around an airfoil at high angles of attack *Z. Flugwiss. Weltraumforsch.* **7** 392–400

[6] Basset BA1888 *Treatise on Hydrodynamics* (Cambridge: Deighton, Bell and Co.)

[7] Lilley GM, Westley R, Yates AH and Busing JR 1953 Some aspects of noise from supersonic aircraft *J. Royal Aeronautical Soc. College of Aeronautics Report* **71** 396-414

[8] Roohani H and Skews BW 2008 Unsteady aerodynamic effects experienced by aerofoils during acceleration and retardation *Proc. I. Mech.E.* **222** 631–636

[9] Roohani H and Skews BW 2009 The influence of acceleration and deceleration on shock wave movement on and around airfoils in transonic flight *Shock Waves* **19** 297–305

[10] Roohani H *Aerodynamic effects of accelerating objects in air* 2010 PhD Thesis (Johannesburg: University of the Witwatersrand)

Unfolding the fast neutron energy distribution of a NE230 deuterated liquid scintillator detector using the MAXED code

M S Herbert
Physics Department University of the Western Cape, Robert Sobukwe Road, Bellville, 7535, South Africa

msherbert@uwc.ca.za

Abstract. A NE230 deuterated liquid scintillator detector has been used to measure the neutron fluence energy distribution in air of a neutron beam of energy up to ~ 48 MeV. The fluence energy distribution was obtained from measurement of the pulse height distribution by the NE230 detector using the Bayesian unfolding code MAXED with a response matrix that was determined experimentally. The unfolded fluence energy distribution obtained is compared with the fluence energy distribution of the beam measured with a NE213 natural hydrogen liquid scintillator detector using the time-of-flight method.

1. Introduction

Detailed knowledge of the neutron fluence energy distributions are useful in basic research and applications, particularly in applications such as neutron radiotherapy for the treatment of cancer [1]; radiobiology, studying the biological effectiveness of neutrons [2] and radiation protection at nuclear facilities [3]. In neutron radiotherapy, this knowledge is important to calculate the energy distributions of secondary charged particles, and to characterize the radiation quality and absorbed dose both inside and near to the area under treatment. Measuring these neutron fluence energy distributions pose a challenge at energies above 20 MeV since reaction cross-sections above 20 MeV have either not been measured or are not correctly calculated by present nuclear models.. In principle, these fluence energy distributions can either be calculated by Monte Carlo Methods or measured experimentally [4].

There are a variety of methods that can be used to measure neutron beam fluence energy distributions [5]. Of these methods the time-of-flight is the most widely, especially if pulsed beam is available. However, for measurements in water (simulating human tissue) the neutron flight path is undefined, alternative methods need then to be used, such as those based on unfolding analyses [1]. In unfolding analyses the fluence energy distribution has to be unfolded from the corresponding pulse height pulse height distribution that is recorded from the detector which results from the neutron interactions in the detector medium. The fluence energy distribution $\Phi(E)$, the pulse height distribution $\frac{dN}{dH}$, and the detector response $R(H,E)$, are related through the Fredholm integral equation of the first type as below [6] such that;

$$\frac{dN}{dH} = \int R(H,E)\phi(E)dE \qquad (1)$$

When pulse height distribution is recorded by a multichannel analyzer, Equation. (1) takes the discrete form;

$$N_i = \sum_j R_{ij}\phi_j \tag{2}$$

where *Ni (i=1, 2, ..., n)* is the recorded counts in the *ith* channel, *Φj (j=1, 2, ..., m)* is the radiation fluence in the *jth* energy interval, and *Rij* is the response matrix coupling the *ith* pulse height interval with the *jth* energy interval. Equation (2) can be transformed into the matrix notation such that;

$$N = R\phi \tag{3}$$

Where $\boldsymbol{N}=(N1, N2, ..., Nn)^T$, $\boldsymbol{\Phi}=(\Phi 1, \Phi 2, ..., \Phi m)^T$, $\boldsymbol{R}$ is the response matrix with size of $n \times m$. To obtain the neutron fluence energy distribution from the measured pulse height distribution equation (3) need to be inverted. Several mathematical methods and computing algorithms have been used to solve equation (3), such as least-squares [7], Monte Carlo Methods [8], genetic algorithm [9], and populated artificial neural networks in recently years [10]. Using any one of the mathematical methods to obtain the fluence energy distribution from the measured pulse height distribution requires reliable knowledge of the detector response matrix. Monte Carlo codes (eg. MCNPX) which may be used to simulate the detector response matrix, are limited to neutron energies below 20 MeV, since the non-elastic contribution to the reaction cross-sections is significant and have either not been measured or are not correctly calculated by present nuclear models.

This paper presents the overall procedure of measuring and unfolding the fast neutron fluence energy distribution of a neutron beam of energy up to ~ 48 MeV, in air, with NE230 deuterated liquid scintillator detector, using an experimentally determined response matrix. The unfolded fluence energy distribution obtained is compared with the fluence energy distribution of the beam measured with a NE213 natural hydrogen liquid scintillator detector using the time-of-flight method [11].

2. Experimental procedures and Data analysis

Experiments were conducted at the neutron time of flight facility at the iThemba LABS in Faure, outside Cape Town, South Africa. Fig. 1 is a schematic diagram showing the details of the beam line in the neutron vault, the shielding in the experimental area and the positions of the detectors including the neutron monitor.

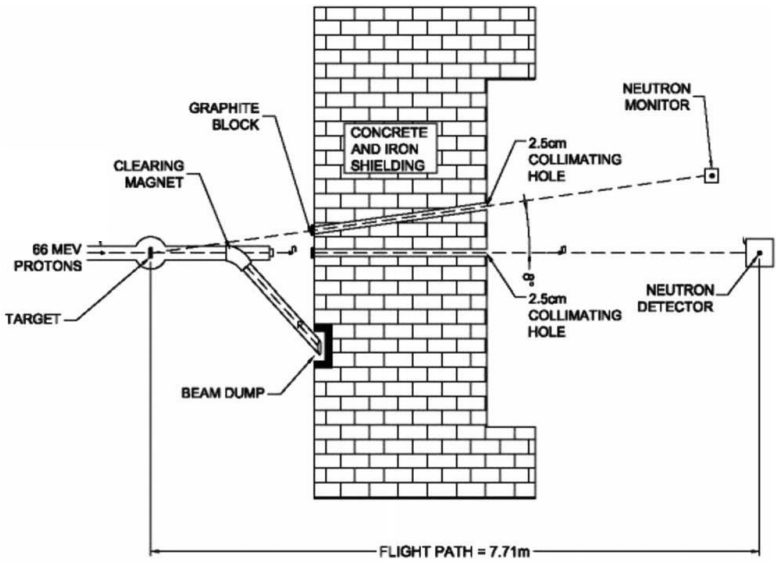


Figure 1. Schematic diagram of the neutron beam line, shielding and detector.

Neutrons were produced by bombarding either a Li metal target (thickness 1 mm) or a Be metal target (thickness 10 mm) or a graphite target (thickness 10 mm) with a pulsed beam of 66 MeV protons from the iThemba LABS time-of-flight facility. A 2 m thick shielding wall (concrete and iron) separated the experimental area from the target. A circular aperture (25 mm diameter) in the wall provided a collimated neutron beam at angle 0° to the proton beam direction. The neutron beam profile measured at the end of the neutron flight path of 7.7 m was found to be uniform within 5% over a circular area of diameter 50 mm. Measurements of the neutron beams were taken with either the NE230 detector (25 mm diameter x 25 mm) or a reference detector NE213 detector (50 mm diameter x 50 mm) at a distance of 7.7 m away from the target. Both the detectors were equipped with LINK pulse shape discriminator units to suppress gamma rays and to select only events such as n-p elastic scattering or n-d elastic scattering and all heavier particles resulting from neutron interaction in the detector mediums [1 and 12].

All runs were normalised to the same number of neutrons measured by the neutron monitor. After applying the event selection procedures, data were analysed as follows. For the generation of the experimental response matrix, two-parameter distributions of events as a function of pulse height L and neutron time-of-flight T were obtained from the combined two parameter (LT) data obtained using the Li, Be and C targets. The neutron energy, E was determined from T. The pulse height response functions of the NE230 detector were determined at 29 equally spaced 2 MeV neutron energies ranging from 10 MeV to 66 MeV from the (LT) data. The response functions were then combined to form a response matrix of dimensions 29 (E) X 104 (L) as described in Brooks *et al.*[13]. The NE230 detector detection efficiency $\varepsilon_d(E)$ for each of the 29 energies were determined from the ratio of counts recorded (in the same neutron beam) by the NE230 detector and NE213 reference detector, respectively. The efficiency $\varepsilon_d(E)$ is given by the product of this ratio and the known detection efficiency $\varepsilon_p(E)$ of the NE213 detector, which was determined by reference to the n-p elastic scattering cross-section, as described in Klein and Brooks [5].

Applying pulse selection in the off-line analyses the pulse height distribution for deuterons and heavier charge particles resulting from the interactions of neutron beam of energy up to ~ 48 MeV (neutrons produced by a 66 MeV proton beam on a graphite target) with the NE230 detector medium was determined. The Bayesian unfolding code MAXED [14] using the experimentally determined response matrix was use to determine the neutron fluence energy distribution from the pulse height distribution. The MAXED code use an unfolding algorithm based on maximum entropy principle to problems of data analysis and require prior information in the form of a prior energy distribution to obtained solution energy distribution to equation (3). The code is available from the PTB Braunchweig, Germany. The unfolded fluence energy distribution obtained is compared with the fluence energy distribution of the beam measured with a NE213 natural hydrogen liquid scintillator detector using the time-of-flight method.

3. Results and discussion

In this work the pulse height distribution was unfolded with the MAXED code, using the experimentally determined response matrix with either the time-of-flight prior energy distribution containing a lot of information about the solution energy distribution or the flat prior energy distribution which contain no information about the solution energy distribution, as shown in figure 2. Figure 3 shows the results of the pulse height distribution measured (blue curve), together with refolded MAXED fit (red curve) from the unfolding using the time-of-flight prior energy distribution. The refolded MAXED fit agrees well with the measured pulse height distribution. A similar result was obtained from the unfolding using the flat prior energy distribution. Figure 4 shows results of the neutron fluence energy distributions obtained from the MAXED unfolding, using (a) the time-of-flight prior energy distribution (black solid circles) and (b) flat prior energy distribution (blue solid circles). In each frame the unfolded neutron fluence energy distribution obtained is compared with the fluence energy distribution (black solid line histogram) obtained by the time-of-flight method.

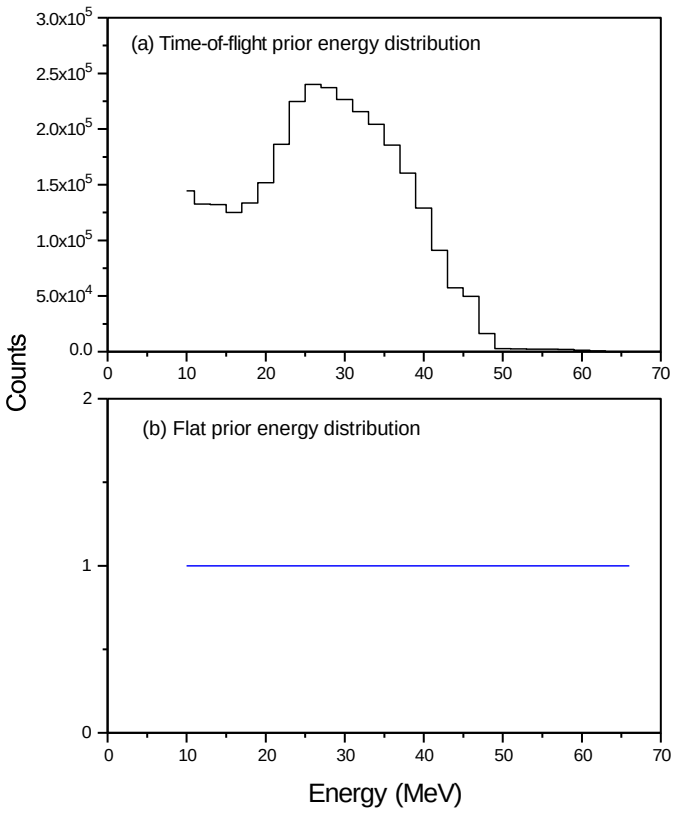


Figure 2. (a) Time-of-flight prior distribution and (b) flat prior distribution used in unfolding

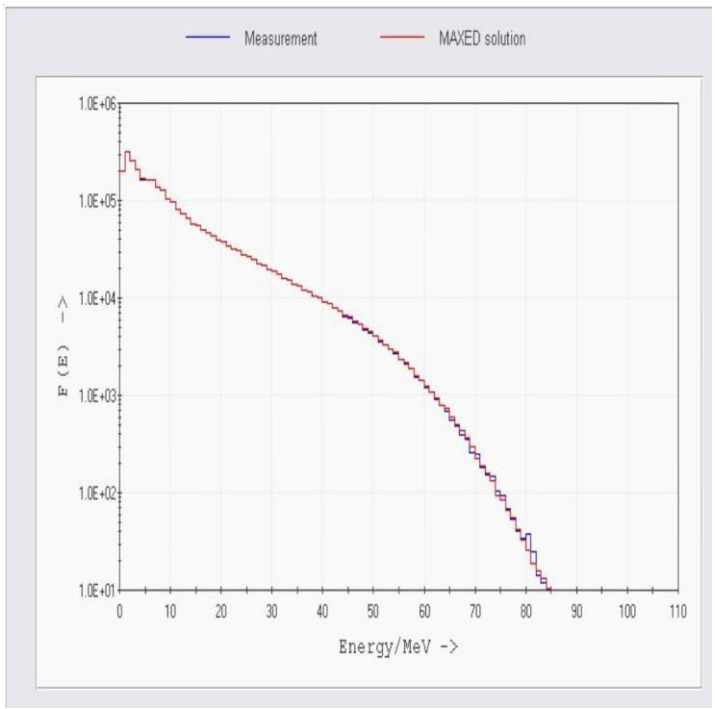


Figure 3. Measured pulse height distribution together with the refolded MAXED fit using the time-of-flight prior energy distribution.

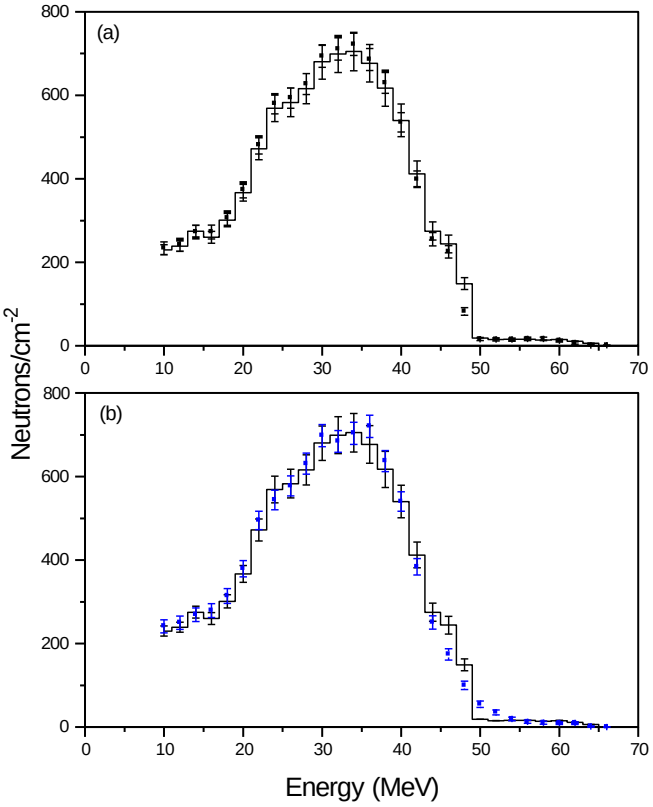


Figure 4. (a) Comparisons of fluence energy distributions (solid black circles) obtained with unfolding using time-of-flight prior energy distribution and fluence energy distributions (solid blue circles) obtained from unfolding using flat prior energy distribution with fluence energy distribution (solid line histogram) obtained with the time-of-fight method

Both fluence energy distributions obtained by unfolding agrees well with the fluence energy distribution obtained with the time of flight method. However, it can be seen from the plots that at high energies the unfolded fluence energy distribution obtained by unfolding using the time-of-flight prior energy distribution agrees better with the fluence energy distribution using the time-of-flight method. This better agreement obtained at higher energies with the time-of-flight prior energy distribution might be attribute to the fact that the prior information compensate for the response matrix which are less sensitive at higher energies due to the wall effect [1].

4. Conclusions

The work done demonstrates that the unfolding code MAXED with an experimentally determined response matrix can reliably determine the fast neutron fluence energy distribution of a neutron beam of energy up to ~ 48 MeV, in air. In future this work will be extended to increase the sensitivity of the response matrix at higher energy. Also the unfolding procedures will be use to investigate how well the unfolding can be use to determine the fast fluence energy distributions of quasi-mono- energetic beams at neutron energies above 20 MeV.

5. References

[1] Herbert M S, Brooks F D, Allie M S, Buffler A, Nchodu M R , Makupula S A, Jones D T L, Langen K M 2007 *Radiat. Prot. Dosim* Vol. **126**, No. 1–4, pp. 346–349

[2] Nolte R, Mhlbradt K, Meulders J, Stephan G, Haney M and Schmid E. *Radiation and environmental biophysics*, **44** (3), 201−209.

[3] Taniguchi S, Nakamura T, Nunomiya T, Iwase H, Yonai S, Sasaki S, Rokni S, Liu, J, Kase K and Roesler S 2003 *Nuc.Instru. and Meth* **A503**, 595

[4] Bohm T, Deluca P, Cox L, Maughan R, Jones D and Lennox A 1999 *Med. Phys.* **26,** 783−792.

[5] Brooks F and Klein H 2002 *Nucl. Instr. Meth.* **A476**, 1−1.
[6] Knoll G F 2010 *Radiation Detection and Measurement* fourth edition. Hoboken New Jersey: John Wiley & Sons, Inc.
[7] Koohi-Fayegh, Green S, Scott M C 2001 *Nucl. Instrum. Meth.* A **460:** 391−440
[8] Sanna R, Obrien K 1971 *Nucl. Instr. Meth.* **91**: 573−576
[9] Mukherjee B. BONDI-97 1999 *Nucl. Instr. Meth. A* **432**: 305-312
[10] Avdic S, Pozzi S A, Protopopescu V 2006 *Nucl. Instr. Meth. A* **565:** 742−752
[11] Herbert, M 2014 *Rad Prot Dosim* pp. 1-6.
[12] Adams J and White G 1978. *Nucl. Instr. and Meth* **156** 459−476
[13] Brooks F D, Buffler A, Allie M S, Herbert M S, Nchodu M R, Jones D T L, Smit F D, Nolte R, Dangendorf V 2007 *Rad Prot Dosim* Vol. **126**, No. 1–4, pp. 218–222
[14] Reginatto M, Goldhagen P. and von Neumann S 2002 *Nucl. Instr. Meth. A* **476** 342

Fibre-to-the-Hut Technology: A solution for Cheap Access for High Speed-Optical Network in South Africa

G M Isoe[1], E K Rotich[1], R R G Gamatham[2], A W R Leitch[1] and T B Gibbon[1]

[1] Optical Fibre Research Unit, Nelson Mandela Metropolitan University,
P O Box 77000, Port Elizabeth, 6031, South Africa Tel: +27 41 504 2141

[2] The Square Kilometre Array Project (SKA) South Africa

E-mail: George.Isoe@nmmu.ac.za

Abstract. Fibre-to-the-Home (FTTH) is a technology where optical fibre networks are deployed from a central access point to individual homes to provide high-speed broadband access. FTTH has extensively been deployed in many countries with high population density within large cities and urban centres and high per capital income in Asian and Europe. However, African countries are still facing some challenges like uneven population distribution with isolated remote villages and socio-economic challenges. This hinders direct implementation of traditional FTTH solutions in Africa. It is for these reasons that we specially customize the traditional FTTH, based on challenges facing Africa and design a Fibre-to-the-Hut (FTTHut) optical network to suit the African scenario. The use of vertical cavity surface emitting Laser (VCSEL) within a Raman amplified optical fibre framework to support FTTHut technology in South Africa is studied. VCSELs offer high bandwidth at low drive currents, while fibre Raman amplifiers offer longer amplification spans. We therefore investigate experimentally the noise figure (NF) and optical signal to noise ratio (OSNR) performance of fibre Raman amplifier (FRA) using a directly modulated 1550 nm VCSEL as the signal source. Two pumping techniques namely co- and counter pumping were employed. An OSNR of 6.8 dB and 6.4 dB was achieved for co- and counter pumping schemes, respectively, for 25 km SMF-Reach. An OSNR of 4.5 dB and 4.3 dB was attained for 50 km fibre for co- and counter pumping respectively. A NF of -1.3dB and -0.7dB was achieved for co- and counter pumping schemes, respectively, for 25 km fibre at 23 dB pump power. The NF also increased with increase in fibre. This work is valuable in providing South Africa with increased Hut-to-Hut broadband access especially in long-reach networks serving rural populations at reasonable low cost.

1. Introduction

The demand for internet services by residential homes in South Africa has risen rapidly [1]. End users are demanding for a fast, reliable and affordable communication network that can support broadcast television, radio, voice telephony, video-based multimedia, peer-to-peer file transfer, high definition multimedia on-line gaming among many other services. Previously, a limited number of carriers monopolised access to the only major cable serving the country, driving the access cost high. However, the landing of four international submarine fibre-optic cables in South Africa between 2009 and 2012 brought down the cost of international bandwidth dramatically. As a result, fibre to the home (FTTH) optical networks is becoming a reality in South Africa due to its network transparency and bandwidth advantage [2].

As a matter of fact, a number of companies in South Africa have committed to initiate the roll out of FTTH in some major cities namely Johannesburg, Cape Town, Port Elizabeth, Pretoria, Durban and Bloemfontein. These cities have been singled out because they make up the most developed and highly populated urban areas in South Africa; hence the FTTH models applied in developed countries overseas can be followed and implemented to benefit people living in these cities. However, the larger part of the South African population lives in rural areas, where the social and geographical dynamics differ significantly as these areas are associated with uneven population distribution with isolated remote villages [3]. The rural population has also been marginalized from access to facilities, as well as isolated from some basic services. This hinders the implementation of traditional FTTH solutions in rural South Africa thus a customized, high cost sensitive solution is urgently required.

It is for these reasons that we specially customize the FTTH based on these challenges and design a Fibre to the Hut (FTTHut) optical network to suit the larger part of the South African population who live in the rural areas [4]. FTTHut is expected to act as an alternative to electrical signal based technologies such as digital subscriber line techniques (DSL, ADSL and VDSL) and cable modem which are increasing in speed but at the cost of reduced network reach.

The FTTHut approach factors in all the constraints involved. In our approach, we introduce VCSEL technology and Raman amplification in the Access network to lower the transmission cost as well as increase network reach. VCSEL laser sources were preferred because they offer a cheap solution to access networks since they are relatively of low cost. Additionally, these devices consume very low power as they only operate in the mA region.

2. Theory

VCSELs are semiconductor sources with a monolithic laser resonator. The emitted light leaves the device in a direction perpendicular to the surface. The cavity is made up of two semi-conductors Bragg mirrors between which there is an active region with numerous quantum wells and a total thickness of only a few micrometres [5]. A lot of study on VCSELs application in fibre-optic networks and as optical interconnects has so far been carried out [6, 7, 8]. VCSELs have drawn much attention because of their; compatibility, low cost wafer scale fabrication and testing techniques, high-volume, lower-cost manufacturing and compatibility with most active optical devices. These devices also possess some other special features that make them attractive for fibre-optic systems. These include a circularly shaped output beam for high coupling efficiency, high modulation bandwidths at low current levels, single mode operation, low power consumption and the potential to produce integrated modules and arrays on wafer [8].

However, Frequency chirp and chromatic dispersion is the main limiting factor for VCSELs. Frequency chirping is the instantaneous change of the central optical frequency υ in response to variations in optical power. The instantaneous frequency chirp is expressed as [9, 10].

$$\Delta\nu(t) = -\frac{\alpha}{4\pi}\left(\frac{d}{dt}InP(t) + kP(t)\right) \tag{1}$$

where P(t) is the instantaneous optical power, and α is the laser linewidth enhancement factor and the k parameter are constants. The k parameter is related to the non-linear gain and depends on the geometry of the device.

3. Research Design

A complete experimental scheme is illustrated in Fig.1. A 10Gbps 1550 nm VCSEL is directly modulated with a non-return to zero pseudo-random binary sequence (PRBS) of pattern length of 2^7-1. The reach needs to be extended due to geographical fluctuations, and uneven population distributions within remote villages, at the same time maintain the signal integrity. This calls for proper signal amplification thus Raman amplification was adopted due to its low noise and remarkable signal recovery. The signal source was then coupled with a Raman pump operating at a wavelength of 1450 nm with a maximum output power of 25 dBm, then transmitted over an ITU-T G. 655 fibre at varied

lengths and the recovered signal analyzed using an optical oscilloscope and a power meter. Two pumping techniques namely co-pumping and counter pumping were adopted.

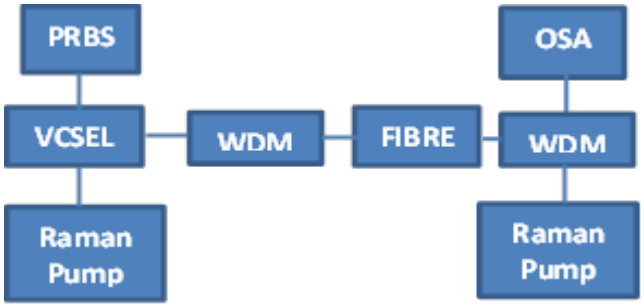


Figure 1: Raman amplified optical transmission

Their performances were then investigated based on the OSNR and NF so as to determine the best pumping technique for extended reach application. An ITU-T G.655 optical fibre manufactured by the Optical Fiber Solutions (OFS) Company was used. G.655 fibre is well suited for long transmissions because of its low attenuation, low dispersion and flexibility for dispersion management purposes.

4. Results

The OSNR behaviour for co and counter pumping techniques at different pump powers is shown in fig 2**.**

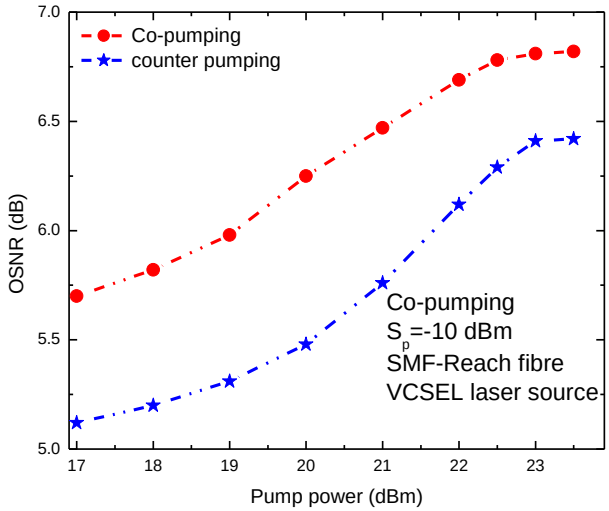


Figure 2: OSNR variation with pump power for co-pumping and counter pumping schemes of a 25 km ITU-T G.655 fibre.

The OSNR increases with increase in pump power in both pumping techniques. However, at 23 dBm pump power, the OSNR flattens due to gain saturation. Co-pumping shows superior OSNR performance of 0.45 dB than counter pumping at 23dBm pump power.

From Fig 2, Co-pumping shows a better OSNR performance than counter pumping. This is because for co-pumping, the signal is co-propagated with the pump. This ensures that much of the pump power is transferred to the signal to achieve higher amplification levels.

Fig 3 shows the OSNR performance of a 25 km and 50 km ITU-T G.655 fibre lengths at co-pumping scheme.

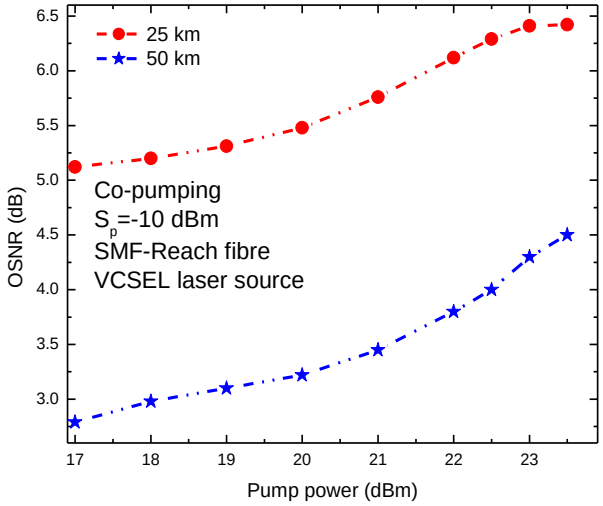


Figure 3: OSNR as a function of pump power at different fibre lengths for co-pumping scheme.

An OSNR of 6.41 dB and 4.50 dB is recorded for a 25 km and 50 km fibre ITU-T G.655 fibre respectively, at 23 dBm pump power. The OSNR degradation with increase in fibre length is attributed to ASE noise accumulation within the transmission fibre.

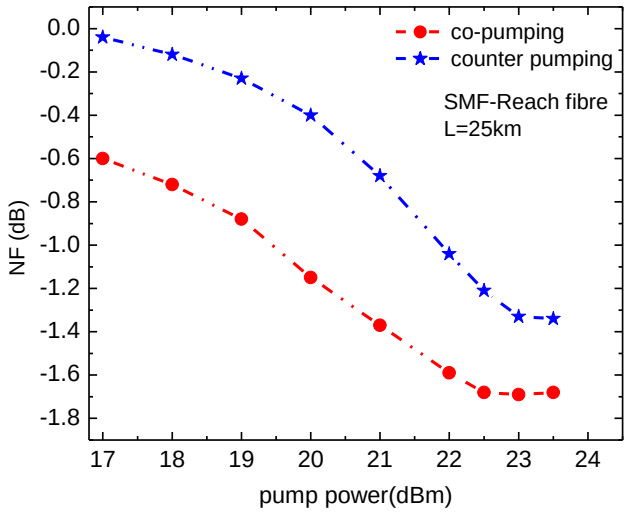


Figure 4:NF variations with pump power for co-pumping and counter pumping schemes.

Fig.4 shows the noise figure (NF) performance of the two pump techniques. Co-pumping shows a better NF performance than counter pumping. This is due to the fact that for co-pumping scenario, both the pump power and the signal power propagate in the same direction. This in turn results to more pump to signal power transfer hence higher amplification levels are attained. Secondly, the ratio between the pump power and the signal power is maintained almost at the same level throughout the transmission span. For the case of counter pumping, the signal propagates counter to the pump direction. This limits the pump to signal power transfer thus leaving much of the pump power unutilized.

5. Conclusion

We have presented results on the OSNR and NF performance analysis of a Raman amplified optical network using a directly modulated 4.25Gbps 1550 nm VCSEL. Signal degradation in optical fibres during Raman amplification depends on the fibre length, pump power and the pumping scheme used.

The OSNR decreases with an increase in fibre length. This decrease is more pronounced in counter pumping than in co-pumping scheme.

From our results, the ASE noise accumulation was minimal in the co- pumping thus accounting for the higher OSNR recorded in this pumping scheme. This makes co-pumping superior in noise performance compared with the counter pumping. We therefore recommend co-pumping technique for the FTTHut extended reach applications. This work is valuable in providing FTTHut project in South Africa with a VCSEL technology and suitable signal recovery technique through forward Raman amplification for increased broadband access in the extended reach networks to serve the rural population at a very reasonable cost.

6. Acknowledgment

We are grateful for Research Funding and support from: Telkom, Dartcom, Ingoma, CISCO, DST, CSIR, NLC, NRF, THRIP, ALC and scholarship funding from SKA/NRF.

References

[1] http://www.budde.com.au/Research/South-Africa-Broadband-Market-Insights-and-Statistics.html#sthash.qUpcHpCD.dpuf

[2] Tim P S, "Update on fibre deployments in South Africa & drivers and inhibitors," FTTH Conference, turbine hall, Johannesburg 28-29 October, 2014

[3] Linard C, Gilbert M, Snow R W, Noor A M and Tatem A J, "Population Distribution Settlement Patterns and Accessibility across Africa in 2010," Plos One, vol. **27**, issue 2, e31743, 2012

[4] Romeo G, Enoch R, Andrew L and Timothy G, "Fibre-to-the-Hut: Research into Tailored FTTH Solutions for Africa," AFRICON, pp.1-5, 9-12 Sept. 2013

[5] Paschotta, "Encyclopaedia of Laser Physics and Technology," Straus: Morlenbach, R 2008

[6] Rotich E K, Kourouma H, Chabata T V, Gamatham R R G, Leitch A W R and Gibbon T B, "Optimizing VCSEL Transmission for Longer Reach in Optical Access Networks," Proc. of the 16th annual Southern Africa Telecommunication Networks and Applications Conference (SATNAC) Spier, Stellenbosch, Western Cape, South Africa, pp. 27-30, 1-4th Sept ISBN: 978-0-620-57883-7, 2013

[7] Gibbon T B, Prince K, Neumeyr C, Rönneberg E, Ortsiefer M and Monroy I T, "10 Gb/s 1550 nm VCSEL transmission over 23.6 km single mode fibre with no dispersion compensation and no injection locking for WDM PONs," Proc. Opt. Fib. Comm. Conf. & Expos. and the National Fib. Opt. Eng. Conf., San Diego, CA, USA, 2010

[8] Karim A, Bjorlin S, Piprek J and Bowers J E, "Long-wavelength vertical-cavity lasers and amplifiers," IEEE *J. Sel. Top. Quant. Electron.*, vol. **6,** pp. 1244, 2000

[9] Bjerkan L, Rgyset A, Hafskjaer L and Myhre D, "Measurement of Laser Parameters for Simulation of High-speed Fibre Optic Systems," *J. of Lightw. Tech.*, vol. **14**, no. 5, pp. 839-850, 1996

[10] Koch T L, Linke R A, "Effect of nonlinear gain reduction on semiconductor laser wavelength chirping," *Appl. Phys. Lett.*, vol.**48**, no.10, pp. 613-615, 1986

Time of Crossing in Pulsed Eddy Current Signals

B Kibirige
Department of Physics and Engineering,
University of Zululand, Private bag x 1001, KwaDlangezwa, 3886

E-mail: kibirigeb@unizulu.ac.za

Abstract. Signals picked up from pulsed eddy current systems used to evaluate aluminium specimen exhibit interesting trends among which are points of intersection in the time-voltage curves. This study investigated the relationship between these picked up signal trends observed in the presence of a specimen to those observed in its absence to establish if they could be used to gauge small (μm) differences in thickness around a nominal thickness in thin aluminium sheets. Six aluminium specimen sheets, all at a temperature of 30 °C, and with thicknesses varied between 0.5 mm and 3 mm in steps of 0.5 mm, were used in the study. The locus of the signal picked up in the presence of a specimen always reached steady state conditions at a later time than that of the signal picked up in the absence of a specimen resulting in a time of crossing observed for each specimen thickness. This TOC was used to establish a relationship between time and the specimen thickness. The established relationship could be used to gauge differences in thickness around a nominal thickness in thin aluminium sheets.

1. Introduction

Pulsed eddy currents (PEC) have been considered for non-destructive evaluation of various materials as far back as the 1950s [1]. Since pulsed waves are made up of a continuum of sinusoidal waves of various frequencies, they contain information that can only be acquired by the use of many sinusoidal scans of conventional eddy currents (CEC) [2]. Circulating (eddy) currents are generated in bulk materials that are in close proximity with time varying magnetic fields as postulated in Maxwell's equations. These fields, when generated as a result of a controlled electric current flowing through a coil (drive-coil) of known dimensions, can generate predictable eddy currents within an aluminium bulk material of known thickness. The effect of the generate eddy currents picked up by another coil (pick-up coil) results in a current signal within the pick-up coil that depicts the bulk characteristics of the aluminium specimen. It is this locus of the pick-up coil signal that is observed and investigated in this study.

Crossing points have been observed in picked up pulse eddy current signals by various researchers. One search example is the lift off point of intersection (LOI). The term "lift-off" refers to the distance between the coil probe and the evaluated specimen. Figure 1 gives an example of LOI first observed by Waidelich and Haung [3]. In their investigations, they concluded that at a specific time, during a cycle of the pickup signal, there will be found a crossing point for any parameter in the system, such as lift-off, thickness of the specimen, conductivity of the specimen and permeability of the specimen. The LOI in figure 1 was obtained by carrying out experiments with all other parameters held constant while the distance between the probe and the specimen were varied in steps of 1 mm [4]. This point of intersection points to the fact that at the time of its happening, the effect of lift off is effectively eliminated. By time gating and collecting information on thickness, Giguere et el [5], effectively used LOI to characterise material loss. LOI has also been used, to map out corrosion in a specimen [6], to detect cracks under material fasteners [7] and to quantify defects and mark defect locations in multi-layered specimen [8].

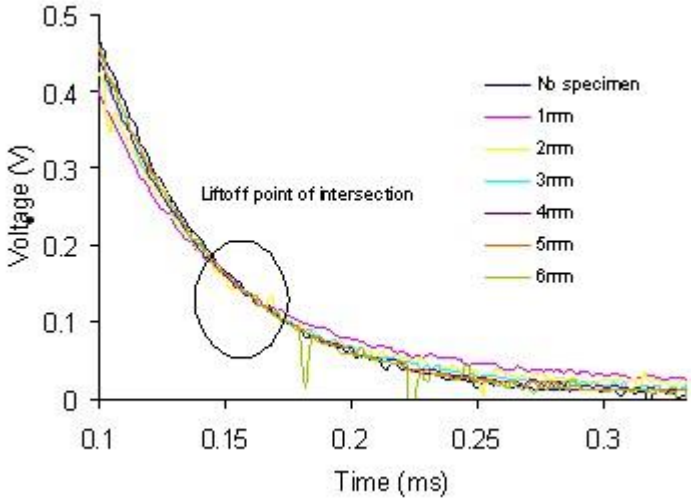


Figure 1. Lift-off point of intersection obtained with data collected with lift-off varied in steps of 1 mm.

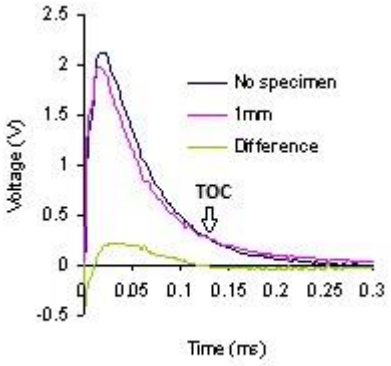


Figure 2. Pick-up coil signal.

Figure 2 shows the crossing point between the signal received in the absence of a specimen and that received in its presence now referred to as the TOC. The exponential nature of the discharge locus of the pickup coil signals in figure 2, suggests a difference in time constant in the decaying pick up coil voltage signals received in the absence of the specimen and that received in its presence. The PEC system can be modelled as an electric circuit model and this was developed in the Matlab and Simulink environment [4]. The system is an interaction between three subsystems: the drive coil, the pickup coil and the specimen. The system is represented by a set of energy storage parameters that were lumped together and reduced to a single resistance in series with an inductor for each of the sub systems. The lumped parameters R_s and L_s for a specimen with a thickness, y are expressed as [4]*:*

$$R_s = \frac{\rho_{s1} l_{ed}}{\delta(t)\left(\frac{1}{\exp\left(-y/\delta(t)\right)}-1\right) r_0 \ln\frac{r_2}{r_1}} \tag{1}$$

$$L_s = \frac{\mu_s \pi r_1^{\,2}}{2\left(\delta(t)\left(\exp\left(y/\delta(t)\right)-1\right)+r_o \ln\frac{r_2}{r_1}\right)} \tag{2}$$

Since the presence of the signal changes the locus of the pickup coil signal, a closer look at the expressions stipulated to represent the specimen electrical properties, shows that the resistance and inductance of the specimen change with a change in the thickness, y, of the specimen but at the same time they are affected by the skin depth, δ (t), which is a function of time (t) [4]. This means that the time constant, τ of the decaying locus changes with time. Also the loci in figure 2 suggest the signal received in the absence of the specimen has a shorter time constant compared to that received in the presence of the 1 mm-thick specimen. The two signals will therefore always have a crossing point at a time referred to as the time of crossing (TOC). In this study, the TOC for different specimen thicknesses were investigated experimentally and by simulation, to establish whether small differences, to the order of microns, in thicknesses of a specimen can be determined from a Time-Thickness characteristic.

2. Experimental procedure

Figure 3 shows the experimental set up used in the study. A Sampo 1617 function generator, with an internal resistance of 50 Ω, was connected through a signal conditioning circuit to the eddy current reflection probe. The probe was placed on the tested specimen and the voltage signal across the pickup circuit was collected using a PC30FA, 11 channel data collection card, with an input range of -5 to +5 volts and an input impedance of 500 MΩ and data were stored on a desk top computer.

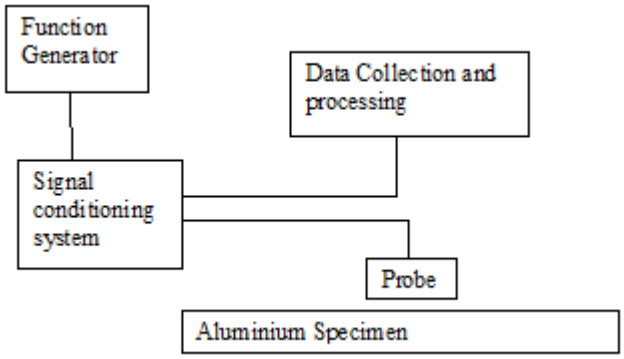


Figure 3. Block diagram of experimental set up.

In order to analyse the data presented for the time of crossing at 30 °C, pick-up coil signals were collected for aluminium specimens of different thicknesses that varied from 0.5 to 3 mm in steps of 0.5 mm. Figure 2 is an example of the pick-up coil signal collected by the PC30FA data collection card. An example of the difference signal obtained by subtracting the signal received in the presence of a specimen from that received in the absence of the specimen (the reference signal) is included in the diagram. Difference signals were obtained for each of the specimen thicknesses. Data for each specimen was processed to obtain four different points of crossing for each specimen thickness, under similar conditions were recorded in a table.

Since variations in specimen thickness to the order of µm were not easily emulated in the experimental setup, a simulation study of the pulsed eddy current system modeled in Matlab and Simulink environment [4] was used. The results of the simulation study carried out at a specimen nominal thickness of 2 mm were recorded in table.

3. Results and Discussion

3.1. Results

Figures 4 and 5 show graphical representations of the difference signals obtained from the experiments

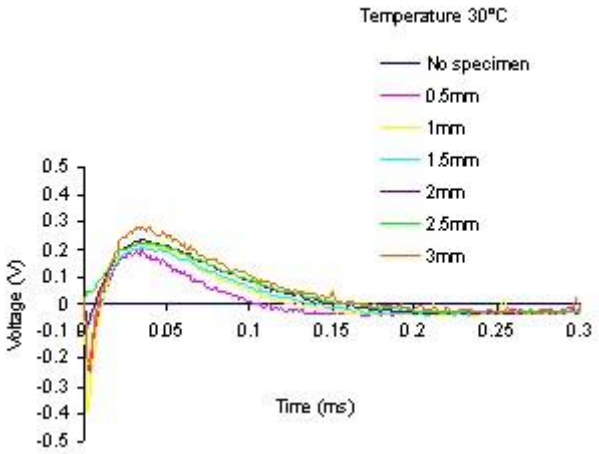


Figure 4. Difference signals.

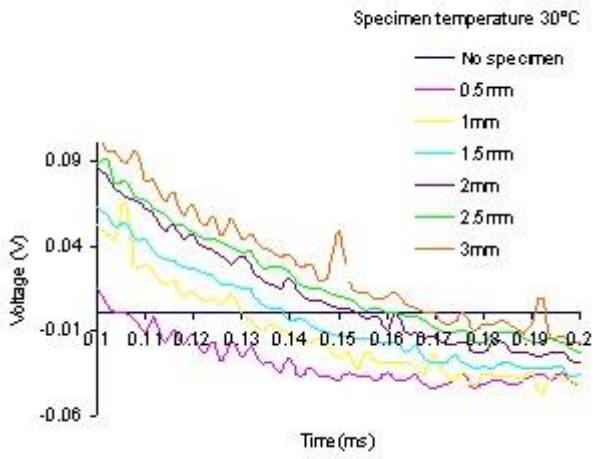


Figure 5. Zero Crossing is Time of crossing (TOC).

and Table 1 the averages calculated from the experimental data collected for the TOC at the specified specimen thicknesses.

Table 1. Averaging Time of crossing for the specimen thickness considered in the experiment.

Thickness (mm)	t1 (ms)	t2 (ms)	t3 (ms)	t4 (ms)	**tave (ms)**
0.5	0.098	0.098	0.102	0.102	0.1
1.0	0.118	0.118	0.118	0.122	0.119
1.5	0.132	0.136	0.134	0.138	0.135
2.0	0.136	0.142	0.14	0.14	0.14
2.5	0.152	0.156	0.158	0.162	0.157
3.0	0.168	0.17	0.172	0.17	0.17

Figure 6 shows the plot corresponding to the relationship between TOC and the thickness of the specimen (Time-Thickness characteristic). The 4th order polynomial included in the diagram was found to be the best trend that fitted the data.

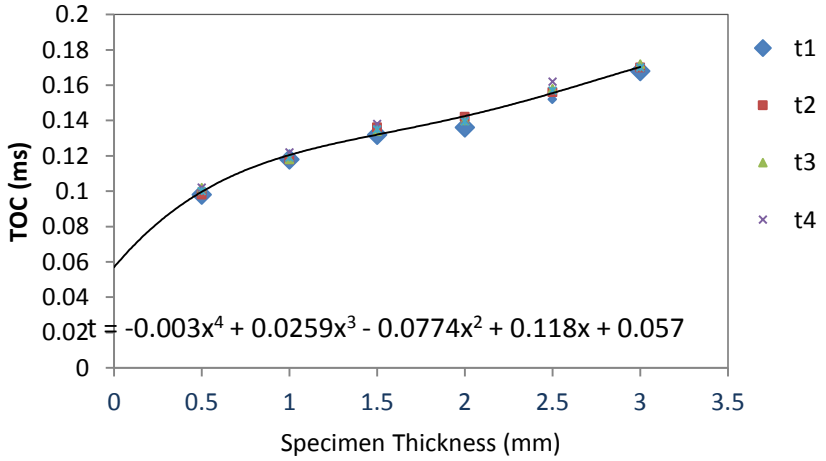


Figure 5. Experimental results: Thickness versus Time of Crossing (TOC).

Table 2 is a record of the simulated differences in TOC and their corresponding differences in specimen thickness at a nominal thickness of 2 mm.

Table 2. Simulated and calculated results for difference in Time of crossing, Δt, versus difference specimen thickness, Δx, at nominal thickness of 2 mm.

TOC Difference Δx (ms)	Thickness Difference Δx (μm)	
	Simulated	Calculated
0.00010	-4	- 4.3
0.00007	4	3.02

3.2. Discussion

Equation (1) is a possible representation of the trend lines for the graphical results shown in figure 6. It is a 4th order polynomial forecasted backward to include thicknesses below 0.5 mm. It shows that despite the fact that the time of crossing increases with increasing specimen thickness, it does not do so linearly. This could be attributed to the rather involved interaction between the electromagnetic effects of the three subsystems that make up the PEC system as discussed in the introduction.

3.2.1. Time constants and amplitudes.

Dividing the period of interest of the received Pickup Signal into sections, exponential trends were obtained for the varied specimen thicknesses. The start time of each section was set at 0 ms and results of the decaying voltage relationship for each signal $Ve^{-t/\tau}$ obtained for the various specimen thicknesses are shown in table 3.

Table 3. Time gated pickup coil signal ($Ve^{-t/\tau}$).

Thickness (mm)	Time gate 0 to 0.12 ms	Time gate 0.12 to 0.24 ms
0.0	$2.4756e^{-19535t}$	$0.2463e^{-20052t}$
0.5	$2.1231e^{-18096t}$	$0.2599e^{-14855t}$
1.0	$2.081e^{-18801t}$	$0.2358e^{-12983t}$
1.5	$2.1007e^{-19532t}$	$0.2113e^{-12134t}$
2.0	$2.1355e^{-19802t}$	$0.2052e^{-11963t}$

It is noted that as the signal decays toward a steady state value of zero volts, the values of the time constants in the two time gated sections considered are not the same. As predicted from the model, the resistance and inductance are function of the skin depth, δ (t) which varies with time. This could be the reason for the time changing time constant τ.

3.2.2. Time constants and amplitudes.

Equation (3) the relationship between the TOC and specimen thickness is not linear. However, it is possible to linearize the trend at a point by considering small changes about the point.

$$t = -0.003x^4 + 0.0259x^3 - 0.0774x^2 + 0.118x + 0.057 \quad (3)$$

At any point (x_0, t_0) on the curve shown in figure 6, there exist another point ($\boldsymbol{x_0 + \Delta x,\ t_0 + \Delta t}$), due to a small change in thickness, Δx. The resulting change in the time of crossing, Δt, is determined by considering the gradient at point (x_0, t_0) which is expressed as:

$$\Delta t = \left(\frac{dt(x)}{dx}\right)_{x_0} \Delta x \qquad (4)$$

Differentiating equation (1) and substituting for $\boldsymbol{x}$ at the nominal thickness x_0 give:

$$\Delta t = (-0.012x_0^3 + 0.0777x_0^2 - 0.1548x_0 + 0.118)\Delta x \qquad (5)$$

The small change in thickness, Δx can then be determined from a linear expression given in equation (6):

$$\Delta x = \Delta t/(-0.012x_0^3 + 0.0777x_0^2 - 0.1548x_0 + 0.118) \qquad (6)$$

Changes in TOC about a nominal thickness of 2 mm were obtained for 4 µm thickness differences in the specimen. The simulated changes were compared with those calculated using equation (6) and results are shown in table 2. Although these values are quite close, the differences between them could suggest that the 4 µm change in thickness was out of the range of the linearization used for the estimate. The results, however, could be used confidently in a colour coded scale as an alert for thickness differences within the 4 µm range about the 2 mm nominal thickness.

4. Conclusion

The PEC study carried out here suggests that a relationship between the TOC and varying specimen thickness, could be used to indicate differences in thickness of aluminium specimen of up to the order of microns. A closer look at the PEC system pickup coil signal trend showed that the time constant τ, of a signal received at a given specimen thickness, changed with increasing time within a period. This could be a possible reason for the nonlinearity of the TOC Time-Thickness characteristic. Although the relationship between the TOC and the specimen thickness is not linear, small changes about a nominal thickness could be estimated by using the differential solution of the actual relationship. The results could be used confidently in a colour coded scale as an alert for thickness differences within the 4 µm range about the 2 mm nominal thickness.

5. Acknowledgements

University of Zululand generously supported this research by providing time and financial assistance.

6. References

[1] Waidelich DL 1970 *Pulsed Eddy currents* (Research Techniques in Non-destructive Testing, Academic press) p 383

[2] CAN/CGSB-48 1986 *Advanced Manual for Eddy Current Test Methods* (Canadian General Standards Board, Canada) 14-M86

[3] Waidelich DL and Huang SC 1972 *The use of crossing points in pulsed eddy currents* (Materials evaluation) p 20

[4] B Kibirige 2012 *An investigation of pulsed eddy currents for the development of a profile gauging system suitable for aluminium hot rolling mills,* (mobile.wiredspace.wits.ac.za)

[5] Giguere S, Lepine BA and Dubios JM S 2002 *Pulsed Eddy Current Technology: Characterizing material Loss with Gap and lift-off Variation*, (Research in Non-destructive Evaluation 13) p 119

[6] Giguere JRS and Dubios JMS 2002 *Pulsed Eddy Current: Finding corrosion independently of transducer lift-off* (Review of progress in QNDE 19) p 449

[7] Giguere, JRS, Lepine BA and Dubios JMS 2003 *Detection ofcracks beneath rivets via pulsed eddy current technique (*Review of Progress in QNDE 21) p 1968.

[8] Lepine BA., Giguere JRS, Forsyth DS, Chahbaz A and Dubios JMS 2003 *Interpretetion of pulsed eddy current signals for locating and quantifying metal loss in skin lap splices* (Review of Progress in QNDE 21) p 415

Electronic tracking system for quantum cryptography and radio telecommunication

M Mariola[1], Y Ismail[1], A Mirza[1], F Petruccione[1,2]

[1]University of KwaZulu-Natal, Westville Campus, Durban, South Africa
[2]National Institute for Theoretical Physics (KZN), Durban, South Africa

E-mail: Mariolam@ukzn.ac.za

Abstract. In optical and radio communications the transmitter and receiver should be aligned in order to transmit and receive the data. The perfect alignment in optical communication is vital to ensure a low bit error rate and in radio communications, a correct alignment between the antennas maximizes the power received. Quantum cryptography in free-space permits to share an encryption key between a transmitter and receiver, where the bits of the key are sent in the form of polaraised single photons. The electronic system presented below is able to align the transmitter and receiver as well as the polarization bases of the transmitter and receiver using a polarised laser beacon transmitted form the receiver. The electronic system measures the intensity of the laser beacon in order to align the polarization basis and through the knowledge of three geographical coordinates it is possible to adjust the transmitter in the direction of the receiver.

1. Introduction

Quantum cryptography is topical of late due its influences in optical communication. Quantum cryptography is a process of ensuring the security of information between numerous entities. The generic quantum crypto scheme consists of a transmitter and receiver, which are conventionally known as Alice and Bob respectively, as well as some medium to transfer information known as a quantum channel. Information is transferred in the form of quantum carriers or more specifically single photons. In the holistic scheme Alice transmits a string of photons to Bob through a medium that could either be free-space or fibre. This process is followed by a measurement and sifting process which results in the distribution of a raw key. Further privacy amplification performed on the raw key results in the procurement of a final key which can be used to encrypt and decrypt information. The quantum carriers used to generate the key can be encoded through polarisation or phase. For free-space communication it is preferable to encode using the polarisation state of light due to the negligence of the change of state under the influence of turbulence. The advantage quantum cryptography has over classical systems is that it is governed by the laws of quantum mechanics specifically by the Heisenberg Uncertainty principle and the no-cloning theorem [1], hence any external interferences leaves traces in the system which can be traced.

The first quantum cryptography protocol proposed by Bennet and Brassard, known as BB84 [2], utilised single photons encoded in the polarisation state. This protocol makes use of two non-orthogonal polarisation bases, namely the rectilinear and diagonal bases, to prepare and measure the transmitted photons. In the case where the quantum channel is a free-space

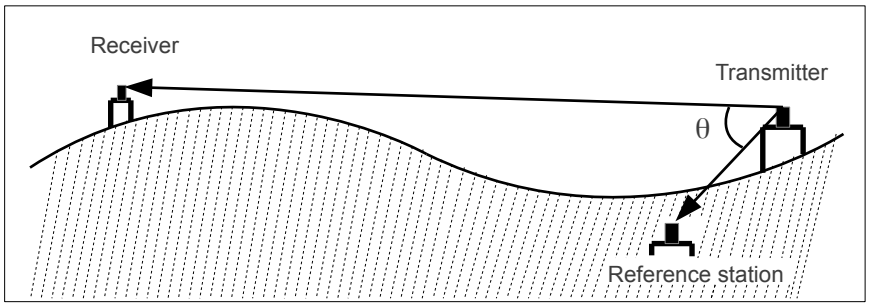


Figure 1. Scheme of the tracking. The coordinates of the receiver and the reference station, are transmitted to the transmitter. The reference station is located near the transmitter. The laser beacon is directed towards the reference station by a manual or automatic control and by processing the coordinates is able to know the direction of rotation and the angle.

medium with a line of sight, the alignment of the polarisation bases of the transmitter and receiver becomes crucial for the successful implementation of the key distribution process. If the polarisation bases of the transmitter and receiver are not aligned, the Quantum Bit Error Rate (QBER) increases. For the procurement of a final key, the QBER should not exceed a threshold percentage dependent on the protocol being implemented. Experimentally if the misalignment angle is lower than 20° the QBER that is obtained is acceptable to identify the presence of interference within the quantum channel [3].

The Electronic system must be able to align the bases with an error lower than 20° since for optical communication it is necessary that the system is able to align the transmitter in the direction of the receiver and vice-versa. In the proposed system a polarised laser beacon directed from the receiver to the transmitter is considered. The laser beacon carries the information about the polarisation bases direction of the receiver and can be exploited to collimate the transmitter and receiver with a minimum error. This system can be used as a coarse alignment for optical communication and fine alignment for radio communication.

2. Tracking

Classical alignment systems use the geographical coordinates of the transmitter and the receiver and include the Hall effect and the direction of the Earth magnetic field to coarsely align two points namely the transmitter and receiver. Unfortunately the terrestrial magnetic field is not homogeneous due to the geology of the Earth, as for example the Reunion Island magnetic field results in a deviation with respect to the north pole greater than 25° [4]. Other systems include exploiting the stars as a reference frame, however in this case it is necessary to know the map of the sky for each geographical position for a particular time.

In our system three geographical positions were used. The first geographical position is indicative of the position of the transmitter. The second position, called the reference station, is located near the transmitter and is used as a reference point. The third and last position is referred to the receiver. The coordinates of the reference station and receiver, are transmitted to the transmitter side. Once the electronic device receives the coordinates, using vectorial calculus, it is possible to determine the angle θ between the line of sight of the transmitter-reference station and that of the transmitter-receiver as shown in Figure 1. A microcontroller is used to calculate the angle θ using single precision from which the errors, due to the approximation is less than 0.5° as shown in Table 1. The microcontroller has a limited number of digital pins to control other external devices such as the controls for the motors of the optics. For the aforementioned reason, a stack of shift registers is used for serial to parallel conversion hence using a limited

Table 1. Test 1: Coordinates of the transmitter, receiver and reference station on the south hemisphere; **Test 2**: Coordinates of the transmitter, receiver and reference station on the north hemisphere; λ, μ is the latitude and longitude of the transmitter TX, reference station Rf and the receiver RX; $\theta_{\mathbf{G}}$ angle between the direction transmitter-reference station and transmitter-receiver measured by Google maps; $\theta_{\mathbf{m}}$:angle between the direction transmitter-reference station and transmitter-receiver calculated by the micro-controller; ϵ_θ is the difference between the calculate and the measured angle; **d**: distance between the transmitter and receiver in km.

	Test 1	Test 2
λ_{TX}	-29.060006944	41.43288333
μ_{TX}	27.24506667	12.83350833
λ_{Rf}	-29.05515000	41.43427500
μ_{Rf}	27.23995833	12.83227778
λ_{RX}	-29.00026667	41.48340278
μ_{RX}	27.30960833	13.32911667
θ_G	85.6100	115.6700
θ_m	85.8719	115.7889
ϵ_θ	0.2619	0.1189
$d(km)$	9.1357	41.7860

number of digital pins of the microcontroller. The system is now able to control the motors required to turn the mount of the laser and the optics in order to measure the single photons.

3. Polarisation basis alignment

The laser beacon is polarised with a direction of 45° which respect to the vertical polarisation basis used to receive the single photons. The receiver directs the laser beacon to the transmitter in order to align its vertical polarisation basis with the vertical polarization base used to transmit the single photons. At the transmitter the direction of polarisation of the laser beacon is measured. The vertical polarisation basis of the transmitter should also be tilted by 45° with respect to the direction of polarisation of the laser beacon. The bases are aligned when the transmitter measures 45° with respect to the polarisation of the laser beacon as shown in Figure 2. Once the tilt angle is acquired, the transmitter turns the polarisation bases. The core optics

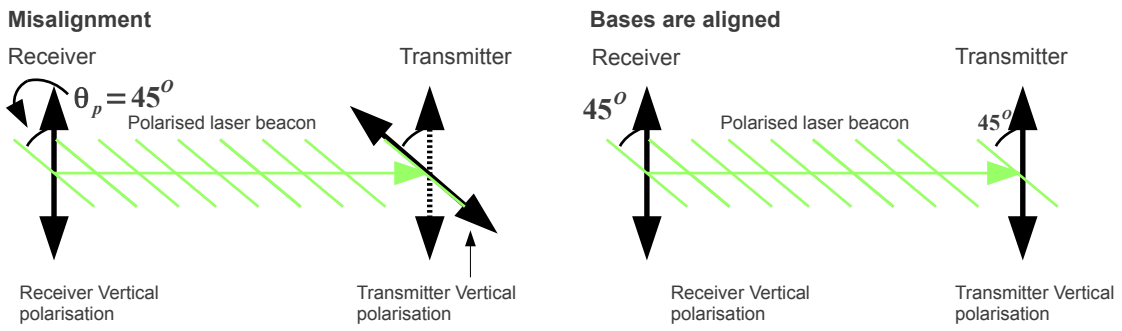


Figure 2. The receiver transmits the polarised laser beacon to the transmitter. The transmitter turns its bases when the polarisation of the laser beacon is 45° which respect to vertical polarisation bases.

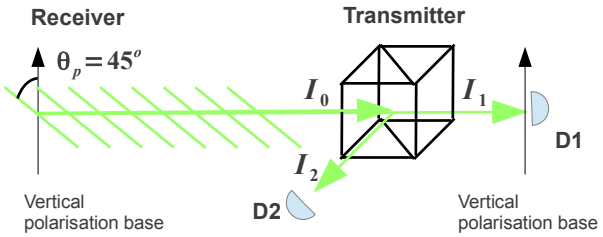


Figure 3. The PBS splits the laser beacon into two paths with the intensity I_1 and I_2 respectively. When the system measures $I_1 = I_2$ the polarisation bases are aligned.

used to measure the tilt angle of the polarisation alignment is a Polarising Beam Splitter (PBS) [5]. As shown in Figure 3, the intensity received by the transmitter is I_o. The PBS splits the intensity I_0 into two paths according to the following equation:

$$I_1 = I_0 \cos^2(\theta_p), \tag{1}$$

$$I_2 = I_0 \sin^2(\theta_p). \tag{2}$$

The tilt angle θ_p can be calculated using the equation:

$$\theta_p = \arctan \sqrt{\frac{I_2}{I_1}}. \tag{3}$$

Equations (1), (2) and (3) follow the ideal behavior of the PBS. Since the PBS is optimized for an angle $\theta_p = 45°$, it is necessary to characterize the PBS to find an error function $\Delta\theta_p(I_1, I_2)$ calculated form the normalized values of I_1 and I_2. The tilt angle θ_p is calculated by Equation (4).

$$\theta_p = \arctan \sqrt{\frac{I_2}{I_1}} + \Delta\theta_p(I_1, I_2). \tag{4}$$

The signals I_1 and I_2 are measured using an electronic device that amplifies the signal from the photodiodes D1 and D2 that subsequently are used in a microcontroller unit that calculates the tilt angle using the Equation (4). To simulate the electronic circuits, the intensities I_1 and I_2 were measured for a total power $I_{tot} = I_1 + I_2 = 1.95mW$ and for an angle $0° \leq \theta_p \leq 90°$. The results are shown in Figure 4.

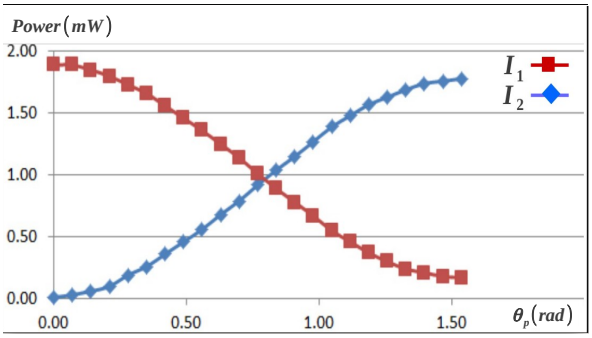


Figure 4. Measurement collected for the electronic simulation for a tilt angle $0° \leq \theta_p \leq 90°$ for a total power of $1.95mW$.

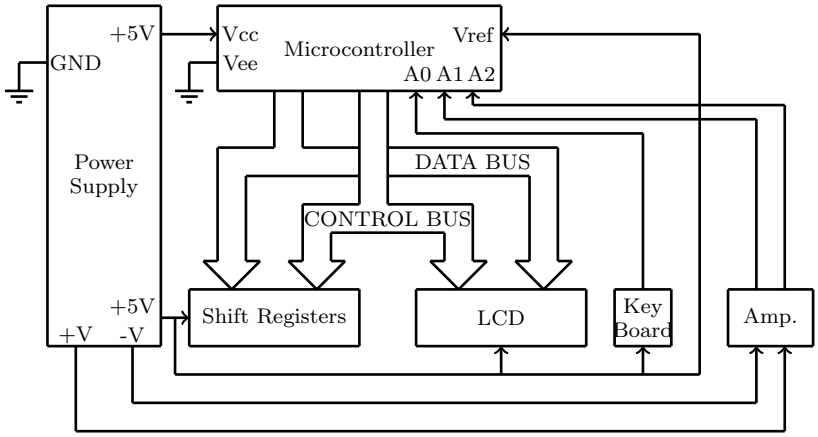
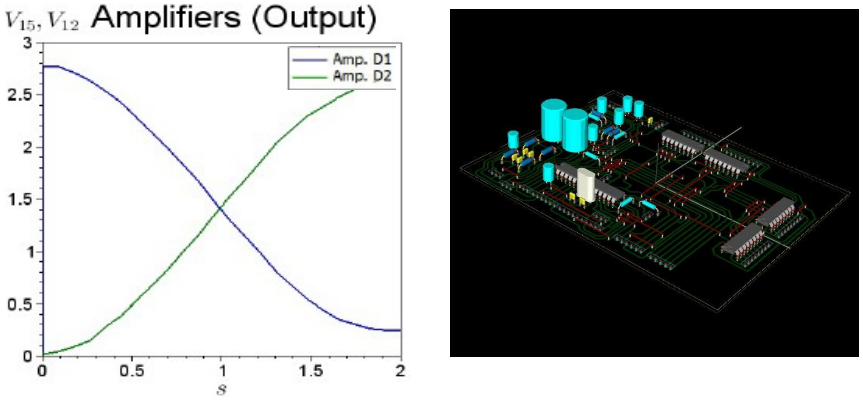


Figure 5. Scheme block of the portable tracking system. The microcontroller has an independent power source of $5V$. The CONTROL BUS and the DATA BUS are unidirectional and are used for the Shift register and for the LCD.

4. The electronics

The electronic devices as shown in Figure 5, consist of a Power supply with two 5 V outputs to supply the microcontroller unit, the shift register stack to control the mechanics of alignment and the LCD display. Two outputs (+V and -V) are used to supply the amplifiers. The signals I_1 and I_2 are amplified through the amplifier block (Amp.) and the amplified signals are sent to the analog to digital converter of the microcontroller indicated with the letter A. These signals are used to calculate the tilt angle of the polarisation bases. The keyboard is used to input the geographical coordinates of the transmitter, receiver and the reference station to the microcontroller Atmel Atmega328-P [6] in order to align the direction of the transmitter and the receiver. To simulate the dynamics of the system, the data are collected for a total power $I_{tot} = 1.95mW$ and for $0° \leq \theta_p \leq 90°$, as explained in section 3. For the simulation the software Ngspice [7] was used and the rotation velocity of polarisation was $45°/s$. The simulation in Figure 6(a) shows that the amplification follows the of the intensity collected from the photodiode D1 and D2. The electronic system proposed in this manuscript was also built



(a) Amplified signals at 1.95mW

(b) 3D rendering of the PCB

Figure 6. (a) Amplifier response for the signal collected from the photodiode D1 and D2. (b) PCB of the elecronic tracking system without the amplifiers.

and tested for a line of sight between the transmitter and receiver. The 3D rendering of the Printed Circuit Board (PCB) is shown in Figure 6(b).

5. Conclusion

In this paper, we presented a tracking system for quantum cryptography and radio telecommunication. Using this system, the transmitter is able to be directed towards the receiver and vice versa. The error of alignment is less than one degree and this fact permits to say that the system can be considered as a fine alignment system for radio telecommunication. In optical communication more accuracy is required and for that reason the system proposed should be used in cooperation with another fine alignment subsystem. The alignment of the polarisation bases used for quantum cryptography does not require future improvement.

Acknowledgments

This work is based on research supported by South African Research Chair Initiative of the Department of science and Technology and National Research Foundation.

References

[1] Weigert S 2009 *Compendium of Quantum Physics* ed Greenberger D, Hentschel K and Weinert F (Springer Berlin Heidelberg) pp 404–405 ISBN 978-3-540-70622-9 URL `http://dx.doi.org/10.1007/978-3-540-70626-7_124`

[2] Bennett C H, Brassard G *et al.* 1984 *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* vol 175 (New York)

[3] D'Ambrosio V, Nagali E, Walborn S P, Aolita L, Slussarenko S, Marrucci L and Sciarrino F 2012 *Nature communications* **3** 961

[4] Mariola M 2015 *Free-Space Communication in Quantum Key Distribution* Ph.D. thesis University of KwaZulu-Natal, School of Chemestry and physics

[5] M Mariola, A Mirza and F Petruccione, Polarization alignment system for quantum key distribution, in Proceedings of SAIP2014, the 59th Annual Conference of the South African Institute of Physics, edited by Chris Engelbrecht and Steven Karataglidis (University of Johannesburg, 2015), pp. 528 - 533. ISBN: 978-0-620-65391-6. Available online at http://events.saip.org.za

[6] Atmel 8-bit avr microcontroller with 4/8/16/32k bytes in-system programmable flash URL `www.atmel.com/images/doc8161.pdf`

[7] Ngspice URL `http://ngspice.sourceforge.net/download.html`

Open-Source electronic board designed in South Africa for Africa

M Mariola[1], F Petruccione [1,2]

[1]University of KwaZulu-Natal, Westville Campus, Durban, South Africa,
[2]National Institute for Theoretical Physics, South Africa

E-mail: Mariolam@ukzn.ac.za

Abstract. Several experiments require acquisition devices to read and elaborate the signals from the external environment. These devices can be interfaced with a computer through the serial communication. Because the continuous evolution of electronics changes the communication standard, e.g. from the standard RS-232 to USB, it is necessary to use an external adapter to connect laboratory equipment to a computer. Laboratory equipment is often expensive and not affordable for schools or universities in developing countries. Through the use of Open Source it is possible to design experimental equipment, without license fees and build proper electronic devices using inexpensive components. In this article we present an electronic board for prototyping using the Arduino features, called AFRICHINO. This board was developed based on our research experience, and represents a synthesis of what is necessary to have a complete experimental board.

1. Introduction

In the scientific world, physical quantities must be measured and collected to describe and understand natural phenomena. Most of the physical parameters e.g. temperature, light intensities and many others can be transduced as electric signals through a sensor. To measure and store this signal to be subsequently elaborated several proprietary devices can be used. Some of this laboratory equipment is expensive for schools located in development countries or for students that want to repeat some experiments to improve their personal knowledge. Some specific measurement devices can be used in conjunction with a personal computer through the serial protocol RS-232. In the latter case the last generation of personal computer use the USB port, and a converter USB-RS232 is required.

Following the Open-Source philosophy, in this manuscript we propose a board, called Africhino built based on of another famous Open-Source board called Arduino UNO 3R ® [1]. The core of the board is a micro-controller ATMEGA-328P. The micro-controller has six pins linked with a successive approximation Analog to Digital Converter (ADC) and fourteen digital I/O ports. The analog pins, can be used as a digital input output pin. The clock frequency is 16 MHz and due to the nature of the ADC, the time of conversion depends on the frequency of the micro-controller timing. In Africhino the micro-controller drive a TFT monitor with an SD card on board, a USB to RS232 converter and an external Static Random Access Memory (SRAM) of 32 kb. The board can be programmed using the Integrated Development Environment (IDE) of Arduino and, can be eventually used as an Arduino UNO 3R. This board can be considered as a Swiss army knife for electronic prototyping and electric measures.

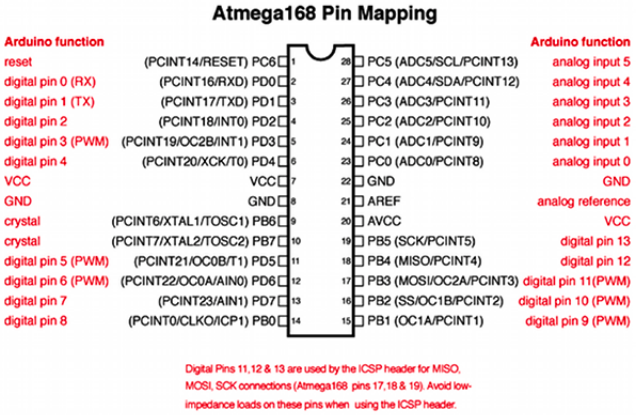


Figure 1. The Pinout of Arduino and the micro-controller Atmega-328 [2].

2. The Hardware

To use the Arduino ® IDE, the connections of the micro-controller with the electronics components follow the pinout of Arduino. The pinout of Arduino with respect to the micro-controller, are described in Figure 1. The different devices of the board can share a particular pin of the micro-controller. In order to avoid the conflicts of the different peripherals throught the switch-dip showed in Figure 2, it is possible to exclude one or more devices, for example the external SRAM when the screen TFT is used. If Africhino must be used as an Arduino ®, all the switches must be off. Arduino requires 5V but several devices, using in conjunction with, needs 3.3 volt and for this reason an external voltage regulator LM2937ET-3.32 is used for this task. Africhino is designed to be compatible with Arduino Uno board such as several open-source boards.

2.1. USB to RS232

Several laboratory equipment has the serial port RS232 to communicate with other devices such as a computer. Unfortunately, recent computers adopt the USB port however Africhino has a USB to RS-232 converter. The serial cable of the instruments can be connected with Africhino, and the data can be stored on the SD card, displayed on the TFT screen or transfer directly to the computer. The circuits use an IC circuit MAX233CPP connected with the digital pin 2 and 3 [3]. The data, from the RS232, can be used from the micro-controller, or through the digital pin 0 and 1, connected with the serial to USB converter in order to transfer data to the

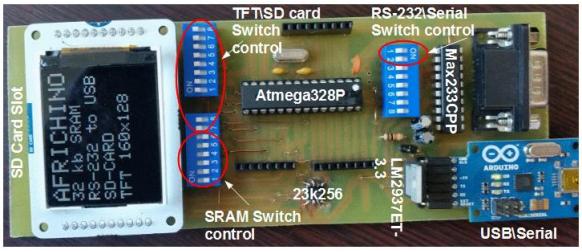


Figure 2. Overview of the entire board. The USB-Serial converter is an Arduino device. The USB to Serial converter is indicated on the board as USB\Serial. The red circles indicates the switch-dip to activate or deactivate the different parts of the Board.

computer. For this function the open-source library NewSoftSerial must be used [4].

2.2. External memory SRAM
The micro-controller ATMEGA328P has 32 kb of flash memory to store the programs and only 2kb of SRAM to store the value of the variables. The program loaded in the flash memory is loaded in the static RAM during the execution. If the memory needed exceeds the maximum amount, the micro-controller does not work properly or some values contained in the variable changes. Arduino contains the proper instructions to write the variable in flash memory, however datasheet guarantees the rewriting of the flash memory, no more than 10000 times. To use the external memory, the library SpiRam is necessary [5]. The RAM works at 3.3 V and a generic bi-directional level shift voltage is necessary to communicate with the RAM at the aforementioned voltage of 3.3 V.

2.3. SD card
SD card mounted on the TFT monitor can be used to store data in a text file from the serial port or from the external sensors. The library of the SD card is native in Arduino IDE.

2.4. TFT monitor
The analog signal converted by the internal ADC of the micro controller can be displayed on the TFT in a graphical or numerical form. The TFT library can be downloaded from the official Arduino website or numerous forum [6].

2.5. The electronic scheme
The micro-controller requires a USB to serial converter in order to be programmed and communicate with a computer. The converter is indicated on the scheme as connector P2 as shown in Figure 3. The transmission pin (TX) is connected with the pin number 0 (RX). The receiver pin of the converter (RX) is connected with the pin 1 (TX). The 5 V for the board is taken from the converter, indicated with 5V in the scheme. The integrated circuit LM2937ET-3.3 provides the 3.3V for other devices. The corresponding pins of Arduino are shown by the red characters depicted on Figure 1.

3. How to use the board
Africhino is a versatile device to follow several experiments since it is easy to apply the analog and digital signal from the transducers or other digital devices. In this section it is shown how Africhino can print the data on the TFT screen and the possibility to generate random signals.

3.1. Generation of random curve
Suppose that it is required to address a periodical signal $f(t)$ to a device with period T defined as:

$$f(t) = 2.5 + 2.5 \sin t, \quad \text{for} \quad 0 \leq t \leq \frac{T}{2}, \tag{1}$$

and

$$f(t) = 2.5 - t, \quad \text{for} \quad \frac{T}{2} \leq t \leq T. \tag{2}$$

The curve can be built using a software such as Excel whereby each sample of the curve it is necessary to calculate the corresponding binary code. In this case the decimal number are utilised. For the 5V used, in conjunction with the 8 bit digital to analog converter, the minimum value of the curve is 0 which correspond to the value 0, and the maximum value in output is 4.98

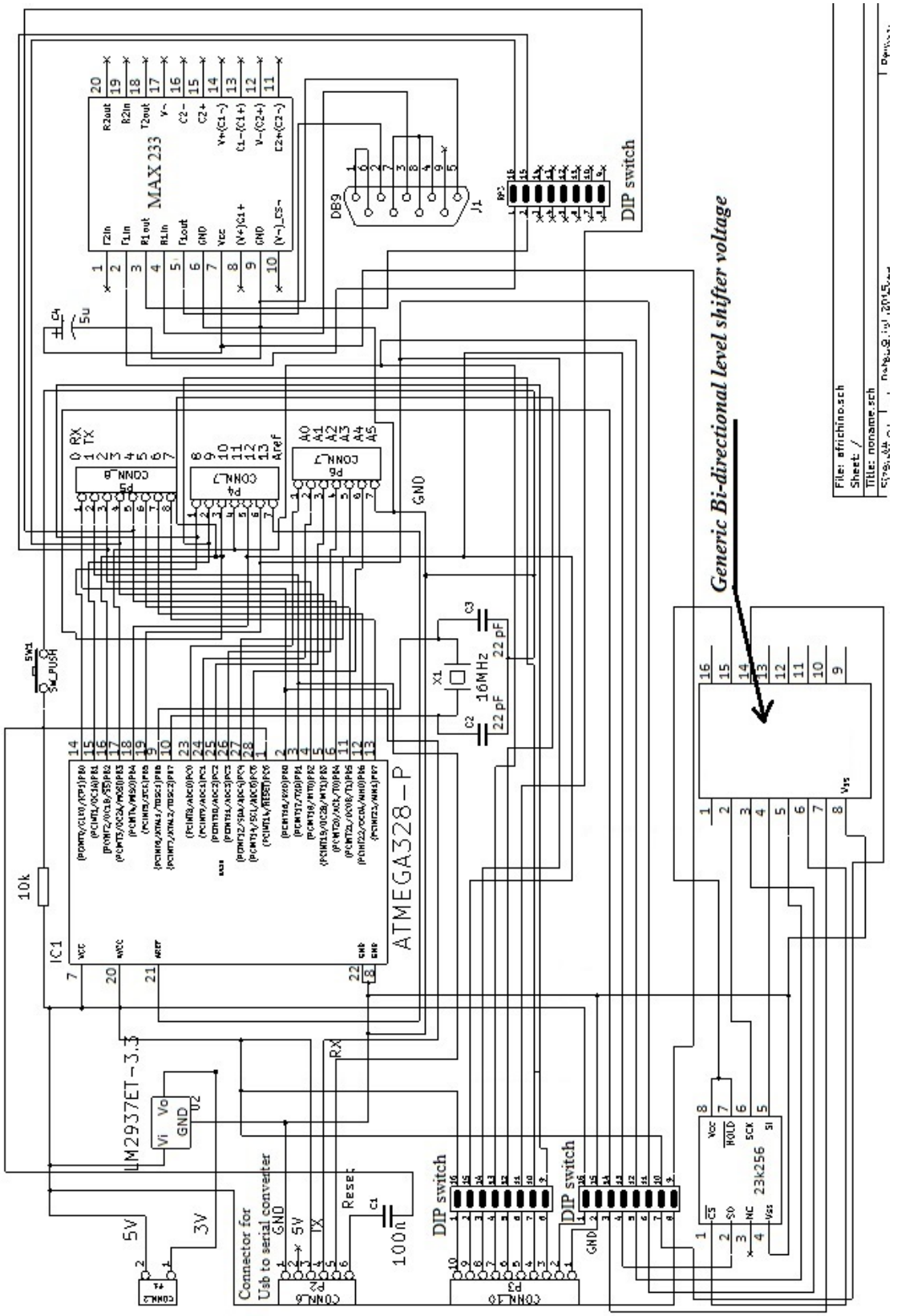


Figure 3. Electronic scheme of Africhino.

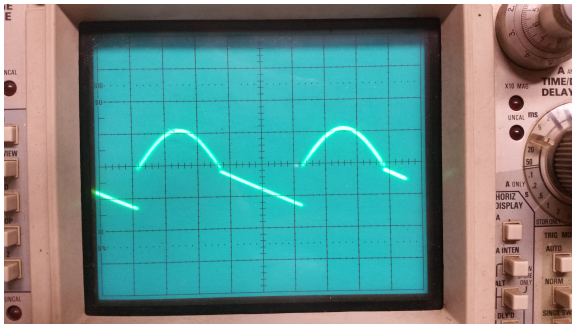


Figure 4. The curve generated by Africhino following the concept of Direct Digital Synthesis (DDS) [7].

V which in binary correspond to the decimal value 255. The desired output value in volt can be converted by the integer approximation of the function $N(V_{out})$ calculated by the Equation 3:

$$N = V_{out}\frac{255}{5}, \tag{3}$$

where V_{out} is the desire output voltage. Each sample is sent in output using a particular delay in order to change the frequency of the reconstructed signal. For this experiment, according to Figure 1 the pin used ranges from 0 to 7. The curve mentioned in this section was measured by an oscilloscope as shown in Figure 4.

3.2. Visualize the data on the TFT screen

Using the analog input of the micro-controller it is possible to visualize the evolution of a signal in time, as shown in Figure 5 by the numerically generated sinusoidal curve linked with africhino. Using the native TFT library of the Arduino IDE and by some modification of the algorithm it is possible to visualize the signal.

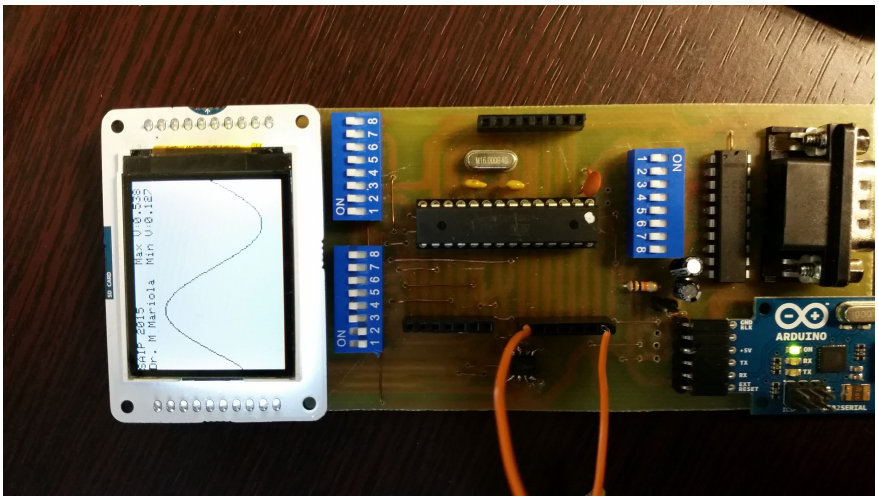


Figure 5. The figure shows the evolution of an analog signal generated by the DDS philosophy by an Arduino board coupled with a digital analog converter.

 ISBN: 978-0-620-70714-5

4. conclusion

With this technical work we propose a multipurpose electronic board, useful for research and teaching. This low cost board can be used as a multipurpose instrument in schools located in developing areas, where a lot of instruments are not affordable. The circuit is easy to build, and each component is visible for the student. The students have the possibility to improve their skill by building the board, or develop other feature that are not found on this board. Raspberry Pi ® and Arduino ®, was developed to teach programming and electronics, similarly Africhino was developed for the same purpose.

5. 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. Special thank to Dr. Y. Ismail for the editing help and encouragment.

References

[1] www.arduino.cc
[2] www.arduino.cc/en/Hacking/PinMapping168
[3] http://www.chim.unipr.it/crocoite/homemadelab/ardurs232.html
[4] http://arduiniana.org/libraries/newsoftserial/
[5] http://playground.arduino.cc/Main/SpiRAM
[6] http://forum.arduino.cc/index.php?PHPSESSID=vct72vgindhgkebkva99cfba41action=dlattach;topic=192072.0;attach=56590
[7] Vankka J 2000 *Direct Digital Synthesis:Theory, Design and Application* Ph.D. thesis Helsinki University of Thechnology, Department of Electrical and Communications Engineering

Analysis of homogeneity in thin film photovoltaic modules using large area light beam induced current (LA-LBIC) measurements

M Okullo[*1], F J Vorster[1], E E van Dyk[1] and J L Crozier[1]

[1]Department of Physics, Nelson Mandela Metropolitan University, Port Elizabeth, 6031, South Africa

*Correspondence author's email: s214229785@nmmu.ac.za

Abstract. In this study a large-area light beam induced current (LA-LBIC) measurement system is used to analyse homogeneity in photo-current in thin film PV modules. A laser light source mounted onto a motorized x-y scanning stage is used to raster scan while illuminating a PV module point-by-point. The measured PV module current output is mapped and presented as an LBIC image. As a consequence of the large area deposition process and the laser scribing to form individual cells in thin film PV modules, the module may be prone to inhomogeneities that can cause a reduction in performance. The accurate interpretation of the LBIC data depends on knowledge of the operating bias voltage of each cell. In this work the module is scanned at various constant bias voltages. The experimental setup used is presented and LA-LBIC results discussed. Electroluminescence imaging of the scanned modules is also used to verify the results of the LA-LBIC mapping. In addition, the limitations of the LA-LBIC technique are highlighted and possible solution to obtain meaningful results from the LA-LBIC technique is presented.

Key words: LA-LBIC, Thin film, PV Modules, Homogeneity

1. Introduction

Film thickness of the semiconductor absorbing layer in thin film Photovoltaic (PV) module cells is one of the key parameters that influence the module electrical performance. The effects of the absorbing layer thickness on the performance of solar cells have been reported [1-4]. The production method for thin film Photovoltaic (PV) modules, where a thin layer of semiconductor material is deposited over a large area makes it difficult to maintain material thickness uniformity over the entire area. When series connected module cells have varying film thicknesses, their electrical properties may become inhomogeneous affecting the module current output. Inhomogeneity in electrical performance may also arise from the presence of defects in the active layers [5].

Different techniques such as solar LBIC, electroluminescence (EL) and thermoreflectance imaging have been reported as some of the techniques to investigate the nature of defects and their localization in solar cells and modules [6-8]. Structural, point defects or impurities in the solar cell may introduce localized intermediate energy levels in the band gap of semiconductor materials. These levels act as deep level traps for charge carriers. Material defects in thin film modules may result from the laser scribing process during module manufacture [9]. In this work, the use of large area light beam induced current (LA-LBIC) technique is used to investigate the homogeneity in photo-current in PV modules.

Further the correlation between the changes in the photo-current with the nature and type of defects is explored.

2. Experimental

2.1. Equipment

The LA-LBIC measurement system used for the non-destructive investigation of PV module performance is shown in Figure 1. The system allows for investigation of a module area up to 1.3 m x 1.3 m. The system consists of a motorized x-y scanning stage with a mounted light beam source, a source measure unit (SMU) and a control computer. The entire system is controlled by a custom-built labview program.

The amorphous silicon (a-Si) and copper indium diselenide (CIS) PV module used in this study are the ones manufactured by monolithic integration. The CIS module is manufactured by Siemens while the manufacturer of the a-Si module is not indicated.

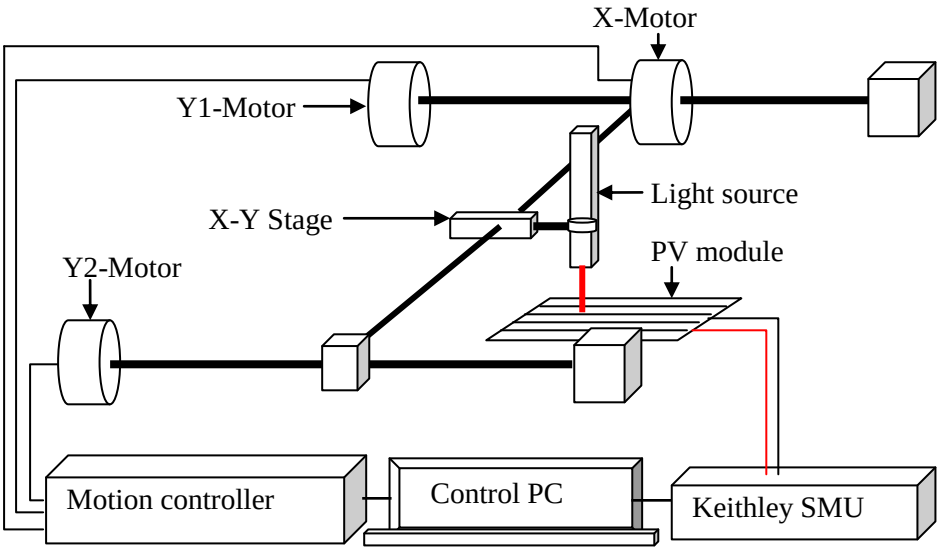
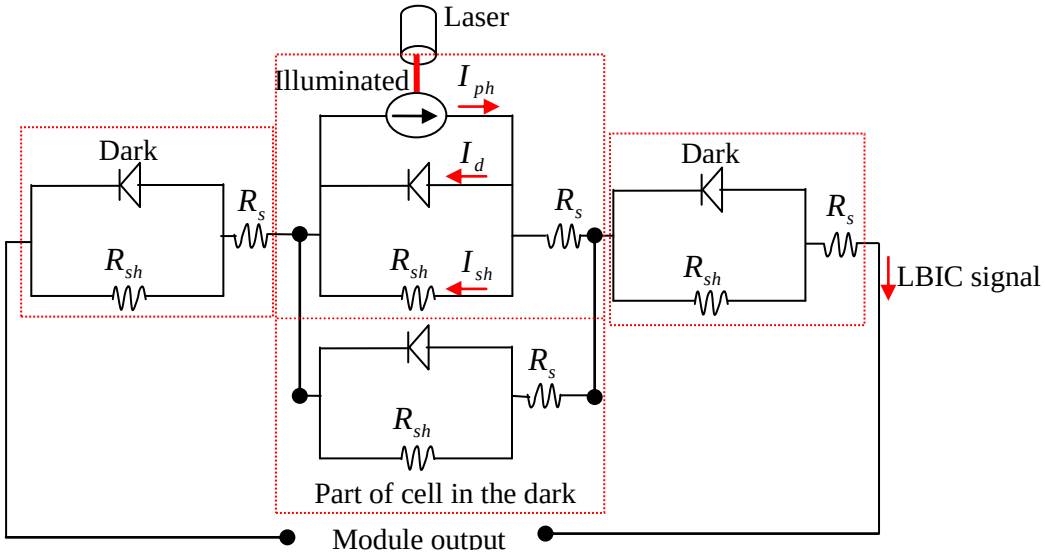


Figure 1: Schematic diagram of LA-LBIC measurement system

2.2. LBIC principle

In a PV module, the cells are connected in series. As a result, the current generated by the laser beam that is incident at a point on a cell has to pass through all the other series connected cells. Figure 2 illustrate the flow of induced current from one cell in a module made of three series connected cells. The LBIC signal depends on the laser induced current at a point on a particular cell as well as the combined constant series and shunt resistance of all the remaining series connected cells in the module. The dark leakage current through the pn junction of the dark cells can be considered negligible. The variation in LBIC signal during the scanning process thus originates from the cell being scanned and not from the rest of the module cells that are essentially in the dark.



I_{ph} is the photo-generated current, I_d is the dark current, R_{sh} is the shunt resistance and R_s is the series

Figure 2: Shows the current flowing through R_{sh} and R_s of the dark cells.

2.3. Measurements

A laser light source (633 nm, 5 mW) mounted on an x-y stage is used to spot illuminate the module point-by-point. A source measure unit was used to keep the module at a fixed bias voltage while measuring the module current output. Measurements are performed on the module in the dark.

Since the performance of the individual cells in a module is likely to vary, cell mismatch in modules become very probable. When a module with mismatched cells is illuminated, the operating points of individual cells get shifted along the module's current-voltage (I-V) curve in order to ensure equal laser beam induced current to flow in the series connected string. Figure 3 illustrate the shift in the operating points of the module cells during scanning. When a point on a cell is illuminated, the illuminated cell is forced into forward bias which subsequently reverse bias the non-illuminated cells. The changing laser beam induced current causes the operating point of individual cells to change during the scanning process while the sum of the individual cell voltages remains constant, equal to pre-selected module bias set by the SMU. The operating points of individual scanned cells therefore depend on the selected module bias and on the sum of the reverse bias voltages of remaining cells in the dark. The operating point range of individual cells may therefore be vastly different, making it difficult to compare LBIC signals between cells of a module. LBIC signals within an individual cell can however be compared due to the fact that the influence from neighbouring cells in the dark is considered constant. By changing the pre-set module bias voltage between scans and comparing the LBIC signals between cells, valuable insight may be gained on the operating point of scanned cells as well as the origin of LBIC features.

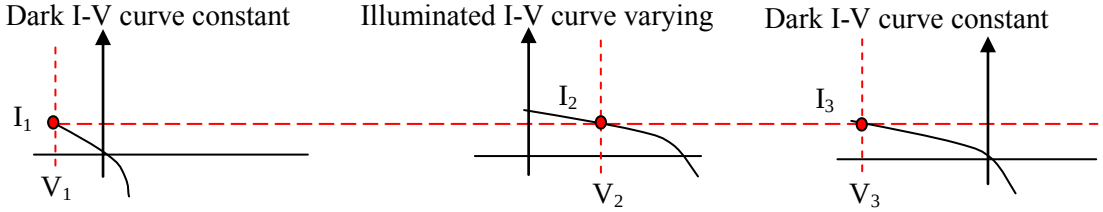


Figure 3: Shifts in operating points on the I-V curves of three module cells during scanning.

3. Results and Discussion

The LA-LBIC measurement results for small areas of a-Si and CIS module are presented. The complete a-Si module has an area of 45 cm x 30 cm and is made up of 29 series connected cells while the CIS has an area of 36 cm x 28 cm, of 50 series connected cells. Figure 4 shows photographs of the scanned areas of the modules.

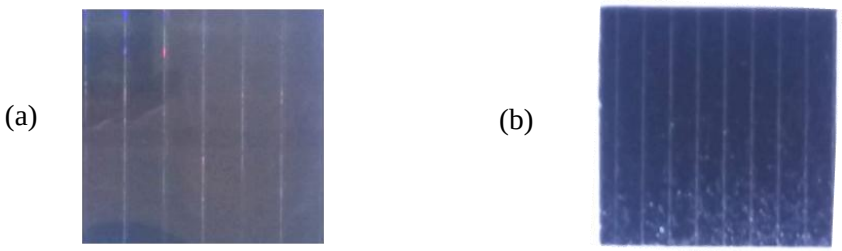


Figure 4: Photograph of part of (a) a-Si module of area 6 cm x 6 cm (b) CIS module of area 5 cm x 5 cm

The modules were scanned at various constant module bias voltage levels using a laser with a 633 nm wavelength, a power of 5 mW and an illumination spot diameter of approximately 1 mm. Figure 5 shows the photo-response maps of the a-Si module at different module bias voltages. Apart from cell 3 and 4 from the left, the rest of the cells show a reduction in photo-response with increase in forward biasing, which is typical of the diode electrical behaviour along the I-V curve. Cell 3 and 4 however show an increase in photo-response with increase in forward bias indicating an unexpected non-linear response. Since the power of the laser beam and the temperature of the module are assumed to remain constant, the I-V curves of all cells should remain constant while the laser beam illuminates a particular spot on cell 3. The slight overall increase in LBIC signal (module current) can thus only result from a drop in the bias voltage of the (current producing) cell 3 while the sum of the dark cell's bias increases. This non-linear behaviour is contradictory and is the subject of further investigations. The response of this cell to forward biasing is similar to the defect stretching from the left and covering five cells of the module shown in Figure 5. This defect could result from a variation in deposition layer thickness. Little variation in current is observed within cells and indicates a relative uniformity of the deposited material properties over the cell surface. The large variation in current between cells are due to differences in individual cell biases and can be ascribed to large differences in series and shunt resistances between cells.

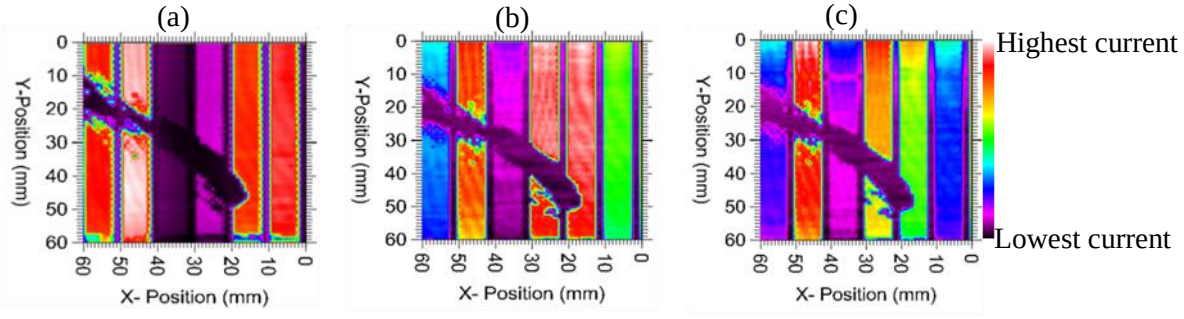


Figure 5: Photo-response map of a-Si module at module bias of (a) 0 V (b) 3 V (c) 5 V

The LBIC scan of the CIS module shown in Figure 6(a), shows a line of electrical disconnection across the cells. A line scan extracted from the LBIC map in Figure 6(a) at position x = 15 mm is shown in Figure 6(c), indicating a drop in current from 3.61 µA to 0.26 µA as a result of the disconnection. This electrical disconnection is verified by the electroluminescence (EL) image in

Figure 6(b) and is indicated by arrows. EL measurement is performed by forward biasing the module using a d.c power supply and taking its image by use of a CCD camera. When a module is forward biased, radiative recombination of carriers results, leading to emission of light. Significant variations in performance both between and within individual cells are observed in this particular module. Large variations in performance within individual PV cells are usually uncommon, and for this particular case, it is a result of the disconnecting feature.

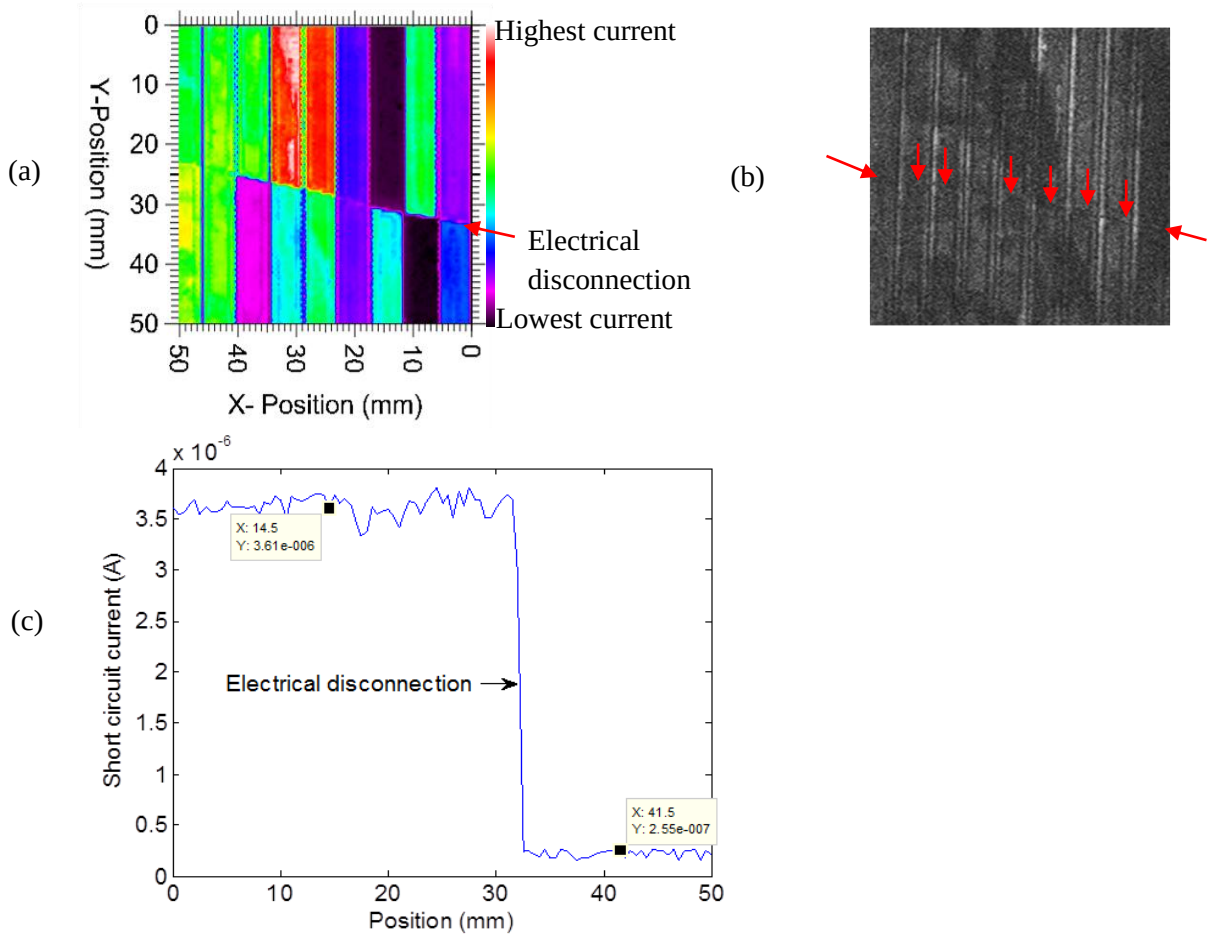


Figure 6: (a) LBIC map of CIS module (b) EL image (c) Line scan at position x = 15 mm in (a)

4. Challenges of LA-LBIC and possible solution

The variations in individual dark cell voltages due to the differences in cell resistances complicate the interpretation of LBIC data since the conditions are bound to vary from cell to cell during scanning.

To minimise this effect, the use of background light biasing is proposed. This method has been reported previously [10] and entails placing the scanned cell under a limiting condition. To achieve this, background illumination is applied to the whole module while the cell being scanned is covered with a bandpass filter that allows the scanning laser beam's light to pass through. This is expected to yield more reliable results as the cell being scanned is the weakest cell and hence determines the module current output and LBIC signal.

5. Conclusion

Homogeneity in photo-current within, and between cells as well as defects in a-Si and CIS thin film photovoltaic modules were investigated using a LA-LBIC measurement system. As a consequence of variations in I-V operating points between series connected cells, small material differences between cells are greatly amplified. Within a cell, the scanned points are essentially parallel connected and variations are less dramatic, since the whole cell's bias remains constant. The uniformity of the LBIC signal within cells gives an indication of the homogeneity of the deposited thin film material whereas the dramatic differences between cells points to big differences in the contact formation between cells. The electroluminescence technique was found to complement the LA-LBIC measurement technique in module defect investigations. The complications involving interpretation of results obtained by scanning the modules in the dark at a fixed module voltage were discussed and possible solution presented. Performing LA-LBIC scans on a cell placed under a complete limiting condition was proposed to be the most viable solution to acquire reliable signals from modules while scanning at a fixed module voltage. Variation in photo-current between cells and defects are regarded as the common performance hindrances in thin film PV modules.

References

[1] H.Y. Zhang, W.G. Wang, H. Liu, R. Wang, Y.M. Chen, Z.W. Wang, Effects of TiO_2 film thickness on photovoltaic properties of dye-sensitized solar cell and its enhanced performance by graphene combination, Materials Research Bulletin, 49 (2014) 126-131.

[2] M.M. Islam, S. Ishizuka, A. Yamada, K. Matsubara, S. Niki, T. Sakurai, K. Akimoto, Thickness study of Al:ZnO film for application as a window layer in $Cu(In_{1-x}Ga_x)Se_2$ thin film solar cell, Applied Surface Science, 257 (2011) 4026-4030.

[3] L. Li, D.E. Zhao, D.S. Jiang, Z.S. Liu, P. Chen, L.L. Wu, L.C. Le, H. Wang, H. Yang, The effects of InGaN layer thickness on the performance of InGaN/p-i-n solar cells, Chinese Physics B, 22 (2013) 1-4.

[4] E.Q.B. Macabebe, C.J. Sheppard, E.E. van Dyk, Device and performance parameters of Cu(In,Ga)(Se,S)2-based solar cells with varying i-ZnO layer thickness, Physica B: Condensed Matter, 404 (2009) 4466-4469.

[5] D.K. Schroder, Semiconductor material and device characterisation, Third ed., John Wiley &Sons, Hoboken, New Jersey, 2006.

[6] F.J. Vorster, E.E. van Dyk, Bias-dependent high saturation solar LBIC scanning of solar cells, Solar Energy Materials and Solar Cells, 91 (2007) 871-876.

[7] A. Gerber, V. Huhn, T.M.H. Tran, M. Siegloch, Y. Augarten, B.E. Pieters, U. Rau, Advanced large area characterization of thin-film solar modules by electroluminescence and thermography imaging techniques, Solar Energy Materials and Solar Cells, 135 (2015) 35-42.

[8] D. Kendig, G.B. Alers, A. Shakouri, Thermoreflectance imaging of defects in thin-film solar cells, in: Reliability Physics Symposium (IRPS), 2010 IEEE International, 2010, pp. 499-502.

[9] C.Y. Tsai, C.Y. Tsai, Development of Amorphous/Microcrystalline Silicon Tandem Thin-Film Solar Modules with Low Output Voltage, High Energy Yield, Low Light-Induced Degradation, and High Damp-Heat Reliability, Journal of Nanomaterials, 2014 (2014) 10.

[10] P. Vorasayan, T.R. Betts, R. Gottschalg, Limited laser beam induced current measurement: a tool for analysing integrated photovoltaic modules, Measurement Science and Technology, 22 (2011) 1-7.

Enhancing light absorption and lifetime stability of organic solar cells using pentacene encapsulation

F Otieno[1,3], K Kamalakannan [2], M Airo[2] and D Wamwangi[1,3]

[1]Material Physics Research Institute, School of Physics, University of the Witwatersrand, Private Bag 3, Wits, 2050, Johannesburg, South Africa
[2] School of Chemistry, University of the Witwatersrand, Private Bag 3, Wits, 2050, Republic of South Africa
[3] Materials for Energy Research group, University of the Witwatersrand, Private Bag 3, 2050 Wits, Johannesburg, South Africa

Email: Francis.Otieno@students.wits.ac.za

ABSTRACT. In the present work we compare the performance of a pristine P3HT:PCBM based bulk heterojunction solar cell with a device incorporated with pentacene encapsulated Ag nanoparticles. Encapsulation has been achieved through thermal evaporation of a 22 nm pentacene thin film. Introduction of pristine Ag nanoparticles revealed better overall performance with a 22% increase in photo-conversion efficiency. This is attributed to the localized surface Plasmon effect by Ag nano-islands. A study of device lifetime showed that despite the efficiency enhancement, the device drastically degraded upon exposure to air. Encapsulating Ag nanoparticles with pentacene increased light absorption and device stability but traded off photo-conversion efficiency. This long lifetime could be as a result of silver passivation by pentacene which minimizes oxidation of Silver nanoparticles and etching of the ITO electrode by the acidic PEDOT:PSS.

1. Introduction

The increasing demand for energy in this era of escalating energy prices as well as global warming resulting from the extensive use of fossil fuels has prompted intensive research on green energy as potential remedies to providing clean, cheap and inexhaustible energy. Photovoltaic offers a promising alternative due to the intrinsic characteristics of the energy source since solar energy is abundant, free and green [1-2].

Among the numerous photovoltaic options, organic solar cells made from conjugated polymers and hybrid materials continue to attract a lot of attention. This is driven by the low cost of high volume production of polymers, their ultrafast optoelectronic response and generation of charge carriers at the donor – acceptor interface. In addition, these materials can be fabricated on flexible substrates and their band gaps can be tuned depending on the length and functional group of polymers used for synthesis [3].

Despite these attractive properties their low photo-conversion efficiency (10-12%) and instability to air remain their greatest drawbacks to making them competitive with grid parity. Furthermore the tradeoff between their short exciton diffusion lengths (10-20 nm) and the high thickness dependent absorption (> 10^5 cm^{-1} within the visible spectrum) constrain the optimal thickness of the active layer

to a few nanometers [4]. Traditionally, polymer based solar cells have a short life expectancy compared to their inorganic counterparts [5]. Degradation of organic solar cells is attributed to the oxidative damage of the photoactive layer associated with device illumination in the presence of oxygen. Furthermore, the chemical interactions at the ITO electrode/organic interface have also caused device degradation. This mechanism of degradation is attributed to the hygroscopic and acidic nature of PEDOT:PSS which etches out the ITO layer. Concerted research efforts continue to be employed to curb device degradation [6-8].

Incorporating Ag nanoparticles in the device as sub-wavelength scattering elements can couple and trap propagating plane waves into an absorbing inorganic - polymer thin film. Furthermore as sub-wavelength antennas, they have coupled the plasmonic near-field to the semiconductor for increased effective absorption cross-section [9]. Silver shows stronger plasmonic properties in the visible optical spectrum with optical cross-section [10]. In this work we explore the effect of encapsulating Ag nanoparticles with a high mobility conjugated polymer on the optical, plasmonic characteristics and the electrical properties of the device. Ag nano-islands have been sputtered on ITO coated glass substrates in Argon ambient and at room temperature. The islands were later annealed in Argon to modify nanoparticles geometry and distribution. Pentacene was then thermally evaporated on the Ag surface. Our results show an intriguing but unexplored finding that apart from enhanced absorption, pentacene encapsulated devices showed a slowed degradation of the solar cell device in which they are incorporated into.

2. Experimental details

Ag nanoparticles have been deposited using RF magnetron sputtering onto ITO coated glass substrates at room temperature. The chamber was evacuated to a base pressure of about 2×10^{-5} mbar while sputtering was performed under Argon with a working gas pressure of 2.3×10^{-3} mbar. A 76 mm diameter Ag disk (99.99% purity) was used as the cathode for sputtering at an RF power of 15W for 10 seconds. The 'as deposited' Ag nano-islands of 10 nm thickness were annealed at 250 °C in Argon ambient for 15 mins. The sputtered Ag islands were loaded to a thermal evaporator evacuated to a base pressure of 1.5×10^{-5} mbar for pentacene deposition. Pentacene was heated to 170 - 200 °C to yield a deposition rate of 1.6-2.1 $Å s^{-1}$. The nanoparticles on ITO acting as the substrate were kept at room temperature during the evaporation of pentacene. The deposition rate was monitored using the quartz crystal monitor that showed the film thickness of 22 nm.

Optical absorption measurements of the encapsulated annealed Ag islands were made using a Cary 500 UV–Visible Spectrometer. The topography and nano-particle morphology of the films were studied using atomic force microscopy (AFM) in tapping mode. These nanoparticles were then incorporated in solar cell device. The procedures for device assembly have been explained in detail elsewhere [11]. Annealing of the reference and surface Plasmon based devices was carried out at 80°C for 15 mins under Argon ambient to remove organic polymer solvent and any residual water. The thicknesses of the PEDOT:PSS layer and P3HT:PCBM blend was measured by Surface profilometer. The current density–voltage (J–V) characteristics were obtained using HP 4141B source measure unit under 70 mW/cm^2 illumination (AM 1.5G) over time period of 166 hours of exposure to air ambient.

3. Results and discussion

3.1 Morphology studies by AFM

Figure 1 (a)-(b) and (c) - (d) depict the surface topography AFM image of annealed Ag nanoparticles before and after encapsulation using thermally evaporated pentacene, respectively.

From the morphology studies using Veeco Di 300 AFM, pentacene thin film on Ag nanoparticles shows rough surface with voids. The rms of the nanoparticles annealed at 250^{o}C increased by 6 nm This molecular voids may trap or scatter charge carriers. Craig et al. reported that the environment

with relaxed molecules surrounding a void is energetically more stable than that in a perfect crystal [12].

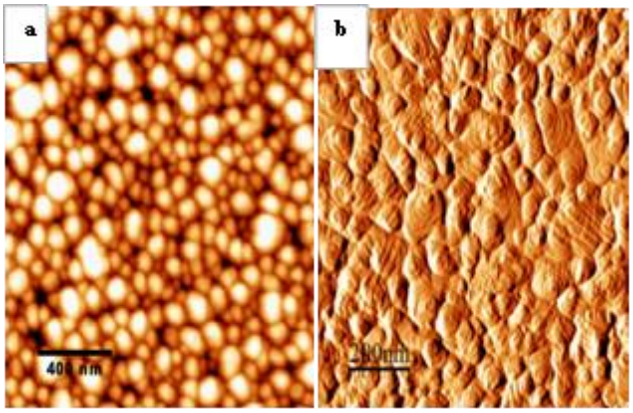


Figure 1(a). AFM image of RF sputtered Ag nanoparticles annealed at 250^{o}C in Argon ambience (b) 22.2 nm evaporated pentacene on the annealed nanoparticles.

However, the deformed lattice near a void may change the molecular electronic states in the defect region. Such distortion in the crystalline lattice may lead to formation of traps for holes or electrons [13]. These might be the cause of reduced current density of the device despite the enhanced optical absorption resulting from coupling of surface plasmons with pentacene shown in Figure 2.

3.2 Absorption spectra

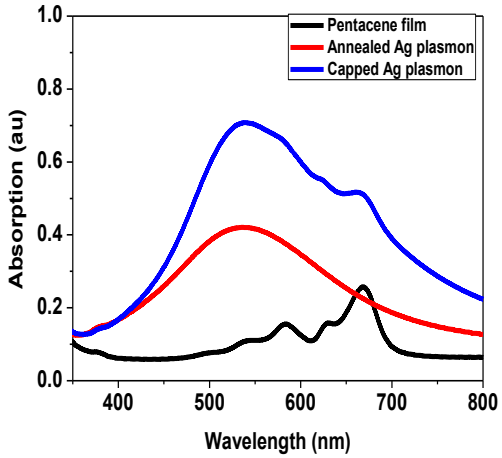


Figure 2. Optical absorption spectra of annealed Ag nanoparticles, pentacene, and pentacene encapsulated Ag nanoparticles (T_{ann} = 250 °C).

Pentacene thin films experience ultra-fast singlet and triplet exciton excitations due to polymorphism. Furthermore singlet fission into two triplets can occur within the visible region.
From the UV-Vis measurements shown in figure 2, pentacene thin film has four main different absorption peaks with energies of 1.85 eV (λ = 671 nm) , 1.97 eV (λ = 630 nm), 2.12 eV (λ = 585 nm), and 2.27 eV (λ= 545 nm). Since the excited states in pentacene thin film are singlet states, the excitation at λ = 671 nm is assigned to the S_1 state while that at λ = 630 nm corresponds to the second level of the Davydov doublet after Davydov splitting [14]. The absorption bands at λ = 545 nm (2.27 eV) correspond to an electron-hole pair, situated on two neighboring molecules of adjacent unit cells. The absorption band at λ = 585 nm is a charge transfer state at the sites of the two molecules within a unit cell [15]. The coupling of pentacene with Ag nanoparticles produced the enhanced absorption.

This enhancement is essential in the performance of the organic solar cell device. Spin coating of P3HT:PCBM blend on pentacene creates a different dielectric environment with enhanced optical absorption (fig. 3).

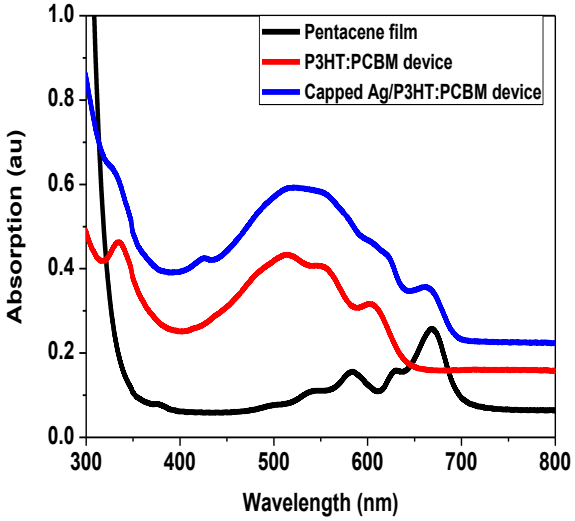


Figure 3. Absorption spectra of PEDOT:PSS, P3HT:PCBM on top of annealed Ag nanoparticles encapsulated with pentacene thin film.

The enhanced absorption at a wavelength of 670 nm extends the absorption spectrum of the active layer to a higher wavelength to provide multiple absorption peaks. However, the conductivity of the encapsulated layer remains a fundamental parameter in the determination in charge transport within the device. From figure 3, pentacene introduces a contact barrier of ~0.8 eV. This could cause hole trapping at the PEDOT:PSS-pentacene interface leading to reduced Jsc as observed in the J-V characteristic curve in figure 4.

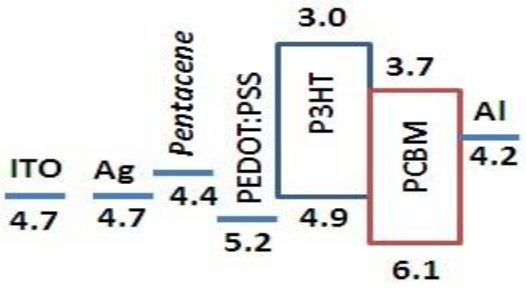


Figure 4. Band energy diagram for ITO/Ag nanoparticles /Pentacene/PEDOT:PSS/P3HT:PCBM/Al device.

3.3 J-V Characteristics curve

Despite the clear absorption enhancement upon encapsulation of Ag nanoparticles using pentacene, the Jsc, and Voc reduced by 55% and 5%, respectively. The reduction in Voc and Jsc lowers the PCE but improves device stability when exposed to air ambience as shown figure 6. The degradation rate in Jsc reduces upon encapsulation from 28.27 to 9.22 μA/cm^{2}min while Voc degrades slowly from 2.4 to 1.8 to mV/ min.

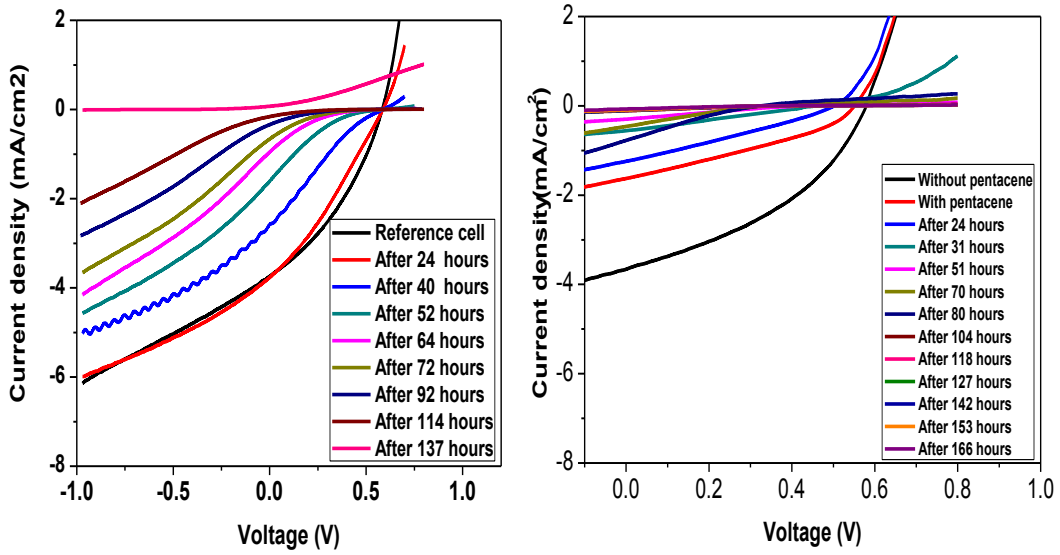


Figure 5. J-V characteristics for ITO/PEDOT:PSS/P3HT:PCBM/Al solar device degrading with time in air under AM 1.5 G filtered spectral illumination at incident intensity, 70 mW/cm^2.

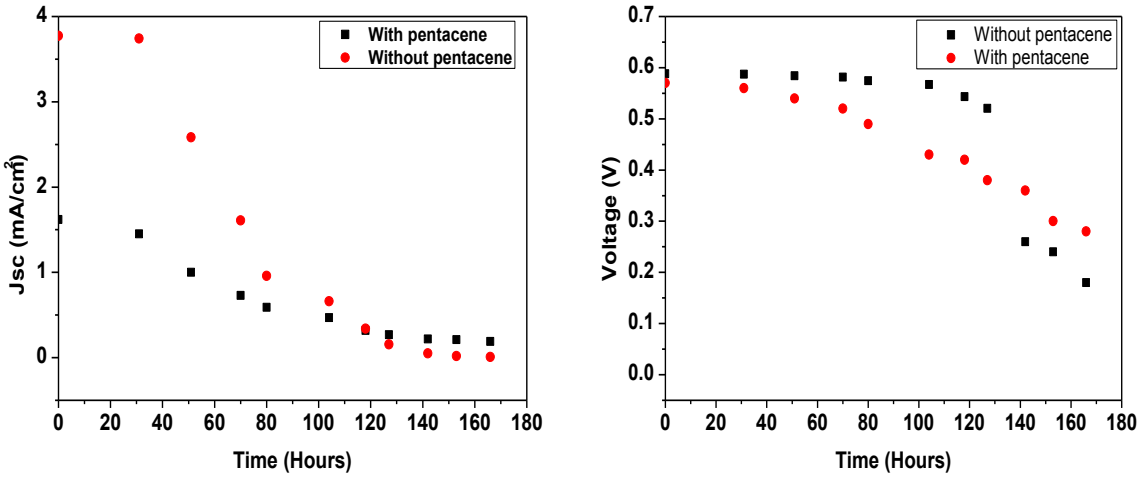


Figure 6. Performance of encapsulated device as a function of exposure time to ambient air.

Analysis of Jsc shows that devices without pentacene encapsulation acquired same value as the one with encapsulation after 118 hours despite its high original Jsc upon fabrication, thus encapsulation reduces the degradation rate of Jsc and Voc by 67% and 25%, respectively (fig. 6). Beyond this time, the encapsulated plasmonic device showed a greater Jsc. A similar trend was seen with the Voc where after 140 hours, encapsulated plasmonic device recorded a higher Voc compared to the device without encapsulation. The long electrical activity upon encapsulation could be as a result of pentacene forming a thin layer separating PEDOT:PSS from ITO Pentacene preventing etching of ITO layer by the PEDOT:PSS. The encapsulation of nanoparticles could also be used for passivation of the silver surface hence reducing the rate of oxidation of the nanoparticles that could lead to reduced plasmonic absorption. However, such passivation only requires a very thin layer in order not reduce the field around the nanoparticles as well as molecular self-assembly to enhance charge transport due to better π orbital alignment [10].

The reduced Jsc compared to plasmonic devices without encapsulation can be attributed to formation of a Schottky barrier between pentacene and Ag nanoparticles resulting in a large contact resistance between the ITO electrode and the active layer due to mismatch of their Fermi (HOMO) energy level.

4. Conclusion

ITO/Ag/PEDOT:PSS/P3HT:PCBM/Al bulk heterojunction solar cells was prepared with encapsulated Ag nanoparticles compared to that without encapsulation. The device performance was monitored in air ambient. The UV-Vis measurement has shown an enhanced optical absorbance of the device upon encapsulation of the Ag nanoparticles. The device performance was evident over a long period of time (166 hours) while the un-encapsulated showed no performance beyond 114 hours. However, the Jsc reduced upon Ag encapsulation. This could have been as a result of the HOMO mismatch of Pentacene and the PEDOT:PSS which is likely to cause charge recombination due to hole trapping at this interface. The low values of photocurrent in an encapsulated polymer cells despite the presence multiple absorption peaks and prolonged lifetime call for the device architecture modification to reduce the contact barrier to charge transport.

Acknowledgement

The authors would like to thank the University of the Witwatersrand, Material Physics Research Institute, School of Physics; the MMU facilities; Wits and Material Energy Research Group (MERG) for funding and NRF for funding of the research facilities.

References

[1] Asif M, Muneer T. Energy supply, its demand and security issues for developed and emerging economies. Renewable and Sustainable Energy Reviews. 2007 Sep 30;11(7):1388-413.

[2] Muneer T, Asif M, Munawwar S. Sustainable production of solar electricity with particular reference to the Indian economy. Renewable and Sustainable Energy Reviews. 2005 Oct 31;9(5):444-73.

[3] Sun SS, Sariciftci NS, editors. Organic photovoltaics: mechanisms, materials, and devices. CRC press; 2005 Mar 29.

[4] Sariciftci NS, Smilowitz L, Heeger AJ, Wudl F. Photoinduced electron transfer from a conducting polymer to buckminsterfullerene. Science. 1992 Nov 27;258(5087):1474-6.

[5] Norrman K, Larsen NB, Krebs FC. Lifetimes of organic photovoltaics: combining chemical and physical characterisation techniques to study degradation mechanisms. Solar energy materials and solar cells. 2006 Nov 6;90(17):2793-814.

[6] Reese MO, Morfa AJ, White MS, Kopidakis N, Shaheen SE, Rumbles G, Ginley DS. Pathways for the degradation of organic photovoltaic P3HT: PCBM based devices. Solar Energy Materials and Solar Cells. 2008 Jul 31;92(7):746-52.

[7] Kawano K, Pacios R, Poplavskyy D, Nelson J, Bradley DD, Durrant JR. Degradation of organic solar cells due to air exposure. Solar energy materials and solar cells. 2006 Dec 15;90(20):3520-30.

[8] Frohne H, Shaheen SE, Brabec CJ, Müller DC, Sariciftci NS, Meerholz K. Influence of the anodic work function on the performance of organic solar cells. ChemPhysChem. 2002 Sep 16;3(9):795-9.

[9] Atwater HA, Polman A. Plasmonics for improved photovoltaic devices. Nature materials. 2010 Mar 1;9(3):205-13.

[10] Sachan R, Ramos V, Malasi A, Yadavali S, Bartley B, Garcia H, Duscher G, Kalyanaraman R. Oxidation-Resistant Silver Nanostructures for Ultrastable Plasmonic Applications. Advanced Materials. 2013 Apr 11;25(14):2045-50.

[11] Kim MS. Understanding organic photovoltaic cells: Electrode, nanostructure, reliability, and performance. ProQuest; 2009.

[12] Craig DP, Markey BR, Griewank AO. Structural disturbance and lattice relaxation in molecular crystal packing calculations. Chemical Physics Letters. 1979 Apr 1;62(2):223-9.

[13] Seo S. Structural and electronic properties of pentacene at organic-inorganic interfaces. 2010.

[14] Davydov A S. (1964). The theory of molecular excitons. *Physics-Uspekhi*,*7*(2), 145-178.

[15] Fiebig M. *Spatially resolved electronic and optoelectronic measurements of pentacene thin film transistors* (Doctoral dissertation, lmu).

Division G – Theoretical and Computational Physics

Density functional theory on a lattice: Particle number dependence of the exchange-correlation potential.

Kossi Amouzouvi and Daniel Joubert

National Institute for Theoretical Physics, School of Physics and Mandelstam Institute for Theoretical Physics, University of the Witwatersrand, Johannesburg, Wits 2050, South Africa.

E-mail: kossi@aims.edu.gh

Abstract. In Kohn-Sham Density Functional Theory, the interacting system is mapped onto a fictitious independent particle system. In an ensemble continuous particle number formulation the exchange-correlation contribution to the potential of the independent particle system has a discontinuity as a function of particle number at integer particle numbers. This discontinuity is equal to the difference between the fundamental gap of the interacting system and the independent particle system. As the exchange-correlation depends on the Kohn-Sham potential, we numerically investigate the exact Kohn-Sham potential as a function of particle number (either integer or fractional) for a finite dimensional Hubbard model and compare the exact results to a local density approximation to the exchange-correlation functional.

1. Introduction

First introduced by Thomas and Fermi [1] in 1927 and placed on a formally sound base by Hohenberg and Kohn [2] and then formulated in a more practical form by Kohn and Sham [3] in the mid 1960's, Density Functional Theory (DFT) turned out to be a powerful tool for investigating the electronic structure of atoms, molecules or in general a many-electron system. It is mainly used in the material sciences, physics and chemistry. Discussions around DFT formalism, and even doubt about the existence of a well defined exchange-correlation (XC) potential [4] used to explained the problem of calculating the fundamental band gaps [5, 6] encountered in insulator and semiconductor materials, is ongoing. The exact DFT functionals are not known and in practice approximations have to be made to implement DFT. For models where the interacting many particle Hamiltonian can be solved exactly, an exact numerical solution for the fictitious single particle Kohn-Sham DFT equations can be determined and formal properties of DFT can then be investigated. Assessing the formalism and gaining insight into the properties of DFT can be assisted by investigating solvable models. Lattice-DFT [7, 8] is probably one of the best models where DFT theories can be accurately assessed. The simplest of these is the one-dimensional Hubbard model [9−11]. Although it is not possible to find a general analytic solution, there exist analytic solutions for special cases. For small systems Exact Diagonalization of the Hubbard model is feasible and that is the approach we chose. It is instructive to work out the effective Kohn-Sham potential for the Hubbard model self-consistently, to check its dependence on the number of particles (specifically for integer particle

numbers) and furthermore to investigate in the exchange-correlation potential as a function of fractional particle numbers in an ensemble formalism.

We focus, in Section 2, on the fundamentals of Lattice-DFT. In Section 3 the numerical approaches we used to probe the effective Kohn-Sham potential in terms of the number of electrons in the lattice using an ensemble DFT formalism are discussed. The corresponding results followed by their analysis are presented in Section 4. A thorough summary is included in Section 5.

2. Background

From here on N represents the number of sites of a finite one-dimensional Hubbard model, N_e is an integer number of electrons in the lattice whereas J is a fractional number of electrons. In the ensemble version of DFT fractional particle numbers J are well defined [12]. The simplest ensemble formulation arises when $E_0(N_e) - E_0(N_e + 1) = A$, the electron affinity, is less or equal to $E_0(N_e - 1) - E_0(N_e) = I$, the ionization potential. If this condition, the so-called convexity condition [13, 14], is satisfied, the energy for fractional particle numbers $J = (1-\alpha)N_e + \alpha(Ne+1)$ or simply $J = N_e + \alpha$, interpolates linearly between integer particle numbers: $E(J) = (1-\alpha)E(N_e) + \alpha E(N_e+1)$. In a similar way the ensemble ground density $\rho(J)$ interpolates linearly between integer particle numbers as $\rho(J) = (1-\alpha)\rho(N_e) + \alpha\rho(N_e+1)$. The chemical potential,

$$\mu = \frac{\delta E(J)}{\delta N_e} \tag{1}$$

and the functional derivative of $E(J)$ with respect to the charge density, can therefore have a discontinuity at integer particle numbers. Depending on whether $J \geq N_e$ or $J \leq N_e$, $\mu = E(N_e+1) - E(N_e) = A$ or $\mu = E(N_e) - E(N_e-1) = I$, respectively.

We adopt the linear chain single band Hubbard Model to describe the physics on a lattice. For an external site potential $v_{i\sigma}$, the Hamiltonian is given by:

$$\hat{H} = \sum_{1\leq i,j\leq N,\sigma} t_{ij} C^{\dagger}_{j\sigma} C_{i\sigma} + \sum_{1\leq i\leq N} u_i n_{i\uparrow} n_{i\downarrow} + \sum_{1\leq i\leq N,\sigma} v_{i\sigma} n_{i\sigma}. \tag{2}$$

Here σ is the electron spin and $C^{\dagger}_{i\sigma}$ and $C_{i\sigma}$ are respectively the Fermi creation and annihilation operators acting on spin σ electrons on site i or on the left or right vacuum. Finally, we denote by $n_{i\sigma}$ the operator $C^{\dagger}_{i\sigma}C_{i\sigma}$. The hopping matrix $\left(t_{ij} = \langle\, \phi_j \mid -\frac{\hbar^2}{2m}\nabla^2 \mid \phi_i \,\rangle\right)_{1\leq i,j\leq N}$ [15] conveys the same idea of motion of particles in a general system. For simplicity, we limited the hopping amplitude up to the first neighbor on the lattice and fix them to $-t$. This means $t_{ij} = -t\delta_{(i\pm1)j}$. The on-site coulomb repulsion is carried by u_i and reflects the Hartree-Fock potential. We fixed the repulsion term to a constant u and only considered spinless external potentials ($v_{i\sigma} = v_i$). With the resulting Hamiltonian

$$\hat{H} = -t\sum_{\sigma}\left[C^{\dagger}_{2\sigma}C_{1\sigma} + \sum_{i=2}^{N-1}\left(C^{\dagger}_{(i-1)\sigma}C_{i\sigma} + C^{\dagger}_{(i+1)\sigma}C_{i\sigma}\right) + C^{\dagger}_{(N-1)\sigma}C_{N\sigma}\right] + u\sum_{1\leq i\leq N} n_{i\uparrow}n_{i\downarrow} + \sum_{1\leq i\leq N,\sigma} v_i n_{i\sigma}. \tag{3}$$

we solved the finite chain Hubbard Model using Exact Diagonalization (ED). This works well for small systems, but even for a few tenths of sites in the model, the size of the matrix makes exact diagonalization a challenge. The non interacting, or KS system, is addressed in the Hubbard model by setting the on-site coulomb potential to zero and replacing the on-site potential v_i, by the KS potential $v_{ks,i}$.

3. Method

We start by solving the finite interacting Hubbard model exactly and determine the ground state density. Then, starting with a guess for the KS potential, we self-consistently determine the KS-potential that yields the same density as the corresponding interacting system. This procedure determines the KS-potential to within a constant only, since potentials that differ by a constant yield the same density. However, the highest occupied KS single particle energy ϵ_{max} must satisfy $\mu(J,v) = \epsilon_{max}(v_{ks})$ where $\mu(J,v)$ is the chemical potential of the interacting J-particle system with external potential v [13]. This condition determines the KS-potential uniquely.

For comparison, we follow reference [16] and define a local density approximation (LDA) for an infinite homogeneous lattice system: the site ground state; Hartree and kinetic energies are approximated by

$$e(t,u,n_i) = \begin{cases} \frac{-2t\beta(t,u)}{\pi} \sin\left(\frac{\pi}{\beta(t,u)} n_i\right) & \text{if } n_i \leq 1 \\ \frac{-2t\beta(t,u)}{\pi} \sin\left(\frac{\pi}{\beta(t,u)}(2-n_i)\right) + u(n_i-1) & \text{otherwise ,} \end{cases} \tag{4}$$

$$e_H = \frac{u}{4} n_i^2 \qquad \text{and} \qquad e_T = e(t,0,n_i). \tag{5}$$

where β is the solution of : $-\frac{2t\beta}{\pi} \sin\left(\frac{\pi}{\beta}\right) = -4t \int_0^\infty dx \frac{J_0(x)J_1(x)}{x(1+exp(x\frac{u}{2t}))}$.

With $V_{ks,i}^{LDA} = \frac{\delta}{\delta n}(e - e_T)\Big|_{n=n_i}$, we can easily evaluate the KS potential for a given density.

4. Results and analysis

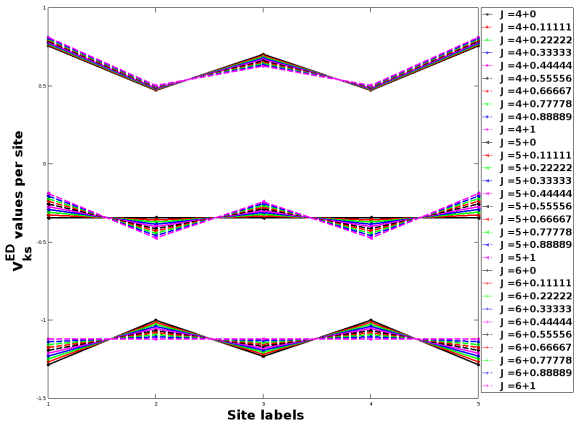


Figure 1. Exact KS potential as a function of particle number for particles in an infinite square well.

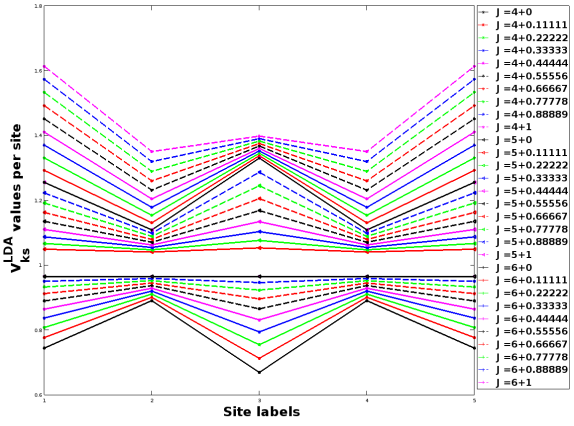


Figure 2. KS potential in the LDA approximation as a function of particle number for particles in an infinity square well.

Let us consider a 5 site lattice with a constant potential for our application. This corresponds to a infinite potential square well. We set the on-site coulomb potential u to 2 and the hopping term t to 1 in arbitrary units. In Figure 1 results for the exact KS potential are shown as a function of particle number. Figure 2 shows results for the LDA. There are spatially independent

jumps in the exact KS potential whenever the particle number passes through an integer as can clearly be seen in Figure 3. The LDA potential is continuous as a function of particle number as shown in Figure. 2. There is a rapid change in the LDA KS-potential near $J = 5$, at half band filling. This is a result of the discontinuity in the derivative of the energy of the infinite homogeneous Hubbard model at half filling (see Eq. 4) and is unrelated to the underlying properties of DFT. We note that, not surprisingly, both the exact and the LDA KS-potentials follow the evolution of the charge density (Figure 4), but the numerical values are quite different.

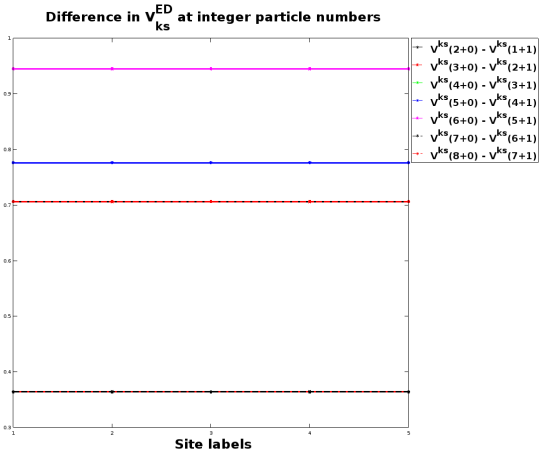


Figure 3. Difference in the exact KS potentials at integer particle numbers for J approaching an integer from above and below.

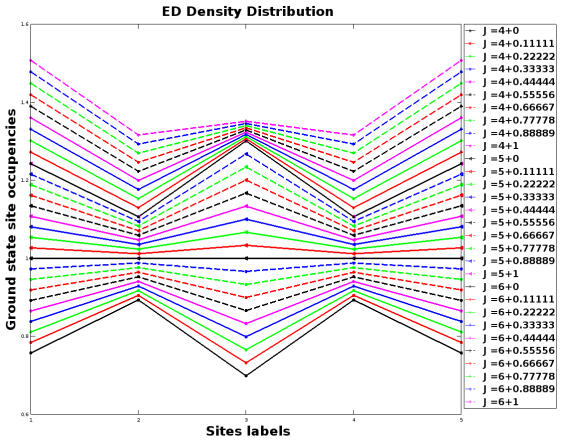


Figure 4. Variation of the ground state density.

5. Conclusion and Outlook

The discontinuity in the (functional) derivative of the exchange-correlation plus Hartree energy was confirmed for a finite Hubbard model. This is seen in the exact KS potential which changes continuously for particle numbers between integers, but for which there is a spatially independent discrete jump as the particle number passes through an integer. For comparison, the charge density is continuous and the approximate LDA-KS potential is continuous at integer particle numbers. A site derivative discontinuity in the LDA at site density $n_i = 1$ gives rise to a rapid change in the LDA-KS potential at half filling, but the LDA-KS potential is a continuous function of particle number. The discontinuity in the KS potential at integer particle numbers is at the heart of the problem of determining the fundamental gap from KS calculations [13]. As a further study we intend to explore an ensemble definition of the exchange-correlation energy to construct approximations to the KS potential that includes a discontinuity at integer particle numbers.

References

[1] Thomas L H 1927 *Mathematical Proceedings of the Cambridge Philosophical Society* vol 23 (Cambridge Univ Press) pp 542–548
[2] Hohenberg P and Kohn W 1964 *Physical review* **136** B864
[3] Kohn W and Sham L J 1965 *Physical Review* **140** A1133
[4] Zahariev F E and Wang Y A 2004 *Physical Review A* **70** 042503
[5] Sham L and Schlüter M 1983 *Physical Review Letters* **51** 1888
[6] Brosco V, Ying Z J and Lorenzana J 2013 *Scientific reports* **3**

[7] Schönhammer K, Gunnarsson O and Noack R 1995 *Physical Review B* **52** 2504
[8] Schonhammer K and Gunnarsson O 1987 *Journal of Physics C: Solid State Physics* **20** 3675
[9] Hubbard J 1963 *Proc. R. Soc. (London)* A **276**
[10] Essler F H, Frahm H, Göhmann F, Klümper A and Korepin V E 2005 *The one-dimensional Hubbard model* (Cambridge University Press)
[11] Schindlmayr A and Godby R 1995 *Physical Review B* **51** 10427
[12] Joubert D P 2012 *The Journal of chemical physics* **136** 174113
[13] Perdew J P, Parr R G, Levy M and Balduz Jr J L 1982 *Physical Review Letters* **49** 1691
[14] Sagvolden E and Perdew J P 2008 *Physical Review A* **77** 012517
[15] Jafari S A *Iranian Journal of Physics Research* **8** 113–120
[16] Capelle K, Lima N, Silva M and Oliveira L 2003 *The Fundamentals of Electron Density, Density Matrix and Density Functional Theory in Atoms, Molecules and the Solid State* (Springer) pp 145–168

Non-universality in a constrained period doubling route to chaos for Rössler's system

André E Botha[1] **and Wynand Dednam**[1,2]

[1]Department of Physics, University of South Africa, Science Campus, Private Bag X6, Florida 1710, South Africa

[2]Departamento de Fisica Aplicada Universidad de Alicante, San Vicente del Raspeig, E-03690 Alicante, Spain

E-mail: bothaae@unisa.ac.za

Abstract. Some important questions concerning the existence of periodic orbits in chaotic systems were investigated for the 3-dimensional Rössler system at different parameter values. While previous studies have classified the periodic orbits by varying only one or two of the three parameters at a time, the present study made use of optimization to find a continuous curve in parameter space corresponding to periodic orbits through one specific point in state space. The set of periodic orbits through the specific point underwent an interesting series of bifurcations that was different to the unconstrained canonical route described by Feigenbaum. It was concluded that the idea of classifying periodic orbits by grouping them into sets that pass through specific points of state space may provide further insight into clustering of orbits and the apparent relationship between the point dimension in parameter space and the Lyapunov dimension of the chaotic attractor at the corresponding parameters. Constrained stable periodic orbits may also have several practical applications, such as electrical timing circuits.

1. Introduction

State spaces of systems exhibiting low-dimensional chaos are replete with unstable periodic orbits that exist in the midst of chaotic attractors. Ever since this realization, which only took place in the 1990s, attempts have been made to characterize chaotic attractors via the properties (such as the characteristic exponents) of unstable periodic orbits [1]. In addition to being of a fundamental importance to understanding of chaotic dynamics, the properties of periodic orbits in chaotic systems is also of practical importance, since several methods have been devised by which periodic orbits may be extracted directly from experimentally measured time series [2]. The detection of periodic orbits in time series data is a good test for determinism and is also a starting point for many current methods of chaos control [3].

Recently Botha and Dednam [4] conjectured that periodic orbits exist through any point in the state space of the well-known 3-dimensional Rössler system [5]. The validity of their conjecture was demonstrated by using an optimized shooting method [6] to find periodic orbits corresponding to 10 000 randomly chosen initial conditions (x_0', y_0', z_0'), where x_0', y_0', $z_0' \in [-100, 100]$. (See Ref. [4] for details.) For all the tested initial conditions, the optimized shooting method was able to determine at least one periodic orbit through each initial condition, without any exceptions. Furthermore, it was noted that two or more distinct sets of the optimized parameters (giving periodic orbits) could be found for the same initial condition. In the present

article, the consequences of this multiplicity of parameter sets corresponding to a periodic orbit through one particular point will be explored. For simplicity, in our subsequent discussions, we will refer to all periodic solutions that pass through the same point, simply as a *state*.

One may thus ask, how many sets of parameters (typically) can be found for a given state? Could it perhaps be infinite? To answer this question, we consider the case in which a periodic solution is structurally stable within a given range of parameters. If one changes two of the parameters, say a and b in Rössler's system, the periodic solution should sweep its neighborhood two-dimensionally, unless the direction of sweep is degenerate. Therefore, in general, the three-dimensional space near the periodic solution may be expected to be fully covered by changing only two of the system parameters. For this reason, one may at first think that an infinite set of parameters, for which a periodic solution passes the given state, may exist in a one-parameter family. However, this is not the case. In general, the periodic orbit undergoes a series of saddle-node bifurcations, which follow the well-known Feigenbaum scenario [7, 8], i.e. as one of the parameters is varied, there is a cascade of periodic doubling orbits, eventually leading to fully developed chaos. As an example, consider the bifurcation of a period-(n-1) solution into a period-n solution. Initially the period-(n-1) solution is stable. As the single parameter is varied, it becomes progressively less stable in one direction only, until it becomes unstable, at which point a period-n solution is born. Immediately after this saddle-node bifurcation, the trajectory of the period-n orbit can be thought of as lying on the edge of a Möbius strip containing the unstable period-(n-1) orbit [9]. The unstable manifold of the period-(n-1) orbit lies along the surface of the Möbius strip. Thus in the canonical Feigenbaum scenario, period-n solutions are created by repulsion away from period-(n-1) solutions. One would therefore expect a period-n orbit to have no points in common with the period-(n-1) orbit from which it originated. When only one parameter is varied, this is indeed the case. However, if one considers bifurcations that occur when more than one parameter is allowed to vary at a time, then it is possible to obtain bifurcated orbits with one or more points in common.

2. Model and computational techniques

For the purposes of doing numerical calculations it is useful to write Rössler's system in terms of scaled coordinates. First, let $T > 0$ be some characteristic time in the system and rescale the time t' in terms of an unprimed (dimensionless) coordinate t, where $t' = Tt$. Here T could be the period of a periodic solution or, in the case of a chaotic attractor, an average orbital time. Second, choose a dimensionless scaling factor $\alpha \geq 1$, such that $x' = \alpha x$, $y' = \alpha y$, $z' = \alpha z$. In these unprimed coordinates, the Rössler system is given by

$$\dot{x} = T\,(-y - z)\,, \quad \dot{y} = T\,(x + ay)\,, \quad \dot{z} = T\,(b/\alpha + \alpha xz - cz)\,. \tag{1}$$

The form of (1) has several advantages. First, the scaled coordinates x, y, and z never become too large, since one can always increase the value of α to ensure that they remain bounded below a certain threshold. This boundedness of the coordinates facilitates more accurate numerical integration at large values of the primed coordinates. Second, (1) can be integrated conveniently over exactly one period T by integrating t over the closed interval $[0, 1]$. Moreover, the integration can be done without *a priori* knowledge of T, which is determined later, together with the other parameters, via Levenberg-Marquardt optimization.

In the present work we have modified the integration scheme that was previously used in our optimized shooting method. Instead of the 5th order (fixed time step) Runga-Kutta method which we used previously, we now employ the Hairer and Wanner implementation of the Dormand-Prince embedded method of order (4)5 [10]. For the same global accuracy, the latter method is significantly faster because it employs adaptive time stepping. It also has the advantage of allowing control over the local relative and absolute errors with dense output.

3. Results

The idea of clustering of points in the parameter space, corresponding to periodic orbits of (1), has been considered previously [11–16]. These works considered different methods for generating peculiar "swallow" or "shrimp" shaped regions of periodic orbits embedded in chaotic regions in two-parameter space, usually by fixing one of the parameters and then varying the other two. The "swallows" or "shrimps" then emerged when the maximum Lyapunov exponent was plotted as a function of the varying parameters. The "swallow" or "shrimp" shaped regions appeared as universal features of certain continuous chaotic flows [11, 15, 16].

In the present work we establish the distribution of periodic orbits as functions of the parameters. We therefore consider a random set of initial conditions which is constructed to have a uniform average density of 100 points per unit volume of state space, i.e. although the points are randomly chosen, they are chosen so that the point dimension (See equation 11 of Ref. [2].) of any initial condition is close to 3, which is the dimension of the state space. We then find the set of parameters corresponding to all possible periodic orbits through the chosen set of initial conditions. These periodic orbits were found by simultaneously optimizing the three control parameters and the period. In figures 1(a) and (b) we show two-dimensional projections of the parameter space of periodic points. In this case we considered 10800 initial conditions (in the positive half space $z_0 > 0$) of a $6 \times 6 \times 6$ cube centered on the origin. Only orbits with $1 < T < 50$ were sought.

As can be seen in figure 1, the periodic points generally appear to be most densely packed nearer to the origin, where the system in fact becomes integrable [17]. However, by calculating the point dimension of a selection of points we have shown that the distribution is non-uniform and that the point dimension of the parameter space is quite close to the Lyapunov dimension of Rössler's attractor at typical parameter values (See, for example, Ref. [18].) This result may just be coincidental, but it does warrant further investigation. At present we cannot postulate a reason why the point dimension in parameter space is close to the Lyapunov dimension of the attractor.

We conclude this section by noting that the found parameter points were all consistent with previous theoretical considerations that preclude the existence of periodic orbits in certain regions of the parameter space [13]. For example, in the present case, periodic orbits may only exist for $c^2 - 4ab \geq 0$, and with either (i) all parameters having the same sign (as in figure 1), (ii) $a < 0$, $b < 0$, $c > 0$ and $ab/c \geq c - 1/a - a$ (not shown), (iii) $a > 0$, $b > 0$, $c < 0$ and $ab/c \leq c - 1/a - a$ (not shown).

As explained in the introduction, the bifurcations that occur on orbits passing through the same point are fundamentally different from standard period-doubling bifurcations. To obtain such a sequence of bifurcations we arbitrarily chose some initial condition $\boldsymbol{x}_0$ and repeatedly searched for periodic orbits through $\boldsymbol{x}_0$, by using our optimized shooting method to optimize the parameters a, b and period T . The third system parameter was not part of the optimization. After obtaining the first periodic orbit corresponding to (a_1, b_1, c_1, T_1), we increased the value of c by an amount $\Delta c = 0.001$ and used the parameter values $a_1 \pm r_{11}, b_1 \pm r_{12}$ and period $T_1 \pm r_{13}$ as the initial guess for searching for the next periodic orbit. Here the r_{1j} $(j = 1, 2, 3)$ were random numbers ranging from just under ten percent to just over ten percent of previously found values of a, b and T. For example, after finding a_1, the initial guess for the next a was randomly chosen from the open interval $(a_1 - a_1/10, a_1 + a_1/10)$. Similarly for b and T. This procedure was repeated for $c \in [1, 5]$ and it produced about 10 000 orbits. As the value of c was systematically increased, the method sometimes failed to find a periodic orbit without exceeding the numerical tolerance of the residual, which was required to be less than 10^{-12} in this calculation. To proceed, we then guessed a period of twice the maximum value that was obtained at the last successful value of c. With the initial guess for the period doubled, the method was then once again able to find new periodic orbits and the values for c could once

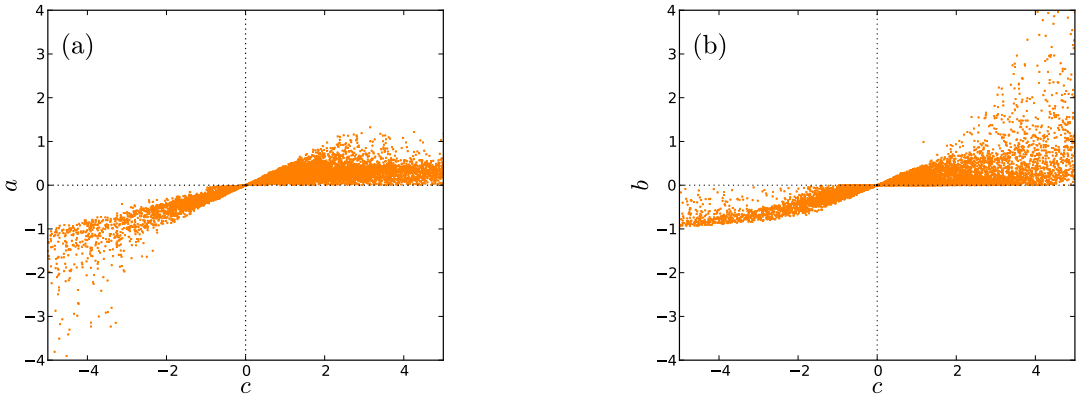


Figure 1. Clustering of points in parameter space, corresponding to periodic orbits of a uniformly dense set of initial conditions. (a) the ac-projection and (b) the bc-projection of the parameter space. Each figure contains 32243 points, with the highest density of points occurring closer to the origin, as discussed in the main text.

again be increased systematically. To ensure that all bifurcations were found, we repeated this procedure several times for systematically increasing and decreasing values of c. In this way we obtained a complete set of periodic orbits (with $1 < T < 50$) passing through $\boldsymbol{x}_0$.

In the remainder of this section we will discuss the periodic orbits through one particular initial condition, which we have randomly chosen to be $\boldsymbol{x}_0 = (1.483, -5.182, 0.0548)$. On the left hand axis of figure 2 are shown the parameters a and b as functions of c, as well as the stability μ as a function of all three parameters. More precisely, μ is the modulus of the maximum non-trivial Floquet multiplier, which has been calculated here according to the method developed by Lust [19]. Accordingly, a periodic solution is asymptotically stable if the modulus of each Floquet multiplier, except the trivial one, is strictly less than one. Otherwise, if one or more of the multipliers is greater than one in modulus, the solution is asymptotically unstable. In figure 2 we have plotted the unstable orbits as dashed lines. The scale on the right hand axis gives the periods of the orbits, which are plotted as a thicker green solid or dashed line. For clarity of presentation we have not included all sequences of periodic solutions that were found.

One interesting new feature of figure 2 is the appearance of unstable periodic orbits, apparently out of nowhere. Moreover, at certain values of c (for example, $c = 2.921$, 3.877 and 4.117) one can see that stable orbits become neutrally stable ($\mu = 1$). At such points there are subharmonic bifurcations from unstable to stable (or vice versa) orbits. Moreover, by looking at the full spectrum of Floquet multipliers immediately before and after the bifurcation, we observe that the bifurcation is not of the usual saddle-node type. Thus the period doubling bifurcations that occur in situations where one parameter is varied without any constraints, do not appear to be *generic*, i.e. their basic character is here seen to be altered by the perturbations that are necessary to maintain the orbit precisely through one point. This is an interesting new aspect of the bifurcations found here.

Lastly, we would like to emphasize that not all of the transitions that can be seen in figure 2, from one type of periodicity to another, are in fact bifurcations. In some cases (most notably at $c \approx 2.2$ and 3.75) the transitions occur for discontinuous values of a and/or b. When viewing figure 2, one thus has to bear in mind that, by definition, bifurcations only occur for infinitesimal change in one or more of the parameters. The inset to figure 2 also reminds us that the unstable orbits may in general lose stability in a variety of ways. In this bifurcation diagram, which

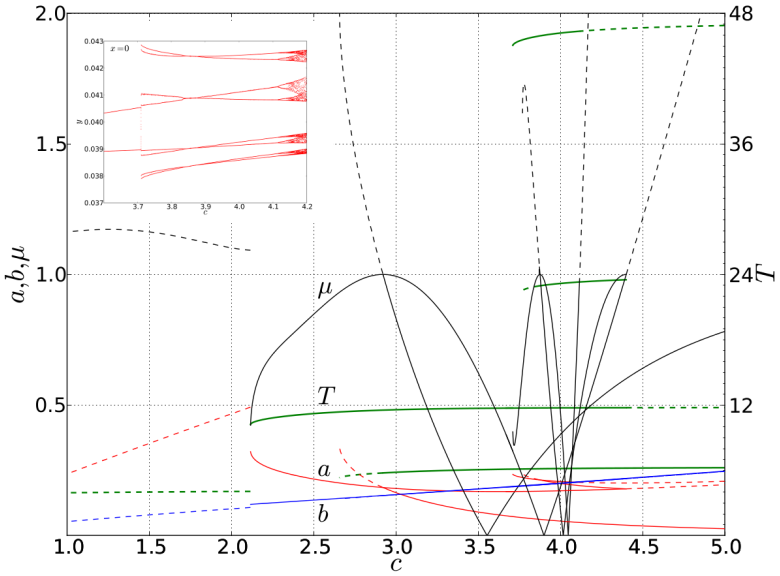


Figure 2. Sequences of periodic solutions through the arbitrary initial condition $\boldsymbol{x}_0 = (1.483, -5.182, 0.0548)$. The left hand axis shows the parameters a (red) and b (blue) as a function of the third control parameter c. Also plotted as a function of c is the largest Floquet multiplier μ (black). The right hand axis shows the period T (green) of the periodic orbit, also as a function of c. Solid lines correspond to stable periodic orbits, while the unstable orbits are indicated by dashed lines. The inset shows the bifurcation diagram obtained from the Poincaré section $x = 0$ for the orbit with $T \approx 47$ and over the range $c \in [3.6, 4.2]$. See text for details.

is a Poincaré section on the plane $x = 0$, we see that the initially stable period-8 orbit (with $T \approx 47$), first decays to a stable period-16 orbits (if they were not stable it would not appear on the bifurcation diagram) and then into chaotic attractors (which can be see to the far right of the diagram), as c is increased from roughly 4.1 to 4.2. Also, it is obvious that traditional bifurcation diagrams, such as the one in the inset, would not obey Feigenbaums universal scaling laws, simply because these law apply only to situations in which one parameter is varied at a time. In the present situation, all three parameters are varied at once.

4. Conclusion

The constrained period doubling route to chaos which we have described for Rössler's system is relevant to a wide class of systems, the dynamics of which can all be described in terms of an underlying unimodal map [1]. Here we have suggested an alternative way of looking at periodic orbits. Instead of fixing one of the system parameters and then considering the remaining two-dimensional parameter space, we have considered a specific point in the state space and varied all three parameters to obtain all periodic orbits passing through the specific point. Our results demonstrate that the periodic orbits correspond to curves in the parameter space, which are mostly smooth curves, i.e. there are only a small number of discontinuities. This is an unexpected result. Furthermore, our numerical investigations have revealed that period doubling bifurcations may not be generic, as was previously thought. We have obtained the curious result that the point dimension of points in the parameter space, corresponding to a uniform distribution of initial conditions in the state space, appears to be almost the same as

the Lyapunov dimension of the attractor at typical parameter values.

Other than being of mathematical interest, our results may also have some practical applications. There are in fact many practical situations where one may be interested in imposing constraints on a system. For example, the Rössler system itself has been implemented in electrical circuits for use in encryption, studies of synchronization, and of course chaos control [20]. These days systems of nonlinear equations are routinely implemented (mainly for use in chaos encryption) in field programmable gate arrays (FPGAs), which offer several advantages over analog electrical circuits. (See, for example, Ref. [21].) There is thus a real possibility of accurately varying the parameters of a system by changing, for example, resistances and capacitances of the constituent electrical components. As our results demonstrate, a careful choice of the system parameters may be used to engineer stable periodic signals of essentially any periodicity, while always retaining a specific value of the periodic signal. Such signals are important for making timing devices, where very specific values of a periodic signal may be required to trigger certain parts of more complicated electrical or FPGA circuits.

Acknowledgments
AEB would like to thank Craig Thompson for his informal participation in this research. WD thanks the National Research Foundation of South Africa for financial assistance in the form of a Scarce Skills Masters Scholarship (Grant Unique Number 92138).

References

[1] Hilborn R C 2000 *Chaos and Nonlinear Dynamics: An Introduction* 2nd ed (New York: Oxford University Press)
[2] So P, Ott E, Sauer T *et al.* 1997 *Phys. Rev. E* **55** 5398
[3] Ott E and Spano M 1995 *Physics Today* **48** 34
[4] Botha A E and Dednam W 2014 *Proceedings of the 59th Annual Conference of the SAIP* ed Karataglidis S and Engelbrecht C (7-11 July, University of Johannesburg)
[5] Rössler O E 1976 *Phys. Lett. A* **57** 397
[6] Dednam W and Botha A E 2015 *Engineering with Computers* **31** 749
[7] Feigenbaum M J 1978 *Journal of Statistical Physics* **19** 25
[8] Feigenbaum M J 1979 *Journal of Statistical Physics* **21** 669
[9] Ott E 1993 *Chaos in Dynamical Systems* (Toronto: Cambridge University Press)
[10] Hairer E, Nørsett S P and Wanner G 2008 *Solving Ordinary Differential Equations I: Nonstiff Problems* (Berlin: Springer-Verlag)
[11] Gomez F, Stoop R L and Stoop R 2014 *Bioinformatics* **30** 2486
[12] Castro V *et al.* 2007 *Int. J. Bifurcation Chaos* **17** 965
[13] Starkov K E and Starkov K K 2007 *Chaos, Solitons and Fractals* **33** 1445
[14] Gallas J A C 2010 *Int. J. Bifurcation Chaos* **20** 197
[15] Zou Y, Donner R V, Donges J, Marwan N and Kurths J 2010 *Chaos* **20** 043130
[16] Prants W and Rech P 2013 *Physica Scripta* **88** 015001
[17] Teryokhin M T and Panfilova T L 1999 *Russian Mathematics* **43** 66
[18] Wolf A, Swift J B, Swinney H L and Vastano J A 1985 *Physica D* **16** 285
[19] Lust K 2001 *Int. J. Bifurcation Chaos* **11** 2389
[20] See, for example, the website called Glen's stuff: last accessed 10/07/2015 *Online at http://www.glensstuff.com/rosslerattractor/rossler.htm*
[21] Qi G, Wang Z and Guo Y 2012 *Int. J. Bifurcation Chaos* **22** 1250287

Simulating cyclic loading of atomic-sized gold tips on gold surfaces via classical molecular dynamics and density functional theory transport calculations

W Dednam[1,2], C Sabater[2,4], M A Fernandez[2], C Untiedt[2], J J Palacios[3], M J Caturla[2]

1 Department of Physics, Science Campus, University of South Africa, Private Bag X6, Florida Park 1710, South Africa
2 Departamento de Fisica Aplicada, Universidad de Alicante, San Vicente del Raspeig, E-03690 Alicante, Spain
3 Departamento de Fisica de la Materia Condensada, Universidad Autonoma de Madrid, Cantoblanco, Madrid 28049, Spain
[4] Huygens-Kamerlingh Onnes Laboratorium, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, Netherlands

E-mail: 36127809@mylife.unisa.ac.za

Abstract. The ability to probe atomic-level interactions via scanning tunneling microscopy and other techniques has led to great interest in contact formation between atomic-sized metal electrodes. In the present work, it is demonstrated by atomistic simulations that atomic-sized gold surfaces can be mechanically annealed, or work hardened, until they are sharp and stable. We study the effect of the initial geometrical configuration of the tip on the process of formation of a stable structure between a tip and a flat gold surface, by using two different initial shapes of the tip in classical molecular dynamics simulations. Density functional theory transport calculations are performed to obtain the conductance of the simulated structures and compared directly with experimental measurements. Our results show that both configurations lead to stable structures after repeated cyclic loading and that the number of cycles required to reach stability depends on the initial configuration. The calculated conductance values at first contact between the tips and surface are consistent with the best available experimental data.

1. Introduction

The desire to manufacture smaller electronic devices and overcome constraints on the size and number of transistors that can be packed onto an integrated circuit [1], has perpetuated active research into various alternatives to well-established silicon technology[2–4]. One such avenue of research deals with atomic-sized conductors [5], in which fundamental processes such as bonding and the manifestation of friction at the atomic level become important. The scanning tunneling microscope (STM) or mechanically controllable break junction (MCBJ) [5, 6] are typically used for studying these systems. Both can be used to realize atomic-sized contacts between two surfaces experimentally. The accuracy of both techniques is made possible through the use of piezoelectric ceramics with the ability to displace, with subnanometer precision, one surface relative to the other [5]. A bias voltage is routinely applied across the two surfaces in order to measure the current as a function of the relative displacement between the surfaces. Small bias

voltages are used to ensure that the surfaces are in equilibrium at all times and consequently, that the electron transport across the atomic-sized contact region between the surfaces is adiabatic [5]. Under such conditions the conductance G ($= 1/R$, the inverse of resistance) across the electrodes can be obtained directly from the measured current, by dividing it by the bias voltage. The conductance instead of the resistance is measured because it exhibits quantization under certain circumstances in metallic nano-contacts, among other atomic-sized systems of various geometries and made from different materials [7, 8]. Conductance is usually expressed in units of $G_0 = 2e^2/h$, the quantum of conductance, with 2 being the multiplicity of the electron's spin, e its charge and h Planck's constant.

Nanocontacts have been studied extensively both experimentally and theoretically. (See Ref. [5] and the references therein.) Such systems exhibit many interesting phenomena. On the one hand, some metals such as Au, can form long monoatomic chains during stretching and before breaking of a nanojunction [9, 10]. This behaviour was first identified through observation of long plateaus in the conductance traces immediately before breaking [11]. Long monoatomic gold chains have also been observed directly in transmission electron microscopy measurements [12]. On the other hand, the conductance through the nanocontact during breaking or contact formation can exhibit very unique behaviours that depend on the type of metal as well as the geometry of the contacts [13, 14]. Interestingly, the geometries of the tips immediately before and after contact can be inferred from the conductance values obtained [15] by means of classical molecular dynamics (CMD) simulations [16, 17] and density functional theory (DFT) transport calculations [18, 19]. In this work we focus on how the geometry of the electrodes will influence the values of measured conductance in the particular case of gold contacts.

2. Computational methods

In previous works [20, 21] we showed that it was possible to model, via CMD simulations and DFT transport calculations, the repeated reforming and breaking of contact between two atomic-sized gold tips obtained from a ruptured gold nanojunction, which gave rise to very stable configurations of the tips in a process known as mechanical annealing (or work hardening). The results exhibited good agreement with the best available experimental data [15]. Here the same approach is taken to model the interaction between a sharp tip and a flat surface (mimicking the STM setup). This configuration is very different from the one we used previously: a narrow neck (mimicking the MCBJ setup). Thus, we aim to explore the influence of electrode geometries on the structures formed after contact formation. Moreover, experimental measurements of conductance between sharp tips and various surface geometries have been performed by the group of Kröger and others (see Ref. [14] and references therein).

We employ the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [22, 23] to run our CMD simulations. The empirically and computationally fitted embedded-atom-model (EAM) potential [24] serves to describe the forces between the atoms in our CMD simulations. Furthermore, in order to simulate work hardening of atomic-sized gold surfaces, we employ a simple but widely used type of EAM potential that has been fitted to various basic material properties of gold [25].

It is first necessary to select a particular tip configuration before we can run simulations. We take two different approaches. In the first approach a cone of near perfect symmetry, generated directly by means of LAMMPS and oriented along the (001) crystallographic direction, has been used (see figure 1). In figure 1(a), the apex atom on the cone-shaped tip approaches an absorbed atom (adatom) on a (111)-oriented surface from directly above (these two atoms will be the first to make contact and so have been colored brown for clarity). This mixed crystal orientation has been chosen based on the fact that (111)-oriented surfaces are the most stable energetically, and thus a tip with a (001) crystallographic orientation remains relatively stable during cyclic loading since (111)-oriented crystal planes form on its oblique faces. In addition, we placed an

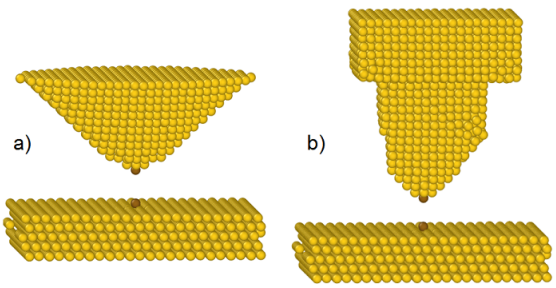


Figure 1. (Colour online) The initial configurations used in the simulations: a) a nearly perfectly shaped conical tip above an adatom on a clean surface, and b) a tip that was obtained by repeated rupture and contact formation of a notched nanowire, above the same surface with the adatom. Both tips are oriented along the (001) crystallographic direction, while the surface has a (111) orientation.

adatom on the surface directly beneath the lowest atom on the tip because this geometrical arrangement has been of particular interest in a number of works (see Ref. [14] and references therein).

Experimentally, however, in order to make a very stable tip and before performing the intended measurements, the tip is prepared by repeated contact formation [26]. To mimic this process, we have generated a second type of geometry of the tip (see figure 1(b)) using simulations similar to those performed previously in our group [15, 20, 21]. We started with a notched (001)-oriented nanowire and performed repeated rupture and contact formation on it to a maximum number of 6 atoms in the minimum cross-section of the constriction. (see Ref. [21] for details).

During the simulations performed in this work, the internal motion of the atoms in the first two top and bottom layers of the initial structures in figure 1 are frozen. The atoms in the top two layers of the tip are then displaced towards (during contact formation) or away (during rupture) from the surface at 0.004Å every picosecond, the remaining atoms responding dynamically to the rigid motion of the top layers. The velocity (equivalent to 0.4m/s) is about four orders of magnitude below the speed of sound in gold (3000m/s) [27], and combined with the low simulation temperature of 4.2K ensures that longer time-scale diffusive processes are negligible. The temperature is kept constant during the simulations by a Nose-Hoover thermostat [28, 29].

To demonstrate that mechanical annealing occurs upon repeated cyclic loading of the tips on the surface, the vertical distance between the first two atoms on the tip and surface to come into contact during approach, is plotted as a function of the bulk tip displacement. Contact is judged to occur when these two atoms come within a distance halfway between first and second nearest neighbours in a perfect gold FCC lattice (about 3.48Å). When the resulting minimum distance traces begin to overlap reproducibly, the surfaces are deemed to be stable and sharp, and the process of work hardening complete. Since experimental conductance traces only exhibit reproducibility up to contact conductances of around $5G_0$ during cyclic loading [20], the contact formation phases of cyclic loading in our simulations only continue until there are 6 atoms in the minimum cross-section. (See Ref. [21] for details on how the number of atoms in the minimum cross-section of the contact structures are determined). This follows from the observation that a monovalent gold atom contributes exactly one eigenchannel to transmission across atomic-sized gold junctions [5, 18, 19].

By definition, there are no electrons in CMD simulations: their effects are contained in the interatomic potentials described above. Thus, complementary methods, such as DFT transport calculations [18, 19] are required to make the analogy with experiments complete. So, to permit a comparison with experimental conductance values at first contact between the tips and surface, the DFT-based transport code ANT.G [30, 31] was employed. Such calculations are very computationally expensive. Therefore, to save time we have selected only 500 of the 5000 atoms from our CMD configurations, since the few atoms in the minimum cross-section

determine the transmission coefficient that is obtained from quantum transport calculations [18, 19]. On the other hand, to make the calculation more exact, an eleven electron basis set has been assigned to between 20 and 40 atoms in the minimum cross-section of the structures employed.

3. Results and discussion

Figure 2 plots minimum distance versus bulk tip displacement after 20 cycles of work hardening of the two initial configurations shown in figure 1. In figure 2(a) we see that it takes only three cycles before the traces overlap in a highly reproducible manner, demonstrating that a sharp and stable tip has formed. In figure 2(b), one additional cycle of contact formation is needed to achieve overlapping traces. These results are consistent with the cyclic loading performed on the wide (001)-oriented notched wire in Ref. [21]. The insets in figure 2 show the first-contact structures for two contact formation cycles (cycle 10 and 20 in both cases), once the tips and surfaces are sharp and stable. In figure 2(a), the first-contact structure resembles a monomer with 4 atoms in the layers immediately above and below it. In figure 2(b), however, the situation is more complicated with more than one atom in the minimum cross-section in both insets. The initial shape of the tip clearly plays a role in the arrangements of atoms in the minimum cross-section at first contact between tips and surfaces.

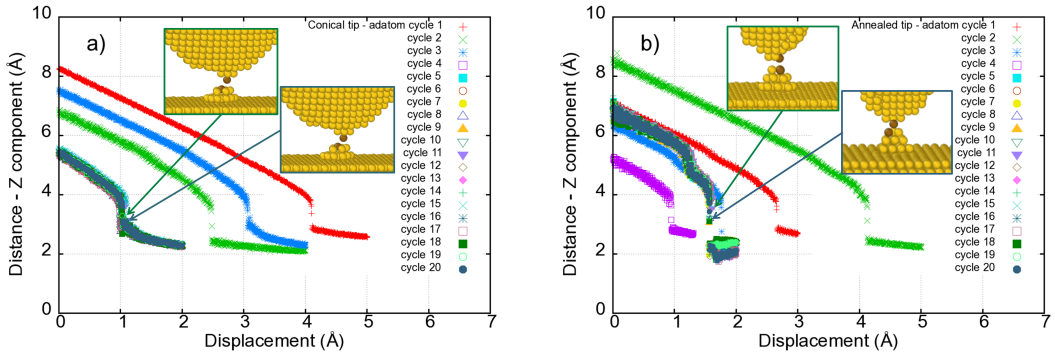


Figure 2. (Colour on-line) Minimum distance between the tips and surface versus bulk tip displacement for: (a) a cone-shaped tip oriented along the (001) direction, with its lowest atom positioned directly above an adatom in a three-atom hollow on the (111) surface, and (b), a (001)-oriented tip obtained by (simulated) cyclic loading of a notched nanowire, with its lowest atom positioned directly above the same surface and adatom as in (a).

It bears mentioning that the CMD simulations of a tip on a surface do not result in indentation through the surface, since in general the STM tip has to be harder than the substrate for this to occur [5, 32, 33]. Instead, after contact a tip is formed on top of the surface as shown in the insets in figure 2. The formation of a small mound on a gold surface after cyclic loading of a STM tip on a clean gold surface has actually been observed experimentally (see figure 3). Figure 3(a) shows the surface before contact with the STM tip and figure 3(b) after cyclic loading (the crystallographic orientations of the tip and surface are unknown in this case) [34]. Although the temporal and spatial resolutions of the experiments and our simulations differ greatly, the overall phenomenon is the same.

Previously in our conductance calculations, we used a valence basis set of only one electron per gold atom (6s atomic orbital). However, this led to an anomalously high value of around $1.60G_0$ for the monomeric gold contact obtained in Ref. [21]. In the present work, an eleven

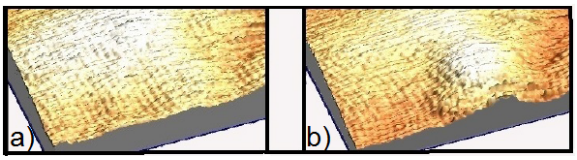


Figure 3. (Colour online) The formation of a mound on a gold surface (20nm x 10nm in dimension) before (a) and after (b) repeated contact formation at $4.2K$ between a STM tip and a clean surface, both made of 99.995 percent gold (of unknown crystallographic orientation) [34].

electron basis set (5d sub shell + 6s sub shell) was assigned to 10 atoms, including the monomer, in the contact region. An improved conductance of $1.18G_0$ was obtained in this way, which is within the range expected for this type of single-atom junction [15]. Applying the 11 electron basis to an even larger number of atoms in the constriction did not affect the outcome of the calculation appreciably. Hence, the eleven electron basis set has been assigned to between 20 and 40 atoms in the minimum cross-section of the structures employed in conductance calculations. As a result, we obtained a value between $1.0-1.20G_0$ for the monomers in the insets in figure 2(a) and around $1.70G_0$ in the case of the structures in the insets in figure 2(b).

4. Conclusion

It is demonstrated that repeated contact formation between gold surfaces mimicking the STM geometry also lead to stable and sharpened tips. In particular, the tips deposit atoms on the surface as cyclic loading proceeds, forming gold mounds resembling sharp tips. This confirms the equivalence between the STM and MCBJ experimental techniques for pure gold nanojunctions.

In the case of the symmetrical tip in figure 1(a), the stable configuration that is obtained after cyclic loading resembles a monomer, and conductance calculations on two first-contact simulation snapshots from two different contact formation cycles yield values within the range expected for monomeric contacts ($0.76 - 1.20G_0$) [15]. The less symmetrical tip in figure 1(b), however, does not produce a simple first-contact geometry. Various atoms overlap directly at the moment of first contact. Nevertheless, the conductance at first contact is consistent with less probable conductance values seen in histograms constructed from hundreds of thousands of conductance traces [13].

Acknowledgments

WD thanks AE Botha for helpful discussions. Financial assistance of the National Research Foundation of South Africa towards this research is hereby acknowledged by WD. This work is supported by the Spanish government through grant FIS2013-47328, the Conselleria d'Educació of the Generalitat Valenciana through grant PROMETEO/2012/011, and by European Union structural funds and the Comunidad de Madrid Program P2013/MIT-2850. It is also part of the research programme of the Foundation for Fundamental Research on Matter (FOM), which is financially supported by the Netherlands Organisation for Scientific Research (NWO).

References

[1] Moore G E 1965 *Electronics Magazine* **38** 114

[2] Lu Y, Huang J Y, Wang C, Sun S and Lou J 2010 *Nature Nanotechnology* **5** 218

[3] Shulaker M M, Hills G, Patil N, Wei H, Chen H Y, Wong H S P and Mitra S 2011 *Nature* **501** 526

[4] Radisavljevic B, Radenovic A, Brivio J, Giacometti V and Kis A 2011 *Nature nanotechnology* **6** 147

[5] Agraït N, Levy-Yeyati A and van Ruitenbeek J M 2003 *Phys. Rep.* **377** 81

[6] Binnig G, Roher H, Gerber C and Weibel E 1982 *Phys. Rev. Lett.* **49** 57

[7] van Wees B J *et al.* 1988 *Phys. Rev. Lett.* **60** 848

[8] Agraït N, Rodrigo J G and Vieira S 1993 *Phys. Rev. B* **47** 12345

[9] Krans J M, van Ruitenbeek J M, Fisun V V, Yanson I K and de Jongh L J 1995 *Nature* **375** 767

[10] Smit R, Untiedt C, Yanson A I and van Ruitenbeek J 2001 *Phys. Rev. Lett.* **87** 266102

[11] Yanson A I, Bollinger G R, van den Brom H E, Agraït N and van Ruitenbeek J M 1998 *Nature* **214** 783

[12] Ohnishi H, Kondo Y and Takayanagi K 1998 *Nature* **395** 780

[13] Untiedt C *et al.* 2007 *Phys. Rev. Lett.* **98** 206801

[14] Kröger J, Néel N and Limot L 2008 *Journal of Physics: Condensed Matter* **20** 223001

[15] Sabater C, Caturla M J, Palacios J J and Untiedt C 2013 *Nanoscale Research Letters* **8** 257

[16] Frenkel D and Smit B 2002 *Understanding Molecular Simulation: From Algorithms to Applications* (New York: Academic Press)

[17] Rapaport D C 2004 *The Art of Molecular Dynamics Simulation* (New York: Cambridge University Press)

[18] Palacios J J, Pérez-Jiménez A J, Louis E, SanFabián E and Vergés J A 2001 *Phys. Rev. B* **64** 115411

[19] Palacios J J, Pérez-Jiménez A J, Louis E, SanFabián E and Vergés J A 2002 *Phys. Rev. B* **66** 035322

[20] Sabater C, Untiedt C, Palacios J J and Caturla M J 2012 *Phys. Rev. Lett.* **108** 205502

[21] Dednam W, Sabater C, Fernandez M A, Untiedt C, Palacios J J and Caturla M J 2015 *Journal of Physics: Conference Series* **574** 012045

[22] Plimpton S 1995 *J. Comp. Phys.* **117** 1

[23] Lammps: Large-scale atomic/molecular massively parallel simulator `http://lammps.sandia.gov` accessed: 2015-05-10

[24] Daw M and Baskes M 1983 *Phys. Rev. Lett.* **50** 1285

[25] Zhou X W *et al.* 2001 *Acta Materialia* **49** 4005

[26] Untiedt C, Rubio G, Vieira S and Agraït N 1997 *Phys. Rev. B* **56** 2154

[27] Dreher M, Pauly F, Heurich J, Cuevas J C, Scheer E and Nielaba P 2005 *Phys. Rev. B* **72** 075435

[28] Nosé S 1984 *Molecular Physics* **52** 255

[29] Hoover W G 1985 *Phys. Rev. A* **31** 1695

[30] Frisch M J *et al.* Gaussian 09 Revision a.01 Gaussian Inc. Wallingford Conneticut 2009

[31] Palacios J J *et al.* Alicante atomistic computation applied to nanotransport `http://http://dfa.ua.es/en/invest/condens/Alacant/index.html` accessed: 2015-05-26

[32] Pascual J I, Mendez J, Gomez-Herrero J, Baro A M and Garcia N 1993 *Phys. Rev. Lett.* **71** 1852

[33] Landman U, Luedtke W D, Burnham N A and Jr R C 1990 *Science* **248** 454

[34] Sabater C 2013 *Theoretical and experimental study of electronic transport and structure in atomic-sized contacts* Ph.D. thesis University of Alicante, Spain

Quasi-normal Modes for Spin-3/2 Fields

Gerhard Erwin Harmsen

National Institute for Theoretical Physics, School of Physics and Mandelstam Institute for Theoretical Physics, University of the Witwatersrand, Johannesburg, Wits 2050, South Africa

E-mail: gerhard.harmsen5@gmail.com

Abstract. With the Large Hadron Collider already able to produce collisions with an energy of 8 TeV, the formation of higher dimensional black holes may soon be possible. In order to determine if we are detecting these higher dimensional black holes we need to have a theoretical understanding of what the signatures of such black holes could be. As such we shall discuss quasi-normal modes (QNMs) for spin-3/2 fields as they travel through a black hole background. We will begin by studying possible QNMs for N-dimensional Schwarzschild black holes, we will then calculate allowed QNMs for N-dimensional Reissner-Nordström black holes. We will use the Wentzel-Kramers-Brillouin approximation to determine the QNMs for the two types of black holes described above.

1. Introduction

There are currently many ways to indirectly detect black holes [1]. One possible direct way of detecting a black hole is through gravitational wave detection. These are oscillations in the space time and are emitted by black holes when they are perturbed through either black hole-black hole collisions, stellar collapse or other matter interactions. These perturbations cause non-radial oscillations on the surface of a black hole and are called quasi-normal modes (QNMs). These QNMs in turn effect the space-time around a black hole and result in quasi-normal frequencies being emitted from black holes, these frequencies would be detected as gravitational waves, or particles [2]. What is interesting is that the allowed quasi-normal frequencies of a black hole are directly related to the characteristics of the black holes. For instance, the allowed quasi-normal frequencies of a Schwarzschild black hole are given by the mass of the black hole, but for a Kerr-Newman black hole they are determined by the angular momentum, charge and mass of the black hole [2]. This means that we could not only directly detect black holes by detecting their gravitational emission, but we could also get an idea of the parameters of the black hole by using this. A more elaborate discussion on detecting these QNMs is given in reference [3]. In recent years the study of higher dimensional general relativity, and by extension the study of higher dimensional black holes, has become a subject of increasing interest. One of the reasons for this interest is that string theory requires more than four dimensions in order to properly account for gravity, in fact in string theory a 5-dimensional black hole is required to correctly account for the entropy of a black hole [4]. Another reason for the study of higher dimensional black holes is that Anti de-Sitter/Conformal field theory (AdS/CFT) correspondence tells us that the dynamics of an N-dimensional black hole correspond to those of an $N-1$ dimensional quantum field theory [5]. A more elaborate discussion on the reasons for studying N-dimensional black holes is given in reference [6]. In order to calculate the allowed QNMs for the various types

of black holes we will use methods such as the Wentzel-Kramers-Brillouin (WKB) method [7] and the improved asymptotic iteration method (AIM) [8]. We shall first give a brief overview of black holes and how to mathematically construct N-dimensional black holes. We then give a brief discussion of QNMs and then give an overview of the WKB method and improved AIM. We will provide a brief overview of what we have done so far and give examples of possible future work.

2. Black Holes

Black holes have been described as the hydrogen atom of general relativity, since they are relatively simple general relativistic objects which display all of the properties predicted by Einstein's field equations. Hence a better understanding of black holes would allow us to better understand general relativity, just as the hydrogen atom can help in our understanding of quantum mechanics. In this paper we will set c = 1.

2.1. Schwarzchild Black holes

The Schwarzschild metric can be used to describe the space time surrounding any non-rotating objects in a vacuum [9]. It is given as:

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dx^2 + r^2 \left(d\theta^2 + \sin(\theta) d\phi^2\right), \tag{1}$$

where $G = 7.426 \times 10^{-28} m/kg$, the mass ($M$) denotes the mass of the object and r denotes the radial distance between a point in the space time and the center of that object. This metric is valid for all values of r greater than zero. We can describe the space time surrounding a black hole by looking at systems for which the value of $2GM$ is large compared to r. When r approaches $2GM$ our metric will produce a singularity, where this singularity is in fact a "coordinate singularity"; which means that there exists a problem with the coordinate system we have chosen rather than there being any actual physical singularity at that point in space-time [9]. The point $r = 2GM$ is in fact the radial location of the event horizon of the black hole. These types of black holes are the simplest types of black holes and provide the simplest case for studying QNMs.

2.2. Reissner-Nordström black holes

Reissner-Nordström black holes are a more general case of the Schwarszchild black hole, that is, they are non-rotating electrically charged black holes. Hence the metric is similar to that of the Schwarszchild black hole, and is given as follows [10]:

$$ds^2 = -\left(1 - \frac{2GM}{r} + \frac{Q^2}{r^2}\right) dt^2 + \left(1 - \frac{2GM}{r} + \frac{Q^2}{r^2}\right)^{-1} dr^2 + r^2 \left(d\theta^2 + \sin(\theta) d\phi^2\right), \tag{2}$$

where G, M and r represent the same quantities as they did for the Schwarzchild metric. Q represents the electric charge of the black hole.

2.3. N-dimensional metrics

In the above metrics we have temporal, radial and spherical components, denoted by dt^2, dr^2 and $d\theta^2$ and $d\phi^2$ respectively. The spherical components of these 4-dimensional metrics are given by the 2-sphere metric. We can rewrite the metrics as follows:

$$ds^2 = -f(r) dt^2 + \frac{1}{f(r)} dr^2 + r^2 d\Omega_2^2, \tag{3}$$

where $f(r)$ denotes the radial functions associated to the temporal and radial components, and $d\Omega_2^2$ is the metric for the 2-sphere. For N-dimensional black holes we can rewrite the metric as follows

$$ds_N^2 = -f(r)dt^2 + f(r)^{-1}dr^2 + r^2 d\Omega_{N-2}^2, \tag{4}$$

where $f(r) = 1 - 2GM/(r)^{N-3}$, for the case of an N-dimensional Schwazrchild black hole, or $f(r) = 1 - 2GM/(r)^{N-3} + Q^2/(r)^{2N-6}$, for the case of an N-dimensional Reissner-Nordström black hole, and where $d\Omega_N^2 = d\theta^2 + \sin(\theta)^2 d\Omega_{N-1}$ with $d\Omega_{N-1}$ the metric for the $N-1$-sphere [11, 12, 13].

3. Quasi-Normal Modes

As stated before, QNMs are the reaction to a black hole being perturbed in some way. A simple explanation of what a QNM is can be described by the tapping of a wine glass. If we tap the wine glass we are in fact perturbing the glass. This perturbation causes the glass to resonate at specific frequencies. The allowed frequencies are determined by the composition of the glass, and even the contents within the glass. If the glass continued to resonate forever then we would be generating a normal mode. Since the glass does stop resonating we say it is a QNM. So when studying QNMs we are studying standing waves with some damping term attached. If we consider our wine glass example we should note that wine glasses with different amounts of water in them ring at different frequencies, so we could, in theory, determine the amount of water in a glass by the sound it makes when we tap the glass. A similar idea can be used for the emitted quasi-normal frequencies of a black hole. Although with a black hole energy is being radiated, not as sound, but rather as disruptions in the space time which we call gravitational waves.

3.1. A Mathematical description of QNMs

In order to develop a more mathematical understanding of QNMs we can look at the mathematical formula for standing waves since, as stated above, QNMs are merely damped standing waves. So we can represent these QNMs as follows [14]:

$$\frac{d^2}{dx^2}\Psi - \frac{d^2}{dt^2}\Psi - V(x)\Psi = 0. \tag{5}$$

With Ψ describing some wavelike function where x and t denote space and time coordinates respectively. $V(x)$ is some x-dependent potential, the 4-dimensional potential is presented in Eq.(16). We can solve the equation as we would a standing wave problem, and so we can assume the following time dependency [3]:

$$\Psi(x,t) = e^{-i\omega t}\phi(x), \tag{6}$$

where ω is the frequency of the wave and is proportional to the energy eigenvalue. Plugging this into our equation for a QNM we get:

$$\frac{d^2\phi(x)}{dx^2} - (\omega^2 + V(x))\phi(x) = 0. \tag{7}$$

This is the general form for the QNMs we will be studying in this paper. In order to solve this equation we require specific boundary terms, similar to those required for solving the standing wave equation. The boundary conditions that we will impose are:

$$\begin{aligned} &V \to 0; x \to \infty, \\ &V \to 0; x \to -\infty. \end{aligned} \tag{8}$$

This means that particles fall into the black hole at the horizon, and those which move to infinity are no longer influenced by the black hole. In order to determine if a frequency of a field is in fact a QNM we need to ensure that it obeys these requirements. In this project we will use both the WKB method, to 6th order [7], and the improved AIM [8] to determine the solutions to our QNM equation.

4. Approximation Methods

4.1. WKB Approximation

The WKB method is a well known approximation, and is usually used in quantum mechanics to determine the solutions of Schrödinger equations in systems with non-constant potentials. The WKB method can be used to determine the approximate solution to second order differential equations. A brief example of how the method is used to solve second order differential equations is given below. We can use the WKB approximation to solve problems of the following form [15]:

$$\epsilon^2 \frac{d^2y}{dx^2} = Q(x)y, \tag{9}$$

where $\epsilon \ll 1$ and positive. Using the WKB method we get a solution of the following form,

$$\begin{aligned} &\epsilon^2 \left(\left(A_0''(x) + \epsilon A_1''(x) + ...\right) + 2\left(A_0'(x) + \epsilon A_1'(x) + ...\right) \frac{iu'(x)}{\epsilon} \right) + \\ &\epsilon^2 \left[\left(A_0(x) + \epsilon A_1(x) + ...\right) \left(\frac{iu''(x)}{\epsilon} - \left(\frac{u'(x)}{\epsilon}\right)^2 - Q(x)\left(A_0(x) + \epsilon A_1(x) + ...\right) \right) \right] = 0. \end{aligned} \tag{10}$$

Where $A_0(x) + \epsilon A_1(x) + \epsilon^2 A_2(x) + ...$ is a series expansion of $A(x, \epsilon)$ with primes denoting derivatives with respect to x. [16]. We obtain solutions for $A_0(x), A_1(x), A_2(x), ...$ by grouping the term by powers of ϵ.

We can use the WKB method to study black hole perturbations since these equations reduce to the form of a wave equation, similar to the example above [3]. This equation, with the known potential, is then solved using the the WKB approximation. The cases for Schwarzschild black holes and Reissner-Nordström black holes has been studied extensively [8], however these studies have not considered the case of spin-3/2 fields.

4.2. Improved AIM

The theory of this method is given in reference [17]. The theorem of this method is given below: Given λ_0 and s_0 in $C_\infty(a, b)$, then the differential

$$y'' = \lambda_0(x)y' + s_0(x)y, \tag{11}$$

has the general solution of:

$$y(x) = exp\left(-\int^x \alpha dt\right) \left[C_2 + C_1 \int^x exp\left(\int^t \left(\lambda_0(\tau) + 2\alpha(\tau)\right) d\tau\right)\right], \tag{12}$$

for some $n > 0$.

$$\frac{s_n}{\lambda_n} = \frac{s_{n-1}}{\lambda_{n-1}} \equiv \alpha, \tag{13}$$

where $\lambda_k = \lambda'_{k-1} + s_{k-1} + \lambda_0\lambda_{k-1}$ and $s_k = s'_{k-1} + s_0\lambda_{k-1}$ for $k = 1, 2, ..., n$. We again solve for the QNMs by first obtaining the potential energy term for a spin-3/2 particle in a black hole background, and then use the AIM approximation to obtain the allowed QNMs.

5. Rarita-Schwinger fields in 4D Schwarzschild background

For this case we have only considered massless Rarita-Schwinger fields,spin-3/2 particles, where to determine how these fields propagate through space time we used the Rarita-Schwinger equation, given as [18],

$$\gamma^{\mu\nu\alpha}\nabla_\nu\Psi_\alpha = 0, \tag{14}$$

where,

$$\gamma^{\mu\nu\alpha} = \gamma^\mu\gamma^\nu\gamma^\alpha - \gamma^\mu g^{\nu\alpha} + \gamma^\nu g^{\mu\alpha} - \gamma^\alpha g^{\mu\nu}. \tag{15}$$

Using this equation we can determine the equations of motion for a massless Rarita-Schwinger field. Which in turn gives us the potential function for these particles given as [18]

$$V_{1,2} = \pm f(r)\frac{dW}{dr} + W^2, \tag{16}$$

with,

$$W = \frac{\left(j-\frac{1}{2}\right)\left(j+\frac{1}{2}\right)\left(j+\frac{3}{2}\right)\sqrt{f(r)}}{r\left(\left(j+\frac{1}{2}\right)^2 - f(r)\right)}, \tag{17}$$

$f(r) = ((r-2M)/r)$ and $j = 3/2, 5/2, 7/2, ...$

In Table 1 we present our results for the 4-dimensional Schwarzschild black hole, using the 3rd order WKB approximation as given in reference [19], and 6th order as given in reference [7]. We compare this to results obtain using the AIM method as given in reference [17]. The QNMs calculated agree across the methods, within their numerical uncertainties, to those obtained in reference [20].

Table 1. Low-lying ($n \leq l$, with $l = j - 3/2$) gravitino quasinormal mode frequencies using the WKB and the AIM methods [18].

		WKB		AIM
l	n	3rd Order	6th Order	150 iterations
0	0	0.3087 - 0.0902i	0.3113 - 0.0902i	0.3108 - 0.0899i
1	0	0.5295 - 0.0938i	0.5300 - 0.0938i	0.5301 - 0.0937i
1	1	0.5103 - 0.2858i	0.5114 - 0.2854i	0.5119 - 0.2863i
2	0	0.7346 - 0.0949i	0.7348 - 0.0949i	0.7348 - 0.0949i
2	1	0.7206 - 0.2870i	0.7210 - 0.2869i	0.7211 - 0.2871i
2	2	0.6960 - 0.4844i	0.6953 - 0.4855i	0.6892 - 0.4834i
3	0	0.9343 - 0.0954i	0.9344 - 0.0954i	0.9344 - 0.0954i
3	1	0.9233 - 0.2876i	0.9235 - 0.2876i	0.9235 - 0.2876i
3	2	0.9031 - 0.4835i	0.9026 - 0.4840i	0.9026 - 0.4840i
3	3	0.8759 - 0.6835i	0.8733 - 0.6870i	0.8733 - 0.6870i
4	0	1.1315 - 0.0956i	1.1315 - 0.0956i	1.1315 - 0.0956i
4	1	1.1224 - 0.2879i	1.1225 - 0.2879i	1.1225 - 0.2879i
4	2	1.1053 - 0.4828i	1.1050 - 0.4831i	1.1050 - 0.4831i
4	3	1.0817 - 0.6812i	1.0798 - 0.6830i	1.0798 - 0.6830i
4	4	1.0530 - 0.8828i	1.0485 - 0.8891i	1.0485 - 0.8891i

6. Concluding remarks

QNMs are very sensitive to the initial conditions which created them, therefore is it necessary to calculate them using very precise methods. We have done this by using the WKB method to 6th order and taking 150 iteration of the AIM.

As stated previously the real part of our QNMs represents their frequencies, which is proportional to their energies, and the imaginary parts represent how damped our QNMs are. In Table 1 we can clearly see that by increasing the angular quantum number, l, we increase the frequency, and therefore energy, of our emitted QNMs. Note also that an increase in the mode number, n, results in a decrease of the frequency. We also see that an increase in l results in an increase in the strength of the damping associated to the QNMs. This means that lower n modes are more unlikely to be detected compared to the higher n modes, with the emission spectra being comprised predominately of our more energetic QNMs, with smaller dampening terms.

Since we have shown that our method can reproduce the values for the 4-dimensional case, we are working on a solution for the N-dimensional Schwarzschild metric and will then investigate the allowed QNMs for Rarita-Schwinger fields in a Reissner-Nordstöm black hole background, both in four dimensions and the N-dimensional case. We would then like to investigate black holes in AdS space times, as then we could begin investigating the connection between N-dimensional black holes and the dynamics of $N-1$ CFT, where a full review of this is given in reference [5].

Acknowledgments

I would like to thank Prof. Alan Cornell and Prof. Hing-Tong Cho for imparting some of their knowledge to me on this topic. I would also like to acknowledge the NRF and NITheP for funding me and this research.

References

[1] Novello M, Visser M and Volovik G E 2002 *Artificial black holes* vol 5 (World Scientific)
[2] Leaver E W 1985 *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **402** 285–298
[3] Kokkotas K D and Schmidt B G 1999 *Living Rev. Rel* **2** 262
[4] Strominger A and Vafa C 1996 *Physics Letters B* **379** 99–104
[5] Aharony O, Gubser S S, Maldacena J, Ooguri H and Oz Y 2000 *Physics Reports* **323** 183–386
[6] Emparan R and Reall H S 2008 *Living Rev. Rel* **11** 0801–3471
[7] Konoplya R 2003 *Physical Review D* **68** 024018
[8] Cho H, Cornell A, Doukas J and Naylor W 2010 *Classical and Quantum Gravity* **27** 155004
[9] MooreThomas 2013 *A General Relativity Workbook* (University Science Books)
[10] Gao C J and Zhang S N 2004 *Physics Letters B* **595** 28–35
[11] Ishibashi A and Kodama H 2003 *Progress of theoretical physics* **110** 901–919
[12] Konoplya R and Zhidenko A 2014 *Physical Review D* **89** 024011
[13] Camporesi R and Higuchi A 1996 *Journal of Geometry and Physics* **20** 1–18
[14] Vishveshwara C 1970
[15] Griffiths D J and Harris E G 1995 *Introduction to quantum mechanics* vol 2 (Prentice Hall New Jersey)
[16] 2015 *Quasi-normal modes for Spin-3/2 Fields*
[17] Ciftci H, Hall R L and Saad N 2003 *Journal of Physics A: Mathematical and General* **36** 11807
[18] Chen C H, Cho H, Cornell A, Harmsen G and Naylor W 2015 (*Preprint* `1504.02579`)
[19] Iyer S 1987 *Physical Review D* **35** 3632
[20] Piedra O P F 2011 *Int.J.Mod.Phys.* **D20** 93–109 (*Preprint* `1006.3327`)

Thomas Rotation and Quantum Entanglement.

JM Hartman[1], SH Connell[1] and F Petruccione[2]

1. University of Johannesburg, Johannesburg, South Africa
2. University of Kwa-Zulu Natal, South Africa

E-mail: hartman.jonathan2@gmail.com

Abstract. The composition of two non-linear boosts on a particle in Minkowski space-time are not commutative. This non-commutativity has the result that the Lorentz transformation formed from the composition is not a pure boost but rather, a combination of a boost and a rotation. The rotation in this Lorentz transformation is called the Wigner rotation. When there are changes in velocity, as in an acceleration, then the Wigner rotations due to these changes add up to form Thomas precession. In curved space-time, the Thomas precession combines with a geometric effect caused by the gravitationally curved space-time to produce a geodetic effect. In this work we present how the Thomas precession affects the correlation between the spins of entangled particles and propose a way to detect forces acting on entangled particles by looking at how the Thomas precession degrades the entanglement correlation. Since the Thomas precession is a purely kinematical effect, it could potentially be used to detect any kind of force, including gravity (in the Newtonian or weak field limit). We present the results that we have so far.

1. Introduction

While Bell's theorem is well known and was originally presented in John Bell's 1964 paper as a way to empirically test the EPR paradox, this was only for non-relativistic quantum mechanics. It was only fairly recently that people have started investigating possible relativistic effects on entanglement and EPR correlations beginning with Czachor [1] in 1997. Czacor found that relativistic effects of massive particles effect the correlation between entangled massive particles and depends on the velocity of the particles in the simple case where both particles are moving in the same direction. The origin of this apparent deviation from the classical violation of Bell's inequality appears to be a length contraction when measuring the spin correlations in the laboratory frame.

Later authors then also started revisiting this issue. In 2003, Terashima and Ueda [2] showed that an observer moving perpendicular to the motion of the entangled particles with opposite spins, moving in opposite directions, will observe a degradation of the entanglement correlation when measured in the usual directions as opposed to an observer that's in the frame of the source. However, they concluded that the entanglement and quantum information is still preserved using the reasoning that one can still get the maximal violation of Bell's inequality if the spins are measured in compensating directions, so the situation is still non-local. The apparent decoherence effect found by Terashima and Ueda is due to a Wigner rotation as measured in the moving frame. It is interesting that they also calculated such an effect for massless entangled particles moving at the speed light c. In 2004, Lee and Chang-Young [3] calculated the general

case, combining Czachor's result with that of Terashima and Ueda and derived the directions in which maximal violation of Bell's inequality occurs in vector form.

Some authors also investigated quantum communication in accelerated frames [4] [5]. However, these authors didn't investigate quantum correlations but rather specific forms of quantum communication such as teleportation. There have also even been recent works investigating how curved space-time affects quantum correlations [6] [7].

2. Maximal violation of Bell's inequality of entangled massive relativistic particles. In order to understand the Wigner rotation and how it affects the correlations between entangled relativistic particles, let's consider the setup as described by Terashima and Ueda [2]. In their setup, they considered 2 particles moving in opposite directions from a source in the x and $-x$ directions respectively and the 2 observers, Alice and Bob, moving at the same velocity, perpendicular to the direction of motion of the 2 particles. A Wigner rotation is a spacial rotation in the coordinates between 2 reference frames when the transformation to the new frame is the result of a succession of 2 non-linear Lorentz boosts. In this case, the first boost is a Lorentz boost from the rest frame of each particle respectively to the rest frame of source, where the particle is moving at a constant velocity $\vec{v}$. The second boost, is a boost from that frame, perpendicular to the motion of the particles, to get to the observer frame of Alice and Bob. Terashima and Ueda found that if, in the rest frame of the source, the 2 particles are in the singlet state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(\left|\vec{p}_+, \frac{1}{2}\right\rangle\left|\vec{p}_-, -\frac{1}{2}\right\rangle - \left|\vec{p}_+, -\frac{1}{2}\right\rangle\left|\vec{p}_-, \frac{1}{2}\right\rangle\right), \tag{1}$$

then applying the unitary operator $U(\Lambda)$ corresponding to the second boost $\Lambda^{\mu}{}_{\nu}$, they get

$$\begin{aligned} U(\Lambda)\,|\Psi\rangle &= \frac{1}{\sqrt{2}}\left[\cos\delta\left(\left|\vec{\Lambda p}_+, \frac{1}{2}\right\rangle\left|\vec{\Lambda p}_-, -\frac{1}{2}\right\rangle - \left|\vec{\Lambda p}_+, -\frac{1}{2}\right\rangle\left|\vec{\Lambda p}_-, \frac{1}{2}\right\rangle\right)\right. \\ &+ \left.\sin\delta\left(\left|\vec{\Lambda p}_+, \frac{1}{2}\right\rangle\left|\vec{\Lambda p}_-, -\frac{1}{2}\right\rangle + \left|\vec{\Lambda p}_+, -\frac{1}{2}\right\rangle\left|\vec{\Lambda p}_-, \frac{1}{2}\right\rangle\right)\right], \end{aligned} \tag{2}$$

where Λp represents the momenta as measured by Alice and Bob and δ is the Wigner angle given by

$$\tan\delta = \frac{\sinh\xi\sinh\chi}{\cosh\xi + \cosh\chi}, \tag{3}$$

where ξ is the rapidity of the rapidity of the particles in the rest frame of the source, whose velocity is given by $\frac{v}{c} = \tanh\xi$ and χ is the rapidity of the observer frame (from which Alice and Bob measure the 2 particles) with respect to the rest frame of the source, with it's relative velocity given similarly as $\frac{V}{c} = \tanh\chi$. In order to find out how this affects Bell's theorem, Terashima and Ueda made use of the CHSH observable

$$C(a, a', b, b') \equiv \left\langle \hat{a}\otimes\hat{b}\right\rangle + \left\langle \hat{a}\otimes\hat{b}'\right\rangle + \left\langle \hat{a}'\otimes\hat{b}\right\rangle - \left\langle \hat{a}'\otimes\hat{b}'\right\rangle, \tag{4}$$

where the CHSH-inequality is given classically by $C(a, a', b, b') \leqslant 2$ but quantum mechanics predicts this observable to have a maximum value of $C(a, a', b, b') = 2\sqrt{2}$. Terashima and Ueda showed that

$$C(a, a', b, b') = 2\sqrt{2}\cos^2\delta, \tag{5}$$

in their setup if the vector set is chosen such that the observable would have the maximum value $2\sqrt{2}$ in the non-relativistic case and thus the inequality may not be violated depending on the relative velocities between the 2 frames and the boosts. They did reason, however, that because the difference is by an angle, a different vector set can be chosen so that the maximum violation of Bell's inequality can still be recovered even in their setup as one merely has to rotate the original vectors though the Wigner angle δ.

Terashima and Ueda however, did not define a relativistic spin observable. One of the very first papers on the EPR paradox for relativistic particles did though. This is what Czachor did in a paper published in 1997 [1]. He considered an even simpler setup to Terashima and Ueda though, where both entangled particles were moving in the same direction. Using his relativistic spin observable, he showed some kind of degradation in the correlation even occurs not only when the particles are moving parallel to each other but also when they're moving in the same direction. In this case though, the difference is due to Lorentz contraction in the direction of motion and is not a Wigner angle. Czachor concluded that when the entangled particles are relativistic, then the Bell inequalities may not be violated. However, later on Lee and Chang-Young showed that in this case too, the maximal violation can again be recovered, simply by choosing a different vector set [3]. They also added this difference as a correction on Terashima and Ueda's result to get a more complicated formula

$$\begin{aligned} C(a,a',b,b') &= \frac{2}{\sqrt{1+\sin^2\theta_\Lambda+\cosh^2\eta\cos^2\theta_\Lambda}} \\ &+ \frac{2}{\sqrt{1+\sin^2\theta_\Lambda+\cosh^2\eta\cos^2\theta_\Lambda}\sqrt{\sin^2\theta_\Lambda+\cosh^2\eta\cos^2\theta_\Lambda}} \\ &\times \{[(\cosh\eta\cos^2\theta_\Lambda-\sin^2\theta_\Lambda)^2-(1+\cosh\eta)^2\sin^2\theta_\Lambda\cos^2\theta_\Lambda]\cos 2\delta \\ &- (1+\cosh\eta)(\cosh\eta\cos^2\theta_\Lambda-\sin^2\theta_\Lambda)\sin 2\theta_\Lambda\sin 2\delta\}, \end{aligned} \tag{6}$$

where the angle θ_Λ is the rotation to the direction of the momentum $\vec{\Lambda p}$ and $\tanh\eta = \frac{\sqrt{(\tanh^2\xi+\sinh^2\chi)}}{\cosh\chi}$.

3. Corrected Bell observables for massive relativistic particles

In the setup considered by Terashima and Ueda, one only has to rotate the vectors in the vector set that give the maximal violation of Bell's inequality by the Wigner angle δ in order to get the corrected vector set that regains the maximal violation. Lee and Chang-Young, however, derived the following components of the corrected vectors $\vec{a}_c$ and $\vec{b}_c$ as used in the joint-spin measurement in their paper,

$$a_{cz} = \frac{\bar{a}_z}{\sqrt{[F_a(1+\cosh\eta)\sin\theta_\Lambda\cos\theta_\Lambda-(\sin^2\theta_\Lambda-\cosh\eta\cos^2\theta_\Lambda)]^2-\bar{a}_z^2\sinh^2\eta(F_a\sin\theta_\Lambda+\cos\theta_\Lambda)^2}}, \tag{7}$$

$$a_{cx} = \bar{a}_x\sqrt{1+a_{cz}^2\sinh^2\eta(F_a\sin\theta_\Lambda+\cos\theta_\Lambda)^2}, \tag{8}$$

$$a_{cy} = a_y\sqrt{1+a_{cz}^2\sinh^2\eta(F_a\sin\theta_\Lambda+\cos\theta_\Lambda)^2}, \tag{9}$$

$$b_{cz} = \frac{\bar{b}_z}{\sqrt{[F_b(1+\cosh\eta)\sin\theta_\Lambda\cos\theta_\Lambda - (\sin^2\theta_\Lambda - \cosh\eta\cos^2\theta_\Lambda)]^2 - \bar{b}_z^2\sinh^2\eta(F_b\sin\theta_\Lambda - \cos\theta_\Lambda)^2}}, \tag{10}$$

$$b_{cx} = b_x\sqrt{1 + b_{cz}^2\sinh^2\eta(F_b\sin\theta_\Lambda - \cos\theta_\Lambda)^2}, \tag{11}$$

and

$$b_{cy} = b_y\sqrt{1 + b_{cz}^2\sinh^2\eta(F_b\sin\theta_\Lambda - \cos\theta_\Lambda)^2}, \tag{12}$$

where

$$F_a = \frac{(1+\cosh\eta)\tan\theta_\Lambda - f_a(\tan^2\theta_\Lambda - \cosh\eta)}{(1+\cosh\eta\tan^2\theta_\Lambda) - f_a(1+\cosh\eta)\tan\theta_\Lambda}, \tag{13}$$

and

$$F_b = -\frac{(1+\cosh\eta)\tan\theta_\Lambda + f_b(\tan^2\theta_\Lambda - \cosh\eta)}{(1-\cosh\eta\tan^2\theta_\Lambda) + f_b(1+\cosh\eta)\tan\theta_\Lambda}, \tag{14}$$

where $f_a \equiv \frac{\bar{a}_x}{\bar{a}_z}$ and $f_b \equiv \frac{\bar{b}_x}{\bar{b}_z}$.

4. Addition of accelerations to the Wigner angle and Thomas precession

As shown by Czachor, the vector basis required for the maximal violation of Bell inequalities does change even when both entangled particles are moving in the same direction. However, this difference can be made to be zero if the spin-directions are measured perpendicularly to the direction of motion. So, for now at least, in our description of including acceleration, we will simply impose an acceleration on Terashima and Ueda's setup. An acceleration can be thought of as successive additions of small velocities. So in Terashima and Ueda's setup, let's consider the particles to be accelerating in their direction of motion. This acceleration could then be thought of as many additions of velocities. So using the same symbols as in previous sections, if we take $\vec{v}$ and $-\vec{v}$ as the initial velocities of the particles in opposite directions, as measured in the rest frame of the source and assume that they're also accelerating in opposite directions then this can be expressed as

$$\frac{v \oplus \Delta v}{c} = \tanh(\xi + \Delta\xi), \tag{15}$$

where $\oplus$ denotes the addition of velocities using the relativistic addition of velocities formula and the value for the other particle is just minus the above formula. The total change in velocity Δv is given by $\Delta v = \bigoplus_i^N \Delta a_i \Delta t$ where each $\frac{\Delta a_i \Delta t}{c} = \tanh(\Delta a_{Pi}\Delta\tau)$. Here Δa_i are the relativistic accelerations and Δa_{Pi} are the proper accelerations respectively corresponding to the relativistic time Δt as measured in the rest frame of the source and the proper time $\Delta\tau$ as measured in the rest frame of the particles. This gives $\frac{\Delta v}{c} = \tanh\left(\sum_{i=1}^N \Delta a_{Pi}\Delta\tau\right)$ and $\Delta\xi = \sum_{i=1}^N \Delta a_{Pi}\Delta\tau$. As $N \to \infty$, $\Delta\tau \to 0$, therefore $\lim_{N\to\infty}\Delta\xi = \int_0^\tau a_P(\tau)d\tau$.

The reason for choosing the acceleration as being in the same or opposite to the direction of motion of the particles is because the addition of all these infinitesimal velocities, and thus the acceleration, would not change their direction of motion as measured in the rest frame of the source, which is why equation (15) is valid. Thus this is a simple case. However, when applying the second boost to the observer frame from where Alice and Bob do their measurement of the

particles, the acceleration does change the direction of motion as measured in this frame, where the change in Wigner angle is given by

$$\tan(\delta_0 + \Delta\delta) = \frac{\sinh(\xi + \Delta\xi)\sinh\chi}{\cosh(\xi + \Delta\xi) + \cosh\chi}. \tag{16}$$

Each velocity gained or lost over every infinitesimal unit of time adds a little bit or takes away a little bit from the Wigner angle such that a precession is observed. This is the Thomas precession and is also responsible for the spin-orbit coupling of an electron in a hydrogen atom. However, it has only recently been realised, since the first paper on the subject in the '90's, that Thomas precession also affects correlations between entangled relativistic particles. It was probably known known before but it was since then that the subject started to be studied in more detail. Others have also studied quantum communication in accelerated frames but those papers were studying the effects on specific phenomena like quantum teleportation, not the pure effects of acceleration on it's own. Our motivation for doing this is the possibility of using such differences to measure forces between particles and investigating whether such an approach would be more sensitive than measuring them directly, such as looking for a loss in energy. If it were indeed a more sensitive way to measure and observe forces, then it may be possible to exploit the entanglement between particles in order to detect forces that would otherwise be too small to measure.

One example that we did think of is the possibility of directly measuring the gravitational force on the femto-metre scale in the Newtonian approximation. We did this by developing a very simple model of how a Newtonian gravitational force of a decaying heavy particle would affect the daughter particles. Due to the weakness of gravity, one might think that the gravitational force of elementary particles would be extremely tiny. We found however, that on the scales we're looking, the gravitational acceleration of the elementary particles are surprisingly large and this is due to the inverse square law. Unfortunately, however, changes in velocity due to the gravitational force are still extremely tiny (on the order of about 10^{-25}) and though the entanglement correlations were found to be more sensitive than measuring such changes directly in some cases, this higher sensitivity may still not be enough to detect something as weak as that. This is an extreme case though and we can still imagine some more practical applications we weak forces may be able to be detected. We've only so far looked at Bell observables and not tried applications of optimal entanglement witnesses which also get used by the method of quantum state tomography, which are likely to be an approach which is more sensitive anyway. We will not give the full derivation of this gravitational model here as it is being prepared for publishing in a larger journal. Let's instead consider the much simpler case of a uniform acceleration. In this case $\frac{\Delta v}{c} = \tanh(a_P\tau)$ and therefore, equation (16) will be given as

$$\tan(\delta_0 + \Delta\delta) = \frac{\sinh(\xi + a_P\tau)\sinh\chi}{\cosh(\xi + a_P\tau) + \cosh\chi}, \tag{17}$$

where a_P is the proper acceleration and τ is the total proper time. As you can see, if you can measure the total change in Wigner rotation $\Delta\delta$, then in principle you can calculate the proper acceleration a_P and therefore measure the forces acting on the particles. The same is true of non-uniform accelerations as in the case of the gravitational force that we spoke of above. In principle, all one has to do to measure the difference in the Wigner rotation is to use the Bell observable, just measure the angle by which one has to rotate the detector in order to get the maximal violation, so this difference is directly measurable in principle.

5. Conclusion

In this paper we have given a neat summary of previous work on quantum entanglement between relativistic particles and their correlations. Czachor found that Lorentz contraction affects the

angle of the spins of the 2 entangled particles as measured by moving observers even when the particles are moving in the same direction. Terashima and Ueda found that the Wigner angle resulting from a succession of 2 boosts gives rise to a measurable difference in entanglement correlations. Lee and Chang-Young combined the results of these previous papers and gave their own derivation of the corrected vector set required to recover the maximal violation of Bell's inequalities. Finally we added our own input by adding an acceleration to Terashima and Ueda's setup and discussed some possible applications of this. We conclude that it may be possible to use these differences in entanglement correlations to detect forces between particles but this is not certain yet. More investigation with regards to this point is required.

References

[1] Czachor M 1997 *Phys. Rev. A* **55**(1) 72–77 URL http://link.aps.org/doi/10.1103/PhysRevA.55.72
[2] Terashima H and Ueda M 2003 *Quantum Information & Computation* **3** 224–228
[3] Lee D and Chang-Young E 2004 *New Journal of Physics* **6** 67 URL http://stacks.iop.org/1367-2630/6/i=1/a=067
[4] Van Enk S J and Rudolph T 2003 *Quantum Info. Comput.* **3** 423–430 ISSN 1533-7146 URL http://dl.acm.org/citation.cfm?id=2011544.2011547
[5] Alsing P M and Milburn G J 2003 *Phys. Rev. Lett.* **91**(18) 180404 URL http://link.aps.org/doi/10.1103/PhysRevLett.91.180404
[6] Terashima H and Ueda M 2004 *Phys. Rev. A* **69**(3) 032113 URL http://link.aps.org/doi/10.1103/PhysRevA.69.032113
[7] Ralph T C, Milburn G J and Downes T 2009 *Phys. Rev. A* **79**(2) 022121 URL http://link.aps.org/doi/10.1103/PhysRevA.79.022121

Security of quantum key distribution

Mhlambululi Mafu

Department of Physics and Astronomy, Botswana International University of Science and Technology, P/Bag 16, Palapye, Botswana

E-mail: mafum@biust.ac.bw

Abstract. Quantum cryptography forms one of the most mature fields of information theory. The goal of quantum cryptography is to create a secret key between authorized parties. In this work, we explain the role played by quantum mechanics in cryptographic tasks and also investigate how secure is quantum cryptography. More importantly, we show by a simple security proof that for any state sent by the sender, the eavesdropper can only guess the output state with a probability that will allow her not to learn more than half of the classical Shannon information shared between the legitimate parties. This implies that with high probability, the shared key is secure.

1. Introduction

Quantum key distribution (QKD), one aspect of quantum cryptography, provides a secure method for distributing cryptographic keys between two parties conventionally known as Alice (sender) and Bob (receiver), who are connected by a quantum channel and an authenticated classical channel in the presence of an extremely competent malicious party, an eavesdropper, Eve [1]. The security of a QKD protocol is mainly based on the laws of quantum mechanics, which state that (i) one cannot make a measurement without perturbing the system unless the quantum state is compatible with the measurement. If there is no disturbance in the system, then no measurement was made, which implies that there was no eavesdropping. Therefore, Eve cannot intercept the information being transmitted in the communication channel without introducing disturbances that would reveal her presence; this is also known as quantum indeterminacy, (ii) it is impossible to duplicate an unknown quantum state with perfect fidelity. This means that Eve cannot intercept the channel and get hold of the quantum system, make a copy of the system and the copy to Bob without being detected. Therefore, quantum mechanics guarantees that the two parties can exchange a secret key securely because the key always remains uncompromised.

Based on Wiesner's idea of conjugate coding [2], Bennett and Brassard in 1984 proposed a first established and operable QKD protocol now commonly known as the BB84 protocol [3]. In 1991, Ekert [4] extended the idea by introducing quantum entanglement and the violation of Bell's theorem [5]. Since then, several protocols have been proposed by both theorists and experimentalists. These include: Bennett 1992 (B92) [6], six state [7]; Phoenix, Barnett and Chefles 2000 (PBC00) [8], the Scarani, Acín, Ribordy, Gisin 2004 (SARG04) protocol [9]. These protocols belong to a family called Discrete-Variable (DV) protocols. However, there exists another family of protocols called continuous-variable protocols and Distributed-Phase-Reference (DPR) protocols [10].

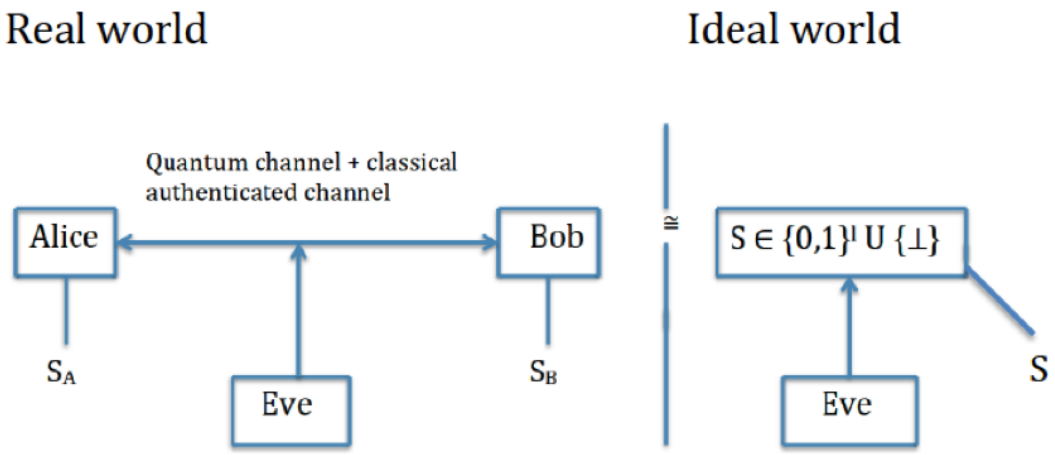


Figure 1. Comparison between what happens in a real and ideal quantum cryptographic world. Alice and Bob use the quantum and classical authenticated channel in the presence of Eve. At the end of the communication, in the real world Alice and Bob share two correlated secret keys S_A and S_B, respectively. In an ideal world, the access of Eve is broken; therefore Alice and Bob share a perfect secret key S.

The aim of this work is to present a simple security proof for a quantum protocol based on measurements on a maximally entangled state. In particular, we demonstrate how the laws of quantum mechanics afford security especially which properties are important in providing security for QKD protocols. This article is organized as follows. In section II we briefly describe the quantum communication procedure. In section III, we provide a short review of QKD security. In section IV, we give a description of the operation principle for our proposed entanglement-based protocol, which we are going to study. In this section we also outline the security requirements for QKD. Our main result is that the success guessing probability, p for the eavesdropper to guess the state sent by Alice or received by Bob will always result in Eve gaining less than half of the information being transmitted i.e., $H(p) = Pr[G = A|E] \leq 1/2$, where $H(p) = -p\log_2 p - (1-p)\log_2(1-p)$ is the classical Shannon information and G is the guess for output A (Alice) when given E (Eve). This means that the eavesdropper can only learn less of the transmitted information and this forbids her from trying to reconstruct the original message shared by the legitimate parties with high accuracy. This implies that the exchanged secret key is always secure. Lastly, section V is the conclusion.

2. Quantum communication procedure

Alice and Bob first use the quantum channel to distribute quantum states and then apply a quantum key distillation scheme to generate a common string of secret correlated data which are later transformed into a secret key. The eavesdropper can freely interact with the transmitted states while the two parties communicate and try to extract information. However, Eve can only perform the most general attack allowed by the laws of quantum mechanics. The quantum channel is used to transmit quantum signals while the classical channel is used to transmit classical information. The classical channel is authenticated so that Eve cannot learn the information that is being transmitted.

In a real world, at the end of the protocol, Alice outputs the key S_A while Bob outputs the key S_B. The output keys must be identical, but because of the presence of an eavesdropper and errors in the channel, the keys are almost identical. However, in the ideal world, Eve's access of

the key is detected and also there are no errors in the communication channel, therefore Alice and Bob generate a perfect secret key S which is of length l. This is shown in Figure 1. This perfect secret key is then used for sending private messages by means of the one-time pad.

3. Review of QKD security

In the last two decades, a lot of progress has been realized in the study of QKD security. Today, the unconditional security i.e., security guaranteed in an information-theoretical sense has been established for many protocols. The first unconditional security proof of QKD was proposed by Mayers in 1996 [11]. Since then, various techniques for proving the security of QKD protocols have been developed [10]. The security proofs generally depend on the construction of the protocol and also on its practical implementation. For example, the unconditional security proofs for the BB84 based protocols have long since been realized [12]. This is mainly because they share a common property of being symmetrical. On the side, the security proofs for the class of DPR protocols still remain unrealized [10, 13], mainly because their construction and encoding deviates from the usual symmetry that exist in BB84-type based protocols. Moreover, the previous security proofs could provide bounds only in the asymptotic limit of infinitely long keys, which is not realistic. But recently, the tools for studying QKD security in the finite-size limit have now become available [14]. This has been followed by various studies on security in the finite-size limit [14, 15, 16, 17, 18, 19, 20, 21]. In these papers, it was shown that the bits which are processed in QKD are indeed of finite length.

However, one of the greatest challenges that still remain in QKD implementations is a mismatch between the theoretical security proofs to real devices. This is because several assumptions are usually made when proving the security of QKD protocols. These assumptions are; devices do what they are supposed to do (according to a specified model) and not more, there should be access to perfect or almost perfect randomness (locally), there should be no side-channels and quantum theory is correct.

In order for a QKD protocol to be secure, it has to satisfy a number of security requirements. These requirements are [22];

a) correctness - a QKD protocol is called ε_{cor}-correct if, for any strategy by the eavesdropper $Pr[S_A \neq S_B] \leq \varepsilon_{\text{cor}}$, where S_A and S_B are Alice's and Bob's output classical keys, respectively.

b) secrecy - if $S \neq \perp$, then S is uniform $\{0,1\}^l$ and independent of Eve.

c) Robustness - a QKD protocol is said to be "robust" if it's guaranteed that it does not abort as long as the eavesdropper is inactive. When an eavesdropper is inactive, the protocol would continue to generate a secret key, otherwise if an adversary tampers with the quantum channel, the protocol recognises the attack and aborts the computation of the key.

d) Finally, a QKD is secure if it is correct and secret, that a protocol is ε-secure, if it is ε_{cor}-correct and ε_{sec} with $\varepsilon_{\text{cor}} + \varepsilon_{\text{sec}} \leq \varepsilon$.

4. Operation of our proposed QKD protocol

A source prepares and distributes a maximally entangled quantum state where one system is sent to Alice and another to Bob. This is shown in Figure 2. Alice and Bob then perform measurements in two mutually unbiased bases on their system respectively. In the absence of an eavesdropper, if they measure in the same basis they obtain perfectly correlated outcomes, which are completely random. The three parties will then share a quantum state $|\psi\rangle_{ABE}$. An example of this protocol is the E91 protocol [4].

If the authorized parties notice some errors in Bob's measurements, this implies that Eve has measured some of the photon polarizations. Therefore, QKD is secure because either of the following happens; if the error rate observed by Alice and Bob is lower than a critical

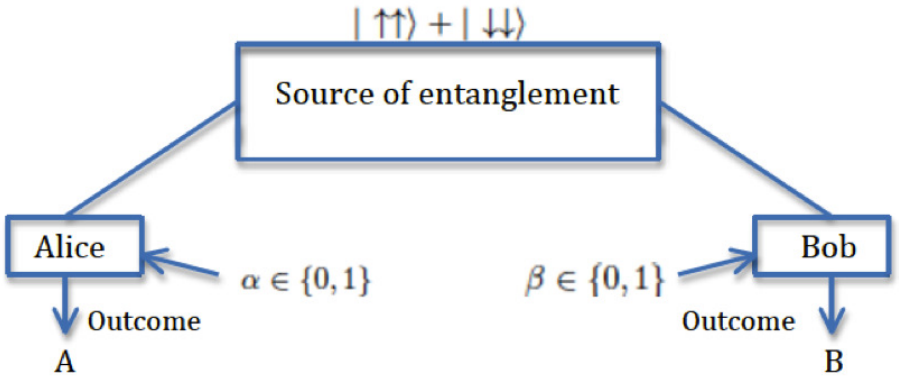


Figure 2. The operation principle of the proposed QKD protocol. An entanglement source produces a pair of entangled signals, which are randomly measured in certain bases chosen by Alice and Bob separately. Alice and Bob generate outcomes A and B respectively.

	Pr[A=G]	Pr[B=G]	
$\phi_\alpha = 0$	p	p	$\beta = 0$
$\phi_\alpha = \delta$		$\geq p - \delta^2$	$\beta = \delta$
$\phi_\alpha = 2\delta$	$\geq p - 2\delta^2$		
$\phi_\alpha = 3\delta$		$\geq p - 3\delta^2$	
$\phi_\alpha = \frac{1}{2}$		$\geq p - \frac{1}{2}\frac{\delta^2}{\delta}$	$\delta = \frac{1}{2}$

Table 1. Example of transmission of qubits between Alice and Bob showing some various possibilities and the result of the inferred bits. The probability that the eavesdropper makes a correct guess on the output held by Alice and Bob is written as p=[A=G] and p=[B=G], respectively, and δ is any value between 0 to 1.

value usually referred to as quantum-bit-error rate (QBER), in which case a secret key can be extracted by using techniques of classical information theory. However, if the error rate is larger than QBER, Alice and Bob throw their data away and never use them to encode any message. Therefore, the eavesdropper is prevented from learning any messages being communicated from Alice to Bob.

Our proposed protocol is executed by the following steps:

a) Alice chooses to measure photons in a certain basis and also the measurement direction of the polarisation e.g., Alice chooses ϕ_α and Bob chooses ϕ_β.

b) Repeat this experiment many times and check whether the statistics are compatible with the law of physics $p = \cos^2(\frac{\phi_\alpha - \phi_\beta}{2})$, where the angle ϕ_α and ϕ_β denotes the measurement direction of the polarisation [23].

c) If the statistics are compatible, then they may choose a particular basis $\phi_\alpha = \phi_\beta = 0$ and take $S_A = A$ and $S_B = B$, if not then $S_A = S_B = \perp$ i.e., they abort the protocol.

Theorem : Let G: guess for output A or B (on input $\phi_\alpha = 0$). We prove that for the classical random variable α, β and ϵ corresponding respectively to Alice, Bob and Eve's measurement outcomes, the joint entropy between Alice and Eve is always less than half, i.e., $I(\alpha, \epsilon) \leq 1/2$.

Proof: In the protocol, Alice and Bob test the presence of an eavesdropper by publicly comparing polarizations of a random subset of the photons on which they think they should

agree. The probability that a photon sent by Alice is detected by Bob is p= Pr[A$\neq$ B]=$\cos^2(\frac{\phi_\alpha - \phi_\beta}{2})$. This means that Pr[A$\neq$ B| $\phi_\alpha = 0, \phi_\beta = \delta] = \delta^2$. In Table 1, if ϕ_α and ϕ_β=0, then Pr[A=G]=Pr[B=G]=p. However, if $\phi_\alpha = \phi_\beta = \delta$, then the probability of choosing Pr[B=G] is $\geq p - \delta^2$ while the Pr[A=G] becomes $1 - p$. This can be generalized for $\phi_\alpha = 2\delta$ and $\phi_\alpha = 3\delta$.

As mentioned above, let α, β and ϵ be the classical random variables obtained by Alice, Bob and Eve, respectively, when they perform measurements on their quantum systems. The joint probability of the distribution for all the parties is expressed as $P(\alpha, \beta, \epsilon)$. By using only error correction and privacy amplification, Alice and Bob can extract a sent key from $P(\alpha, \beta, \epsilon)$ if and only if

$$I(\alpha) \geq (\alpha, \epsilon) \tag{1}$$

or

$$I(\alpha, \beta) \geq I(\beta, \epsilon), \tag{2}$$

where $I(\alpha, \beta)$=$H(\alpha) - H(\alpha|\beta)$ is the mutual information between Alice and Bob and $H(\cdot)$ is the Shannon entropy. Physically, this means that Bob must possess more information about Alice's bits than Eve does.

For such a source, the preparation quality [18] is given by

$$q = \max_{\epsilon,\beta}\{|\langle\epsilon|\beta\rangle|\}, \tag{3}$$

where $|\epsilon\rangle$ and $|\beta\rangle$ are the eigenvalues corresponding to α and β then,

$$I(\alpha, \epsilon) + I(\alpha, \beta) \leq 2\log_2(Nq), \tag{4}$$

where $I(\alpha, \epsilon) = H(\alpha) - H(\alpha|\epsilon)$ and $I(\alpha, \beta) = H(\alpha) - H(\alpha|\beta)$ are the entropies that correspond to the probability of the eigenvalues α priori to and deduced from any measurement by Eve and Bob, respectively, N is the dimension of the Hilbert space and in this case, $N = 2^n$. So, it follows that

$$I(\alpha, \epsilon) + I(\alpha, \beta) \leq \log_2(2^n 2^{-n/2}) = n. \tag{5}$$

Therefore, one can deduce that the secret key rate is obtained when $I(\alpha, \beta) \geq n/2$. Since, $I(\alpha, \beta) = H(\alpha) - H(\alpha|\beta)$, then

$$I(\alpha, \beta) = n[1 - p\log_2 p - (1 - p)\log_2(1 - p)]. \tag{6}$$

which gives us the sufficient condition

$$p\log_2 p + (1 - p)\log_2(1 - p) \leq 1/2, \tag{7}$$

on the error rate p. Because a key can only be extracted if $I(\alpha, \beta) \geq I(\beta, \epsilon)$, it follows that $I(\beta, \epsilon) \leq 1/2$ and this together with Equation (7) satisfies our theorem. Thus, the amount of information that Eve can gain about Bob's or Alice's bit is always less than half. A similar result has also been demonstrated in Ref [24]. This demonstrates that always, the eavesdropper has some limited knowledge of knowing the output from Alice or from Bob. Therefore, QKD provides a kind of security that is very secure.

5. Conclusion

We have demonstrated the principle of operation of QKD. We have shown how one can use the properties of the laws of quantum mechanics to allow the legitimate parties to share a secret key. In particular, we have shown that the eavesdropper cannot guess the output or outcome from Alice and gain more than half of the information being transmitted. This means that the key generated by quantum cryptography is always secure, thus showing the power of quantum mechanics in securing information.

References

[1] Gisin N, Ribordy G, Tittel W and Zbinden H 2002 *Rev. Mod. Phys.* **74** 145–195
[2] Wiesner S 1983 *ACM Sigact News* **15** 78–88
[3] Bennett C, Brassard G *et al.* 1984 *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* vol 175 (Bangalore, India)
[4] Ekert A 1991 *Physical Review Letters* **67** 661–663
[5] Bell J 1964 *Physics* **1** 195–200
[6] Bennett C H 1992 *Phys. Rev. Lett.* **68** 3121–3124
[7] Bruß D 1998 *Phys. Rev. Lett.* **81** 3018–3021
[8] Phoenix S J, Barnett S M and Chefles A 2000 *Journal of Modern Optics* **47** 507–516
[9] Scarani V, Acín A, Ribordy G and Gisin N 2004 *Phys. Rev. Lett.* **92** 057901
[10] Scarani V, Bechmann-Pasquinucci H, Cerf N J, Dušek M, Lütkenhaus N and Peev M 2009 *Rev. Mod. Phys.* **81** 1301–1350
[11] Mayers D 1996 *J. ACM* **48** 351–406 ISSN 0004-5411
[12] Shor P and Preskill J 2000 *Physical Review Letters* **85** 441–444
[13] Mafu M, Marais A and Petruccione F 2014 *Appl. Math* **8** 2769–2773
[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] Mafu M, Garapo K and Petruccione F 2014 *Phys. Rev. A* **90**(3) 032308
[21] Zhou C, Bao W S, Zhang H l, Li H W, Wang Y, Li Y and Wang X 2015 *Phys. Rev. A* **91**(2) 022313
[22] Renner R 2008 *International Journal of Quantum Information* **6** 1–127
[23] Hughes R J, Buttler W T, Kwiat P G, Luther G G, Morgan G L, Nordholt J E, Peterson C G and Simmons C M 1997 *AeroSense'97* (International Society for Optics and Photonics) pp 2–11
[24] Bennett C, Bessette F, Brassard G, Salvail L and Smolin J 1992 *Journal of Cryptology* **5** 3–28

Security of the Bennett 1992 quantum key distribution protocol in the presence of noise

Mhlambululi Mafu[1], Makhamisa Senekane[2], Kevin Garapo[2] and Francesco Petruccione[2,3]

[1] Department of Physics and Astronomy, Botswana International University of Science and Technology, P/Bag 16, Palapye, Botswana
[2] Centre for Quantum Technology, School of Chemistry and Physics, University of KwaZulu-Natal, P/Bag X54001 Durban, South Africa
[3] National Institute for Theoretical Physics (NITheP-KZN), P/Bag X54001 Durban, South Africa

E-mail: mafum@biust.ac.bw, makhamisa12@gmail.com, petruccione@ukzn.ac.za

Abstract. Quantum key distribution allows two parties, Alice and Bob to generate a secret key in the presence of an eavesdropper, Eve [Gisin N, Ribordy G, Tittel W and Zbinden H 2002 *Rev. Mod. Phys.* **74** 145-195]. It promises the legitimate parties to exchange private information by means of provably-secure protocols. The security is based solely on the quantum mechanical laws of physics. Since quantum key distribution is at the level of implementation and since these protocols usually operate in some noisy channels, we investigate how the addition of noise in the communication channel affects the secret key generation rates. The effect of noise for four-state and six-state protocols has been already studied [Mertz M, Kampermann H, Shadman Z and Bruß D 2013 *Phys. Rev. A* **87(4)** 042312]. Here, we investigate the behavior of the secret key when one adds some noise before classical processing for the two-state Bennett 1992 quantum key distribution protocol.

1. Introduction

Quantum key distribution (QKD) uses quantum mechanical concepts such as the quantum no-cloning theorem, Heisenberg's uncertainty principle and quantum entanglement to guarantee a secure communication between two ligitimate communicating parties Alice (the sender) and Bob (the receiver) in such a way that the presence of an eavesdropper (Eve) could be revealed. The first QKD protocol was proposed in 1984 by C. Bennett and G. Brassard, and is known as the BB84 protocol [1]. Another major QKD protocol was proposed by Artur Ekert in 1991, and is known as the E91 protocol [2].

QKD protocols can be divided into two major classes, namely discrete-variable and continuous variable QKD protocols [3]. This manuscript only investigates a class of discrete-variable QKD protocols. This class can be classified into two schemes, namely prepare and measure (P&M) and entanglement-based schemes. The BB84 protocol falls in the former scheme, while the E91 protocol is in the latter. Other notable P&M protocols are the B92 protocol proposed by C. Bennett in 1992 [4], SARG04 proposed by V. Scarani, A. Acin, G. Ribordy and N. Gisin in 2004 [5] and a six-state QKD protocol proposed by D. Bruß [6]. Security proofs for these P&M protocols have been reported [7, 8, 9, 9, 10, 11, 12, 13]. Addittionally, in Ref. [14], security of

both four-state and six-state QKD protocols in the presence of noise is reported. However, no work has been reported on the security analysis of the two-state B92 protocol in the presence of noise. This is despite the fact that the B92 protocol, being a two-state protocol, uses least resources and hence is the simplest to implement. In this manuscript, we analyze the security of the B92 protocol when the quantum noise is added to the quantum communication.

The remainder of this manuscript is structured as follows. The next section provides a background information on the B92 protocol. Section 3 provides a primer on quantum channels. This is followed by Section 4, which provides security analysis for the B92 QKD protocol. The last section concludes this manuscript.

2. The B92 QKD protocol

The B92 protocol was proposed by C. Bennett in 1992 [4]. It is a simplified version of the original BB84 protocol, and it only uses two non-orthogonal states. The protocol is then realized through unambiguous discrimination of these two non-orthogonal states [15, 16, 17]. Since this protocol uses only two states, as opposed to the four-state BB84 protocol, it uses less resources, hence it is simpler to implement. The two states used for this protocol can be denoted as

$$|\psi_0\rangle = |0\rangle, \tag{1}$$

for the bit-value 0 and

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \tag{2}$$

for the bit-value 1.

The B92 protocol uses two communication channels, namely the quantum channel and the classical channel. To initialize the protocol, Alice generates a random bit sequence. The operation of this protocol using the quantum channel can be summarized as follows [16]:

- Alice encodes each bit generated into a qubit, encoding 0 on $|\psi_0\rangle$ and 1 on $|\psi_1\rangle$.
- Alice then sends the resulting qubit sequence to Bob.
- Bob in return applies unambiguous state discrimination on each qubit he receives from Alice.

Once the quantum communication is finished, Alice and Bob then communicate over the classical channel to complete the protocol [17]. The classical communication procedure can be summarized as thus:

- Bob communicates to Alice about the instances where unambiguous discrimination succeeded.
- They both (Alice and Bob) then keep those bits which correspond to instances where unambiguous discrimination succeeded, and discard the rest. The remaining bits then form a raw key.
- They then compare some of their bits to detect the presence of an eavesdropper. The presence of an eavesropper would result in an error rate of 35% or more [16]. If an eavesdopper is detected, the protocol is aborted, else the protocol proceeds.
- Both Alice and Bob use an error correction technique to correct some errors left in the raw key.
- Finally, Alice and Bob use a privacy amplification scheme to considerably reduce the amount of information that an eavesdropper might have about the key. The result of this is a secret key for both Alice and Bob.

3. A primer on quantum channels

A quantum channel is a quantum operation (mapping) with the following properties [18, 19]:

- linearity,
- trace-preservation, and
- complete positivity.

3.1. Bit-flip channel

A bit flip channel applies the identity operator $\mathbf{I}$ with probability p and Pauli $\mathbf{X}$ (σ_x) operation to a quantum state with probability 1-p. For a quantum state ρ, the bit-flip channel transforms the state such that:

$$\rho \mapsto p\mathbf{X}\rho\mathbf{X}^{\dagger} + (1-p)\rho. \tag{3}$$

3.2. Dephasing channel

The dephasing channel is also known as the phase-flip channel. For any quantum state ρ, the phase-flip channel transforms the state such that for probability p:

$$\rho \mapsto (1-p)\rho + p\mathbf{Z}\rho\mathbf{Z}, \tag{4}$$

where $\mathbf{Z}$ is the Pauli σ_z operator.

3.3. Pauli channel

The Pauli channel is the generalization of the bit-flip channel and the dephasing channel. This channel is very important for QKD security analysis, since Eve induces such a channel on the QKD protocol [19]. For a two-dimensional quantum state ρ, the mapping due to the Pauli channel is given as:

$$\rho \mapsto \sum_{i,j=0}^{1} p(i,j)\mathbf{Z}^i\mathbf{X}^j\rho\mathbf{X}^j\mathbf{Z}^i. \tag{5}$$

3.4. Depolarizing channel

This channel is the most pessimistic channel [19]. It maps a given state ρ such that for a probability p:

$$\rho \mapsto (1-p)\rho + p\pi, \tag{6}$$

where π is given as $\frac{\mathbf{I}}{d}$ for a d-dimensional system. Hence, for a two-dimesional quantum system, π is simply given as $\frac{\mathbf{I}}{2}$.

4. Security of the B92 protocol

One of the techniques used to analyze the security of a P&M protocol is to reduce it to an entanglement distillation protocol (EDP). We employ this technique in our security analysis of the B92 QKD protocol in the presence of noise. Additionally, it is worth noting that if

$$|\langle\psi_0|\psi_1\rangle|^2 = \frac{1}{2}, \tag{7}$$

for two non-orthogonal states, the security analysis of the B92 protocol reduces to the analysis of symmetric QKD protocols like the BB84 and the six-state protocols.

In order to realize the B92 protocol, let the state being sent by Alice to Bob be given as

$$|\varphi_j\rangle = \beta|0\rangle + (-1)^j\alpha|1\rangle, \tag{8}$$

where $j = 0$ or 1, $\beta = \sqrt{1-\alpha^2}$, and $0 < \alpha < 1/\sqrt{2}$. Additionally, let $|\varphi'_j\rangle$ be a vector orthonormal to $|\varphi_j\rangle$. During the sifting stage, Alice and Bob discard the signals where Bob measured $|\varphi_j\rangle$ and retain those signals where he measured $|\varphi'_j\rangle$. The joint state between Alice and Bob can be given as

$$|\Psi\rangle = \frac{1}{\sqrt{2}}[|0\rangle_A|\varphi_0\rangle_B + |1\rangle_A|\varphi_1\rangle_B]. \tag{9}$$

Finally, let X be a collection of Alice's signals, Y be Bob's signals and E be a collection of Eve's signals, then the asymptotic secret key rate is given as

$$r = S(X|E) - H(X|Y), \tag{10}$$

where H(.) is Shannon binary entropy and S(.) is von Neumann entropy [18]. For discrete variable X and probability p_i, Shannon binary entropy H(X) is given as

$$H(X) = -\sum_i p_i \log p_i. \tag{11}$$

On the other hand, for a density operator ρ, von Neumann entropy is given as

$$S(X)_\rho = -tr(\rho \log \rho) = -\sum_i \lambda_i \log \lambda i, \tag{12}$$

where λ_i are the eigenvalues of ρ.

Let U_{BE} be Eve's evolution state, then the total state ρ_{ABE} after an action U_{BE} is given by

$$\rho_{ABE} = (\mathbf{I}_A \otimes U_{BE})(|\Psi\rangle\langle\Psi|_{AB} \otimes |X\rangle\langle X|_E). \tag{13}$$

Conventional security analyses adopt a very pessimistic view, where all the disturbances are attributed to Eve's intervention. This limits the key rates that could be generated. In this work, we adopted a different approach as follows. Depending on the nature of the channel, we can discern whether disturbance is due to eavesdropping or due to channel loss. This means that the disturbances due to channel losses would not give Eve any information about the key. This way, key generation rates would be enhanced.

In what follows, we provide the security analyses of the B92 protocol under different scenarios of quantum noise. It is also worth noting that an introduction of noise varies a quantum state ρ, which in turn changes von Neumann entropy S(ρ), since it (von Neumann entropy) is the function of ρ. Because key rate r is dependent on S(ρ), then varying noise (and hence ρ) will also affect the key rate r.

4.1. Security of the B92 protocol using bit-flip channel

When a bit-flip channel is used, the joint state of Alice, Bob and Eve can be derived to be

$$\rho_{ABE} = (\mathbf{I}_A \otimes U_{BE})(\mathbf{I}_A \otimes \mathbf{N}_{bf}\mathbf{I}_E)(|\Psi\rangle\langle\Psi|_{AB} \otimes |X\rangle\langle X|_E), \tag{14}$$

where $|X\rangle$ is Eve's initial state and $\mathbf{N}_{bf}$ is the bit-flip channel noise introduced by Alice, and is given in equation (3). Eve's unitary interaction (U_{BE}) can be given as

$$U_{BE}|0\rangle|X\rangle_E = \sqrt{1-\delta}|0\rangle_B|A\rangle_E + \sqrt{\delta}|1\rangle_B|B\rangle_E \tag{15}$$

$$U_{BE}|1\rangle|X\rangle_E = \sqrt{1-\delta}|1\rangle_B|C\rangle_E + \sqrt{\delta}|0\rangle_B|D\rangle_E, \tag{16}$$

where δ is the disturbance ($\delta = [0, \frac{1}{2}]$), $|X\rangle_E$ is Eve's initial state, and $|A\rangle_E$, $|B\rangle_E$, $|C\rangle_E$ and $|D\rangle_E$ are her states after the interaction. The disturbance corresponds to the QBER introduced

by Eve if the channel is noiseless. Using the argument given in [14], $|A\rangle_E$, $|B\rangle_E$, $|C\rangle_E$ and $|D\rangle_E$ are then determined. They are chosen in a way such that U_{BE} is a unitary operator. For example in the first instance, $\langle X|\langle 0|U^\dagger U|1\rangle|X\rangle = (\sqrt{1-\delta}\langle A|\langle 0| + \sqrt{\delta}\langle B|\langle 1|)(\sqrt{1-\delta}|1\rangle|C\rangle + \sqrt{\delta}|0\rangle|D\rangle)$ where the unitary results in the first constraint $\langle A|D\rangle + \langle B|C\rangle = 0$. One can apply the same unitary transformation U should be applied to the other initial states to get the other constraints in order to evaluate the states $|A\rangle_E$, $|B\rangle_E$, $|C\rangle_E$ and $|D\rangle_E$.

By tracing out Eve's state in ρ_{ABE}, the shared state between Alice and Bob, ρ_{AB}, can is then determined. Finally, by performing local von Neumann measurements on ρ_{ABE}, the state ρ_{XYE} is obtained. By tracing out either Bob's part or Eve's part, states ρ_{XE} and ρ_{XY} are obtained respectively, and these states are then used in Equation(10) to calculate the key rate r.

The procedure explained above is also used to calculate the key rate for the remaining noise scenarios given below. It is worth noting that for different noise scenarios, the formalism would still be the same. The only difference would be introduced by the noise scenario under consideration.

4.2. Security of the B92 protocol using dephasing channel

When Alice deliberately adds the dephasing noise, the joint state of three parties can be given as

$$\rho_{ABE} = (\mathbf{I}_A \otimes U_{BE})(\mathbf{I}_A \otimes \mathbf{N}_{deph}\mathbf{I}_E)(|\Psi\rangle\langle\Psi|_{AB} \otimes |X\rangle\langle X|_E), \tag{17}$$

where U_{BE} is Eve's evolution state, $|X\rangle$ is Eve's initial state, and $\mathbf{N}_{deph}$ is the phase-flip channel noise introduced by Alice, and is given in equation (4).

4.3. Security of the B92 protocol using Pauli channel

For the Pauli channel noise added to the communication channel by Alice, similarly to equations (14) and (17), and with the Pauli channel mapping $\mathbf{N}_{Pauli}$, the joint state of three parties is then given as

$$\rho_{ABE} = (\mathbf{I}_A \otimes U_{BE})(\mathbf{I}_A \otimes \mathbf{N}_{Pauli}\mathbf{I}_E)(|\Psi\rangle\langle\Psi|_{AB} \otimes |X\rangle\langle X|_E). \tag{18}$$

4.4. Security of the B92 protocol using depolarizing channel

Lastly, for the joint state of the depolarizing channel can be derived to be

$$\rho_{ABE} = (\mathbf{I}_A \otimes U_{BE})(\mathbf{I}_A \otimes \mathbf{N}_{dep}\mathbf{I}_E)(|\Psi\rangle\langle\Psi|_{AB} \otimes |X\rangle\langle X|_E), \tag{19}$$

where $\mathbf{N}_{dep}$ is the depolarizing noise added by Alice.

5. Conclusion

We have reported the security of the B92 protocol when the noise is added to the quantum channel. We considered different scenarios of quantum noise, and investigated security and hence robustness of the protocol in the presence of such a noise. Our work was only limited to the asymptotic key analysis. Future work will focus on the finite key analysis. Additionally, addition of noise using concatenated quantum channels deserves an investigation, because this might shed more light on how to minimize Eve's knowledge of the secret key.

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. M.M. acknowledges support from the BIUST.

References

[1] Bennett C H, Brassard G *et al.* 1984 Quantum cryptography: Public key distribution and coin tossing *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* vol 175 (New York)
[2] Ekert A K 1991 *Physical Review Letters* **67** 661–663
[3] Van Assche G 2006 *Quantum cryptography and secret-key distillation* (Cambridge University Press)
[4] Bennett C 1992 *Physical Review Letters* **68** 3121–3124
[5] Scarani V, Acin A, Ribordy G and Gisin N 2004 *Physical Review Letters* **92** 57901
[6] Bruß D *Phys. Rev. Lett.* **81** 3018
[7] Lo H K and Chau H F *Science* **283** 2050–2056
[8] Shor P and Preskill J *Phys. Rev. Lett.* **85** 441
[9] Gottesman D, Lo H K, Lütkenhaus N and Preskill J 2004 Security of quantum key distribution with imperfect devices *Proceedings of IEEE International Symposium on Information Theory* (IEEE) p 136
[10] Christandl M, Renner R and Ekert A *Arxiv Preprint: quant-ph/0402131*
[11] Renner R, Gisin N and Kraus B *Physical Review A* **72** 012332
[12] Scarani V, Bechmann-Pasquinucci H, Cerf N J, Dušek M, Lütkenhaus N and Peev M 2009 *Reviews of modern physics* **81** 1301
[13] Tomamichel M, Lim C C W, Gisin N and Renner R 2012 *Nature Communications* **3** 1
[14] Mertz M, Kampermann H, Shadman Z and Bruß D 2013 *Physical Review A* **87** 042312
[15] Gisin N, Ribordy G, Tittel W and Zbinden H 2002 *Reviews of Modern Physics* **74** 145–195
[16] Bergou J and Hillery M 2013 *Introduction to the theory of quantum information processing* (Springer)
[17] Senekane M, Mirza A, Mafu M and Petruccione F 2012 Realization of B92 qkd protocol using id3100 Clavis^2 system *Proceedings of the 56th SAIP Conference* (SAIP) pp 1–6
[18] Nielsen M and Chuang I 2010 *Quantum computation and quantum information* (Cambridge University Press)
[19] Wilde M 2013 *Quantum information theory* (Cambridge University Press)

One and Two Dimensional Models of Dye Adsorption for application in Dye Sensitized Solar Cells

N.E. Maluta[1,2], N. Mphephu[3], V. Sankaran[1], T.S. Mulaudzi[1], F. Nemangwele[1]

[1]University of Venda, Department of physics, P/Bag X 5050, Thohoyandou, 0950
[2]National Institute for Theoretical Physics (NITheP), Gauteng, South Africa
[3]University of Venda, Department of Mathematics and Applied mathematics, P/Bag X 5050, Thohoyandou, 0950

E-mail: eric.maluta@univen.ac.za

Abstract. The dye sensitized solar cells are currently the subject of intense research in the field of renewable energy as a low-cost photovoltaic device. The light adsorption occurs in dye molecules adsorbed on a highly porous structure of titanium dioxide (TiO_2) porous film. The progress in the efficiency and stability of these solar cells is very low, due to many fundamental aspects of their operation that are still unknown. One process, for which there is limited information, is the time taken to upload the dye on the TiO_2 nanoporous film which acts as a semiconductor. Dye molecule is adsorbed onto a TiO_2 working electrode by dipping it into the dye solution for periods of several hours to several days. However, such long dipping times are not economical for industrial production of DSSCs. The factors controlling this process are not yet fully understood. We have developed a 1D and 2D models based on the Langmuir isotherms to study and understand the diffusion and adsorption of the dye molecules in TiO_2 nanotube films. Our modelling results show that the adsorption of dye into the TiO_2 nanotubes is controlled by the diffusion coefficient, the adsorption-desorption ratio and the initial dye concentration.

1. Introduction

With increasing worldwide energy demand, the need to adopt the renewable energy technologies as an alternative energy sources is very important to developing countries. Among all the renewable energy technologies solar energy is one of the potential candidate to satisfy the global energy needs. One promising photovoltaic technology is the dye sensitized solar cells, which are regarded as a low cost solar cell with high light to energy conversion efficiency [1-2]. Currently the DSSCs based on the TiO_2 nanoparticles films sensitized with ruthenium (Ru) dyes have been reported to achieve a light to energy conversion efficiency of more than 12 %. The dye molecules adsorbed on the TiO_2 surface act as the photon absorber. Thus the dye molecules adsorbed on the TiO_2 working electrode are of importance during the photon absorption and the overall efficiency of the solar cells.

The adsorption of organic and inorganic molecules for the functionalization of titanium surfaces is of particular interest for number of technological applications like photovoltaic, biosensors, water filter, etc. The interaction between the TiO_2 semiconductor surfaces with dye

molecules or other molecules in particular has attracted more attention in the field of renewable energy as application on dye sensitized, Perovskites and quantum dots solar cells [1-3]. In dye sensitized solar cells (DSSCs), the adsorption of dye molecules like ruthenium dye onto the TiO_2 surfaces has relevance in understanding the surface interaction and photon absorption in these types of solar cells. The formation of strong bonds and surface coverage between the dye molecules as light sensitizer and TiO_2 semiconductor is an important factor for allowing efficient charge transfer between the two materials in DSSCs. In DSSCs the experimental procedure used for the adsorption of ruthenium dye molecules on the surface of the TiO_2 is by dipping in the solution of dye molecules, usually the ruthenium complex dye. This experiment is sometimes done in several hours or days [3, 4]. The optimum adsorption time is an extremely important factor during the experiment. It is well known that the time taken during the dye uptake experiment has an impact on the efficiency of the solar cells [4]

2. Computational Method

The present model is intended to simulate the formation of a monolayer on the surface of the TiO_2 film (a general case in which the dye molecules are adsorbed on the surface of TiO_2 film to form a mono layer) and understand the factor controlling this process. The model is based on the Langmuir adsorption isotherm, which uses active site concept in adsorption expression in order to address the reduction of its rate with the coverage of the wall. The model assumes that [4-6]:

(a) Adsorption of dye molecules is controlled by transient diffusion along the film,

(b) The dye molecules stick to the surface of the film and only a monolayer is formed and

(c) There is a constant number of molecules of the adsorbate even after the adsorption.

The concentration of dye molecules inside the film can be calculated using Fick′s second law for cartesian (one dimensional) and cylindrical (two dimensional) [4]. We solved one and two dimensional diffusion equation for dye molecules concentration $c(r, z, t)$, where z is the distance from the second end of the TiO_2 working electrode, r is the radius of the nanotubes and t is the time, for two dimensional models given by [4, 6]:

$$\frac{\partial c}{\partial t} = D\nabla^2 c = D\left[\frac{1}{r}\frac{\partial}{\partial r}\left(\frac{\partial c}{\partial r}\right)\right] + D\frac{\partial c}{\partial z}, \tag{1}$$

Where D is the diffusion coefficient, c is the concentration, r is the distance from the centre of the pore and z is the distance along the pore from the extracting electrode as illustrated on Fig.(1).

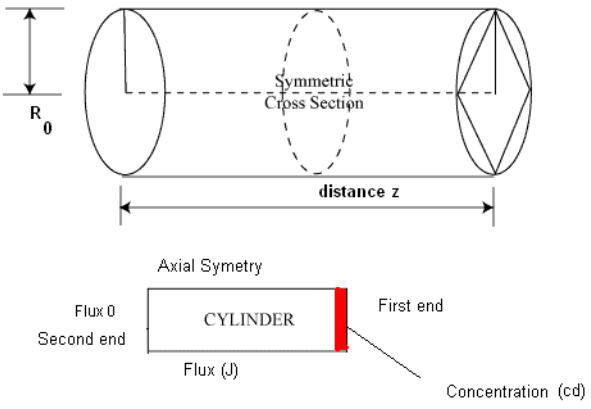


Figure 1. The TiO_2 nanotube represented by a cylinder penetrating the entire TiO_2 film with a bulk concentration of dye molecules at the right hand end of the nanotube.

As defined by Kimmish and Nestle [6] we consider an effective diffusion coefficient which depends on the value of the sticking or material coefficient (k) as

$$D_{eff} = \frac{D}{1+k}. \tag{2}$$

To find the net flux of dye molecules $J(z,t)$ moving towards the pore surface

$$J = -D\frac{\partial c}{\partial z}. \tag{3}$$

We consider the molecules flowing onto the surface and being adsorbed onto to it less the desorbed molecules to give a fractional surface coverage (θ), given by,

$$\theta(z,t) = \frac{\Gamma(z,t)}{\Gamma_{max}}. \tag{4}$$

Where $\Gamma(z,t)$ is the surface coverage a distance z from the electron extracting electrode and Γ_{max} is the maximum surface coverage of the TiO_2 nanotube film. If all molecules a distance l_{mol} out of the surface are adsorbed, where l_{mol} is the length of the dye molecules,

$$\Gamma_{max}\frac{\partial\theta}{\partial t} = k_{ads}\cdot l_{mol}(1-\theta(z,t))c - k_{des}\Gamma(z,t), \tag{5}$$

where k_{ads} and k_{des} are the adsorption and desorption rate of dye molecules respectively.

We assume a quasi-static equilibrium, i.e., the adsorption and desorption take place quickly compared to time taken for the dye molecules to diffuse to the surface, thus the left hand side of the above equation (Eq.(5)) is set to zero. Hence, if $k = \frac{k_{ads}}{k_{des}}$, then

$$k_{ads}\cdot l_{mol}(1-\theta(z,t))c - k_{des}\Gamma(z,t) = 0, \tag{6}$$

So that we have

$$\theta(z,t) = 1 - \frac{1}{1+\frac{c}{c_{con}}} \tag{7}$$

where c_{con} is the constant for Langmuir isotherm given by,

$$c_{con} = \frac{\Gamma(z,t)}{k\ l_{mol}} \tag{8}$$

The total number of dye molecules on the surface can be written as,

$$N_{surf} = 2\pi\Gamma_{max}\int_0^d \theta(z,t)dz \tag{9}$$

The boundary conditions at the curved surface (radial axis) at all times t is given by

$$J(r = R_0, r) = -D\frac{\partial c}{\partial r} \tag{10}$$

In order to achieve this modelling analysis we have employed this model using the MATLAB programing software.

3. Results and discussion

The adsorption process is a complex phenomenon which comprises of the different steps like diffusion of dye molecules and adsorption of the molecules on the surface of the pore. This involves the interaction between the dye molecules and active sites of the adsorptance. Fig.(2) shows the time variation due to the diffusion coefficient. It is clear that the time variation depends on the diffusion coefficient. But if we scale our graph by the characteristic time ($t_z = \frac{d^2}{D}$) where d is the length of the film and D is the diffusion coefficient, the variation of dye uptake on the pore surface at the end is the same as illustrated on Fig.(1). As illustrated by Fig.(3), along the pore the diffusion takes place slower and takes longer than the characteristic time ($t_z = \frac{d^2}{D}$). Diffusion along the pore is slowed down by the adsorption of dye molecules at the pore surface.

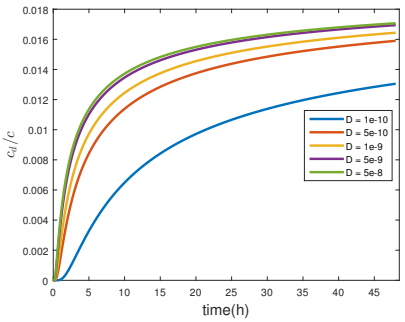


Figure 2. Dependence of surface coverage on different values of diffusion coefficient. The values used in simulation are $k = 1 \times 10^3$, $c_d = 6 \times 10^{21}\mathrm{m}^{-3}$ and $d = 15\mu\mathrm{m}$

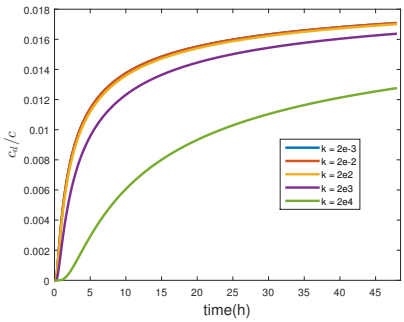


Figure 3. Dependence of surface coverage on different values of material constant. The simulation values are $D = 5 \times 10^{-9}\mathrm{m}^2\mathrm{s}^{-1}$, $c_d = 6 \times 10^{21}\mathrm{m}^{-3}$ and $d = 15\mu\mathrm{m}$

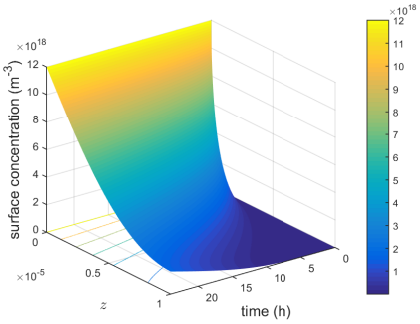


Figure 4. Surface concentrations on the TiO_2 film for 24 hours simulated time. The values used in simulation are $k = 1 \times 10^3$, $D = 6 \times 10^{-9}\mathrm{m}^2\mathrm{s}^{-1}$ and $d = 15\mu\mathrm{m}$

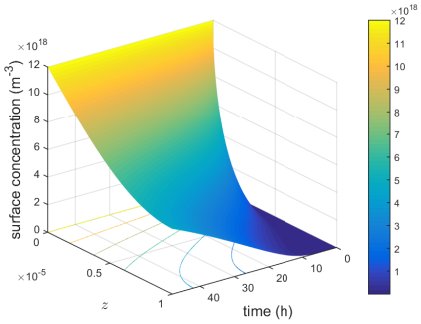


Figure 5. Surface concentrations on the TiO_2 film for 48 hours simulated time. The values used in simulation are $k = 1 \times 10^3$, $D = 6 \times 10^{-9}\mathrm{m}^2\mathrm{s}^{-1}$ and $d = 15\mu\mathrm{m}$

This delay is seen more clearly for large k values. This illustrates the importance of understanding the value of the material sticking coefficient. The duration of the dye loading plays an important role on the surface coverage. It has been reported experimentally that the longer duration for dye loading affects the efficiency of the solar cells [5, 8]. As illustrated in

Fig.(4) and (5), it can be observed that the time of dye loading plays an important role on the surface coverage. The time effect is also observed for the higher values of the material coefficient (k), due to the fact that a higher value of k in our models shows that the dye molecules which interact with the surface will stick on the surface.

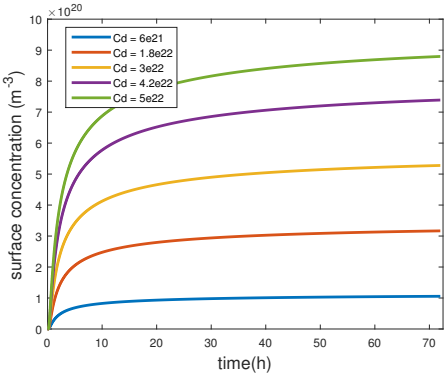


Figure 6. Dependence of surface coverage on different initial dye concentration. The simulation values are $k = 1 \times 10^3$, $D = 5 \times 10^{-9}\mathrm{m}^2\mathrm{s}^{-1}$, and film thickness $d = 15\mu\mathrm{m}$

Since there are interactions between the dye molecules and the surface, the process will be much slower compared to smaller values of k. This delay is not needed in this experiment as it affects the overall efficiency of the solar cells and it's not good for economical production of these solar cells. The effect of increasing the initial dye concentration (c_d) on the surface coverage and its saturation time is illustrated on Fig.(6). It can be observed from Fig.(6) that the surface coverage increases with increase initial dye concentration. When we increase the initial dye concentration we are increasing the probability of more dye molecules to diffuse into the film and hence reducing the saturation time. Similar effect was also reported and observed experimentally by Chou et al. [5].

4. Conclusion

Our model demonstrates that dye molecules can easily take few hours to be fully adsorbed onto the surface of the pore in the nanotube TiO_2 film. The model shows that the time depends on the ratio of adsorption to desorption rates (k). The time scale for diffusion along the pore radius is given by simple expression $t_z = \frac{d^2}{D}$. The initial dye concentration plays an important role on the surface coverage and the dye uploading time.

5. Acknowledgement

This work is financially supported through the program NRF Thuthuka and also by National institute for theoretical Physics (Nithep). We would like to also thank the University of Venda for the support to carry out this research.

References

[1] Banerje S, Gautam R.K, Jaiswal A, Chattopadhyaya M.C and Sharma Y.C 2015 *RSC Advances.* **5**(19), 14425-14440.
[2] Giovanneti R, Zannotti M, Alibabei L, and Ferraro S 2014 *Int. J. Photoenergy.* doi: 10.1155/2014/834269
[3] Yuasa T, Kawakami R, Sato Y, Mori Y, Adachi M T and Yoshikado S 2012 *Sol. Energ. Mat. Sol.* C **102**, 2-7.
[4] Kankare J and Vinokurov I 1999 *Langmuir.* **15**(17),5591-5599.
[5] Chou T.P, Zhang Q and Cao G 2007 *J. Phy. Chem.* C **111**(50), 18804-18811.
[6] Nestle N.F.E.I and Kimmich R 1997 *J. Phy. Chem.* **100**, 12569-12573
[7] da Silva W.P, e Silva C.M and Gomes J.P 2013 *J. Food. Eng.* **117**(3), 417-424.
[8] Ghicov A, Albu S.P, Hahn R, Kim D, Stergiopoulos T, Kunze J and Schmuki P 2009 *Chem. Asian. J.* **4**(4), 520-525

Molecular dynamics studies of Schottky and Frenkel defects in cerium dioxide

Thabiso Lekoko[1], **Thuto Mosuang**[1], **and Erasmus Rammutla**[1]

[1] Department of Physics & Geology, University of Limpopo, Private Bag x1106, Sovenga, 0727, Polokwane, South Africa

Corresponding author: thuto.mosuang@ul.ac.za

Abstract. Schottky and Frenkel defect energies in cerium dioxide are studied using the classical molecular dynamics. Buckingham potentials are used to understand the cerium-cerium, cerium-oxygen and oxygen-oxygen interactions in the bulk and defect structures. The formulation uses the NVT Evans ensemble to obtain the various defect energies. Oxygen and cerium vacancy defect energies relative to bulk cerium dioxide total energies are used to get more insight on cerium dioxide as a catalyst in exhaust systems. The anion Frenkel-pair defect is the most favourable form of intrinsic ionic defect.

1. Introduction

Cerium dioxide (CeO_2), also referred to as ceria, is the oxide of the rare earth metal cerium. Ceria is an organic solid, mono-constituent appearing under the form of a white ivory to pale beige powder. Generally, it may not be considered as hazardous to human health and the ecosystem environment. This oxide has been under investigation for many applications because of its useful properties [1]-[6]. An exclusive property of ceria is that it has high oxygen mobility, meaning that oxygen atoms within a ceria crystal can diffuse at high rates [1]. Another notable property of ceria is that it can give up some of its oxygen atoms in order to regulate oxygen partial pressures in the atmosphere without actual decomposition [2]. Furthermore, it is utilised as non-reactive base for the burning of nuclear fission products. This means that CeO_2 is remarkably resistant to ion radiations unlike the nuclear fuel, urania (UO_2) because they share the same fluorite crystal structure [3].

CeO_2 of variable composition is generally known as ceria; its unique chemical and electronic properties make it an important material for a number of applications. Easy shifting redox equilibrium in ceria makes it an excellent catalyst for various chemical processes, including production and purification of hydrogen, and carbon monoxide removal from the automobile exhaust. Mobility of oxygen vacancies and consequently high ionic conducting makes ceria a promising electrolyte for solid oxide fuel cells [4] [5]. Electronic structure of ceria leads to its use as UV absorber in cosmetic industry and manufacturing of glass windows [6]. It is also used in light harvesting devices and optical displays [6] [7].

Cerium is mainly used as such in formulation as high precision polishing agent for glass products, as catalyst support for other applications and raw materials for the production of glass and ceramics as well as additions in paints and varnishes [8]. The most important application of CeO_2 is the formation of three way catalysts (TWC) for automobile exhaust gas [7] [8].

Experimental and theoretical studies have shown that transition metal oxides like CeO_2 can be reduced through introduction of defects and/or impurities [9]-[11]. Alternatively, transition metals can be supported on an oxide substrate [12] [13]. Clearly these properties can influence positively the

ability of oxides to reduce unwanted gases. It is the objective of this research to introduce Ce and O charged vacancies in the bulk CeO_2 to improve elimination of poisonous gases. This will be accomplished by modelling Schottky and Frenkel related defects type within the fluorite CeO_2 matrix.

2. Computational method

The classical molecular dynamics (MD) calculations using the Daresbury simulation package DL_POLY has been used [14]. In the present study, the Buckingham potentials have been used to accurately reproduce the interactions between the cerium and the oxygen atoms. The ions were allowed to relax using the canonical NVT evans ensemble. The MD used a routine of Verlet leapfrog to integrate the Newton's equations of motion. Simulations were allowed to run with a 10-15 s timestep for a period of 2000 steps and an equilibration after every 200 steps. The ceria supercell is formed of a cube with 4 x 4 x 4 fluorite unit cells, where 256 cerium (Ce) and 512 oxygen (O) atoms are included. Throughout the study, periodic boundary conditions have been functional.

The Buckingham potential [15] is the most common short-range interaction which is given by:

$$V(r_{ij}) = Aexp\left(-{}^{r_{ij}}/_{\rho}\right) - \frac{C}{r_{ij}^6}, \quad (1)$$

where A, ρ, and C are adjustable parameters, chosen in such a way that they reproduce significant physical properties of real materials. The potential parameters used in this study can be found in a paper by Xu et al. [16]. It must be mentioned that unlike in their potential evaluation, the shell model parameters were excluded in this study, which simply implies that formal charges of -2 for oxygen and +4 for cerium ions are being considered. In order to evaluate the quality of this potential radial distribution functions for a 256 CeO_2 supercell, O vacancy in 256 CeO_2 supercell , and Ce vacancy in 256 CeO_2 supercell are calculated. Oxygen (V_O^{2-}) and Cerium (V_{Ce}^{3+}) vacancies are modelled by the removal of a single atom of each in the 256 CeO_2 fluorite supercell. This results in a defected CeO_2 supercell now having either 511 O atoms or 255 Ce atoms for the V_O^{2-} and V_{Ce}^{3+} vacancy defects respectively. The Schottky defect is modelled by a removal of a Ce atom and two O atoms to depict the removal of a complete CeO_2 formula unit. An in depth assessment of these properties is presented in the subsequent section.

3. Results and discussion

3.1 *Structural properties*

Structural effects due to an O vacancy (V_O^{2-}) and Ce vacancy (V_{Ce}^{3+}) have also been explored at a temperature of 300 K and pressure of 0.0 atm. A radial distribution function **g(r)**, has been used to analyse the consistency of the Buckingham potentials in explaining the interactions within the CeO_2 system. The peak positions in **g(r)** allow a fairly accurate prediction of the most probable distances between the atoms. At equilibrium 300 K, atomic positions for the Ce-Ce, bond lengths peaks appear at 3.8 and 5.4 Å in accordance with the first and second coordination respectively relative to any Ce atom. In the same manner, Ce-O bond lengths peaks appear at 2.3 and 4.5 Å in accordance with the first and second nearest neighbours in reference to a Ce atom, while the O-O peak values appear at 2.8 and 3.8 Å relative to any O reference atom. The **g(r)** of a ceria supercell at 300 K is shown in fig. 1, of which agrees with the simulated annealing of small clusters method of Cordatos et al. [17]. An increase in temperature can also contribute to oxgyen atoms vibrations about their lattice sites. If this vibrational energy exceeds the lattice energy a vacancy get created in the crystal lattice [18]. It can also be seen that creating an O^{2-} vacancy in the material tend to allow Ce atoms to cluster, as their first and second neighbour atoms congregate as displayed in fig. 2. More detailed information on calculated first and second nearest neighbour distances and their respective number of atoms is presented in table 1. Introducing an oxygen vacancy (V_O^{2-}) increases the number of first nearest nieghbors Ce atoms around the Ce atom as the second nearst neighbors are reduced. In contrast, simultaneous introduction of a vacancy and an interstitial oxygen ($V_O^{2-} - O_{\text{int}}$) in the ceria matrix results in reduction of first and

second neighbor O atoms around an O atom. These findings corroborate those reported by Campbell and Peden [19] paper.

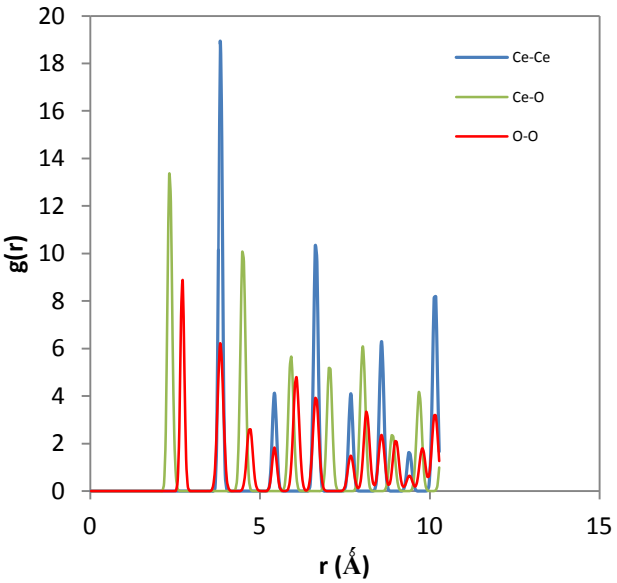


Figure 1. The radial distribution functions of fluorite CeO_2 supercell at 300 K.

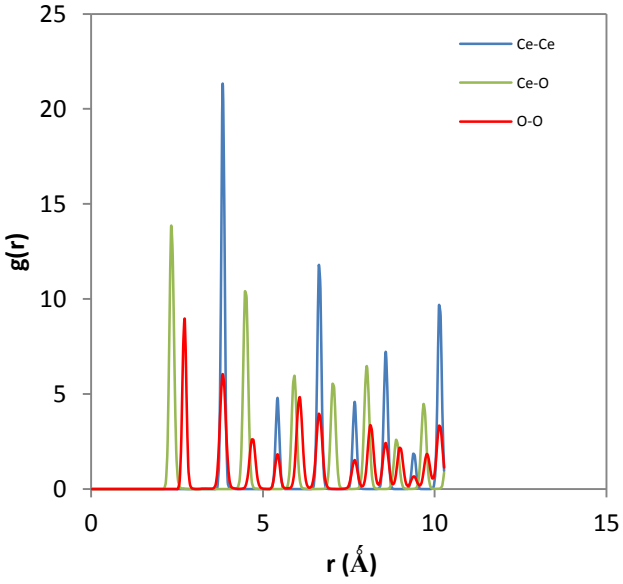


Figure 2. The radial distribution functions of an O vacancy in fluorite CeO_2 supercell at 300 K.

Table 1. Mean first and second nearest neighbour distances (r_1, r_2) and number of atoms (n_1, n_2) in fluorite CeO_2 supercell cell, its O vacancy and an O vacancy-interstitial defect at 300 K.

	Ce-O				Ce-Ce				O-O			
	r_1(Å)	n_1	r_2(Å)	n_2	r_1(Å)	n_1	r_2(Å)	n_2	r_1(Å)	n_1	r_2(Å)	n_2
CeO_2	2.4	2.7	4.5	2.0	3.8	3.8	5.4	2.0	2.8	1.8	3.8	6.0
V_O^{2-}	2.4	2.8	4.5	2.1	3.8	4.3	5.4	1.0	2.8	1.8	3.8	6.0
V_O^{2-}-O_{int}	2.4	2.8	4.6	2.2	3.8	4.5	5.4	1.0	2.8	1.6	3.8	1.2

3.2 *Defect properties*

Investigation of the vacancies, Schottky and Frenkel type defect energies in the fluorite 256 CeO_2 supercell was also performed. The oxygen and cerium vacancies (V_O^{2-} and V_{Ce}^{3+}) in fluorite 256 CeO_2 supercell are created by the removal of either an O^{2-} or Ce^{3+} ions in the supercell. The Schottky defect is obtained by removal of cations and anions in formula units to form vacant lattice sites,

$$\mathrm{Ce_{Ce}} + 2\mathrm{O_O} \;\rightleftharpoons\; V_{Ce}^{3+} + V_O^{2-} + \mathrm{CeO_2}(s). \quad (2)$$

From this reaction the Schottky defect energy is calculated using the equation:

$$E_{\mathrm{Schot}}^{\mathrm{CeO_2}} = \frac{1}{3}\left(E[V_{Ce}^{3+}] + 2E[V_O^{2-}] + E[\mathrm{CeO_2}]\right). \quad (3)$$

In a similar manner, a Frenkel defect is created by the displacement of either an anion or a cation to the local interstitial site [20]. Specifically for the favourite oxygen, $O_O \rightleftharpoons O_{\mathrm{int}}^{2+} + V_O^{2-}$. The anion Frenkel-pair defect energy is calculated from:

$$E_{\mathrm{Frenk}}^{\mathrm{ani}} = E[\mathrm{CeO_2}] - E[O_{\mathrm{int}}^{2+} + V_O^{2-}]. \quad (4)$$

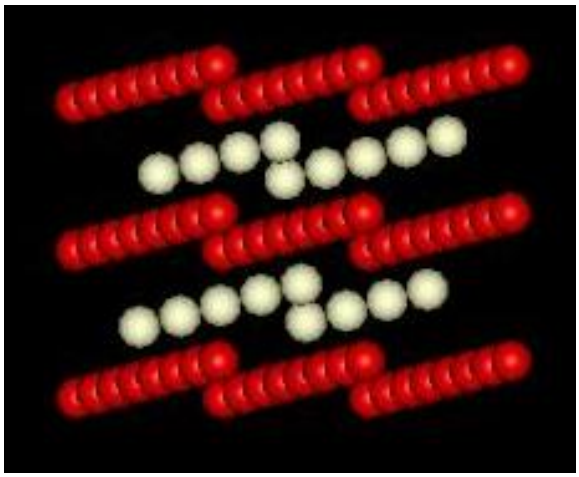


Figure 3. Visualisation of a ceria supercell is organised in a defect free fluorite CeO_2.

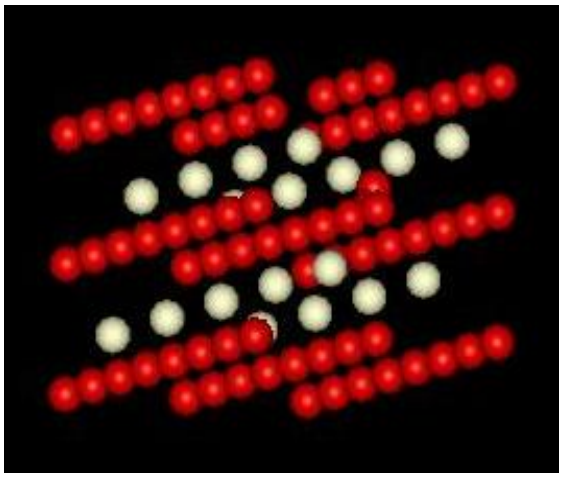


Figure 4. Visualisation of how an O vacancy and O interstitial manifest itself in a fluorite CeO_2 supercell.

Table 1: Defect energies (eV/defect) for all the point, Schottky and Frenkel-pair defects in fluorite CeO_2 using Buckingham potentials in the present work compared with Walsh *et al*. [13].

Defect	Present work	Other calculations
$\boldsymbol{V_O^{2-}}$,	14.66	12.91
$\boldsymbol{V_{Ce}^{3+}}$,	82.74	94.41
Anion Frenkel-pair	-7.37	3.21
Cation Frenkel-pair	2.50	15.94
Schottky defect	2.23	5.79
Defect free	-105.38	-102.87

Two vacancy defects ($V_O^{2-} and\ V_{Ce}^{3+},$), an anion and a cation Frenkel-pair defects, and a Schottky defect have been considered in this investigation. From the present and the Walsh *et al*. [13] calculations, an anion Frenkel-pair defect is the most preferred form of an intrinsic disorder. The anion Frenkel-pair defect according to the present investigation suggest a defect energy of -7.37 eV per defect, whilst the anion Frenkel-pair formation according to Walsh *et al*. [13] suggest a 3.21 eV per

defect. The huge deviation on the values (i.e. -7.37 eV/defect and 3.21 eV/defect) could be owed to different procedures of modelling the Frenkel –pair defects. Walsh *et al*. [13] used the intrinsic V_O^{2-} defect energy and the interstitial oxygen (O_{int}) energy defect to calculate the associated anion Frenkel-pair defect energy $(E_{\mathrm{Frenk}}^{\mathrm{ani}})$. This presentation firstly modelled a collective V_O^{2-} and O_{int} defects in one supercell structure, thereafter their mutual defect energy $(E[O_{\mathrm{int}}^{2+} + V_O^{2-}])$ was calculated. This was subtracted from the defect free fluorite 256 CeO_2 lattice energy $(E[\mathrm{CeO_2}])$ as shown in equation (4).

4. Conclusion

The structural and defect conformation properties of ceria have been studied with help of the Buckingham potentials. ***g*(r)** and ***S*(k)** results demonstrate how the Ce-Ce, Ce-O, and O-O bonds arrange and distribute themselves around a given atom. The nearest neighbour atoms in respect of a given atom can determine the nature of defect in CeO_2. The oxygen vacancy and interstitial defects in the ceria matrix can be explained using the radial distribution functions. Furthermore the calculations show that the anion Frenkel-pair defect has the lowest value as demonstrated by the Walsh *et al*. [13] calculations; suggesting it to occur easily compared to other defect types. This finding promotes the fact that oxygen plays a crucial role in CeO_2 exhaust catalyst.

Acknowledgements

The UL and IBSA are thanked for financial assistance. CHPC is thanked for computational facilities.

References

[1] Gotte A, Hermansson K, Baudin M 2004 *Surface structure* **552** 273.
[2] Zhu WZ, Deevi SC 2003 *Materials science and engineering* A **362** 228.
[3] He H-W, Wu X-Q, Ren W, Shi P, Yao X, Song Z-T 2012 *Ceramics international* **38** S501.
[4] Zhu B, Mat MD 2006 *International Journal of Electrochemical Science* **1** 383.
[5] Gangopadhyay S, Frolov DD, Masunov AE, Seal S 2014 *Journals of alloys and compounds* **584** 199.
[6] Sayle DC, Maicaneanu SA, Watson GW 2002 *Journal of American Chemical Society* **124** 11429.
[7] Gerward L, Staun Olsen J, Petit L, Vaitheeswaran G, Kanchana V, Svane A 2005 *Journal of Alloys and Compounds* **400** 56.
[8] Liu L, Yao Z, Deng Y, Gao F, Liu B, Dong L 2011 *ChemCatChem* **3** 978.
[9] Nakajima A, Yoshihara A, Ishigame M 1994 *Physical Review B* **50** 13297.
[10] Gerward L, Staun Olsen J, Petit L, Vaitheeswaran G, Kanchana V, Svane A 2005 *Journal of Alloys and Compounds* **400** 56.
[11] Chafi Z, Keghouche N, Minot C 2007 *Surface science* **601** 2323.
[12] Sun L-L, Cheng Y, Ji G-F 2010 *Journal of Atomic and Molecular Science* **1** 143.
[13] Walsh A, Woodley SM, Catlow CRA, Sokol AA 2011, *Solid State Ionics*, **184** 52.
[14] Smith W, Forester TR and Todorov IT, 2009 *The DL_POLY 2 User Manual*, STFC Daresbury Laboratory, Darebury, Warrington WA4 4AD Cheshire, UK.
[15] Buckingham RA 1938 *Mathematical and Physical Sciences* **168** 264.
[16] Xu H, Behera RK, Wang Y, Ebrahimi F, Sinnott SB, Wachsman ED, Phillpot SR 2010 *Solid State Ionics* **181** 551.
[17] Cordatos H, Ford D, Gorte RJ 1996 *Journal of Physical Chemistry* **100**18128.
[18] Dutta P, Pal S, Seehra MS, Shi Y, Eyring EM, Ernst RD 2006 *Chemistry of Materials* **18** 5144.
[19] Campbell CT, Peden CHF 2005 *Science* **309** 713.
[20] Guglielmetti A, Chartier, van Brutzel L, Crocombette J-P, Yasuda K, Meis C, Matsumura S 2008 *Nuclear Intsruments and Methods in Physics Research* B **266** 5120.

A phenomenological description of the bath and its effect in photosynthetic light-harvesting systems

JA Nöthling[1], TPJ Krüger[1] and T Mančal[2]

[1] Department of Physics, University of Pretoria, Lynnwood Road, Hatfield 0028, South Africa
[2] Faculty of Mathematics and Physics, Charles University in Prague, Ke Karlovu 5 CZ-12116, Czech Republic

E-mail: janothling@gmail.com

Abstract. It is practically impossible to calculate the dynamics of each degree of freedom in a large, complex quantum system. When only a small part of the system is of interest, the rest of the system can be modelled as a thermal bath consisting of an infinite number of independent harmonic oscillators. In quantum master equations, dynamics of the bath and the coupling of the relevant subsystem to the bath are introduced through bath correlation functions. However, the physical significance of these correlation functions and the microscopic dynamics and interactions of the bath, are not clear from the derivation of the master equations. In this paper we give a microscopic interpretation of the bath and discuss the significance of the correlation functions in the context of photosynthetic light harvesting. The ideas and results are applicable to many other quantum systems as well.

1. Introduction

In many fields of physics and chemistry, the dynamics of a small quantum subsystem contained in a much bigger system need to be calculated. The total system evolves unitarily and, in principle, its dynamics can be calculated exactly. In reality, these systems are far too big and far too complex to allow calculation of the dynamics of all degrees of freedom. Since the dynamics of the subsystem depends on interaction with the rest of the system, considering an isolated subsystem is also not a solution. In general, due to energy exchange between the subsystem and the other degrees of freedom, it is not even possible to describe the subsystem dynamics with a Hamiltonian approach. One of the ways to treat this problem is to partition the total Hamiltonian in three parts [1]: H_S describing the subsystem as if it were an isolated system, H_B describing the rest of the system (bath) as if it were isolated and H_{SB} describing the interaction between the relevant subsystem and the bath. Instead of treating H_B and H_{SB} exactly, the actual bath is substituted with an effective bath having more or less the same effect on the system. A commonly used effective bath consists of an infinite number of independent quantum harmonic oscillators [2]. These harmonic oscillators are coupled linearly (see below) to the relevant subsystem. The harmonic bath causes both fluctuations and dissipation in the relevant subsystem. Although this approach is useful for many physical systems, its application to photosynthetic light harvesting will be discussed to emphasize its applicability.

2. A brief introduction to quantum dynamics in photosynthesis

Photosynthesis can be summarised as a process consisting of two phases. In the first phase, light energy is captured and transported to a photosynthetic reaction center [3]. The second phase, which is initiated at the reaction center, entails the conversion of the captured energy to energy-rich chemical compounds [4]. The capturing and energy conversion phases can be regarded as separate processes. The light-harvesting complexes responsible for capturing light energy consist of pigment molecules embedded in a protein conglomerate [5]. During the light capturing phase, a photon is absorbed and its energy used to excite a pigment molecule from the ground state to an electronic excited state. This excitation is transferred from pigment molecule to pigment molecule until it reaches the photosynthetic reaction center from where the chemical conversion phase is initiated [3]. The quantum efficiency of the first phase is extremely high; often more than 95% [6]! This high efficiency was one of the catalysts for the large interest in theoretical biophysics in recent years. Most of this research field is focused on the quantum mechanical description of energy transfer during the energy-capturing phase of photosynthesis. The light harvesting complexes are very large and intricate macromolecules (see for example [7]). To describe the dynamics of all the degrees of freedom is impossible. In general, we are interested only in the electronic degrees of freedom. However, the protein scaffold is important for energy quenching and its effect on the electronic degrees of freedom has to be considered. For these three reasons the Hamiltonian-partitioning approach outlined above is frequently used.

3. The relevant system[1]

Each of the pigment molecules in a light harvesting complex can be regarded as a two-level system with a ground state and an excited state [8]. Due to the possibility of excitation, each pigment molecule has an associated transition dipole moment. The transition dipole moments of different molecules interact in a similar manner as transient dipole moments in matter do to bring about London dispersion forces. One can say that different pigment molecules are *coupled* to one another. The strength of the coupling depends on the distance and relative orientations of the pigment molecules. The Hamiltonian of the relevant system is therefore:

$$H_S = E_g \left|g\right\rangle \left\langle g\right| + \sum_{i=1}^{N} E_{e_i} \left|e_i\right\rangle \left\langle e_i\right| + \sum_{i \neq j}^{N} V_{ij} \left|e_i\right\rangle \left\langle e_j\right|, \tag{1}$$

where the first term is the ground state energy of the system (all pigment molecules are in their ground states) and the second term is a sum of single-excitation energies (only molecule i is in the excited state). The last term represents the dipole coupling between different pigment molecules.

In what follows, we will mostly consider a single pigment molecule as the relevant system. Extension to more than one pigment molecule does not change the way in which the bath couples to the relevant system.

[1] The word "system" can cause confusion. We will use the term "relevant system" for the electronic degrees of freedom and "bath" for all other degrees of freedom such that the union of relevant system and bath evolves unitarily in time. We will reserve the term "system" exclusively for this union.

4. Interaction of the relevant system with an oscillator

Consider an isolated pigment molecule consisting of two identical subunits. For comparison with Eq. (1), we will call this molecule i.

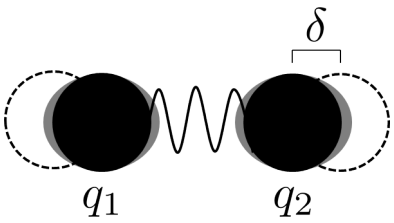


Figure 1. A pigment molecule consisting of two identical subunits. The coordinates of the two subunits measured on the same coordinate axis are given by q_1 and q_2, and δ is the equilibrium shift in each of these coordinates upon excitation of the molecule.

One of the vibrational modes of the molecule constitutes stretching and contracting along the axis connecting the two subunits. Let's consider an effective coordinate $Q = q_1 - q_2$. The potential energy causing the oscillation is now quadratically dependent on Q. Let the minimum of this potential be indicated by the black circles when the molecule is in its ground state. Suppose that the structure of this molecule in its excited state is such that the equilibrium position of each subunit is shifted a distance δ outwards. The potential minimum (in coordinate Q) is therefore shifted by $d = 2\delta$. Because of the large masses of the subunits, one can assume that the oscillations happen on a much slower timescale than the transition from the ground to the excited state in molecule i (adiabatic approximation). To excite molecule i from the vibrational ground level of its electronic ground state, one therefore has to apply more excitation energy than the energy difference $E_{e_i} - E_g$ in Eq. (1). The energetics accompanying the electronic excitation is depicted in figure 2. The parameter d in figure 2, which is equal to 2δ in figure 1, describes how strongly the relevant system is perturbed by (or *coupled* to) the oscillation. This parameter will be useful later again.

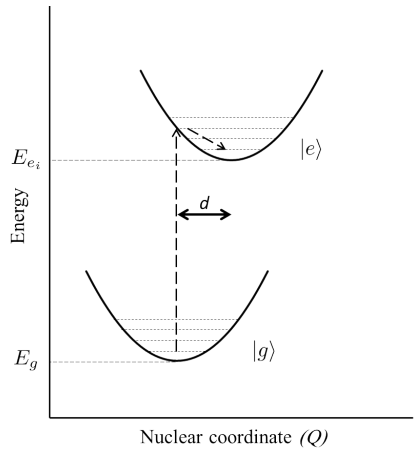


Figure 2. The harmonic oscillator potentials corresponding to the ground and excited states of a pigment molecule with two subunits. E_g and E_{e_i} are the energies of the ground and excited electronic states respectively. As indicated by the dotted line arrows, the molecule is excited to a non-ground vibrational level of the electronic excited state, before relaxing to the ground vibrational level. The parameter d is an indication of the coupling strength between electronic excitation and the molecular vibration.

Although it is very elucidating, the discussed model is an oversimplification: the pigment molecules are complex structures and have numerous intramolecular vibrational modes. In addition, intermolecular vibrations due to other pigments and the protein environment need to be taken into account. One can think of these protein vibrations as damping or driving

oscillations of pigment molecules. To account for all of these vibrations, one can consider the system as an electronic two-level system coupled to a collection of normal harmonic oscillator modes called the bath.

5. The bath

In general, microscopic knowledge of the system is not necessary. It is not necessary, for example, to know the exact positions or configurations of proteins in photosynthetic light harvesting complexes. We can replace the real bath with an infinite set of harmonic oscillators as long as the interaction with the relevant system is the same for the original and the new bath. The Hamiltonian of the system is then

$$H = |g\rangle H_g \langle g| + |e_i\rangle H_e \langle e_i| , \tag{2}$$

with

$$H_g = \sum_l \big[\frac{p_l^2}{2m_l} + \frac{1}{2} m_l \omega_l^2 q_l^2\big]$$

$$H_e = \hbar(E_{e_i} - E_g) + \sum_l \big[\frac{p_l^2}{2m_l} + \frac{1}{2} m_l \omega_l^2 (q_l - d_l)^2\big],$$

where m_l, q_l, p_l and ω_l are the mass, coordinate, momentum and frequency of the l^{th} oscillator and d_l is the distance by which the coordinate equilibrium of oscillator l is shifted upon electronic excitation of the pigment molecule.

Instead of coupling only to a single harmonic oscillator like in figure 2, the relevant system now couples to infinitely many independent harmonic oscillators. This model is known in literature as the Spin-Boson Model [2].

While it is clear from Eq. (2) how the bath is changed when the relevant system is excited, the opposite effect (change of the relevant system due to the bath) is not clear. In Markovian master equations based on the Spin-Boson Hamiltonian, the bath degrees of freedom enter the equation of motion only through two-times correlation functions of the form $C(t) = \langle q_n(t) q_n(0)\rangle$, where q_n is the bath coordinate coupled to the n^{th} pigment molecule and $\langle \ldots \rangle$ indicates ensemble averaging over the bath degrees of freedom. If one does not have any knowledge of the microscopic nature of the bath, it is hard to give a physical interpretation of these correlation functions from the derivation of the quantum master equations.

An equivalent way of thinking about the Spin-Boson Model is the so-called Multimode Brownian Oscillator model [2] (which one can consider as an extended Jaynes-Cummings model). In this model the relevant system is linearly coupled to a finite set of harmonic oscillators, like in Eq. (2). In turn, these harmonic oscillators (which we will call the primary oscillators) are coupled linearly to an infinite set of other harmonic oscillators (which we will call the free oscillators since they are not coupled to the relevant system). A schematic of this model is shown in figure 3. It is important to note that the bath correlation functions enter the master equation description through commutators of the relevant system-bath interaction Hamiltonian. Since the free oscillator coordinates don't appear in this interaction Hamiltonian, the only correlation functions in the master equations are therefore the ones containing the coordinates of the primary oscillators. These correlation functions are statistical parameters comparing the coordinates of the primary harmonic oscillators at two different times. These parameters are determined by the coupling of the primary oscillators to the free bath and the relevant system.

6. Effect of the bath on the relevant system

The bath provides both fluctuations and dissipation in accordance with the Fluctuation-dissipation theorem [1]. The primary oscillators are quantum Brownian oscillators (quantum

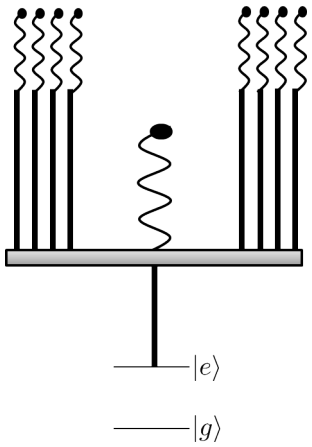


Figure 3. A schematic of the quantum Multimode Brownian Oscillator model. The straight vertical lines indicate linear coupling and the grey bar represents the primary oscillator.

oscillators subject to the fluctuating net force of an infinite set of other oscillators). The thermal fluctuations of the free oscillators cause the primary oscillator coordinates to fluctuate. Through the linear coupling between the primary oscillator coordinates and the relevant system, the excited state energy (and therefore the energy gap between the ground and excited state) of the pigment molecule also fluctuates. When the fluctuating energy gap of one two-level system corresponds to the energy gap of another two-level system, resonance energy transfer between the two-level systems is induced. In this way the bath facilitates energy transfer between pigment molecules.

The bath also acts as dissipator. When energy is transferred from one pigment molecule to a non-ground vibrational level of another pigment, the bath can dissipate the energy difference between the latter vibrational level and the ground level through the coupling to the free bath. Because the free bath has so many degrees of freedom, the energy is never transferred back to the system. This dissipation continues until the system as a whole reaches thermal equilibrium.

In order to show how the effect of bath fluctuation and dissipation on the relevant system can be characterised, let's suppose that, as shown in figure 3, we describe the system with only a single primary oscillator. This single oscillator is the link between the free bath and the relevant system, and the effect of these two subsystems on each other should therefore be channelled through the primary oscillator. All the dynamics of the bath which have an influence on the system can be therefore be obtained by a characterization of the primary oscillator coordinate. The bath correlation functions provide such a characterisation.

To get a better understanding of the meaning of these correlation functions, let us consider a specific correlation function; that of an overdamped Brownian oscillator [2]. This correlation function is often used in simulations because it can be calculated analytically and depends on only three physical parameters: the temperature, bath decorrelation time and the bath reorganisation energy (the energy released when the system relaxes to a ground vibrational state—see figure 2). This correlation function is a complex number for a quantum oscillator. As a very crude approximation, however, one can assume this correlation function (for each pigment molecule) to be an exponentially decreasing function like it is in the classical case. Let us now expand our relevant system to a number of pigment molecules with varying coupling between them. It was noted earlier that $\langle \ldots \rangle$ indicates an ensemble average. Suppose therefore that we prepare a very large number of initially identical systems and measure the product $q_n(t)q_n(0)$ in each of them. The correlation function $C(t) = \langle q_n(t)q_n(0) \rangle$ will then be an average of all those measurements. Assuming the crude approximation for the correlation function discussed above, the ensemble average of $C(t)$ will be non-zero immediately after electronic excitation of the i^{th} pigment molecule in the system. This means that all systems (and relevant systems) in the

ensemble behave similarly and the dynamics of the relevant system (which is also determined by an ensemble average) are driven in a specific direction by the bath. When the correlation functions decayed significantly, different members of the ensemble behave differently and there exist no net drive on the system due to the bath. It is clear now that the bath correlation functions can be thought of as parameters determining the strength and duration of the influence of the bath on the system.

7. Conlusion

We have described how a complex quantum system can be partitioned into a small relevant subsystem and a bath. The Spin-Boson Model was introduced to describe the bath and the interaction of the relevant system with the bath. The effects of the bath (fluctuation and energy dissipation) on the relevant system were illustrated by considering the Multimode Brownian Oscillator model, which is equivalent to the Spin-Boson Model. The statistical importance and the physical interpretation of the bath correlation functions were discussed. The correlation functions characterises the strength and duration of the bath's influence on the relevant system.

Acknowledgments

NIThePand the University of Pretoria are thanked for their financial support.

References

[1] Breuer H P and Petruccione F 2002 *The Theory of Open Quantum Systems* (New York: Oxford University Press)

[2] Mukamel S 1995 *Principles of Nonlinear Optical Spectroscopy* (New York: Oxford University Press)

[3] Anna J M, Scholes G D and Van Grondelle R 2014 *BioScience* **64** 14–25

[4] Chenu A and Scholes G D 2014 *Annu. Rev. Phys. Chem.* **66** 69–96

[5] Van Amerongen H, Valkunas L and van Grondelle R 2000 *Photosynthetic Excitons* (Singapore: World Scientific)

[6] Yen T C and Cheng Y C 2011 *Procedia Chem* **3** 211–21

[7] Croce R and van Amerongen H 2011 *J. Photochem. Photobiol.* B **104** 142 – 53

[8] Valkunas L, Abramavicius D and Mancal T 2013 *Molecular Excitation Dynamics and Relaxation: Quantum Theory and Spectroscopy* (Weinheim, Germany: John Wiley & Sons)

Neutrino mass hierarchy and CP phase measurement using atmospheric neutrino flux

Soebur Razzaque

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

E-mail: srazzaque@uj.ac.za

Abstract. Multi-megaton scale ice or water Cherenkov detectors with relatively low (sub-GeV) threshold energy can accumulate huge statistics of atmospheric neutrino data. With reasonable energy and angular reconstruction efficiency for the neutrino events, these data can be used to establish yet unknown neutrino mass hierarchy with high confidence. Leptonic CP phase can also be measured using atmospheric neutrino flux, once hierarchy is established and uncertainty on the flux and other neutrino parameters are better understood. Following up on previous work on this topic we will present the latest calculation results in light of recent developments.

1. Introduction

Neutrino mass hierarchy and the value of the leptonic CP phase are two outstanding problems in particle physics [1, 2]. Atmospheric neutrino flux measurements in large water/ice detectors can be used to determine the mass hierarchy, i.e. the sign of the atmospheric mass-square-difference $\pm\Delta m^2_{31}$, and the Dirac CP phase δ. Information on hierarchy and different δ values are encoded in the neutrino oscillation probabilities after propagation inside the earth (matter effect) [3, 4] .

It has been found recently that future detectors such as the *Precision IceCube Next Generation Upgrade* (PINGU) and *Oscillation Research with Cosmics in the Abyss* (ORCA) with $\sim$ 3 GeV threshold will have good sensitivity to determining the neutrino mass hierarchy [5, 6, 7]. However, measurement of δ would require a detector with larger effective volume and improved characteristics in the < 2 GeV range and in this context a future upgrade of PINGU (and also of ORCA), called Super-PINGU, was proposed in Ref. [8] with detailed estimate of sensitivity. Here we present latest development in determining mass hierarchy and measuring leptonic CP phase using atmospheric neutrino flux.

2. Neutrino Oscillograms

Atmospheric neutrinos, produced isotropically by cosmic-ray interactions in the upper atmosphere, penetrate through the earth to reach a neutrino detector. The amount of earth's interior mantle and core structure the neutrinos pass through depends on the arrival direction at the detector (zenith angle θ_z). This also determines how much oscillation is induced by the matter effect [9, 10]. Figure 1 shows oscillation probabilities of different neutrino flavors passing through the earth with different trajectories and energy. We use the earth's interior model from ref. [11]. These 2-D plots, also called the oscillograms of the earth [3] are useful to understand

effects of various neutrino oscillation parameters (and eventually determine them), such as the neutrino mass hierarchy and CP phase, on the probabilities.

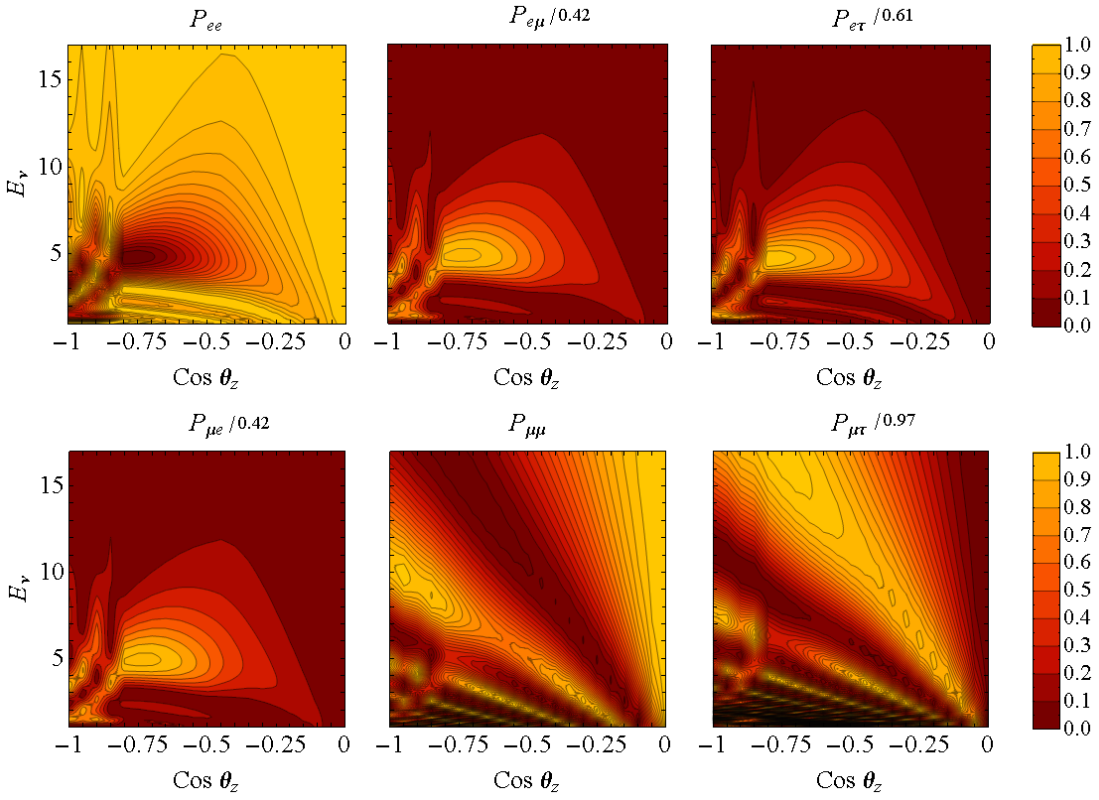


Figure 1. Neutrino oscillation probabilities in the E_ν–$\cos\theta_z$ plane for different oscillation channels with the contours representing lines of equal probabilities. The probabilities are normalized by their maximal values in the parameter space of the panels: $P_{\alpha\beta}/P_{\alpha\beta}^{max}$, with $P_{ee}^{max} = P_{\mu\mu}^{max} = 1$. E_ν is in GeV. Such plots are called neutrino oscillograms of the Earth. From Ref. [5]

3. PINGU and Super-PINGU

We compute atmospheric neutrino event rates for different flavors in the proposed PINGU and Super-PINGU detectors. The PINGU detector [6] will have 40 strings additional to the IceCube DeepCore strings with 60 Digital Optical Modules (DOMs) at 5 m spacing in each string. The effective mass, i.e. density times the effective volume, of PINGU can be parametrized as [8]

$$\begin{aligned} \rho V_{\text{eff},\mu}(E_\nu) &= 3.0\,[\log(E_\nu/\text{GeV})]^{0.61}\ \text{Mt} \\ \rho V_{\text{eff},e}(E_\nu) &= 3.1\,[\log(E_\nu/\text{GeV})]^{0.60}\ \text{Mt}, \end{aligned} \tag{1}$$

respectively for ν_μ and ν_e. For the Super-PINGU detector we use an effective mass (both for ν_μ and ν_e) parameterized as [8]

$$\rho V_{\text{eff}}(E_\nu) = 2.6\,[\log(E_\nu/\text{GeV}) + 1]^{1.32}\ \text{Mt}, \tag{2}$$

which can be realized for a total of 126 strings and 60 DOMs per string. This gives an effective mass of $\sim$ 2.8 Mt at $\sim$ 1-2 GeV, which is 4 times larger than PINGU.

Since the most sensitive energy range for hierarchy determination is $\sim$ 4–14 GeV [5], PINGU can determine hierarchy rather efficiently but with little sensitivity to the CP phase. The energy range most sensitive to the CP phase is below $\sim$ 2 GeV [8], where PINGU has very little sensitivity. A detector such as Super-PINGU with large volume at low energy will be required to measure the CP phase.

4. Distinguishability between hierarchies and CP phases

The number of neutrino events for a particular flavor $\alpha = e, \mu$ with energies and zenith angles in small bins $\Delta(E_\nu)$ and $\Delta(\cos\theta_z)$ marked by subscript ij can be calculated as

$$N_{ij,\alpha} = 2\pi N_A \rho T \int_{\Delta_i \cos\theta_z} d\cos\theta_z \int_{\Delta_j E_\nu} dE_\nu \; V_{\text{eff},\alpha}(E_\nu) d_\alpha(E_\nu, \theta_z). \tag{3}$$

Here T is the exposure time, N_A is the Avogadro's number. The density of events of type α, d_α (the number of events per unit time per target nucleon), is given by $d_\alpha(E_\nu, \theta_z) = \sigma_\alpha \Phi_\alpha + \bar{\sigma}_\alpha \bar{\Phi}_\alpha$ in terms of fluxes at the detector, and $\Phi_\alpha = \Phi^0_\mu P_{\mu\alpha} + \Phi^0_e P_{e\alpha}$ with corresponding oscillation probabilities $P_{\mu\alpha}$ and $P_{e\alpha}$, and νN interaction cross sections σ_α and $\bar{\sigma}_\alpha$. The original muon and electron neutrino fluxes at the production are $\Phi^0_\mu = \Phi^0_\mu(E_\nu, \theta_z)$ and $\Phi^0_e = \Phi^0_e(E_\nu, \theta_z)$. We use the atmospheric flux model in ref. [12] which fits available data rather well.

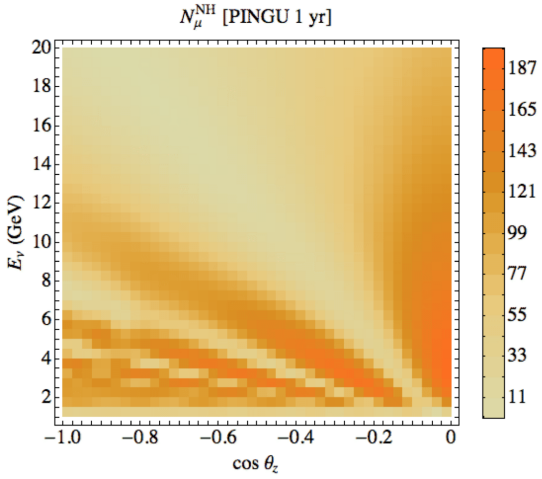


Figure 2. Distribution of $\nu_\mu + \bar{\nu}_\mu$ events/year in PINGU for NH in the E_ν–$\cos\theta_z$ plane. From ref. [5]

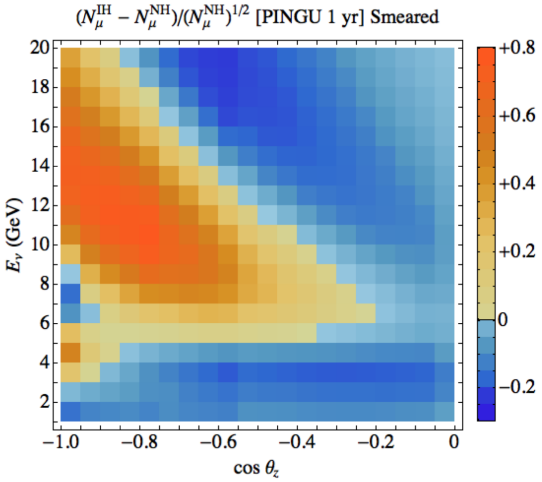


Figure 3. Distribution of S_{ij} for $\nu_\mu + \bar{\nu}_\mu$ events in PINGU between the NH and IH in the E_ν–$\cos\theta_z$ plane. From ref. [5]

Figure 2 shows the distributions of expected $\nu_\mu + \bar{\nu}_\mu$ events/year in PINGU in the E_ν–$\cos\theta_z$ plane. Accumulation huge statistics ($\sim 70,000$/year for PINGU and $\sim 90,000$/year for Super-PINGU) can be used to discriminate between the mass hierarchies and different CP phases.

To determine mass hierarchy with PINGU we compute the distributions of $\nu_\mu + \bar{\nu}_\mu$ and $\nu_e + \bar{\nu}_e$ events for the normal hierarchy (NH) and inverted hierarchy (IH), keeping all other oscillation parameters fixed, and take difference between the distributions in the E_ν–$\cos\theta_z$ plane. Since there are errors associated with reconstructing the true neutrino energy and directions, we smear the ideal distributions with the energy and angular resolution functions of the detector to mimic the real situation. To estimate the sensitivity to IH from NH (which is assumed true) we employ a distinguishability parameter defined as [5]

$$S_{ij}(f) = (N^{\text{IH}}_{ij} - N^{\text{NH}}_{ij})/\sigma_{ij}(f), \tag{4}$$

where $N_{ij}^{\rm NH}$ and $N_{ij}^{\rm IH}$ are the reconstructed number of events in the ij-th bin in the E_ν–$\cos\theta_z$ plane for the NH and IH, respectively, and $\sigma_{ij}^2(f) = N_{ij}^{\rm NH} + (fN_{ij}^{\rm NH})^2$ is the total error in the ij-th bin. Parameter f is a measure of uncorrelated systematic errors [5]. Figure 3 shows the expected distribution of S_{ij}, with $f = 0$ for 1-year PINGU data. The total distinguishability

$$S_\sigma = \sqrt{\sum_{ij} S_{ij}^2} \tag{5}$$

is a quick measure of statistical significance. We found that hierarchies can be determined in this approach with $S_\sigma \sim 3$–4 within 3 years of PINGU operation [5].

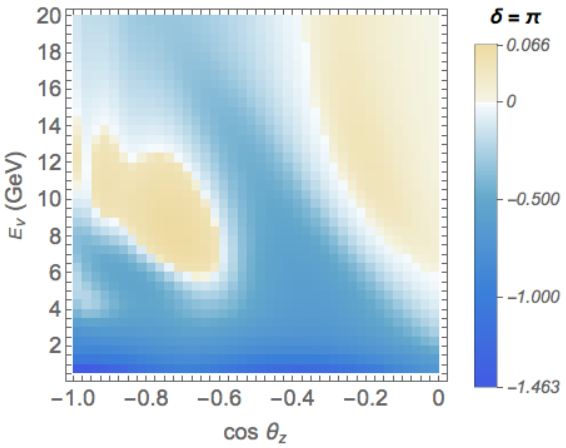


Figure 4. Distribution of S_{ij} for $\nu_\mu + \bar{\nu}_\mu$ events in Super-PINGU between $\delta = \pi$ and $\delta = 0$ in the E_ν–$\cos\theta_z$ plane. From ref. [8]

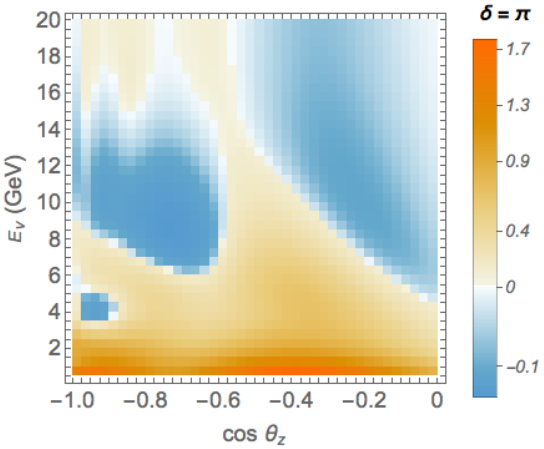


Figure 5. Distribution of S_{ij} for $\nu_e + \bar{\nu}_e$ events in Super-PINGU between $\delta = \pi$ and $\delta = 0$ in the E_ν–$\cos\theta_z$ plane. From ref. [8]

In order to measure the CP phase with atmospheric neutrinos, we calculate neutrino events in Super-PINGU by varying δ and compare with $\delta = 0$, keeping all other oscillation parameters fixed. As we did for hierarchy, we compute the distributions of $\nu_\mu + \bar{\nu}_\mu$ and $\nu_e + \bar{\nu}_e$ events for $\delta = 0$ and $\delta \neq 0$ and take difference between the distributions in the E_ν–$\cos\theta_z$ plane. To estimate the sensitivity of measuring a CP phase different from zero we employ a distinguishability parameter, similar to the hierarchy case, defined as

$$S_{ij}(f) = (N_{ij}^{\delta} - N_{ij}^{0})/\sigma_{ij}(f), \tag{6}$$

where N_{ij}^{δ} and N_{ij}^{0} are the reconstructed number of events in the ij-th bin in the E_ν–$\cos\theta_z$ plane for δ and $\delta = 0$, respectively, and $\sigma_{ij}^2(f) = N_{ij}^{0} + (fN_{ij}^{0})^2$ is the total error in the ij-th bin. The total distinguishability is calculated using the same formula, Eq. (5). Figure 4 shows the distinguishability S_{ij} for $\nu_\mu + \bar{\nu}_\mu$ events with $f = 0$ for $\delta = \pi$ and $\delta = 0$ using 1-year of Super-PINGU data. Normal mass hierarchy is assumed. The shape of the distributions, specially their domain structures, is largely explained as due to grids of solar, atmospheric and interference magic lines in the E_ν–$\cos\theta_z$ plane. The oscillation probabilities are independent of δ along these lines, thus separating regions of same sign distinguishability. Figure 5 shows S_{ij} distributions for $\nu_e + \bar{\nu}_e$ events.

The uncertainties associated with atmospheric neutrino flux, νN cross section, effective volume, etc. affect neutrino event distributions across bins in the E_ν–$\cos\theta_z$ plane. We include

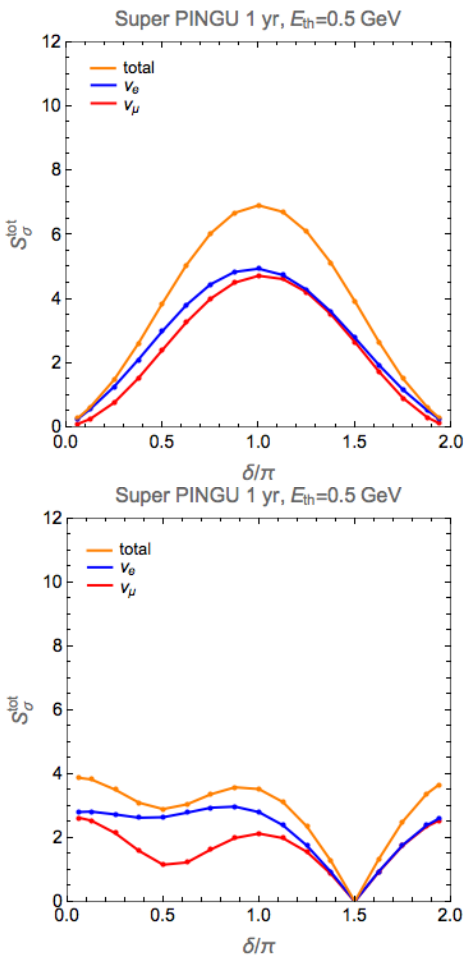


Figure 6. Total distinguishability to measure the CP phase with Super-PINGU 1-year data in the ν_μ and ν_e channels. The top panel is for true $\delta = 0$ and the bottom panel is for true $\delta = 3\pi/2$.

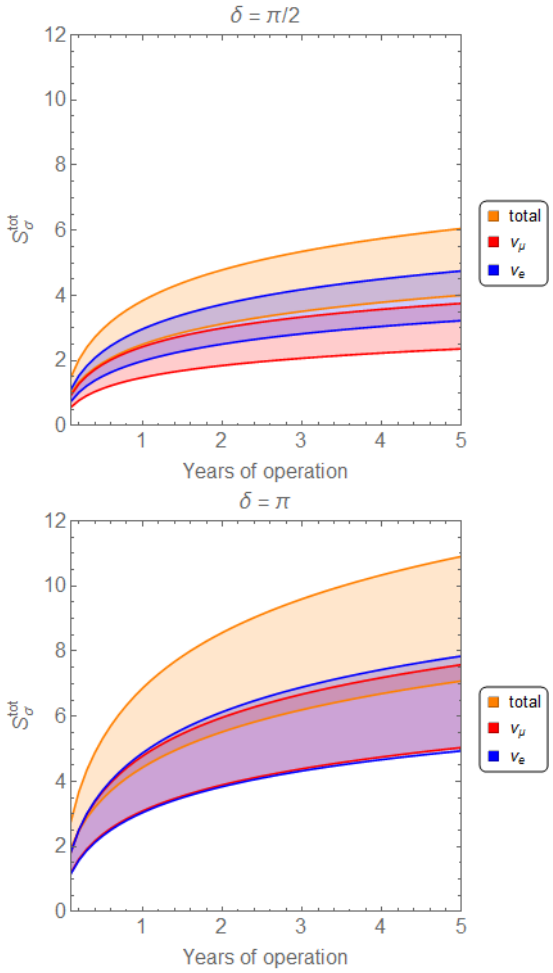


Figure 7. Time to distinguish CP phase $\delta = \pi/2$ (top panel) and $\delta = \pi$ (bottom panel) from $\delta = 0$ using Super-PINGU data, with given level of the total distinguishability $S_\sigma^{\rm tot}$. The shaded bands correspond to different levels of uncertainties.

effects of these correlated uncertainties in our calculation with analogy to the pull method in χ^2 analysis. In particular we minimize the following distinguishability parameter

$$S_\sigma^{tot}(\xi_k) = \left[\sum_{l=e,\mu} \sum_{ij} \frac{[N_{ij,l}^{\delta}(\xi_k) - N_{ij,l}^{0}(\xi_k^{st})]^2}{\sigma_{ij,l}^2} + \sum_k \frac{(\xi_k - \xi_k^{st})^2}{\sigma_k^2} \right]^{1/2}, \tag{7}$$

where ξ_k are the pull variables and ξ_k^{st} are their standard values. The event distributions with varying ξ_k are calculated as

$$N_{ij,l}^{\delta}(\xi_k) = \alpha z_l (E/2\ \text{GeV})^{\eta} [1 + \beta(0.5 + \cos\theta_z)] N_{ij,l}^{\delta}(\xi_k^{st}), \tag{8}$$

where α is the overall normalization factor with the error $\sigma_\alpha = 0.2$, z_l is the flux (flavor) ratio uncertainty ($z_e \equiv 1$ for ν_e events), with the error $\sigma_z = 0.05$; η is the energy tilt parameter with

$\sigma_\eta = 0.1$; β is the zenith angle tilt with $\sigma_\beta = 0.04$. Figure 6 shows the S_σ^{tot} minimized over (ξ_k) for different correlated uncertainties. The upper (lower) panel corresponds to the case when the true CP is $\delta = 0$ $(3\pi/2)$. A threshold energy of 0.5 GeV, $f = 2.5\%$ as well as all correlated uncertainties have been assumed. Note that the contributions of ν_e and ν_μ channels to S_σ^{tot} are comparable.

Figure 7 shows the years of Super-PINGU data required to achieve different levels of distinguishability in case $\delta = \pi/2$ (top panel) and $\delta = \pi$ (bottom panel). The shaded bands correspond to different levels of flavor misidentifications in the detector, i.e., what fraction of $\nu_m u$ events are reconstructed as ν_e events and vice versa. The upper edge of the bands correspond to no flavor misidentification and the lower edge of the bands correspond to 20% flavor misidentification. Preliminary PINGU simulations show that flavor misidentification can be kep at 10%–30% level [6].

5. Summary

In summary, we find that the proposed future PINGU upgrate of the IceCube neutrino observatory will be able to determine the neutrino mass hierarchy using atmospheric neutrino flux, either NH or IH, within $\sim$ 3 years of its operation. The hypothetical Super-PINGU detector will be sensitive to sub-GeV neutrinos and can distinguish the CP phase $\delta \geq \pi/2$ from zero within $\sim$ 5 years of its operation. We took into account various systematic effects in our computation as close as possible to the realistic scenario. Systematic effects dominate in both PINGU and Super-PINGU. Determination of hierarchy with PINGU can be very quick compared to the reactor/accelerator experiments [13]. Measurement of CP phase with Super-PINGU using atmospheric neutrinos can be competitive with long baseline neutrino experiments [14].

Acknowledgments

This work was supported in part by the National Research Foundation (South Africa) Grant No. 87823 (CPRR) to S.R.

References

[1] Branco G C, Felipe, R G and Joaquim F R 2012 *Rev. Mod. Phys.* **84** 515
[2] Nunokawa H, Parke, S J and Valle J W F 2008 *Prog. Part. Nucl. Phys.* **60** 338
[3] Akhmedov E Kh, Maltoni M and Smirnov A Yu 2008 *JHEP* **0806** 072
[4] Mena O, Mocioiu I and Razzaque S 2008 *Phys. Rev. D* **78** 093003
[5] Akhmedov E Kh, Razzaque S and Smirnov A Yu 2013 *JHEP* **1302** 082; *erratum ibid* 2013 *JHEP* **1307** 026
[6] Aartsen M G *et al.* [IceCube-PINGU Collaboration] 2014 Letter of Intent: The Precision IceCube Next Generation Upgrade (PINGU) *Preprint* arXiv:1401.2046 [physics.ins-det].
[7] Katz U F [KM3NeT Collaboration] 2014 The ORCA Option for KM3NeT *Preprint* arXiv:1402.1022 [astro-ph.IM].
[8] Razzaque S and Smirnov A Yu 2015 *JHEP* **1505** 139
[9] Wolfenstein L 1978 *Phys. Rev. D* **17** 2369
[10] Mikheev S P and Smirnov A Yu 1985 *Sov. J. Nucl. Phys.* **42** 913 [*Yad. Fiz.* **42** 1441]
[11] Dziewonski A M and Anderson D L 1981 *Phys. Earth Planet. Interiors* **25** 297
[12] Honda M, Kajita, T, Kasahara, K and Midorikawa, S 1995 *Phys. Rev. D* **52** 4985
[13] Blennow M and Schwetz T 2013 *JHEP* **1309** 089
[14] Adams C *et al.* [LBNE Collaboration] 2013 The Long-Baseline Neutrino Experiment: Exploring Fundamental Symmetries of the Universe *Preprint* arXiv:1307.7335 [hep-ex]

Projection operators in the theory of open quantum systems

V Semin, F Petruccione

University of KwaZulu-Natal and National Institute for Theoretical Physics, University Road, Durban 4001, South Africa

E-mail: semin@ukzn.ac.za

Abstract. We study different forms of projection operators and their application to open quantum systems. In particular, we show that applying a special class of projection operators to open systems may lead to non-linear dynamical equations, while other projection operators always lead to linear equations. We discuss general features of linear and non-linear dynamical equations. The main properties of different projectors are illustrated by examples of a qubit in a thermal bath and two interacting qubits in a common environment.

1. Introduction

The theory of open quantum systems deals with systems interacting with their environment. This interaction leads to continuous interchange of energy and entropy between a system and its environment and, eventually, to relaxation and decoherence in the system. The general strategy to study open systems is based on an assumption that an open system and its environment together form a closed system. Then, the evolution of the large closed system is investigated with the help of traditional methods of quantum mechanics. Finally, tracing out the degrees of freedom of the environment gives the sought-for time evolution of the open system. Mathematically the evolution of an open quantum system is governed by the following expression

$$\rho_S(t) = \mathrm{Tr}_E \left\{ U(t,t_0)\rho_{SE}(t_0)U^\dagger(t,t_0) \right\}, \tag{1}$$

where ρ_S is the density operator of the open system, $\rho_S E$ is the density operator of the open system and the environment, $U(t,t_0)$ is the evolution operator of the large closed system, and Tr_E is a partial trace over degrees of freedom of the environment.

The Eq. (1) provides comprehensive information about the evolution of an open system. However, the Eq. (1) is hardly satisfactory for the following reasons. First, the environment usually contains infinitely many degrees of freedom and, thus, one has to solve an infinite system of differential equations to find the evolution operator. This is impossible in the most cases. Second, one is often interested only a defined set of observables, characterising an open system, and Eq. (1) brings excessive details of the evolution.

In this paper we study open quantum systems with the help of projection operators and show that the projection operator technique allows to overcome some of the limitations of Eq. (1) mentioned above. The paper is organised as follow. In Sec. 2 we provide necessary information about the projection operator technique. In Sec. 3 we discuss some ideas leading to concrete

forms of projection operators. In Sec. 4 and 5 we apply different projection operators to two models of a qubit in thermal bath and two interacting qubits in a common environment. Finally, in Sec. 6 we conclude.

2. Projection operators and dynamical equation

The main idea behind the use projection operators is to divide a mathematical object, such as the density operator, into several parts and to study each part separately. Usually, in the theory of open quantum systems, the density operator contains some relevant information about the open system itself and irrelevant information about the environment. Of course, one wants to study only the relevant information about the system and ignore irrelevant one. Let $\mathcal{P}$ be a projection operator, $\mathcal{P}^2 = \mathcal{P}$, extracting the relevant part of the density operator,

$$\rho_{\rm rel}(t) = \mathcal{P}\rho(t), \tag{2}$$

where $\rho_{\rm rel}(t)$ is the relevant part of the full density operator $\rho(t)$. Let us assume that the following relation holds

$$\frac{\partial}{\partial t}\mathcal{P}\rho(t) = \mathcal{P}\frac{\partial}{\partial t}\rho(t) = \frac{\partial}{\partial t}\rho_{\rm rel}(t). \tag{3}$$

Under the assumption (3) one can build a fruitful theory and successfully describe the relevant part (2) of the density operator. Indeed, acting the projection operator onto both sides of the Liouville equation $\frac{\partial}{\partial t}\rho(t) = -i[H(t), \rho(t)] = \mathcal{L}(t)\rho(t)$, where $H(t)$ is the system Hamiltonian and $\mathcal{L}(t)$ is the Liouville superoperator, one can proceed in the standard way [1] and derive two different types of master equations, namely, the time-convoluted Nakajima-Zwanzig master equation [2, 3] and the so-called time-convolutionless (TCL) master equation [1].

The celebrated Nakajima-Zwanzig master equation has the following general structure

$$\frac{\partial}{\partial t}\mathcal{P}\rho(t) = \mathcal{P}\mathcal{L}(t)\mathcal{P}\rho(t) + \mathcal{P}\mathcal{L}(t)\mathcal{G}(t,t_0)\mathcal{Q}\rho(t_0) + \mathcal{P}\mathcal{L}(t)\int_{t_0}^{t}\mathcal{G}(t,s)\mathcal{Q}\mathcal{L}(s)\mathcal{P}\rho(t)ds, \tag{4}$$

where $\mathcal{G}(t,s) = \exp_-\left[\int_s^t \mathcal{Q}\mathcal{L}(s')ds'\right]$, $\mathcal{Q} = 1 - \mathcal{P}$, and $\exp_\pm$ is chronological (antichronological) exponent, while the TCL master equation reads

$$\frac{\partial}{\partial t}\mathcal{P}\rho(t) = \mathcal{K}(t)\mathcal{P}\rho(t) + \mathcal{I}(t)\mathcal{Q}\rho(t_0), \tag{5}$$

where $\mathcal{K}(t) = \mathcal{P}\mathcal{L}(t)[1 - \Sigma(t)]^{-1}\mathcal{P}$, $\mathcal{I}(t) = \mathcal{P}\mathcal{L}(t)[1 - \Sigma(t)]^{-1}\mathcal{G}(t,t_0)\mathcal{Q}$, $\Sigma(t) = \int_{t_0}^{t}\mathcal{G}(t,s)\mathcal{Q}\mathcal{L}(s)\mathcal{P}G(t,s)ds$, and $G(t,s) = \exp_+\left[-\int_s^t \mathcal{L}(s')ds'\right]$ is the inverse propagator.

Eqs. (4)-(5) encode most types of master equations. Moreover, with an appropriate choice of a projection operator these equations give a powerful tool to study only a relevant part of the density operator. Both above master equations are equivalent. In this paper we deal with the TCL master equation (5).

3. Explicit form of projection operators

There exist several type of the projection operators, which can be divided into two classes, namely, algebraic and functional projection operators. Action of the both type of projection operators extracts only the relevant part of the density operator characterising the system, but the mathematical structure of the relevant density operator (2) is different.

3.1. Functional projection operators

Let us choose a set of relevant dynamical variables or elements of the density operator P_m, characterising the system. This set carries all necessary information about a quantum system. We also assume that the relevant density operator (2) has some specific functional form

$$\rho_{\rm rel}(t) = \mathcal{F}(P_m, t). \tag{6}$$

To avoid contradictions we demand that the following self-consistency condition holds

$$\langle P_m \rangle^t = \langle P_m \rangle^t_{\rm rel} \equiv {\rm Tr}(P_m \rho_{\rm rel}(t)), \tag{7}$$

where the superscript means time-dependence. In other words, we want that the real averages, following from the quantum Heisenberg or Liouville equations, coincide with the averaging with respect to the relevant density operator. Generally, the self-consistency conditions are a system of n algebraic equations, where n is the number of relevant variables, and formally, the real averages $\langle P_m \rangle^t$ are known.

As example, we can form a Gibbs-like relevant density operator in the form

$$\rho_{\rm rel} = \exp[-\sum_m F_m(t) P_m]/{\rm Tr}\exp[-\sum_m F_m(t) P_m], \tag{8}$$

and the parameters $F_m(t)$ are fixed with the help of the self-consistency conditions (7). In this case the parameters F_m have transparent physical meaning as non-equilibrium thermodynamical parameters and the self-consistency conditions (7) connect dynamical and thermodynamical variables.

The Kawasaki-Gunton projection operator which replaces any density operator by the relevant one (6) has the following form

$$\mathcal{P}(t)A = \rho_{\rm rel}(t){\rm Tr}A + \sum_m \Big\{ {\rm Tr}(AP_m) - ({\rm Tr}A)\langle P_m \rangle^t \Big\} \frac{\partial \rho_{\rm rel}(t)}{\partial \langle P_m \rangle^t}. \tag{9}$$

The main properties of the Kawasaki-Gunton projection operator (9) can be found in [4, 5]. Here we notice that the Kawasaki-Gunton projection operator can be used to take into consideration ideas of non-equilibrium thermodynamics. In particular, using the principle of maximal entropy one can choose different types of entropy and build other types of relevant operators distinguished from (8). In such cases, the self-consistency conditions determine another type of generalised thermodynamics.

Notice that Eqs. (4) and (5) with the Kawasaki-Gunton projection operator (9) represent a non-linear system of equations in general. Nevertheless, due to direct connection of dynamics and thermodynamics of quantum systems the method based on the functional projection operator allows to study quantum systems more carefully and to determine more information about the system without additional non-trivial calculation in terms of only relevant variables. This fact may be very useful for some applications.

3.2. Algebraic projection operators

Algebraic projection operators iare used to extract the relevant degree of freedom without any additional assumption about a system. The general form of the algebraic projection operator is

$$\mathcal{P}A = \sum_{ij} {\rm Tr}(E_{ij} A) E^\dagger_{ij} / {\rm Tr} E_{ij} E^\dagger_{ij}, \tag{10}$$

where ${\rm Tr} E_{ij} E^\dagger_{kl} \sim \delta_{ik}\delta_{jl}$ is any set of orthogonal operators.

Particular cases of the projection operator (10) are widely used in different applications, especially in the theory of open quantum system. The most commonly used projection operator [1] has the form

$$\mathcal{P}A = (\mathrm{Tr}_E A) \otimes \rho_E(0), \tag{11}$$

where the partial trace is taken over the environmental degrees of freedom and the $\rho_E(0)$ is some fixed state of the environment. Clearly, that the projection operator (11) is a special case of the general form (10).

Below we show two examples how to use the projection operator formalism.

4. Two-level system in external field

We study this model with the help of the functional projection operator. The Hamiltonian of the system is $H = \omega_0\sigma_z + (\Omega(t)\sigma_+ + \Omega^*(t)\sigma_-) + \sum_j \omega_j b_j^\dagger b_j + \sum_j g_j(\sigma_- b_j^\dagger + \sigma_+ b_j)$, where $\Omega(t) = \frac{\Omega}{2}e^{-i\omega_L t}$ is monochromatic classical field with frequency ω_L and Rabi frequency Ω, $\sigma_\pm = (\sigma_x \pm i\sigma_y)$, and the σ_i are spin matrices. In this case the open system has a finite number of degrees of freedom. As relevant variables we can choose the components of the Bloch vector $\sigma_-, \sigma_+, \sigma_z$. The relevant distribution has the form

$$\rho_{\mathrm{rel}} = \frac{1}{Z_1 Z_2}\exp[-F_1(t)\sigma_+ - F_2(t)\sigma_z - F_3(t)\sigma_-]\exp[-\beta\sum_j \omega_j b_j^\dagger b_j], \tag{12}$$

where $Z_1 = \mathrm{Tr}\exp[-F_1(t)\sigma_+ - F_2(t)\sigma_z - F_3(t)\sigma_-] = 2\cosh(\frac{1}{2}\sqrt{4F_1(t)F_3(t) + F_2(t)^2})$ and $Z_2 = \mathrm{Tr}\exp[-\sum_j \omega_j b_j^\dagger b_j]$. The solution of the self-consistency conditions (7) is [5]

$$F_1(t) = -\langle\sigma_-\rangle^t R(t),\ F_2(t) = -2\langle\sigma_z\rangle^t R(t),\ F_3(t) = F_1^*(t), \tag{13}$$

where we introduced the functions $R(t) = \mathrm{arctanh}(2X)/X$, and $X = \sqrt{|\langle\sigma_-\rangle^t|^2 + (\langle\sigma_z\rangle^t)^2}$. Thus, the non-equilibrium entropy $S = \mathrm{Tr}\rho\log\rho$ of the system has the form

$$S(t) = -2X^2 R(t) + \ln 2 - 1/2\ln(1 - 4X^2) + S_{\mathrm{eq}}, \tag{14}$$

where S_{eq} is the equilibrium entropy of the bath.

The dynamics of the two-level system is governed by the master equation (5). For the considering model the equations up to the second order in the coupling constant are written as [5]

$$\frac{\partial\langle\sigma_z\rangle^t}{\partial t} = i(\Omega\langle\sigma_-\rangle^t - \Omega\langle\sigma_+\rangle^t) - \langle\sigma_z\rangle^t\left(f(t,\beta) + f^*(t,\beta)\right) - \frac{1}{2}\left(f(t) + f^*(t)\right), \tag{15}$$

$$\frac{\partial\langle\sigma_+\rangle^t}{\partial t} = -2i\Omega\langle\sigma_z\rangle^t - f(t,\beta)\langle\sigma_+\rangle^t. \tag{16}$$

In the above equations we introduced the following correlation functions

$$f(t,\beta) = \int_0^t dt' \int_0^\infty d\omega J(\omega)\coth\left(\frac{\beta\omega}{2}\right)e^{i(\omega-\omega_0)t},$$

$$f(t) = f(t,\infty) = \int_0^t dt' \int_0^\infty d\omega J(\omega)e^{i(\omega-\omega_0)t},$$

The dynamical equations (15)-(16) are identical to the usual result derived with the help of the TCL master equation [1] with traditional projection operator (11). For this model the two approaches give the same results. Nevertheless, due to the self-consistency conditions we have additional information about the thermodynamics of the open quantum system.

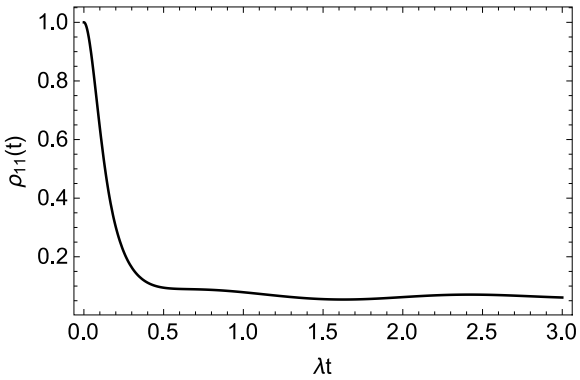


Figure 1. The evolution of the collective excited state $\rho_{11}(t)$ of two interacting qubits in the common thermal bath. Parameters in the system are $V = 0.6\lambda$, $\Omega = 10\lambda$, $\omega = 2\lambda$, $\beta = 0.3$, $\alpha_1 = 0.4 + 0.3i$, $\alpha_2 = 0.5 + 0.2i$

5. Two dipole-dipole interacting qubits

In this section we consider more complicated example. The Hamiltonian for the system is written as

$$H = H_0 + H_{12} + H_B + H_{int}, \tag{17}$$

where $H_0 = \nu \sum_i \sigma_z^i$ is the free qubits Hamiltonian, $H_B = \sum_j \omega_j b_j^\dagger b_j$ is the free Hamiltonian of the bath, $H_{12} = V(\sigma_+^1\sigma_-^2 + \sigma_+^2\sigma_-^1)$ is the Hamiltonian of the qubit interactions and $H_{int} = \sum_i \sum_j g_j b_j \sigma_+^i \alpha_i + \text{h.c.}$ is the qubits-bath interaction Hamiltonian, σ^i is the Pauli matrices for the ith qubit, b_i is the annihilation operator of the jth oscillator in the bath, ν and ω_j are the transition frequency of the qubits and the jth oscillator in the bath, correspondingly, V is the constant of dipole-dipole interaction, g_j is the coupling constant of the qubit and jth oscillator in the bath, α_i are the geometrical factors mark the position of the ith qubit.

First of all transform the Hamiltonian (17) to the interaction picture

$$H_I(t) = e^{i(H_0+H_{12}+H_B)t} H_{int} e^{-i(H_0+H_{12}+H_B)t} = (P_+^1 R(\alpha_2, -\alpha_1) + P_-^1 R(\alpha_2, \alpha_1))\sigma_+^2 B(t) e^{i\nu t} \tag{18}$$
$$+(P_+^2 R(\alpha_1, -\alpha_2) + P_-^2 R(\alpha_1, \alpha_2))\sigma_+^1 B(t) e^{i\nu t} + \text{h.c.} = K(t) e^{i\nu t} B(t) + K^\dagger(t) e^{-i\nu t} B^\dagger(t),$$

where $P_i^+ = \sigma_+^i \sigma_-^i$, $P_i^- = \sigma_-^i \sigma_+^i$ and $R(\alpha, \beta) = \alpha \cos(tV) + i\beta \sin(tV)$, $K(t) = (P_+^1 R(\alpha_2, -\alpha_1) + P_-^1 R(\alpha_2, \alpha_1))\, \sigma_+^2 + (1 \leftrightarrow 2)$, $B(t) = \sum_j g_j b_j$.

The master equation (5) up to the second order with the projection operator (11) for the factorised initial conditions can be written as

$$\dot\rho_S(t) = \int_0^t dt_1 \Big\{ [K^\dagger(t_1)\rho_S(t)K(t) - K(t)K^\dagger(t_1)\rho_S(t)]F$$
$$+ [K(t_1)\rho_S(t)K^\dagger(t) - K^\dagger(t)K(t_1)\rho_S(t)]R + \text{h.c} \Big\}, \tag{19}$$

where $F = \int_0^\infty d\omega J(\omega)(\coth(\beta\omega/2) + 1)/2 e^{i(\nu-\omega)(t-t_1)}$ and $R = \int_0^\infty d\omega J(\omega)(\coth(\beta\omega/2) - 1)/2 e^{-i(\nu-\omega)(t-t_1)}$.

The dynamics of the collective exited state for the spectral density of the bath $J(\omega) = \lambda\omega \exp[-\omega/\Omega]$ following from the TCL master equations is shown in Fig.1. One can see that the exited state relaxes to the equilibrium quite fast and after that the curve oscillates with small amplitude around the equilibrium. These oscillations may be considered as a signature of the non-Markovian evolution.

6. Conclusion

In this paper we studied different types of projection operators in the context of the theory of open quantum systems. We shown We have shown that general structure of a dynamical equation does not depend on concrete form of a projection operator. There exist two possible forms of the projection operators, which we called algebraic and functional. The functional projection operators lead in general to non-linear dynamical equation and enlighten thermodynamical properties of the system. The algebraic projection operators always lead to linear dynamical equations. We illustrated by two examples use of different types of the projection operators and indicated features given by the corresponding techniques.

References

[1] Breuer H-P and Petruccione F 2002 *The Theory of Open Quantum Systems* (Oxford: Oxford University Press)
[2] S. Nakajima, *Prog. Theor. Phys.* 1958 **20**, 948
[3] R. Zwanzig, *J. Chem. Phys.* 1960 **38**, 1338
[4] D. Zubarev, V. Morozov and G.Röpke 1992 *Statistical Mechanics of Non-equilibrium Processes* vol. 1 (Berlin: Akademie Verlag)
[5] V. Semin and F. Petruccione, *Phys. Rev.* A 2014 **90** 052112

Hypothesising the effects of Higgs portal dark matter in particle colliders

Stefan von Buddenbrock

School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africa

E-mail: stef.von.b@cern.ch

Abstract. The Higgs field mass term in the Standard Model is unique. While all of the other interaction terms in the Standard Model are associated with strictly renormalisable dimension 4 operators (and therefore having marginal couplings), the Higgs field mass term has a coupling of dimension 2. This allows us to explore the possibility of the Higgs boson having decay channels consisting of particles being $SU(3) \times SU(2) \times U(1)$ singlets, meaning that they do not interact with any Standard Model particles apart from the Higgs. We could treat these particles as candidates in a field of study which is known as Higgs portal dark matter. In order to test this possibility, a model independent framework has been developed in the form of a Lagrangian consisting of extensions to the Standard Model: a heavy scalar H and a non-interacting dark matter candidate χ. The implications of this model are considered where a Monte Carlo study has been performed on the process $gg \to H \to h\chi\chi$ and hh. Applying this to ATLAS results for comparison with measured Higgs p_T, it is found that a good fit is found for the mass points $m_\chi = 60$ GeV and $m_H = 275$ GeV.

1. Opening the Higgs Portal

Up until this point, the Standard Model (SM) of particle physics has provided the most accurate prediction about the dynamics of elementary particles, by combining manifestations of QCD and Electroweak Yang-Mills theories [1, 2, 3, 4], most notably using the Higgs mechanism [5]. There are, however, obvious shortcomings in the SM. In particular, it is well understood that there are no SM particles which are viable dark matter (DM) candidates – that is, SM particles are all *visible* in that we have the means to detect them directly. The compelling astrophysical evidence for the existence of DM has yet to be explained by a convincing argument in particle physics terms.

In the literature, one can find a plethora of hypotheses with the intention of presenting a theoretically motivated DM candidate as a particle (for a thorough review, see reference [6]). While most of these hypotheses are model dependent, this study presents a framework for a DM candidate which is based on a small class of DM models called *Higgs portal* models [7, 8, 9], which can largely be explained without having to resort to some higher theoretical explanation.

Higgs portal models, examples of which were first comprehensively presented by Brian Patt and Frank Wilczek [10], essentially make use of the Higgs boson's unique coupling to massive particles in order to motivate the possible existence of states which do not interact with SM particles. The Higgs term in the SM Lagrangian is unique in that it does not contain a marginal (or dimensionless) coupling. While all of the other interaction terms contain dimension 4

operators (and therefore marginal couplings), the Higgs term has an operator with dimension 2 forcing its coupling, μ^2, to have dimension 2 also. This means that we could allow the Higgs field to couple to new fields which are SM $SU(3) \times SU(2) \times U(1)$ singlets. SM singlets will not interact with SM particles at all, giving the possibility to add DM candidates[1] which interact only through the Higgs boson.

A key example of a Higgs portal model is that of a complex scalar singlet extension to the SM (CxSM) [7]. The basic idea is that a CxSM adds two new massive scalar degrees of freedom to the SM. One might be thought of as a new scalar boson which interacts with the SM similarly to the Higgs boson, and the other could be thought of as a DM candidate. While there is plenty of rich phenomenology linked to this model (most of which can be found in reference [7]), I only mention this example as a theoretical case study for the framework described below.

2. A Minimal Extension to the Standard Model

In this study we do not consider a theoretical construction of a DM model, but a framework containing two new particles. This is experimentally motivated, since we see experimental results which could be explained most simply be introducing two new particles (this will be elaborated on in the following section).

Firstly, we introduce a Higgs-like boson, which interacts with SM particles in the same way as the Higgs boson and is a CP even scalar, but has a larger mass than the SM Higgs boson, similar to the extra CP-even component introduced when using a 2HDM [11]. We denote this particle by an upper case H, and it is not to be confused with the SM Higgs boson, which we denote by a lower case h.

In addition to this, we introduce a scalar DM candidate denoted by χ. We do not specify any theoretical motivation for χ except that it could come from a model similar to the CxSM described above. In this case, DM denotes any particle which could not be directly observed in a detector, but only be detected through a mismatch in conservation of momentum.

Due to the simplistic nature of this framework, we can very easily write down the Lagrangian for a theory involving these two particles. It would be a completely independent sector from the SM Lagrangian, such that we could write,

$$\mathscr{L} = \mathscr{L}_{\mathrm{SM}} + \mathscr{L}_{\mathrm{BSM}}, \tag{1}$$

where $\mathscr{L}_{\mathrm{BSM}}$ specifies the beyond SM (BSM) additions made to the SM. The BSM sector specifies interactions between h, H, and χ particles as follows:

$$\mathscr{L}_{\mathrm{BSM}} = \mathscr{L}_{\mathrm{K}} + \mathscr{L}_{\mathrm{T}} + \mathscr{L}_{\mathrm{Q}} + \mathscr{L}_{Hgg} + \mathscr{L}_{HVV} + \mathscr{L}_{\mathrm{Y}}, \tag{2}$$

where we have kinetic terms, trilinear interactions, quartic interactions, an effective H to gluon-gluon interaction, a tree level H coupling to the weak bosons V (where V is either Z or $W^{\pm}$), and a Yukawa interaction for H, respectively. These interaction terms could also be found in a typical SM Higgs Lagrangian, barring the coupling to DM.

Omitting the obvious nature of the kinetic term, the form of the other terms are constructed

[1] Brian Patt and Frank Wilczek first referred to this as *phantom matter*, since the term dark matter was only used in an astrophysical context. These days, particle physicists tend to refer to any field which doesn't interact with the SM as dark matter.

as follows [12]:

$$\mathscr{L}_{\mathrm{T}} = -\frac{1}{2}\lambda_{Hhh}Hhh - \frac{1}{2}\lambda_{h\chi\chi}h\chi\chi - \frac{1}{2}\lambda_{H\chi\chi}H\chi\chi, \tag{3}$$

$$\mathscr{L}_{\mathrm{Q}} = -\frac{1}{4}\lambda_{HHhh}H^2h^2 - \frac{1}{4}\lambda_{hh\chi\chi}h^2\chi^2 - \frac{1}{4}\lambda_{HH\chi\chi}H^2\chi^2 - \frac{1}{2}\lambda_{Hh\chi\chi}Hh\chi^2, \tag{4}$$

$$\mathscr{L}_{Hgg} = -\frac{1}{4}\beta_g\ \kappa_{hgg}^{\mathrm{SM}}G_{\mu\nu}G^{\mu\nu}H, \tag{5}$$

$$\mathscr{L}_{HVV} = \beta_V\kappa_{hVV}^{\mathrm{SM}}\ V_\mu V^\mu H, \tag{6}$$

$$\mathscr{L}_{\mathrm{Y}} = -\frac{1}{\sqrt{2}}y_{ttH}\bar{t}tH - \frac{1}{\sqrt{2}}y_{bbH}\bar{b}bH. \tag{7}$$

Other relevant parameters in this framework which appear in the kinetic part of the Lagrangian are m_H and m_χ, the masses of the heavy scalar and the DM candidate, respectively.

While this framework contains a large number of parameters, in practice we are able to fix enough of these to leave just one free parameter. This free parameter is β_g, which is present in equation 5. The practical relevance of this framework is discussed in the following section.

3. Application to LHC Physics

It might be unclear thus far as to what motivation there is for introducing a DM candidate and a heavy scalar, as additional particles to the SM. From an astrophysics point of view, it is obvious that the SM is incomplete and is in need of an extension to explain the DM relic density in the Universe. But more recently, results from the ATLAS and CMS collaborations based at CERN have presented tantalising results which motivate the existence of new particles.

Recent results for di-Higgs resonances have shown excesses in decays coming from double Higgs production. In particular, ATLAS has presented an exciting study in $hh \to \gamma\gamma b\bar{b}$ decays [13] which promotes the possibility of the discovery of a heavy CP even scalar having a mass around 300 GeV. In fact, a few other studies (including some by CMS) also appear to show resonances with a mass between 270 and 300 GeV. These excesses are the motivation for postulating the existence of a heavy scalar.

In addition to this, ATLAS has also released a set of analyses containing kinematic information from Higgs production using only events which are confined to a particular phase space – the so-called fiducial regions. The effect of this is essentially to remove any experimental bias from the results. This has been done comprehensively for both the $h \to \gamma\gamma$ [14] and $h \to ZZ^* \to 4\ell$ [15] decay channels.[2] When we consider the Higgs p_T spectra in these analyses we discover that the current SM prediction does not accurately reproduce the experimental data. Instead, we see an excess of p_T in the region between 20 and 100 GeV (the intermediate region). This boosted p_T could be explained by the Higgs boson recoiling off of a DM candidate.

It is natural, then, to ask if adding these two particles can affect the kinematics of produced Higgs bosons to such an extent that the discrepancies in the data can be explained. To test this question, Monte Carlo simulations were done using `MadGraph 5` [17] based on a model build from the Lagrangian in equation 2 using the Mathematica package `FeynRules` [18]. We generated events using proton-proton collisions having a final state of $h\chi\chi$. The reason for this is to ensure the stability of DM particles. Since a Higgs boson might decay into a pair of χ particles, their mass must be smaller than $m_h/2$. Kinematically, then, we would get a large cross

[2] ATLAS have also released a paper reviewing the combination of both the $\gamma\gamma$ and $ZZ^* \to 4\ell$ decays [16], although we did not use the combination to obtain results, due to the methods they used in combining the two decays. Having said this, the results we obtained would not have been significantly different if we had used the combination study.

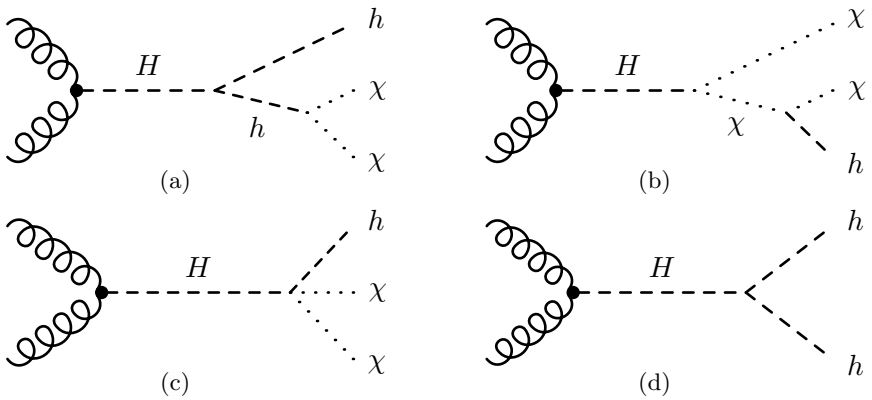


Figure 1. The three dominant Feynman diagrams contributing to Higgs production in proton-proton collisions, when an H s-channel is required. The production of $h\chi\chi$ (a, b, and c) and hh (d) both impart boosted transverse momentum to the produced Higgs bosons.

section contribution from $h\chi\chi$ production[3] (as opposed to $h\chi$ production). The production of hh was also simulated, although this is a small effect compared to $h\chi\chi$ production. These final states were assumed to be dominantly produced through an s-channel H resonance such that the relevant Feynman diagrams which were used for the calculation are those shown in figure 1 – this assumption was checked and shown to be valid.

As can be seen from the diagrams in figure 1, the only relevant couplings which needed to be fixed were λ_{Hhh}, $\lambda_{h\chi\chi}$, $\lambda_{H\chi\chi}$ and $\lambda_{Hh\chi\chi}$. The first coupling, λ_{Hhh}, was fixed by picking a conservative production cross section for $H \to hh$ of 2 pb, coming from the study in reference [13]. The parameters $\lambda_{h\chi\chi}$ and $\lambda_{H\chi\chi}$ were fixed using astrophysical observations, using both the DM relic density [20] and the DM-nuclei inelastic scattering cross sections determined by the LUX experiment [21]. The mass points selected for H and χ were based on experimental indications, and were set at $m_\chi \in \{40, 50, 55, 60\}$ GeV and $m_H \in \{275, 285, 300\}$ GeV. Finally, the coupling $\lambda_{Hh\chi\chi}$ was scanned over by considering three different values for the branching ratio of $H \to h\chi\chi$, those being 50%, 60%, and 70%.

Having fixed all of the relevant parameters relating to the interactions between h, H and χ, the only relevant parameter left free was β_g. Since β_g affects the contributions of all of the diagrams in equal proportion, it was used as a free parameter to fit the model to the ATLAS experimental results in references [14] and [15]. By doing a simultaneous fit for both the $\gamma\gamma$ and $ZZ^* \to 4\ell$ Higgs p_T distributions, we were able to obtain the best fit by setting $m_\chi = 60$ GeV, $m_H = 275$ GeV and $\mathrm{BR}(H \to h\chi\chi) = 60\%$. The best fit (which was obtained by minimising χ^2 over the number of degrees of freedom, $n.d.f$) required a β_g value of 2.07±0.37. It should be noted that the SM prediction does not accurately describe the data, having a $\chi^2/n.d.f$ of 1.59. Using our model with the mass points we have proposed, we are able to reduce this to 0.86, giving a much better fit to the data. For a detailed description of a number of results, the reader is encouraged to look in reference [12].

4. Concluding Remarks

The framework presented is attractive for two obvious reasons. Firstly, it does a very good job at explaining discrepancies which we see in the ATLAS experimental Higgs data. There is a clear correlation between the discrepancies in the data from the $\gamma\gamma$ and the $ZZ^* \to 4\ell$ channels, both

[3] A similar production mechanism has been considered in mono-Higgs searches [19], although without an explicit interest in the importance of hh production.

of which can be explained by the Higgs boson recoiling off of undetected particles. Secondly, the framework is attractive due to the fact that it very elegantly explains two experimental excesses simultaneously. While ATLAS data points to the excess mentioned above, the framework also explains excesses in di-Higgs resonances which are seen by both CMS and ATLAS.

While this short paper has highlighted the explanation of a few experimental results, one could ask what predictions the framework makes and what signatures to be mindful of for future experimental efforts. A possible prediction can be made if equations 6 and 7 are considered. I have not yet made mention of the possible weak boson or jet production mechanisms arising from the addition of the proposed framework. If we assume that the HVV coupling is small, then the possibility for top quark production is not so highly suppressed by the negative interference from the HWW coupling (we see this suppression in the SM). This means that top quark production should occur at a much higher rate than what is predicted by just the SM. In fact, an excess in top quark production has been observed, albeit a small one.

The framework presented is a tantalising and exciting tool to be used for the exploration of new physics at the LHC. While good results are presented in this short paper, it is imperative that more experimental data be accrued in order to clear up any statistical fluctuations. With Run 2 of the LHC currently under-way at the time of writing this paper, we can expect a far more thorough and meaningful analysis to take place in the near future.

Acknowledgements

This work was done in collaboration with the following people: from the University of the Witwatersrand– Alan S. Cornell, Deepak Kar, Mukesh Kumar, Bruce Mellado and Robert G. Reed. From the Harish-Chandra Research Institute in India– Nabarun Chakrabarty, Tanumoy Mandal and Biswarup Mukhopadhyaya. The Claude Leon Foundation and the National Research Foundation (NRF) are acknowledged for their financial support.

References

[1] Yang C N and Mills R L 1954 *Phys. Rev.* **96** 191–195
[2] 't Hooft G 1971 *Nucl. Phys.* **B35** 167–188
[3] Weinberg S 1971 *Phys. Rev. Lett.* **27** 1688–1691
[4] Salam A and Strathdee J A 1972 *Nuovo Cim.* **A11** 397–435
[5] Higgs P W 1964 *Phys. Rev. Lett.* **13** 508–509
[6] Abdallah J *et al.* 2015 *Phys. Dark Univ.* **9-10** 8–23 (*Preprint* 1506.03116)
[7] Gonderinger M, Lim H and Ramsey-Musolf M J 2012 *Phys. Rev. D* **86** 043511 (*Preprint* 1202.1316v2)
[8] Guo W L and Wu Y L 2010 *JHEP* **10** 083 (*Preprint* 1006.2518)
[9] Cai Y, He X G and Ren B 2011 *Phys. Rev.* **D83** 083524 (*Preprint* 1102.1522)
[10] Patt B and Wilczek F 2006 Higgs-field Portal into Hidden Sectors (*Preprint* 0605188v1)
[11] Deshpande N G and Ma E 1978 *Phys. Rev.* **D18** 2574
[12] von Buddenbrock S, Chakrabarty N, Cornell A S, Kar D, Kumar M, Mandal T, Mellado B, Mukhopadhyaya B and Reed R G 2015 (*Preprint* 1506.00612)
[13] ATLAS Collaboration 2015 *Phys. Rev. Lett.* **114** 081802 (*Preprint* 1406.5053v3)
[14] ATLAS Collaboration 2014 *JHEP* **1409** 112 (*Preprint* 1407.4222v2)
[15] ATLAS Collaboration 2014 *Phys. Lett. B* **738** 234–253 (*Preprint* 1408.3226) URL http://arxiv.org/abs/1408.3226http://dx.doi.org/10.1016/j.physletb.2014.09.054
[16] ATLAS Collaboration 2015 (*Preprint* 1504.05833v1)
[17] Alwall J, Herquet M, Maltoni F, Mattelaerc O and Stelzer T 2011 *JHEP* **1106** 128 ISSN 11266708 (*Preprint* 1106.0522)
[18] Alloul A, Fuks B and Sanz V 2014 *JHEP* **1404** 110 ISSN 1029-8479 (*Preprint* 1310.5150)
[19] Carpenter L, Difranzo A, Mulhearn M, Shimmin C, Tulin S and Whiteson D 2014 *Phys. Rev. D* **89** 075017 ISSN 15502368 (*Preprint* 1312.2592)
[20] Planck Collaboration 2014 *Astron. Astrophys.* **571** A16 ISSN 0004-6361 (*Preprint* 1303.5076)
[21] The LUX Collaboration 2014 *Phys. Rev. Lett.* **112** 091303 ISSN 0031-9007, 1079-7114 (*Preprint* 1310.8214v2)