South African Institute of PHYSICS
THE VOICE OF PHYSICS IN SOUTH AFRICA

UNIVERSITY OF ZULULAND
A NODE FOR AFRICAN THOUGHT

The 67th Annual Conference of the South African Institute of Physics (SAIP)

Transforming lives of our communities through Physics

$$i\hbar\frac{\partial}{\partial t}\Psi = H\Psi$$

CONFERENCE PROCEEDINGS

Edited by Prof ARE Prinsloo

A NODE FOR AFRICAN THOUGHT

PROCEEDINGS EDITOR-IN-CHIEF:
Prof Aletta Prinsloo, University of Johannesburg

PUBLISHER:
The South African Institute of Physics (SAIP)

The conference was hosted by the University of Zululand, on 3–7 July 2023. The Proceedings of SAIP 2023, the 67th Annual Conference of the South African Institute of Physics, will be available electronically only on the SAIP website: `www.saip.org.za`.

ISBN: 978-0-7961-3774-6

South African Institute of Physics
CSIR Main Campus, Building 33
Meiring Naude Rd, Pretoria.

version compiled on 2023-12-19 15:26:42+02:00

CONTENTS

EDITORIAL

The University of Zululand hosted the annual South African Institute of Physics (SAIP) conference during 2023. With the COVID-19 pandemic and the limitations associated with it something in the past, the University of Zululand decided to host an in person event. This was done with great success and the delegates commented on the professional running of the event, the wonderful food and the lovely surroundings near the coast. Some papers from this meeting are collected in this peer-reviewed volume. Submissions for the proceedings of SAIP2023 were handled by an Editorial Board headed by an Editor-in-Chief and Associate Editors responsible for submissions in different divisions.

The Editorial Board of the SAIP2023 Proceedings received 106 manuscripts for consideration by the advertised deadline. A total of 88 of these manuscripts met the relevant criteria and were submitted to a full peer-review process involving many individual reviewers. The list of the reviewer names are reflected elsewhere in the document and it is noted that certain reviewers took responsibility for more than one manuscript. The style of these proceedings is that of the (British) Institute of Physics Conference Series, similar to the styling used in previous SAIP Proceedings. Authors were requested to ensure that the defined layout were adhered to in their submitted pdf documents. In the past the review process was initiated with a layout review, followed by a content review. This year the Associate Editors conducted the layout review on each manuscript parallel with the content review. It was noted that there were small deviations between the layout templates available in MSWord and Latex - both of these formats were accepted by the Associate Editors. Manuscripts that deviated considerably from the specified layout specifications, while still broadly appropriate in their composition, were referred back to the authors for layout corrections. This was done together with the content reviews prepared by knowledgeable experts in each field, as well as considering Turnitin reports to ensure that the work is unique and not plagiarized. This year the Editorial Board again aimed to reduce the time between the submissions and publication, with the authors being informed of the outcome of their submissions before the closure for the December holiday and the publication of the document online shortly after that.

The publication of the SAIP Proceedings is highly dependent on the efficiency of the Associate Editors and the goodwill of reviewers from the scientific community in South Africa. The Editor-in-Chief wishes to acknowledge the hard work of the Associate Editors who spent much

time considering the papers and reviewer reports in order to ensure that acceptable academic standards were met during peer-review for the proceedings to be credible. The majority of the content reviews received were done with great care and diligence and to the highest standards. The Editorial Board wishes to voice their sincere thanks to the participating Reviewers for their pro bono work, specifically to those Reviewers that read more than one paper. The meticulous reviewing process described above has ensured that these proceedings contain thoroughly peer-reviewed manuscripts of a high professional standard, which report on novel work that has not been published elsewhere.

This year the Editorial Board again made use of the services of a Technical Associate Editor, Dr Bruno Letarte from NWU. He took responsibility for finalizing the complete document and ensured that it was of a high technical standard. The Editor-in-Chief wish to recognise Dr Letarte's enormous contribution in preparing the neat final document. The Editorial Board appreciate all the hours you dedicated into producing this exceptional document.

The Editor-in-Chief also wishes to recognise and thank Prof Strauss from NWU, Mr Mokhine from the SAIP office and Dr Ceboliyazakha Ndlangamandla from the University of Zululand for their support and help in preparing these proceedings.

Finally, the Editorial Board wishes to thank all of the authors for submitting their research work to this proceedings to undergo the rigorous review process. It is our sincere hope that the final product offered here constitutes a due outcome of their hard work.

MESSAGE FROM THE ORGANISERS

The Organizing Committee would like to take this opportunity to thank all the delegates and their valuable contributions towards the growth of Physics. We also like to convey our appreciation to the South African Institute of Physics (SAIP) for their crucial assistance in the preparation of the conference as well as the running of the conference. The University of Zululand Management under the leadership of Prof Xoliswa Mtose needs to be commended for its unwavering support to ensure that the 2023 SAIP conference is a success story. The reviewed conference proceedings are the consequences of a healthy organization as well as quality papers. The University of Zululand is inarguably and truly a 'Node of African Thought' as we move forward in transforming the lives of our communities through Physics.

LOC 2023

Local Organising Committee

- **Chair** – Prof Thulani Jili
- **Budget and Finance** – Dr Ceboliyazakha Ndlangamandla, Prof Thulani Jili
- **Scientific co-chairs** – Dr Ceboliyazakha Ndlangamandla, Dr Zipho Ngcobo, Prof Sifiso Ntshangase and Dr Makhosonke Dubazane
- **Proceedings** – Dr Ceboliyazakha Ndlangamandla
- **Fundraising and Marketing** – Ms Thobile Mdluli and Mr Sphephelo Khanyile
- **Social functions** – Ms Nelisiwe Chonco and Mr Phathizwe Majozi
- **Winter School coordination** – Dr Linda Mdletshe and Dr Makuhane Sithole
- **Outreach and public engagement** – Dr Getachew Mengiste, Dr Puleng Biyela and Mr Mdumiseni Nxumalo
- **Delegate liaison** – Mrs Ntombifuthi Mothapo
- **Venue coordination** – Mr Prince Mkwae, Mr Sive Noncolela, Mr Charles Thethwayo and Mrs Mpume Cele
- **Catering Winter School** – Dr Puleng Biyela and Ms Thobile Mdluli
- **IT Website** – Mr Mmeli Gwebu, Mr Tebogo Mokine (SAIP) and Mr Sphephelo Khanyile
- **Security** – Mr Richard Lukhele
- **SAIP representative** – Dr Brian Masara

DIVISIONS, EDITORIAL TEAM, AND REVIEWERS

Divisions and Division Chairs

- **Physics of Condensed Matter and Materials**
  – Prof Rudolph Erasmus, University of the Witwatersrand
- **Nuclear, Particle, and Radiation Physics**
  – Dr Edward Khomotso Nkadimeng, University of the Witwatersrand
  – Dr Lindsay Donaldson, iThemba LABS
- **Photonics**
  – Dr Pieter Neethling, Stellenbosch University
- **Astrophysics and Space Science**
  – Dr Geoff Beck, University of the Witwatersrand
  – Dr Katlego Moloto, North-West University
- **Physics for Development, Education, and Outreach**
  – Prof Sam Ramaila, University of Johannesburg
- **Applied Physics**
  – Prof Alan Matthews, University of KwaZulu Natal
- **Theoretical and Computational Physics**
  – Prof Alan Cornell, University of Johannesburg
- **Biophysics**
  – Prof Tjaart Krüger, University of Pretoria
- **Women in Physics**
  – Dr Katekani Shingange, CSIR

Editorial Team

Editor-in-chief:

– **Aletta Prinsloo** is a Professor of Physics in the Department of Physics at the University of Johannesburg. She is an NRF rated research physicist in the field of experimental solid state physics. Her research is focused on the magnetism of chromium-based bulk alloys, thin films and nanomaterials.

Associate Editors:

- **Physics of Condensed Matter and Materials**
  – **Charles Sheppard** is an Associate Professor and a member of the Cr Research Group in the Physics Department at the University of Johannesburg. His current research interest focuses on the various physical properties observed in bulk Cr alloys, Cr thin films, and chrome oxide magnetic nano-materials.

- **Nuclear, Particle, and Radiation Physics**
  – **Mukesh Kumar** is a Senior Lecturer of Physics in the School of Physics at the University of the Witwatersrand. He is an NRF Y-rated research physicist in the field of high energy particle physics. His research is focused on Higgs boson, top quark, and dark matter physics at the Large Hadron Collider (CERN) including the future $e^{-}p$ and $e^{+}e^{-}$ colliders. He is a member of TileCal Speaker committee for ATLAS detector at CERN.

- **Photonics**
  – **Pieter Neethling** is a Senior Lecturer in the Physics Department at Stellenbosch University. He is currently the Director of the Stellenbosch Photonics Institute at Stellenbosch University and the Chairman of the Photonics Division of the SAIP. His research focus is applied laser spectroscopy with applications in chemical and biological systems.

- **Astrophysics and Space Science**
  – **Eugene Engelbrecht** is a Professor of Physics at North-West University, whose research covers topics relevant to the transport of charged particles in turbulent astrophysical plasmas, including both theoretical and observational aspects pertaining to cosmic ray modulation, non-linear diffusion theories, and plasma turbulence.

- **Physics for Development, Education, and Outreach**
  – **Hartmut Winkler** is a Professor of Physics and former Head of the Department of Physics at the University of Johannesburg. He was a past recipient of the Vice-Chancellor Distinguished Teacher Award with an extensive teaching portfolio. His background is in astrophysics, where he has maintained an active interest in the study of the variability of Active Galactic Nuclei. More recently he has also diversified to solar energy research. He is a frequent media commentator on topics pertaining to energy and electricity.

- **Applied Physics**
  – **Thulani Hlatshwayo** is an associate Professor in the Department of Physics in the Faculty of Natural & Agricultural Sciences. His research is focussed on the understanding of the release of radioactive fission products from fuel in the modern nuclear reactors, where chemical vapour deposited (CVD)-SiC is the main barrier to fission products, and on finding alternative materials for nuclear waste storage. Professor Hlatshwayo recently received the Exceptional Young Researchers Award by the University of Pretoria. He is a PIs coordinator for SA-JINR projects in material research and nanoscience and is C2 NRF rated.

- **Theoretical and Computational Physics**
  – **W. A. Horowitz** is an Associate Professor of Physics at the University of Cape Town. Among other honours, Prof Horowitz has received the Claude Leon Merit Award for Early-Career Researchers and the Meiring Naudé Medal for Outstanding Early Career Contributions to Science from the Royal Society of South Africa. Prof Horowitz' research explores the non-trivial emergent many-body properties of the strong force using the methods of perturbative quantum field theory and the AdS/CFT correspondence.

- **Technical**
  – **Bruno Letarte** is a Senior Lecturer at the Centre for Space Research of the North-West University. He specialises in observational astronomy, photometry as well as spectroscopy, with his main interest in stellar astrophysics. He manages the optical telescope at the Nooitgedacht observatory, used to train undergraduate and postgraduate students. He is also the physics subject group leader, what other universities call head of department, on the Potchefstroom campus.

Proceedings Online Administration:

– **Tebogo Mokhine**, South African Institute of Physics

List of Reviewers

- **Dr Hesham Abdelbagi** – University of Pretoria, South Africa
- **Prof Amine Ahriche** – University of Sharjah, United Arab Emirates
- **Dr Mahmood Akbari** – iThemba LABS; Cape Town, South Africa
- **Dr Abdulraoof Ali** – University of Pretoria, South Africa
- **Prof Aroon Beesham** – University of Zululand, South Africa
- **Prof Saphina Biira** – Busitema University, Uganda
- **Dr Gert Botha** – Northumbria University, United Kingdom
- **Dr Pablo Brito Parada** – Imperial College London, United Kingdom
- **Dr Daphney Bucher** – iThemba LABS, South Africa
- **Prof Andy Buffler** – University of Cape Town, South Africa
- **Dr Giacomo Cacciapaglia** – Institut de Physique des 2 Infinis de Lyon, France
- **Prof Emanuela Carleschi** – University of Johannesburg, South Africa
- **Prof Hasani Richard Chauke** – University of Limpopo, South Africa
- **Prof James Chibueze** – UNISA, South Africa
- **Prof Hing-Tong Cho** – Tamkang University, Taiwan
- **Prof Liza Coetzee-Hugo** – University of the Free State, South Africa
- **Prof Alan Cornell** – University of Johannesburg, South Africa
- **Dr Julien Dagbegnon** – Italian Institute of Technology, Italy
- **Prof Aldo Deandrea** – Institut de Physique des 2 Infinis de Lyon, France
- **Prof Mmantsae Diale** – University of Pretoria, South Africa
- **Dr Hector Dlamini** – University of Pretoria, South Africa
- **Dr Lindsay Donaldson** – iThemba LABS, South Africa
- **Prof Bryan Doyle** – University of Johannesburg, South Africa
- **Dr Frederic Effenburger** – Ruhr University Bochum, Germany
- **Prof Eugene Engelbrecht** – North-West University, South Africa
- **Prof Rudolph Erasmus** – University of the Witwatersrand, South Africa
- **Prof Sylvain Fichet** – Universidade Federal do ABC, Brazil
- **Dr Derek Fish** – University of Zululand, South Africa

- **Prof Benjamin Fuks** – Sorbonne Université, France
- **Prof Blassan George** – University of Johannesburg, South Africa
- **Prof Diane Grayson** – University of the Witwatersrand, South Africa
- **Dr Lokesha Handalagere Shankarappa** – Universidade Federal de São Carlos, Brazil
- **Dr Mark Herbert** – University of the Western Cape, South Africa
- **Prof Thulani Hlatshwayo** – University of Pretoria, South Africa
- **Dr Emmanuel Igumbor** – University of Johannesburg, South Africa
- **Dr Yaseera Ismail** – University of KwaZulu-Natal, South Africa
- **Dr Susan Jacobs** – University of Johannesburg, South Africa
- **Dr Pete Jones** – iThemba LABS, South Africa
- **Mr Jano Jonker** – Nelson Mandela Universty, South Africa
- **Dr Abraham Kapim** – Tshwane University of Technology, South Africa
- **Prof Deepak Kar** – University of the Witwatersrand, South Africa
- **Dr Garreth Kemp** – University of Johannesburg, South Africa
- **Dr Bongani Kheswa** – University of Johannesburg, South Africa
- **Dr Mohammed Omer Khojali** – University of Johannesburg, South Africa
- **Dr Kenda Knowles** – University of KwaZulu-Natal, South Africa
- **Prof Lehlohonolo Koao** – University of the Free State, South Africa
- **Dr Isobel Kolbe** – University of the Witwatersrand, South Africa
- **Prof Tjaart Krüger** – University of Pretoria, South Africa
- **Dr Joseph Kuhudzai** – Electric Drive Africa, Zimbabwe
- **Dr Pannan Kyesmen** – University of Agriculture, Nigeria
- **Dr Elena Lawrie** – iThemba LABS, South Africa
- **Prof Raesibe Ledwaba** – University of Limpopo, South Africa
- **Dr Bruno Letarte** – North-West University, South Africa
- **Dr Stefan Lotz** – SANSA, South Africa
- **Prof Ilani Loubser** – North-West University, South Africa
- **Prof Kakoma Luneta** – University of Johannesburg, South Africa
- **Dr Hongze Luo** – Council for Scientific and Industrial Research, South Africa
- **Dr Kelebogile Maaabong** – University of Botswana, Botswana

- **Prof Roy Maartens** – University of the Western Cape, South Africa
- **Prof Khomotso Maenetja** – University of Limpopo, South Africa
- **Dr Siyabonga Majola** – University of Johannesburg, South Africa
- **Dr Peane Maleka** – iThemba LABS, South Africa
- **Prof Johan Malherbe** – University of Pretoria, South Africa
- **Prof Nnditshedzeni Eric Maluta** – University of Venda, South Africa
- **Dr Edwin Mapasha** – University of Pretoria, South Africa
- **Dr Marco Mariola** – University of KwaZulu-Natal, South Africa
- **Dr Mordecai Mashamaite** – University of Limpopo, South Africa
- **Dr Nkanyiso Mbatha** – University of Zululand, South Africa
- **Dr Wendy Mdlalose** – University of Kwazulu-Natal, South Africa
- **Dr Linda Mdletshe** – University of Zululand, South Africa
- **Dr Mofuti Mehlape** – University of Limpopo, South Africa
- **Dr Getachew Mekonnen** – University of Zululand, South Africa
- **Dr Rebecca Mhlongo** – Sefako Makgtho Health Science University, South Africa
- **Prof Simiso Mkhonta** – University of Eswatini, Eswatini
- **Dr Mbuso Mlambo** – Mintek, South Africa
- **Dr Michaela Mlynarikov** – CERN, Switzerland
- **Dr Rosinha Modiba** – Council for Scientific and Industrial Research, South Africa
- **Dr Victor Molefe** – Tshwane University of Technology, South Africa
- **Dr Mahlaga Molepo** – University of the Witwatersrand, South Africa
- **Dr Patience Mthunzi-Kufa** – Council for Scientific and Industrial Research, South Africa
- **Dr Blessed Muchono** – University of Eswatini, Eswatini
- **Dr Sophie Mulaudzi** – University of Venda, South Africa
- **Prof Jeff Murugan** – University of Cape Town, South Africa
- **Prof Deena Naidoo** – University of the Witwatersrand, South Africa
- **Dr Partha Nandi** – Stellenbosch University, South Africa
- **Dr Lethole Ndanduleni** – University of Fort Hare, South Africa
- **Dr Pieter Neethling** – Stellenbosch University, South Africa

- **Dr Vhutshilo Nekhubvi** – University of Venda, South Africa
- **Dr Amore Nel** – SANSA, South Africa
- **Prof Eric Njoroge** – University of Pretoria, South Africa
- **Dr Edward Nkadimeng** – iThemba LABS and University of the Witwatersrand, South Africa
- **Dr Steven Nkosi** – University of Limpopo, South Africa
- **Dr Rendani Nndanganeni** – SANSA, South Africa
- **Dr Volkmar Nolting** – Vaal University of Technology, South Africa
- **Dr Hajar Noshad** – University of Johannesburg, South Africa
- **Dr Petros Ntoahae** – University of Limpopo, South Africa
- **Prof Cloud Nyamere** – Midlands State University, Zimbabwe
- **Dr Nongamso Nyangiwe** – Tshwane University of Technology, South Africa
- **Dr Kingley Obodo** – North-West University, South Africa
- **Prof Amidu Olalekan Mustapha** – Federal University of Agriculture; Abeokuta, Nigeria
- **Dr Saturnin Ombinda-Lemboumba** – Council for Scientific and Industrial Research, South Africa
- **Dr Oluwatayo Racheal Onisuru** – University of Johannesburg, South Africa
- **Dr Alexander Paradzah** – Chinhoyi University of Technology, Zimbabwe
- **Prof Pararajasingham Peratheepan** – Eastern University Sri Lanka, Sri Lanka
- **Prof Aletta Prinsloo** – University of Johannesburg, South Africa
- **Dr Jan-Louis Raath** – SANSA, South Africa
- **Prof Melanie Rademeyer** – University of Pretoria, South Africa
- **Prof Subharthi Ray** – University of KwaZulu-Natal, South Africa
- **Prof Leelakrishna Reddy** – University of Johannesburg, South Africa
- **Dr Arnoux Rossouw** – Joint Institute for Nuclear Research, Russia
- **Prof Charles Sheppard** – University of Johannesburg, South Africa
- **Dr Muzikayise Sikhonde** – University of Cape Town, South Africa
- **Dr Joseph Simfukwe** – The Copperbelt University, Zambia
- **Prof Buyisiwe Sondezi** – University of Johannesburg, South Africa
- **Dr Christine Steenkamp** – Stellenbosch University, South Africa

- **Prof Mark Tame** – Stellenbosch University, South Africa
- **Prof Moise B. Tchoula Tchokonte** – University of the Western Cape, South Africa
- **Dr Thabsile Thabethe** – University of Pretoria, South Africa
- **Prof Christoph Trauernich** – Stellenbosch University, South Africa
- **Dr David Tshwane** – Council for Scientific and Industrial Research, South Africa
- **Dr Aniekan Ukpong** – University of KwaZulu-Natal, South Africa
- **Prof Brandon Van der Ventel** – Stellenbosch University, South Africa
- **Prof Johan van der Walt** – North-West University, South Africa
- **Dr Chani van Niekerk** – University of Johannesburg, South Africa
- **Dr JJ Van Zyl** – Stellenbosch University, South Africa
- **Prof Christo Venter** – North-West University, South Africa
- **Dr Maria Vivien Visaya** – University of Johannesburg, South Africa
- **Dr Robert Warmbier** – University of the Witwatersrand, South Africa
- **Prof Herbert Weigel** – Stellenbosch University, South Africa
- **Dr Jennifer Williams** – Rhodes University, South Africa
- **Prof Hartmut Winkler** – University of Johannesburg, South Africa

Group Photo of Delegates

DIVISION

A

PHYSICS OF CONDENSED MATTER AND MATERIALS

Effects of annealing time on structural, morphology, and optical properties of Zinc Oxide nanoparticles prepared via Sol-gel method

T.C Selema[1*], D.G Tsotetsi[2], and T.D Malevu[1]

1. Department of Physics, Sefako Makgatho Health Sciences University, P.O. Box 94, Medunsa 0204, South Africa
2. Department of Physics, CSET, University of South Africa, Johannesburg, 1710, South Africa

E-mail: **selematebogo53@gmail.com, malevu.td@gmail.com**

Abstract. In this study, Zinc oxide (ZnO) nanoparticles were successfully synthesized using the sol-gel method. To investigate the effect of annealing time on the structural, morphology, and optical properties of ZnO, annealing of the synthesized nanoparticles was further carried out at 550^{o} C for different time intervals (30 min, 1 hour, and 1:30 hour). X-ray diffraction (XRD) analysis confirmed the formation of a crystalline hexagonal wurtzite ZnO structure, which was observed to be more prominent with an increase in annealing time, specifically at 1:30 hour. UV-VIS spectroscopy analysis revealed an improved absorption band with a wavelength of 365 nm, which was redshifted compared to other prepared samples. Additionally, photoluminescence (PL) quenching was observed for the annealed samples, indicating charge transfer that is favorable for solar cell applications. The as-prepared sample showed high PL intensity, possibly due to self-trapped excitons recombination. Finally, the morphologies of the prepared samples were studied using scanning electron microscopy (SEM), revealing a progression of ZnO morphologies from clustered to nanorod structures. Overall, the results suggest that annealing time is a crucial parameter in the synthesis of ZnO nanoparticles, and that optimized annealing conditions can improve their structural, optical, and morphological properties for potential use in solar cell applications.

1. Introduction

Zinc oxide (ZnO) is a low-cost, n-type wurtzite semiconductor with a direct wide bandgap of 3.2-3.3 eV at room temperature [1]. In both the visible and ultraviolet ranges, ZnO possesses good electrical and optical properties, and as a result, it is commonly used in photoelectric displays, solar cells, sensors, and surface acoustic wave devices [2]. Additionally, ZnO has a high exciton energy, which is crucial for the material's optical characteristics. ZnO is environmentally benign and more resistant to breaking down at high temperatures [3]. Many techniques have been employed over the years for the production of ZnO nanoparticles, including spray pyrolysis, chemical vapor deposition (CVD), RF magnetron sputtering, molecular beam epitaxy (MBE), pulsed laser deposition, and sol-gel [4-6]. However, the sol-gel approach has increasingly become the primary method for preparing ZnO nanoparticles in the laboratory because of its simple procedure, straightforward doping, easy-to-form large-area film, and lack of vacuum equipment [7].

Despite its excellent properties, ZnO still suffers from some major drawbacks which restrict its applications. Some of these drawbacks include poor solar light consumption and rapid recombination of photoinduced charge carriers. In recent years, researchers have devoted their attention to improving these drawbacks associated with ZnO to further its potential in different applications. One of the most explored approaches with the aim of improving the material's properties is annealing, which is a heat treatment process that alters the physical, and sometimes chemical properties of a material, leading to an increase in its ductility and a reduction of its hardness, making it more workable [1,2]. Furthermore, annealing has the potential to narrow the bandgap of the ZnO nanomaterial, increasing its absorption wavelength [1,2], and hence allowing it to harness more solar energy for solar cell applications. In this study, we report on the effects of annealing time while keeping the temperature constant at 550^{o}C. This is because ZnO nanoparticles can only be crystallized by heat treatment at over 400^{o}C, so, by annealing at 550°C, the aim was to encourage the growth of ZnO crystallites and enhance their particle size [8]. Other approaches used to modify ZnO nanoparticles include utilizing several methodologies such as coupling with narrow band gap semiconductors, noble metal deposition, surface sensitization by organic dyes, and elemental doping [9].

2. Experimental

2.1. Preparation of ZnO

The ZnO nanoparticles were synthesized using the Sol-gel method. Chemical reagents used during preparation were taken as supplied by the manufacturer without any further purification. 5 g of Zinc acetate dihydrate ($Zn(CH_3COO)_2 \cdot 2H_2O$)(95% extra pure) (Sigma-Aldrich), was reacted with 20 ml of methanol(99.9%) (CH_3OH) (Sigma-Aldrich), this solution was then stirred ultrasonically for 2 hours at room temperature. 2.5 g of sodium hydroxide (NaOH)(99%) (Sigma-Aldrich) was then added to the resultant solution, and the stirring continued for another hour until a viscous gel was formed, Using the blade coating method, the formed gel was then deposited on the glass substrate to make thin films. The thin films were then dried in an oven for 20 minutes at 100^{o} C. Finally, the dried thin films were further divided and classified as the as-prepared sample, and annealed samples of various time intervals (30 min, 1 hour, and 1:30 hour respectively) at a constant temperature of 550^{o} C.

2.2 Characterization of the prepared samples.

The thin films were then characterized using the X-ray diffraction (XRD) technique, Ultraviolet-visible spectroscopy (UV-VIS), as well as Photoluminescence (PL). Lastly, the surface morphology of the thin films was observed using the scanning electron microscopy (SEM) technique.

3. Results and discussion

3.1 X-ray diffraction (XRD)

The diffraction peaks at scattering angle (2θ) of 31.9°, 34.4°, 36.2°, 47.5°, 56.5°, 62.9°, 67.8°, 69.2°,76.8°, and 89.5° correspond to the reflections from (100), (002), (101), (102), (110), (103), (200), (112), and (202) crystal planes respectively as demonstrated in Figure 1. These reflections correspond to the hexagonal wurtzite ZnO structure as confirmed by the JCPDS card number 891397 [9]. The ZnO (As-prepared) sample exhibits little to no peaks in the XRD spectra, which may be the cause of the sample's reduced intensity. This can further be attributed to the presence of smaller or less ordered crystalline domains, which do not generate strong diffraction signals. The amorphous peaks were observed for as-prepared sample. However, as we anneal the sample for 30 min, 1 hour, and 1:30 hour, the diffraction peaks become narrower, exhibiting a higher intensity. This can further suggest that the crystallinity of the material was improved with annealing time.

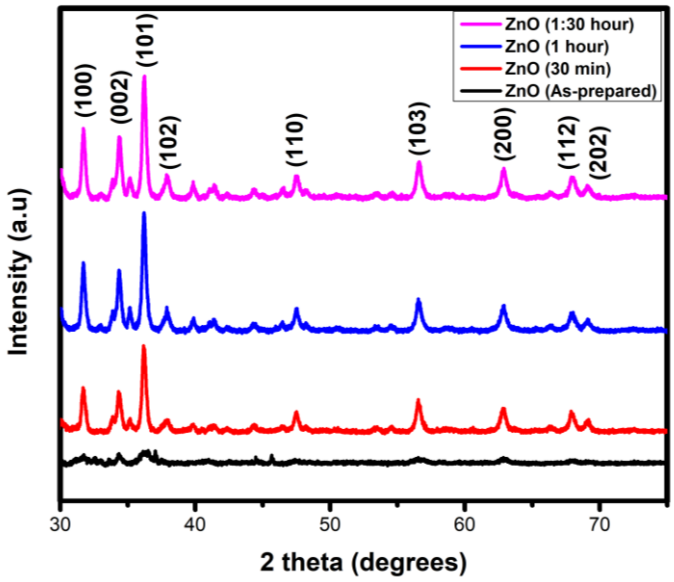


Figure 1. XRD pattern of ZnO As-prepared, ZnO (30 min), ZnO (1 hour), and ZnO (1:30 hour).

The crystallite size of the material was further estimated using the Debye Scherrer equation as stated below:

$$D = \frac{k\lambda}{\beta \cos\theta} \tag{1}$$

where D is the crystallite size, k is the Scherrer constant given by 0.9, λ is the wavelength, and β is the Full-width at Half-maximum (FWHM) of a diffraction line located at angle θ. Table 1 below shows an increase in the crystallite size as a function of annealing time, which further suggests that annealing time is effective in tuning the material's crystal sizes.

Table 1. Crystallite size estimation at the most intense peak (101), peak position (2θ), and the full-width at half maximum (FWHM).

Samples	Peak position (2θ)	FWHM (rad)	Crystallite size (nm)
ZnO (as-prepared)	36.24	1.31	6.38
ZnO (30 min)	36.21	1.15	7.27
ZnO (1 hour)	36.20	0.97	8.62
ZnO (1:30 hour)	36.18	0.82	10.19

3.2 Ultraviolet-visible spectroscopy (UV-VIS)

The UV-VIS spectra show the absorbance of all the prepared samples, with ZnO (As-prepared) having the lowest absorbance. However, as the sample was annealed through the different time intervals, there was a slight red shift in the wavelengths which can be attributed to the electronic transmission of electrons from the valence band to the conduction band, and as a result there is an increase in the absorbance of the material.

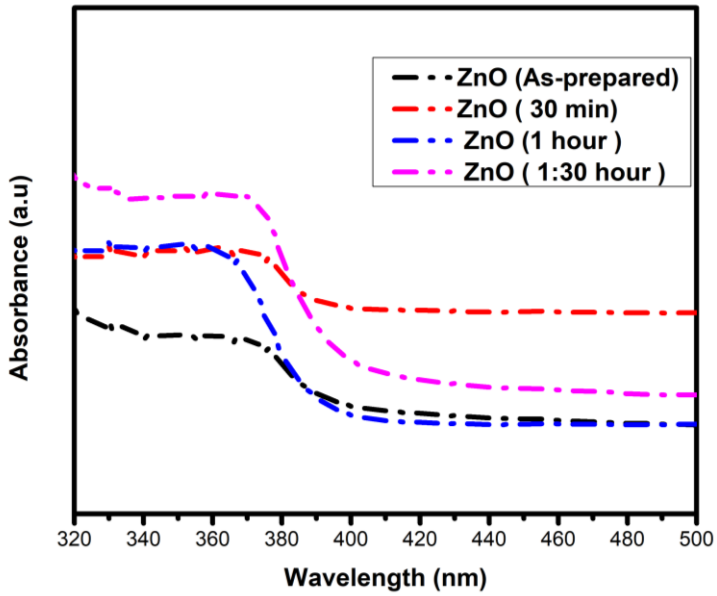


Figure 2. UV-VIS Spectra of ZnO As-prepared, ZnO (30 min), ZnO (1 hour) and ZnO (1:30 hour).

3.3 Photoluminescence (PL)

PL was also used to study the charge transfer of the photogenerated electrons and holes. The figure below exhibits the PL intensities of all the samples excited with 370 nm wavelength. A higher intensity is observed for the ZnO (As-prepared) sample, this is due to the self-trapped exciton recombination which is a combined effect of defect centers generated from oxygen vacancies and particle size. However, when the sample was subjected to a heat treatment, there was a significant decrease in the PL intensity, which suggests that there is charge transference taking place [10]. This further suggests that the modification of annealing time and its derivatives can suppress the recombination of electron-hole-pairs because high intensity implies more drastic recombination of charge carriers.

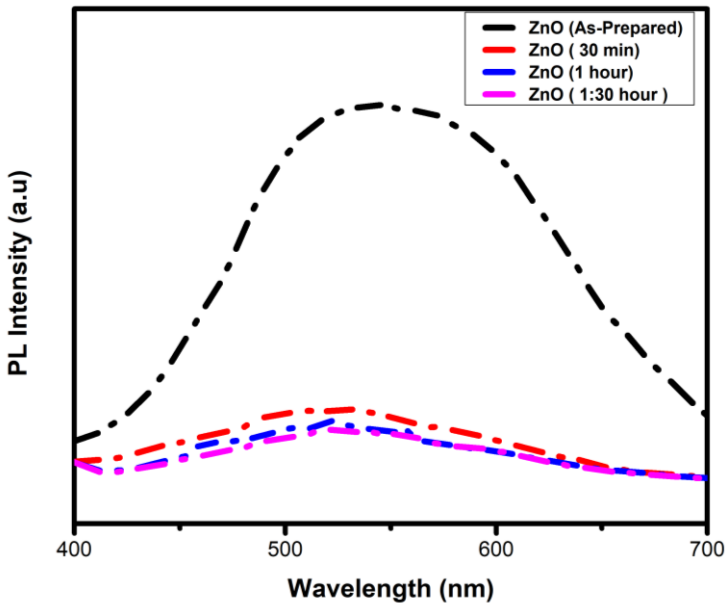


Figure 3. PL spectra of ZnO As-prepared, ZnO (30 min), ZnO (1 hour) and ZnO (1:30 hour)

3.4 Scanning electron microscopy (SEM)

SEM was carried out to observe the surface morphologies of the prepared ZnO nanoparticles. Figure 5(a) shows an agglomeration morphology for ZnO (As-prepared sample), this can be attributed to the presence of residual solvents as depicted on the XRD spectra. Figure 5(b) and (c) shows an evolution from the clustered morphology as a result of annealing. Finally, as the annealing time further increases to 1:30 hour (Figure 5(d)), fully formed nanorod structures are observed, and these nanorod structures are the preferred morphology for solar cell applications. In their study, Nandang Mufti et al. investigated the effect of ZnO nanorods morphology on the electrical properties of perovskite solar cells, and in their findings, they concluded that the performance of perovskite solar cells can be increased by using ZnO nanorods as an electron and hole recombination barrier [11].

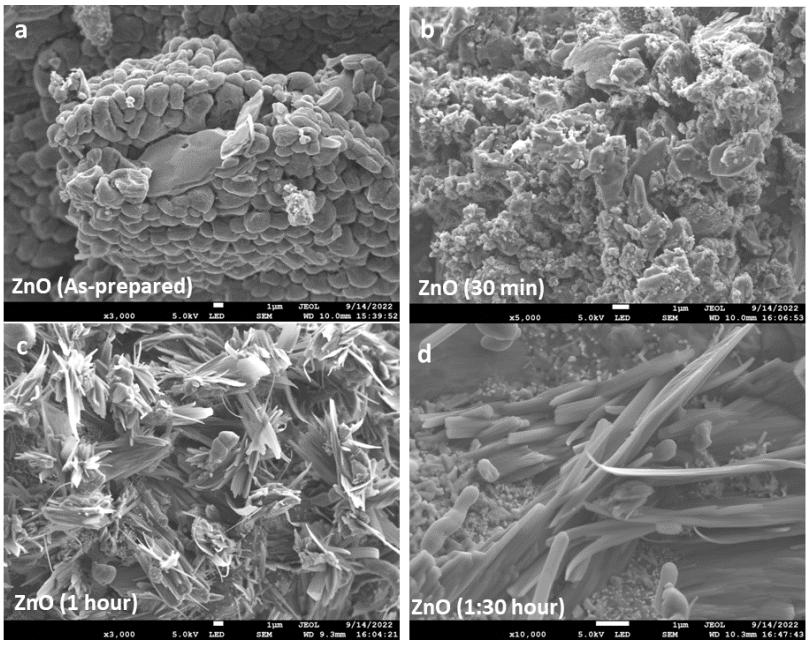


Figure 4. SEM images of ZnO As-prepared, ZnO (30 min), ZnO (1 hour), and ZnO (1:30 hour)

4. Conclusion

ZnO nanoparticles were successfully synthesized using the Sol-gel method. XRD confirmed the presence of a hexagonal structure with an average crystallite size of 8.12 nm. The optical analysis showed an improvement in the absorbance of the material when annealed at a longer time, as well as the presence of charge transfer, which as a result confirmed the recombination of photogenerated charge carriers being suppressed. SEM confirmed the evolution of ZnO morphologies from clustered nanoparticles to the formation of nanorods when annealed for 1:30 hour. The obtained results suggest that annealing time had a positive impact on the material in terms of improving its structure, morphology, and optical properties.

5. Acknowledgements

I would like to acknowledge the National Research Foundation (NRF) for funding this project. Grant number: 146621.

References

[1] Raoufi, D. and Raoufi, T., 2009. The effect of heat treatment on the physical properties of sol–gel derived ZnO thin films. *Applied surface science*, *255*(11), pp.5812-5817.

[2] Li, J., Yang, D. and Zhu, X., 2017. Effects of aging time and annealing temperature on structural and optical properties of sol-gel ZnO thin films. *Aip Advances*, *7*(6).

[3] Shahane, G.S., 2015. Structural, Morphological and Optical Properties of Sol-Gel Spin Coated Al-doped Zinc Oxide Thin Films. *Journal of Nanoscience and Nanoengineering*, *1*(4), pp.259-264.

[4] Kim, S., Nam, G., Yoon, H., Park, H., Choi, H., Kim, J.S., Kim, J.S., Kim, D.Y., Kim, S.O. and Leem, J.Y., 2014. Structural, optical, and electrical properties of ZnO thin films deposited by sol-gel dip-coating process at low temperature. *Electronic Materials Letters*, *10*, pp.869-878.

[5] Yoshino, K., Tanaka, M., Ide, A., Boden, S.A., Bagnall, D.M. and Yoneta, M., 2014. Surface Morphology of Transparent Conductive ZnO Film Grown by DC Sputtering Method. *Advanced materials research*, *894*, pp.403-407.

[6] Kaneva, N.V. and Dushkin, C.D., 2011. Preparation of nanocrystalline thin films of ZnO by sol-gel dip coating. *Bulg Chem Commun*, *43*, *pp*. 259-263.

[7] Santhosh, V.S., Rajendra Babu, K. and Deepa, M., 2014. Influence of Fe dopant concentration and annealing temperature on the structural and optical properties of ZnO thin films deposited by sol–gel method. *Journal of Materials Science: Materials in Electronics*, *25*, pp.224-232.

[8] Shekofteh-Gohari, M., Habibi-Yangjeh, A., Abitorabi, M. and Rouhi, A., 2018. Magnetically separable nanocomposites based on ZnO and their applications in photocatalytic processes: a review. *Critical Reviews in Environmental Science and Technology*, *48*(10-12), pp.806-857.

[9] Seetawan, U., Jugsujinda, S., Seetawan, T., Euvananont, C., Junin, C., Thanachayanont, C., Chainaronk, P. and Amornkitbamrung, V., 2011. Effect of annealing temperature on the crystallography, particle size and thermopower of bulk ZnO. *Solid state sciences*, *13*(8), pp.1599-1603.

[10] Tsotetsi, D, Dhlamini, M, Mbule, P., 2022. Sol-gel derived mesoporous TiO2: Effects of non-ionic co-polymers on the pore size, morphology, specific surface area and optical properties analysis. *Results in Materials,* 14, pp. 100266.

[11] Mufti, N., Idiawati, R., Wisodo, H., Laksono, Y.A., Fuad, A. and Diantoro, M., 2018, September. The Effect of ZnO Nanorods Morphology on Electrical Properties of Perovskite Solar Cells. In *Journal of Physics: Conference Series* (Vol. 1093, No. 1, p. 012028). IOP Publishing.

Kinetic analysis and dosimetric features of thermoluminescence of tanzanite

K A Opoku

Department of Physics and Electronics, Rhodes University, PO 94, Grahamstown, 6140, South Africa

E-mail: **akingsley75@yahoo.com**

Abstract. Tanzanite is a rare gem mineral of high commercial value. It is sensitive to optical and thermal stimulation of luminescence used to study point defects in insulators. Kinetic analysis and dosimetric features of thermoluminescence of tanzanite are reported. A glow curve measured at 1 °C s^{-1} following beta irradiation to 70 Gy shows a high intensity peak at 74 °C and two lower intensity peaks at 138 and 186 °C. The peaks are respectively referred to as I, II and III. The dependence of the peak position on partial heating and on irradiation dose is consistent with first order kinetics. The activation energy for peaks I-III are 0.81, 1.00 and 1.24 eV respectively. The dose response of each peak is sublinear as quantitatively determined using analysis of supralinearity indices.

1. Introduction

Tanzanite is a gem mineral found only at Merelani, Tanzania. The gem has a huge market demand owing to its aesthetic qualities linked to its pleochroism which is evident at different crystallographic orientations. Tanzanite is a type of zoisite [$Ca_2Al_3(Si_2O_7)(SiO_4)O(OH)$], a metamorphic mineral of the epidote group [1,2] and has an orthorhombic crystal structure [3]. The crystal structure is built of octahedral chains connected by single tetrahedra of SiO_4 and double tetrahedra of Si_2O_7 groups.

Zoisite is sensitive to thermal and optical stimulation of luminescence [4,5]. Despite this, there is only one report, namely, that of Chithambo and Folley [6] on the thermoluminescence (TL) of tanzanite. For TL to occur, an insulator or semiconductor is first irradiated with ionizing radiation to induce electron movement into the conduction band. Point defects within the crystal create charge trapping centres. The supply of heat leads to the excitement of trapped electrons into the conduction band followed by their recombination with trapped holes during which luminescence appears [7]. In the study of Chithambo and Folley [6], three TL peaks at 70, 130 and 180 °C were observed in measurements made at 1 °C s^{-1}, after 51 Gy beta irradiation. All peaks were seen to suffer thermal quenching. The TL of peak I ensues from recombination of electron occurring at $[TiO_4\backslash h^+]^0$ sites whereas the TL of peaks II and III appears owing to alkali recombination at $[TiO_4]^-$ sites [6]. Kinetic analysis was conducted only for the prominent peak, that is, peak I and that of peaks II and III was omitted. Fading of peaks II and III, due to delay in readout following irradiation meant that the peaks are not shallow traps, but the activation energy and frequency factor were not reported. To gain deeper insights into the TL of tanzanite, we compute the kinetic parameters and analyze the dose response of all three peaks.

2. Experimental procedure

The sample used was commercially available tanzanite (African Gems and Minerals, Cape Town). Samples were prepared in coarse grain form and used *as received*. TL measurements were made using

a RISØ TL/OSL DA-20 Luminescence Reader. The luminescence was detected by an EMI 9235QB photomultiplier tube through a 7 mm Hoya U-340 filter (transmission band 250 - 390 nm). The sample was irradiated *in-situ* using a $^{90}Sr/^{90}Y$ beta source at a dose rate of 0.10 Gy/s. All measurements were made in a nitrogen atmosphere to ensure good thermal contact between the heater plate and the sample holder. Unless indicated otherwise, all recordings were carried out at 1 °C s^{-1}.

3. Results and discussion

3.1. *Glow curve features*

Figure 1 depicts a glow curve corresponding to 70 Gy. Three peaks at 74 °C, 138 °C and at 186 °C (peaks I-III) are seen. A similar number of peaks and a weaker fourth one were observed by Chithambo and Folley [6]. Thermal cleaning to identify less prominent peaks in the glow curve revealed none.

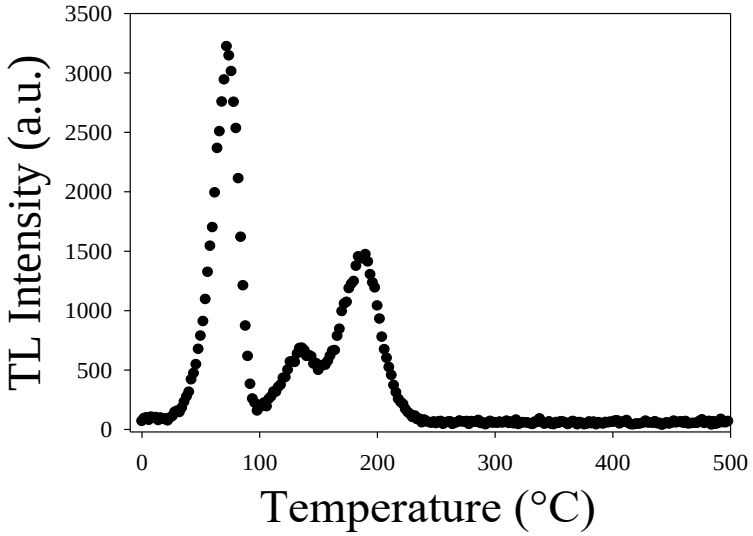


Figure 1. TL glow curve of tanzanite following 70 Gy beta irradiation. Three peaks at 74, 138 and 186 °C can be observed.

3.2. *Analysis for order of kinetics*

The order of kinetics for the peaks was assessed by examining how variable doses and partial heating affect the peak position (T_m). The partial heating procedure (T_m-T_{stop}) monitors the shifts in peak position as the sample undergoes partial heating to a temperature (T_{stop}) each time it is irradiated. For first order kinetics, T_m does not vary with either dose or preheat (T_{stop}) because the initial concentration of trapped charges has no bearing on T_m [7]. For second order kinetics, T_m decreases as dose increases [7]. A graph of T_m against T_{stop} is depicted in figure 2. The position of peaks I-III were determined as 74 ± 1, 139 ± 3 and 191 ± 3 °C. The parameter T_m also shows no dependence on dose at 73 ± 1, 138 ± 3 and 186 ± 2 °C respectively. Therefore all peaks are of first order kinetics.

3.3. *Stability of peaks*

Any change in the TL intensity owing to delay in readout after irradiation was analysed for periods between 0 to 10,000 s following irradiation to 10, 50 and 100 Gy. Figure 3(a) shows the intensity of peak I versus delay in readout corresponding to the three doses. Peak I fades with time for all doses. A similar behaviour was observed before in the literature [6].

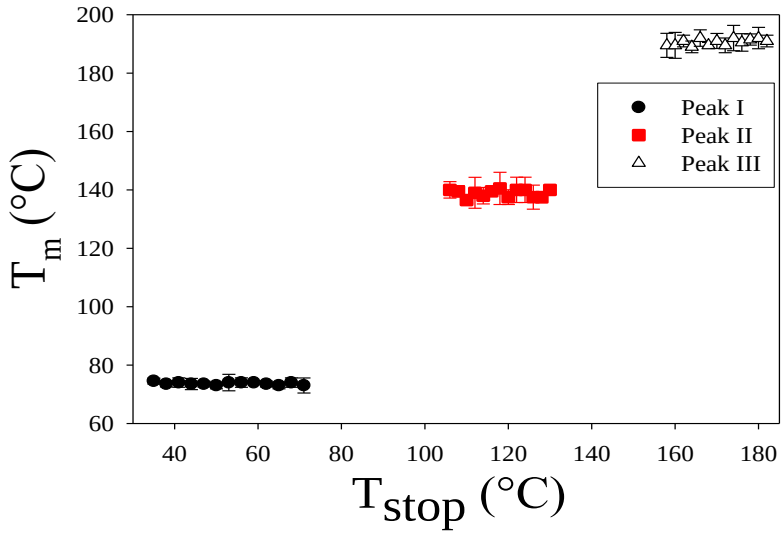


Figure 2. Peak position (T_m) dependence on partial heating (T_{stop}). The peak position remains stable for each peak. The mean of four measurements was evaluated to represent the value of each data point.

An empirical function that best fit the data is given by:

$$I(t) = a + I_0 e^{-\lambda t}, \tag{1}$$

where a is a constant, I_0 is initial intensity, t is delay and λ is decay constant. Figure 3(b) shows the intensity of peak II versus delay. The intensity is stable with time for all doses. A similar result was seen for peak III.

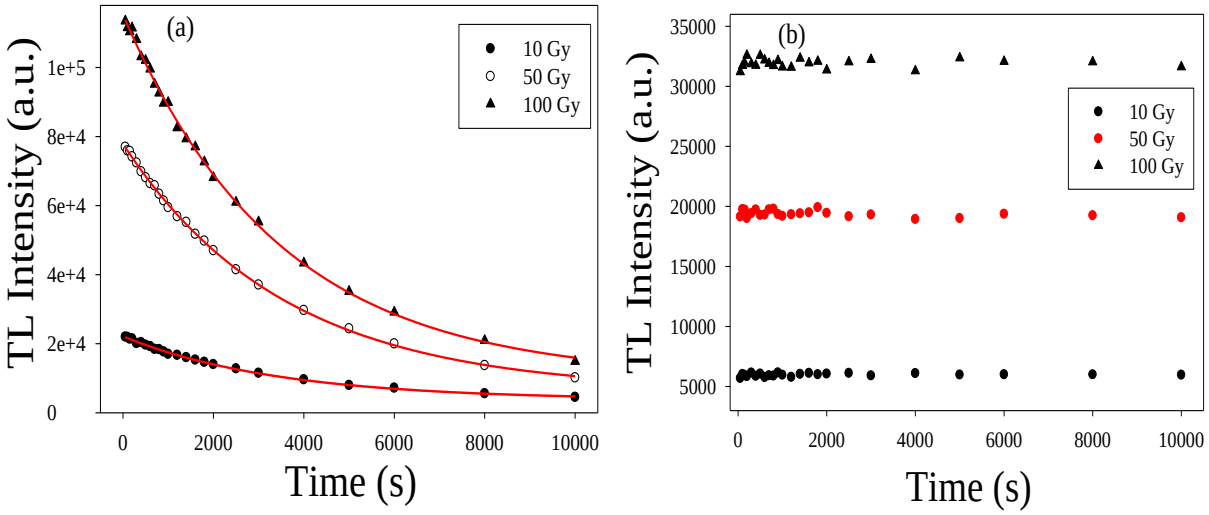


Figure 3. The relationship between intensity and delay in readout after irradiation for peak I (a) and peak II (b). The continuous lines in (a) represent the fits of equation (1).

3.4. *Dose response*

The dose response for all three peaks was analysed for doses from 10 to 200 Gy. The plot in figure 4(a) depicts the dose response for each peak. The intensity increases sublinearly with dose. The data is best fitted with the empirically chosen function:

$$y(D) = a\left(1 - e^{-bD}\right), \tag{2}$$

where a is intensity at saturation level, b is a constant and D is dose. Qualitative analysis of the dose response was carried out with the supralinearity index [7] given as:

$$f(D) = [y(D)/D]/[y(D_1)/D_1], \tag{3}$$

where D_1 represents normalization dose and $y(D_1)$ is a function dependent on the normalization dose. Figure 4(b) depicts the plot of $f(D)$ versus dose for all peaks. In such a test, values of $f(D) < 1$ indicates a sublinear dose response and $f(D) > 1$ shows a supralinearity response [7]. From figure 4(b), $f(D) < 1$ for all peaks which shows that the dose response of all peaks is sublinear.

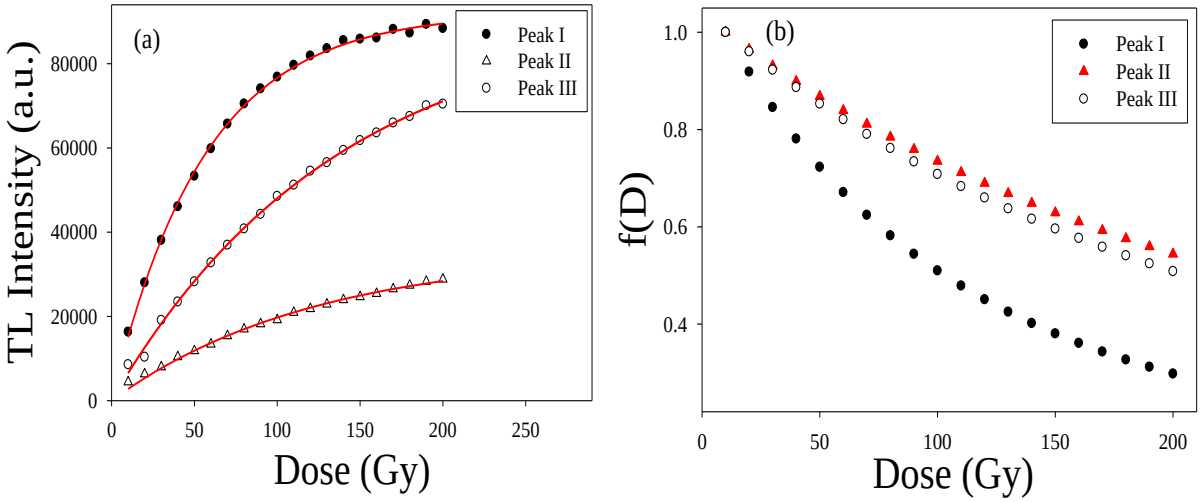


Figure 4. TL dose response for peaks I-III (a), and supralinearity index function $f(D)$ versus dose for peaks I-III (b). The continuous lines in (a) represent the fits of equation (2).

3.5. *Kinetic analysis*

Kinetic parameters such as the activation energy (E), frequency factor (s) and kinetic order (b) provide insight into the luminescence processes. The activation energy is the energy required to detrap an electron into the conduction band. The frequency factor is the number of times per second the electrons interacts with lattice phonons. The parameter b denotes the order of kinetics [7]. Results using various analytical methods used for the kinetic analysis in this work are now presented:

3.5.1. *Curve fitting method (CF).* The glow peaks were fitted with the Kitis et al general order curve fitting equation [8] given as:

$$I(T) = I_m b^{\frac{b}{b-1}} e^{\left(\frac{E}{kT}\frac{T-T_m}{T_m}\right)} \left[1 + (b-1)\frac{2kT_m}{E} + (b-1)\left(1 - \frac{2kT}{E}\right)\frac{T^2}{T_m^2} e^{\left(\frac{E}{kT}\frac{T-T_m}{T_m}\right)}\right]^{-\frac{b}{b-1}}, \tag{4}$$

where I is intensity, T is temperature, k is Boltzmann constant, I_m is maximum intensity and all other parameters retain their usual meanings. The parameter s was evaluated from the expression [7]:

$$s = (\beta E/kT_m^2)e^{(E/kT_m)}, \tag{5}$$

where β is heating rate. As an example, the plot for peak I is depicted in figure 5. The parameter E yielded 0.87 ± 0.02, 0.95 ± 0.07 and 1.13 ± 0.01 eV for peaks I-III respectively. The values of b were 0.96 ± 0.01, 1.2 ± 0.01, and 1.2 ± 0.01 for the respective peaks. Since b for all the peaks are closer to 1,

all peaks are of first order kinetics. From equation (5), s was evaluated as 4.3×10^{11}, 3.4×10^{10} and 1.4×10^{11} s^{-1} respectively. The values of s are comparable to the order of the lattice vibration frequency which is approximately 10^{12} s^{-1}.

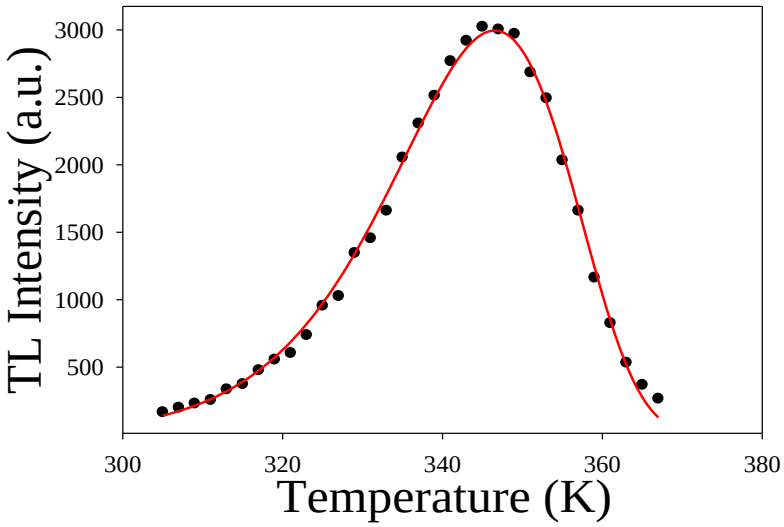


Figure 5. General order curve fitting for peak I. The solid line represents the fit of equation (4).

3.5.2. Variable heating rate method (VHR)

Recording the TL at different heating rates alters the peak position. The TL was recorded at heating rates of 0.5 to 4.5 °C s^{-1} after 70 Gy irradiation each time. The parameters E and s were evaluated using the expression [7]:

$$\ln(T_m^2/\beta) = E/(kT_m) + \ln(E/sk). \quad (6)$$

A plot of $\ln(T_m^2/\beta)$ against $1/(kT_m)$is linear as shown in figure 6. The slope of the plot is equal to E, evaluated as 0.78 ± 0.02, 1.01 ± 0.06 and 1.39 ± 0.11 eV for the respective peaks. The parameter s was evaluated from the y-intercept as 1.6×10^{10}, 1.7×10^{11} and 1.0×10^{14} s^{-1} respectively.

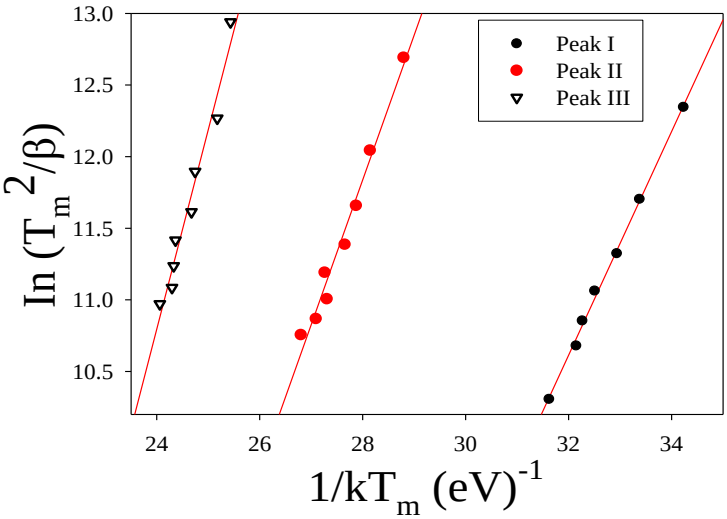


Figure 6. Plot of $\ln(T_m^2/\beta)$ versus $1/(kT_m)$ for peaks I-III. Measurement of the TL was done at different β from 0.5 to 4.5 °C s^{-1} after 70 Gy irradiation each time. The continuous lines are the most linear trends through the data points.

3.5.3. *Whole glow peak method (WGP)*
The integral of the area (A) beneath a peak can be written as a function of temperature as:

$$\ln(I/A^b) = \ln(s/\beta) - E/kT, \quad (7)$$

where a graph of $\ln(I/A^b)$ versus $1/kT$ will yield a linear line for an appropriate value of b [7]. The parameters E and s are computed from the gradient and y-intercept accordingly. Figure 7 depicts such plot for various values of b for peak I. The line with the highest coefficient of determination (r^2) value of 0.9977 corresponds to b=0.9. For peaks II and III, b=1.3 yields the most linear line. The values of E hence found are 0.77 ± 0.01, 1.03 ± 0.02 and 1.20 ± 0.04 eV for the respective peaks. The parameter s was evaluated as 1.3×10^{10} for peak I, 1.7×10^{10} for peak II and 2.8×10^{10} s^{-1} for peak III.

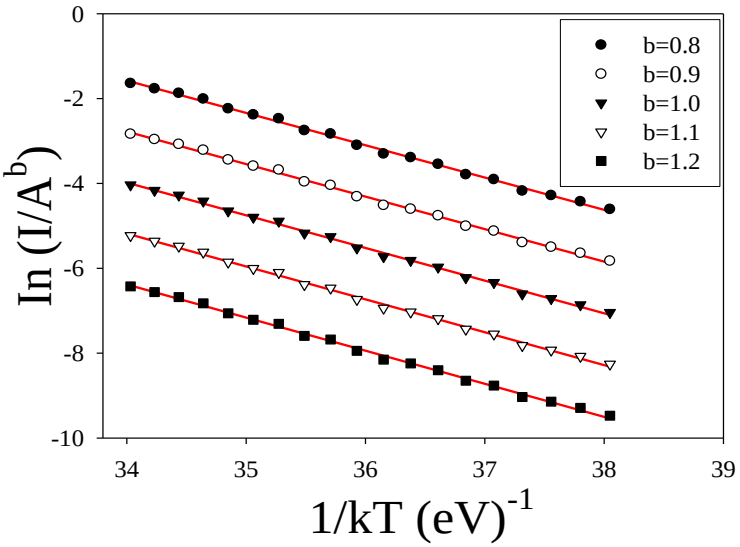


Figure 7. Graph of $\ln(I/A^b)$ against $1/kT$ for peak I. The continuous lines are the most linear trends through the data points.

3.5.4. *Summary of kinetic parameters*
Table 1 depicts the evaluated kinetic parameters for all peaks. The kinetic parameters for peak I are compared with those reported in literature. The parameters compare favourably. Also, the results derived from the different analytical methods are comparable. The mean of the activation energy E is 0.81, 1.00 and 1.24 eV for the respective peaks. The kinetic parameter E for peaks II and III suggest the peaks are not shallow electron traps. This is corroborated by the fact that the peaks do not show any phosphorescence at room temperature. The full meaning of CF, VHR and WGP as used in the table has been specified in sections 3.5.1 to 3.5.3 respectively.

4. Conclusion

We have observed three TL peaks in tanzanite at 74, 138 and 186 °C for a glow curve measured at 1 °C s^{-1} after 70 Gy beta irradiation. All peaks are of first order kinetics. Peak I fades with delay in readout after irradiation, in contrast to peaks II and III which do not. The dose response of peaks I-III is sublinear. The activation energy for peaks I-III are about 0.81, 1.00 and 1.24 eV respectively. The frequency factor is within the order of 10^{11}, 10^{10} and 10^{13} s^{-1} for the respective peaks.

Table 1. Kinetic parameters for peaks I, II and III.

Peaks	Analytical Method	E (eV)	s (s^{-1})	b	Reference
Peak I	CF	0.87 ± 0.02	4.3×10^{11}	0.96 ± 0.01	Figure 5
		0.860 ± 0.004	2.0×10^{11}	0.970 ± 0.003	Chithambo and Folley [6]
	VHR	0.78 ± 0.02	1.6×10^{10}		Figure 6
		0.86 ± 0.03	3.0×10^{11}		Chithambo and Folley [6]
	WGP	0.77 ± 0.01	1.3×10^{10}	0.9	Figure 7
		0.85 ± 0.01	3.0×10^{11}	1.1	Chithambo and Folley [6]
Peak II	CF	0.95 ± 0.07	3.4×10^{10}	1.20 ± 0.01	Current work
	VHR	1.01 ± 0.06	1.7×10^{11}		
	WGP	1.03 ± 0.02	1.7×10^{10}	1.3	
Peak III	CF	1.13 ± 0.01	1.4×10^{11}	1.20 ± 0.01	
	VHR	1.39 ± 0.11	1.0×10^{14}		
	WGP	1.20 ± 0.04	2.8×10^{10}	1.3	

Acknowledgement
This research was conducted under the supervision of Prof. M L Chithambo and was generously funded by Rhodes University and the National Research Foundation of South Africa.

References

[1] Mao Z, Jiang F and Duffy T S 2007 *Am. Min*. **92** 570-576
[2] Pluthametwisute T, Wanthanachaisaeng B, Saiyasombat C and Suthirat C 2020 *Molecules* **25** 3743
[3] Xu J, Zhang D, Fan D, Wu X, Shi F and Zhou W 2019 *Phys. Chem. Miner*. **46** 333–341
[4] Chithambo M L 2021 *J. Lumn*. **234** 117969
[5] Ccallata H J and Watanabe S 2010 *AIP Conf. Proc.* (Santiago) Vol 1265, 391-394
[6] Chithambo M L and Folley D E 2020 *Phys. Condens. Matter* B **598** 412435
[7] Chen R and Pagonis V 2011 *Thermally and optically stimulated luminescence: a simulation approach* (Chichester: John Wiley & Sons)
[8] Kitis G, Gomez-Ros J M and Tuyn J W N 1998 *J. Phys.* D **31** 2636–2641

Characterization of defects in ZnO implanted with Ar^+ ions using positron annihilation technique.

M E Khulu[1], T P Jili[1], M Madhuku[2], C L Ndlangamandla[1] and P Z Ngcobo[1]

[1]Department of Physics, University of Zululand, P/B X1001, KwaDlangezwa, 3886, South Africa

[2]IThemba LABS, Empire Road, Braamfontein, Gauteng, South Africa

E-mail: **musakhulu.mk@gmail.com**

Abstract. ZnO (wurtzite) samples were implanted with Ar^+ ions to generate intrinsic defects within the ZnO samples at fluence range from 1×10^{15} to 3×10^{16} ions.cm^{-2}. Doppler broadening of the annihilation curves were obtained to determine S-parameters which are used to characterize the defects. S-parameters are 0.35795, 0.35809 and 0.39025 for the lowest to the highest fluence, respectively. X-ray diffraction method was employed to determine any structural damage or phase change during the implantation. Positron annihilation spectroscopy shows the formation of anion point defects. Optical absorption measurements suggest the presence of F aggregate centres at 300.0 nm (4.14 eV). Theoretical calculations of the annihilation curves, based on local density approximation as well as generalized gradient approximation, were obtained. Theoretical results compare well with experimental results. Positron annihilation measurements are also correlated with optical absorption results on Ar^+ ion implanted zinc oxide crystal.

1. Introduction

Zinc oxide (ZnO) has drawn much attention because of its properties in device engineering such as transparent ohmic contacts, nano-scale devices as well as in gas sensing [1,2,3,4,]. The ZnO band gap makes it an extremely suitable material for low noise generation, higher breakdown voltages and the ability to withstand large electric fields [5,6].

Positron annihilation technique (PAT) has become a sensitive tool in quantifying defects as well as obtaining the nature of defects. PAT is a non-destructive tool capable of identifying very small to large defects, for example, point to dislocations.

In this research we investigate the radiation damage in ZnO due to 150 keV Ar^+ ion implantation in the fluence range $1 \times 10^{15} - 3 \times 10^{16}$ ions cm^{-2} through the determination of S- and W- parameters. The experimentally obtained Doppler broadening, which is characterized by the S- parameter, compares well with the theoretically determined annihilation curve. The experimental part of the Doppler broadening tells us about the quantity of defects in the material. S-parameter is defined as the area under annihilation centroid divided by the total area, and it is a measure of positron-electron annihilation momentum density at annihilation sites. Positron annihilation results are also correlated with optical absorption measurements in ion implanted zinc oxide materials.

2. Experiment

Pure ZnO crystal samples (space group P63mc) of dimensions 10 mm × 10 mm × 0.5 mm and orientation of (100), were obtained from PAM-XIAMEN in China. ZnO samples were implanted with 150 keV Ar^+ ions in the fluence range 10^{15} to 3×10^{16} ions cm^{-2}. Implantations were performed at iThemba LABS, Gauteng, South Africa. Optical absorption was performed using spectrophotometer to characterize the defects in the ion implanted ZnO target. The spectra of optical absorption were measured at wavelengths between 200 nm and 800 nm.

X ray diffraction (XRD) was used to investigate whether there was any structural damage or phase change that might have occurred during the ion implantation process. XRD method utilizes Bruker D8 Advance fully equipped with Goebel Mirror, a Vantec detector and an Anton Paar XRK 900 chamber. The diffractometer uses Cu-Kα radiation. XRD patterns of implanted ZnO samples in the fluence range $10^{15} - 3 \times 10^{16}$ cm^{-2} were obtained. Rietveld refinements were carried out using TOPAS 4.2.

The experimental arrangement consists of radioactive source, ^{22}Na, whose decay scheme is shown in figure 1. The activity of the source is 20 μCi. The source is sandwiched between two equal ion-implanted ZnO samples. The coincidence positron annihilation experimental circuit arrangement is shown in figure 2.

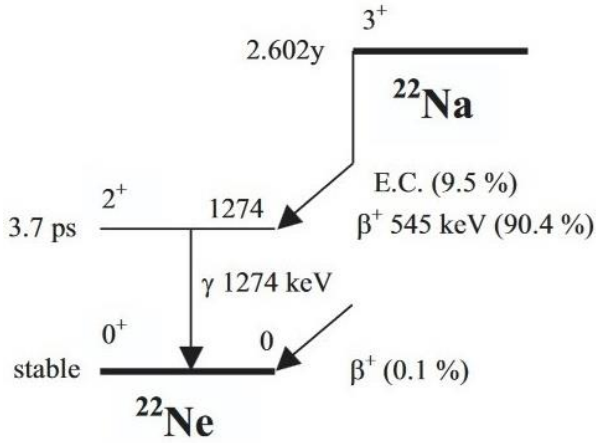


Figure 1. The decay scheme of ^{22}Na into ^{22}Ne with emission of 1.274 MeV photon [7].

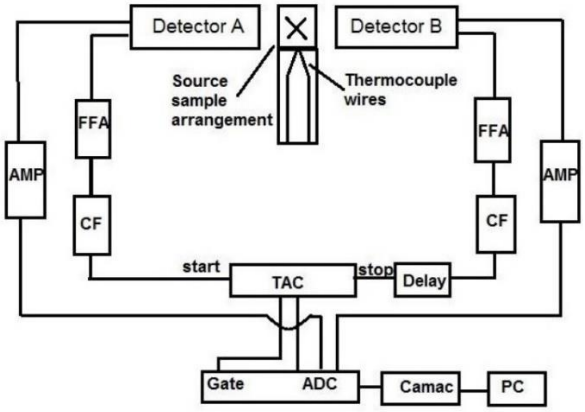


Figure 2. The circuit diagram used for the Doppler broadening and lifetime measurements[12].

Source-sample arrangement was kept at a pressure of about 10^{-5} Torr. Positron annihilation spectra at different fluencies were accumulated at about 2 hours for each run given the low average count rate of 150 s^{-1}. Source correction was also conducted given that between 10 and 15 % of the annihilations of positrons take place in the source material as well as in the foils.

3. Method of calculation

The two-component density functional theory (TCDFT) is needed to solve positron and electron densities. The use of local density approximation (LDA) simplifies the complication of many unknowns in the TCDFT, like the exchange correlation energy and the positron-electron correlation-energy functional but the LDA does not consider the inhomogeneous charge density of the electron structure. The generalized gradient approximation (GGA) is then used because it considers the gradients of the electron density. Furthermore, this can also be simplified by calculating the electron density and Coulomb potential using atomic superposition method (ASM) which simply means superimposing free-atom densities and potentials. For free-atom densities, it is necessary to calculate electron and positron wavefunctions. For positrons [8]:

$$-\frac{1}{2}\nabla^2\psi_i^+(\boldsymbol{r}) + V_{eff}(\boldsymbol{r})\psi_i^+ = \varepsilon_i\psi_i^+(\boldsymbol{r}). \quad (1)$$

where V_{eff} is the effective potential given by:

$$V_{eff}(\boldsymbol{r}) = \phi(\boldsymbol{r}) + \frac{\delta E_{exc}(n_+)}{\delta n_+(\boldsymbol{r})} + \frac{\delta E_c^{e-p}(n_+,n_-)}{\delta n_+(\boldsymbol{r})}, \quad (2)$$

and the Coulombic potential $\phi(r)$ is given by

$$\phi(\boldsymbol{r}) = \int d\boldsymbol{r}' \frac{-n_-(\boldsymbol{r}') + n_+(\boldsymbol{r}') + n_o(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|},$$

where n_o is the charge density in the framework of external potential, E_{exc} is the energy of exchange-correlation and E_c^{e-p} is the electron-positron correlation functional.

The electron and positron densities are then obtained from solving equation (1) to give

$$n_- = \sum_{\varepsilon_i \le \varepsilon_f} |\psi_i(r)|^2 \text{ and } n_+ = \sum_i^{N_+} |\psi_i^+(r)|^2. \quad (3)$$

The sum of electron states is taken from core states to the Fermi level. Normally, we consider one positron, that is, $N_+ = 1$. The annihilations of positrons with core electrons (high momentum) to the valence electrons (low momentum) as well as in the vacancies are then calculated to give electron-positron annihilation momentum density, ρ, as [8]:

$$\rho_j(r) = \pi r_e^2 c u_j^2(0) \left| \int dr e^{-ip\cdot r} \psi_+(r) \psi_j(r) \right|^2, \quad (4)$$

where r_e refers to classical electron radius and u_j^2 the state-dependent enhancement factor.

Equation (4) gives the annihilation curve which is used as a comparison to the annihilation curve obtained experimentally. S-parameter which gives the quantity of defects in the sample, is given as the following ratio [9]:

$$S = \frac{\int_a^b g(p_z) dp_z}{\int_{-\infty}^{+\infty} g(p_z) dp_z} . \quad (5)$$

The numerator in equation (5) refers to the total counts under the annihilation centroid of width b-a in the momentum or energy axis. The denominator in equation (5) refers to the total counts under the annihilation curve.

4. Results and discussion

X-ray diffraction was performed on all Ar^+ implanted single crystal ZnO samples to investigate any structural damage that might have occurred during implantation. The XRD patterns were compared with reference pattern from Crystal Structure Database [10] for peaks identification. Figure 3 clearly shows that there is no phase change, and no structural damage was observed.

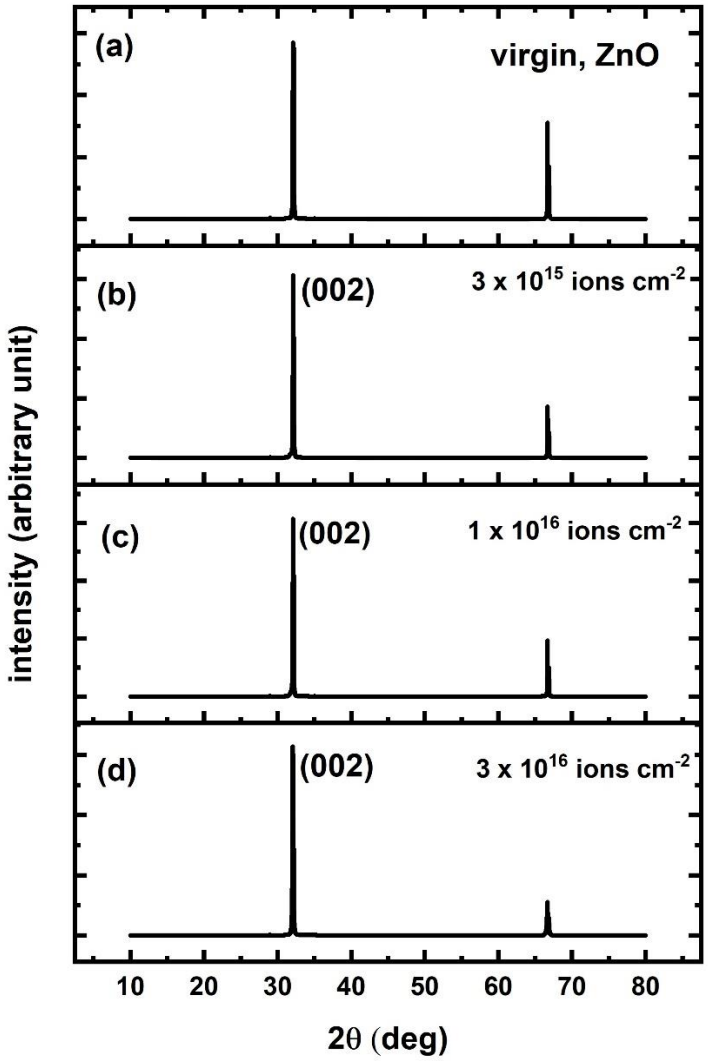


Figure 3. X-ray diffraction pattern of implanted zinc oxide

Optical absorption spectra were measured in all Ar^+ implanted zinc oxide samples. The F centres appeared at a wavelength of 300.0 nm (4.14 eV) in the fluence range between ($1 \times 10^{15} – 3 \times 10^{16}$) ions cm^{-2} as shown in Table 1. It was noticed that the F band grows with the increase in fluence, reaching its maximum at 3×10^{16} ions per square cm. This suggests an anion vacancy which has trapped an electron. The growth is characteristic of ion implanted semiconductor [12].

Table 1. F-band absorbance in all three Ar+ implanted ZnO samples

Fluence (ions cm^{-2})	Wavelength (nm)	Absorbance (arb.units)	Energy (eV)
1×10^{15}	289	0.845	4.30
1×10^{16}	300	4.446	4.14
3×10^{16}	309	4.526	4.02

The annihilation spectra for all implanted samples were analysed and S-parameter was calculated for each spectrum. A typical spectrum is shown in figure 4. Figure 5 shows the calculated S-parameters at all three fluences. S-parameter is a measure of the quantity of defects with a ratio of the centroid to the total area under the annihilation curve around 511 keV.

The calculated positron-electron annihilation momentum densities with core states (high momentum electrons) to low momentum electrons (mostly valence electrons) as well as annihilations in the vacancies are shown in figure 6. MIKA doppler code [11] was used to obtain momentum densities by employing GGA. The localisation of trapped positron in Zn vacancy is shown in figure 7.

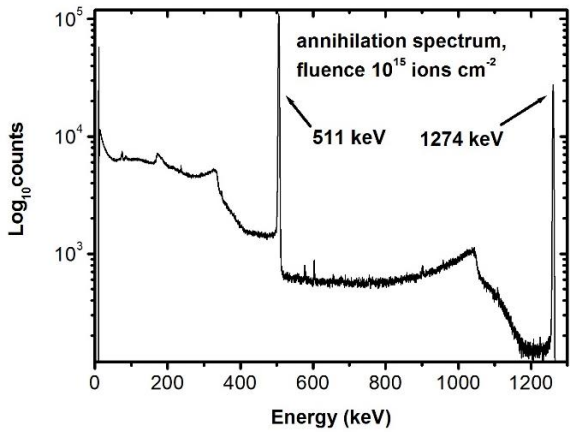


Figure 4. A typical annihilation spectrum showing annihilation line at 511 keV and 1274 keV gamma emitted during a β^+ decay.

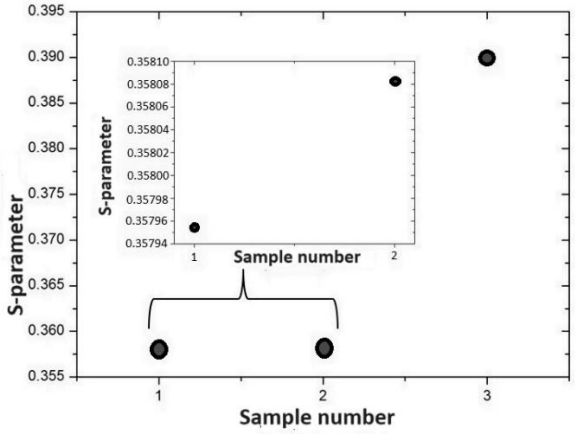


Figure 5. S-parameters calculated from the Doppler broadening of each spectrum around 511 keV. The insert shows the S-parameter difference between 1 and sample 2.

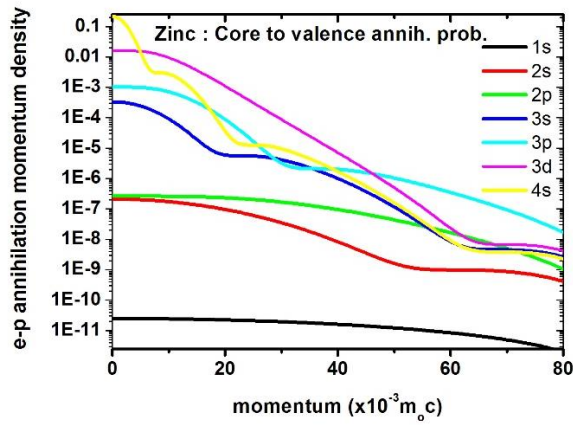


Figure 6. e-p annihilation momentum density clearly shows the dominance of 4s state between $(0 – 4) \times 10^{-3}$ $m_o c$.

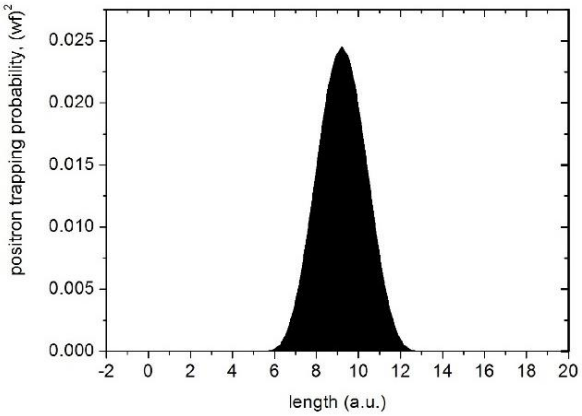


Figure 7. The localization of positron in Zn vacancy. This is measured by the square of positron wave function at the defect site.

The spatial distribution of defects due to ion implantation according to the simulation code TRIM increases with fluence. It is also noted that using TRIM simulation there is a possibility of knock-on effect which is the extension of damage beyond the distance travelled by implanted ions.

5. Conclusion

Positron annihilation technique, which characterizes defects in ZnO implanted with 150 keV Ar^+ ions at fluence range 10^{15} to 3×10^{16} ions cm^{-2}, reveals that positrons annihilate at vacancies as well as in the bulk. It is also interesting to note that part of the annihilations is dominated by 4s – state of Zn and in the range $(0 – 3) \times 10^{-3}$ m_oc, while O 3p – state dominates in the range $(0 – 25) \times 10^{-3}$ m_oc. These annihilations contribute towards the electron-positron momentum density or Doppler broadening. Optical absorption measurements reveal that the positrons annihilate with low momentum electrons at F aggregate centres. The quantity of these aggregates increases as S-parameter increases and showing a large number of aggregates at maximum fluence.

Acknowledgement

The authors would like to thank the NRF and the Research Committee of the University of Zululand for financial support. We also like to thank iThemba LABS for providing the ion implantation facility.

References

[1] Franco MA, Conti PP, Andre RS and Correa DS, 2022 *Sensors and Actuators Reports*, **4** 100100
[2] Li Z, Askim JR and Suslick KS, 2019 *Chem. Rev*. **119** 231-292
[3] Natale CD, Paolesse R, Martinelli E and Capuano R, 2014 *Anal. Chem. Acta* **824** 1-17
[4] Nazemi H, Joseph A, Park J and Emadi A, 2019 *Advanced micro-and nano-gas sensor technology: a review Sensors*, **19** 1285
[5] Ozgur U, Avrutin V and Morkoc H, 2018 *Molecular Beam Epitaxy, Richmond, VA, Elsevier Inc*., 343 – 375
[6] Coleman VA and Jagadish C, 2006 *Zinc Oxide Films and Nanostructures, 1st ed., Canberra, Elsevier* 1 - 20
[7] Prochazka L 2001 *Material Structure* 8 number **2** 55 - 60
[8] Barbiellini B, Puska MJ, Torsti T and Nieminen RM 1995 Phys. Rev. **B51** 7341 – 7344
[9] Hakala T 2001 *Special Assignment*, Helsinki University of Technology, Espoo, Finland
[10] ICDS:FIZ – Karlsruhe-Leibnitz Institute for Information Infrastructure
[11] Schultz PJ and Lynn KG, 1988 *Rev. Mod. Phys*. **60** 701
[12] Sendezera EJ, Davidson AT, Jili PT, Chitambo ML, Anwand W, Brauer G and Nicht E-H, 2002 *Nuclear Instruments and Methods in Physics Researcb* B, **192** 202 - 205

Probing the stability of nickel titanium (100) and (110) surfaces: A DFT study

Vukosi Chauke[1*], David Tshwane[2], Phuti Ngoepe[1] and Hasani Chauke[1]

[1]Materials Modelling Centre, University of Limpopo, Private Bag X1106, Sovenga 0727, South Africa
[2]Next Generation Enterprises and Institutions Cluster, CSIR, PO Box 395, Pretoria, 0001, South Africa

E-mail: **vukosichauke721@gmail.com**

Abstract. Density functional theory was employed in this study to investigate the surface stability of NiTi (110) and (100). Surface energies were analysed, and it was found that the NiTi (110) surface with the lowest surface energy value is more stable than the (100) surface. It was also discovered that the surface energy stability varies with surface termination. In addition, it was also noted that the (100) surface has larger work function value than the (110). The total and partial density of states of (100) and (110) surfaces of NiTi were also computed and analysed. The findings will have a good impact on the understanding of the surface stability of NiTi surfaces.

1. Introduction

NiTi-based intermetallic compounds are known as shape memory materials and are frequently utilized in industries like aerospace, machinery, medical applications, and electronics. This is due to their structure memory effect, super elastic behaviours, high tensile strength, and biocompatibility. However, this NiTi alloy is still a contentious material due to its unstable surface and oxidation during use, more importantly, their restricted surface characteristics and adsorption process [1]. Many fields of scientific investigation and commercial production processes require material surface qualities that are highly predictable and stable [2]

Despite significant experimental efforts to improve the corrosion resistance of NiTi, a detailed understanding of the atomic-scale mechanisms that govern the adsorption and diffusion of oxygen and water on NiTi surfaces is insufficient. This is primarily due to the complex nature of the NiTi alloy, which contains a mixture of titanium and nickel in different phases, resulting in a complex surface morphology [3].

Kulkova *et al*. [4] used fully potential linearized augmented plane-wave approach to investigate electronic structure of low-index surfaces of NiTi and TiPd. It was reported that NiTi (010) has the lowest energy among other surface considered surfaces. First principle computations were used to examine the low-index (001), (110) and (111) surfaces of austenitic NiTi. According to one investigation, under the majority of the Ti-chemical potential, the (110) NiTi surface was the most stable. Also emphasized was the fact that Ni-terminated surfaces were more stable than their Ti-terminated counterparts.

It is essential to study how different crystallographic surfaces of a material behave under various conditions. The properties and stability of crystal surfaces can vary significantly depending on their orientation. By specifying (100) and (110) surfaces in this study indicates that we are investigating the stability of these specific crystal facets of the NiTi material [5]. The research employs DFT, a quantum mechanical computational method, to investigate the electronic and structural properties of Ti-Ni

surfaces and their interactions with oxygen. DFT calculations provide valuable insights into the energetics and stability of surface reactions, making it a powerful tool for studying surface processes at the atomic level. In this work, DFT approach was employed to investigate surface properties of NiTi (100) and (110) surfaces as the lowest miller index. The surface energies, density of states and work function at different slabs were computed to identify the most stable surface.

2. Methodology

All DFT calculations involved in the present study were carried out using the CASTEP code [6]. The Perdew-Burke-Ernzerhof function among generalized gradient approximation (GGA-PBE) was applied to describe the exchange-correlation energy and potential, and the ultra-soft pseudopotential was selected to determine the electrons-ionic cores interactions.

The Monkhorst-Pack scheme was used to generate the k-point grids. A k-point mesh of 4 × 4 × 1 with a cut-off energy of 450 eV was adopted to optimize surfaces. The calculated lattice parameters were a = b = c = 6.014 Å, α = β = γ = 90° for the bulk NiTi. These calculated results were in good agreement with the reported literature [7], proving the reliability of our calculated methods.

Subsequently, a NiTi (110) and (100) for Ti and Ni surface were cleaved from the bulk and different layers symmetrical slab with a vacuum thickness of 30 Å was built to calculate the surface properties using the surface energy formula [8]:

$$\gamma = \frac{(E_{slab} - NE_{bulk})}{2A} \quad (1)$$

E_{surf} and E_{bulk} represent the total energies of the slab and bulk of NiTi. Due to the presence of two symmetrical surfaces surrounding the slab, this equation's 1/2 component is considered, A and N stand for the surface's area and atomic number.

3. Results and discussion

3.1. Surface properties and energies

To elucidate the atomic-level chemical reactions occurring at the surface plane, a dependable model of the NiTi surface was employed, taking into account factors such as the slab thickness and the presence of a vacuum region. The characteristics and composition of these surfaces play a paramount role in governing their properties, where even minor disparities in atomic arrangements can exert substantial effects. As illustrated in Figure 1(a), 1(b) and 1(c), we present the optimized configurations of the Ti (100), Ni (100), and NiTi (110) surface layers, which are delineated by a vacuum region measuring 30 Å. The Ti (100) and Ni (100) surfaces are composed of 5 atomic layers, each terminated by four-fold coordinated titanium and nickel atoms, respectively. In a similar fashion, the NiTi (110) surface is comprised of 5 atomic layers, with terminations consisting of four-fold coordinated titanium and nickel atoms.

Table 1 shows the calculated surface energies for Ti (100), Ni (100) and NiTi (110) surfaces with different terminations. Surface energy is a measure of the energy difference between the atoms or molecules on the surface of a material and those in the bulk [9]. A lower surface energy indicates that the material's surface atoms are more tightly bound and have a higher degree of stability. In this case, NiTi (110) surface has the lowest surface energy among Ti (100) and Ni (100), this suggests that the NiTi (110) surface is more energetically favourable. This can have implications for properties such as corrosion resistance and adhesion.

3.2. Density of states.

3.2.1. Partial density of states.

In the figure 2(a), 2(b) and 2(c), the s-orbital has the smallest contribution to the DOS. It remains relatively low and almost constant across the energy range. The p-orbitals have more significant contribution compared to s-orbital, the p-orbitals play a more substantial role in the electronic states near the Fermi-level and contribute to the metallic bonding of Ti, Ni and NiTi.

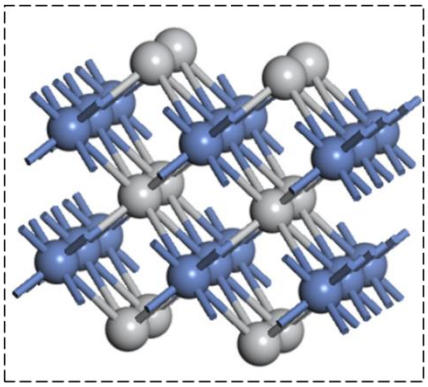


Figure 1(a) Computed atomistic NiTi surface, Ti (100) terminated where grey atoms represent titanium and blue atoms represent nickel.

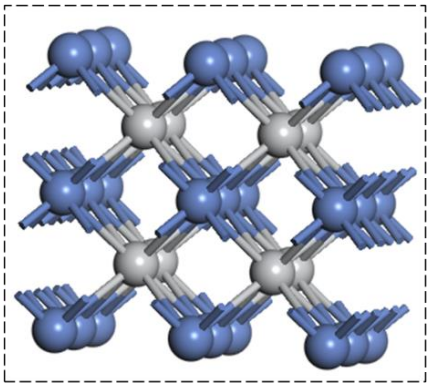


Figure 1(b) Computed atomistic NiTi surface, Ni (100) terminated.

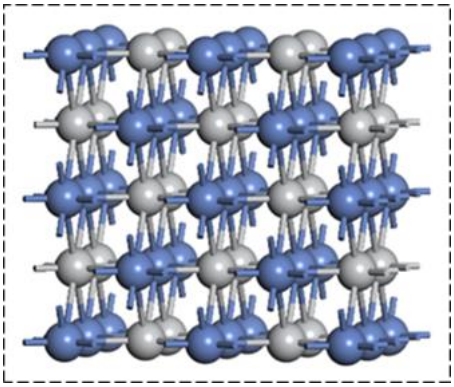


Figure 1(c). Computed atomistic NiTi (110) surface terminated.

Table 1. Calculated surface energy for different terminations.

Surfaces	Terminations	Area ($Å^2$)	$E_{surface}$ ($eV/Å^2$)
(100)	NiTi	486.46	0.00873
	Ti	186.52	1.15110
	Ni	186.52	1.18024
(110)	NiTi	486.46	0.00129
	Ti	186.52	0.00872
	Ni	186.52	0.00691

The d-orbitals dominate the DOS near the Fermi-level with several sharp peaks. This indicate that the d-orbitals play a crucial role in determining the metallic properties and chemical bonding of Ti, Ni and NiTi.

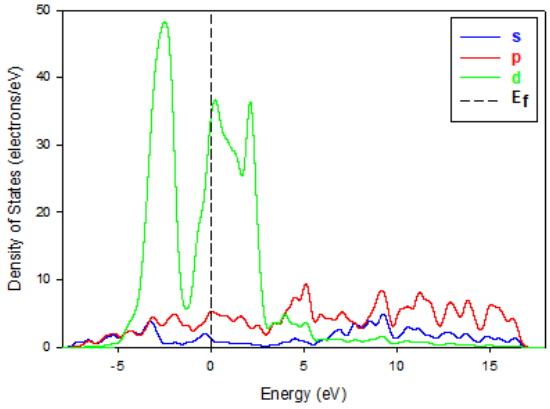


Figure 2(a). Partial density if states for Ti (100) surface.

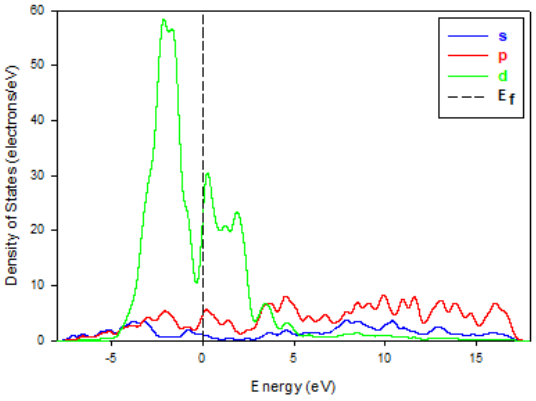


Figure 2(b). Partial density of states for Ni (100) surface.

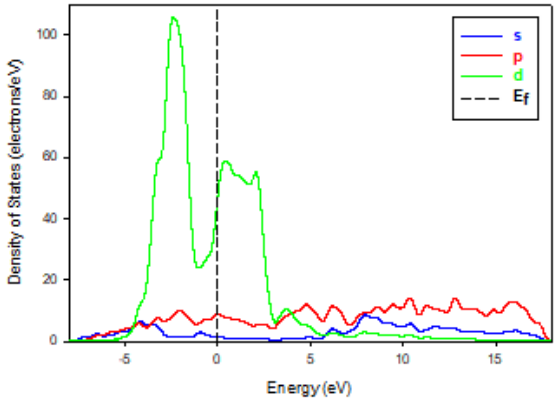


Figure 2(c). Partial density of states for NiTi (110) surface.

3.2.2. Total density of states

The total density of states (TDOS) in figure 3 represent NiTi, Ti and Ni. The TDOS for NiTi is dominant, since it is the resultant DOS of both Ti and Ni. The DOS value indicates the density of electronic states per electron-volt of energy. At the Fermi energy Ef, there is a relative high density of state [10]. This suggests a relatively high availability of energy levels for electrons to occupy near the Fermi level. The DOS is symmetric around the Fermi level, which is a common feature in metallic materials. The peak in DOS around the Fermi level suggests that there are energy levels where electrons are likely to be found, contributing to NiTi's metallic behaviour.

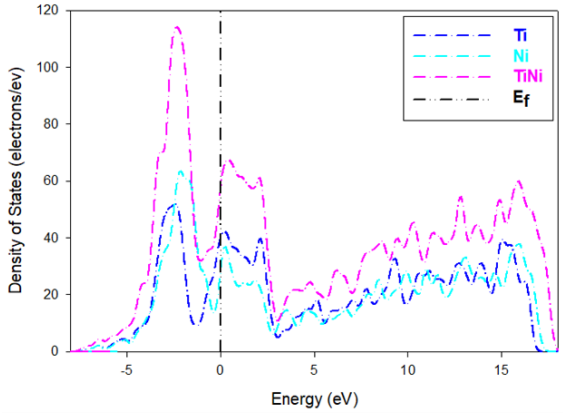


Figure 3. Total density of states for Ti, Ni and NiTi.

3.3. Work function

An important factor to take into account while examining metallic surfaces, also known as electrostatic potential, is the work function (ϕ). The work function, which is defined as the lowest amount of energy required to transport an electron away from a metallic surface in a vacuum, is represented by the formula:

$$\phi = E_{vac.} - E_F \quad (2)$$

where electronic potential energy of the vacuum and Fermi energy level represented by $E_{vac.}$ and E_F, respectively. The work function of a material is the minimum amount of energy required to remove an electron from the surface of the material to a point just outside the surface [11]. The work function provides insights into the reactivity of the material's surface [12]. A higher work function often correlates with lower electron affinity and lower reactivity. Figure 4(a) and figure 4(b) suggests that Ti and Ni are less prone to oxidation or chemical reactions with other substances. Figure 4(c) has the lowest work function value, and hence implying that NiTi (110) is the most reactive. The charge rearrangement of electrons and ions close to the surface mostly determines the electrostatic energy value. Because the work function of a metal surface often depends on its crystallographic orientation, it is possible for various surfaces to have slightly varied work functions [13].

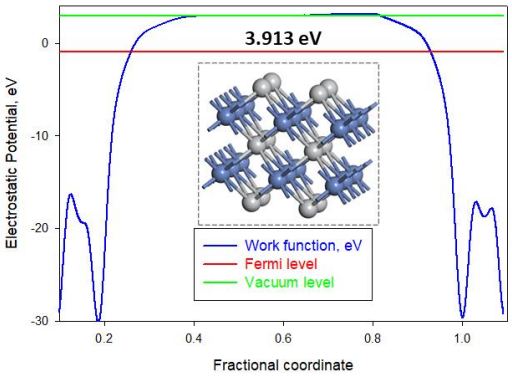


Figure 4(a) Schematic plot of the average profile potentials for Ti (100) surface displaying the work function value.

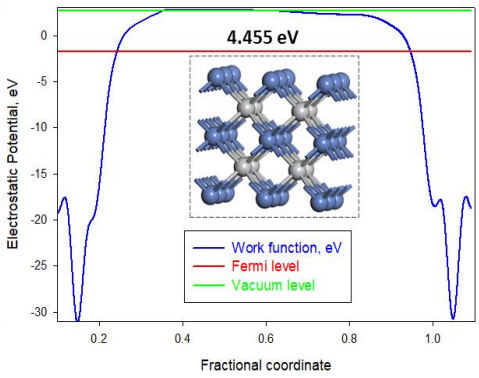


Figure 4(b) Schematic plot of the average profile potentials for Ni (100) surface displaying the work function value.

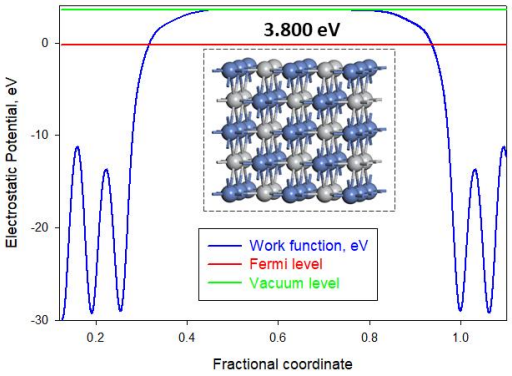


Figure 4(c) Schematic plot of the average profile potentials for NiTi (110) surface displaying the work function value.

4. Conclusion

The electronic and structural properties of NiTi (100) and (110) surfaces were studied using the density functional theory method. The NiTi (110) surface was found to be more thermodynamically stable than the other surfaces, according to surface energy analysis. It was also observed that the surfaces work function differs with surface planes and termination, with NiTi (110) surface metal having the lowest

work function value. Lower work function values typically suggest greater electron availability and higher reactivity. The TDOS plot showed that the material is metallic, as there were states at the Fermi-level.

Acknowledgement
This research was carried out at the University of Limpopo's Materials Modelling Centre (MMC). We acknowledge Centre for High-Performance Computing (CHPC) based in Cape Town for their computing resources.

References

1. Firstov GS, Vitchev RG, Kumar H, Blanpain B, Van Humbeeck J 2002 *Mater.Sci. Eng*, **23**, 4863-4871.
2. Bodner T, Behrendt A, Prax E, Wiesbrock F 2012 *Mater Chem*, **143**, 717.
3. Wang Y 2019 *Journal of Alloys and Compounds*, **6**, 671-678.
4. Kulkova SE, Egorushkin VE, Eremeev SV, Kim JS, Lee G, Koo YM 2004 *Physical rev*, **349**, 342-347.
5. Cheng YL, Wang FH, Shang JX 2016 *Mater sci forum*, **850**, 259-265.
6. Smith J 2021 *Materials Science Review*, **24**, 56-62.
7. Clark SJ, Segall MD, Pickard CJ 2002 *Phys. Cond.Matter*.**14**. 2717–2744.
8. Medveva NI, Enyashin AN, Ivanovskii AL 2011 *Struct. Chem*., **52**, 785-802.
9. Wang J, Wang SQ 2014 *Surf. Sci*., **630**, 216-224.
10. Sun W, Ceder G 2013 *Surface Science*, **617**, 53-59.
11. Kasap S 2005 *Materials and Devices*, **49**, 420-429.
12. Kittel C 1996 *Introduction to Solid State Physics (7th ed.)*, **77**, 563-572.
13. Tshwane DM, Modiba R 2022 *Surface Propert*., **370**, 9005.

Binding nature of fibrin molecules onto Au_{92} and Ag_{92} nanoparticles

MW Makgoba[a], MA Mahladisa[a], MG Matshaba[a], M Gandamipfa[a], N Mhlanga[b], T Mokhena[b], LM Sikhwivhilu[b], TE Mosuang[a]

[a]Department of Physics, University of Limpopo, Private Bag x1106, Sovenga, 0727, South Africa.
[b]Advanced Materials Division/MINTEK, Private Bag X3015, Randburg 2125, Gauteng Province, South Africa.

E-mail: **makgoba.walter94@gmail.com**

Abstract. Density functional theory, Monte Carlo and molecular dynamics simulations were used to study the nature of binding on Au- and Ag-nanoparticles by 1-, 2- and 3-fibrin molecules. In the process, the negative adsorption energies acquired suggest that the fibrin molecules + Au-/Ag-nanoparticles reaction process is exothermic and energetically stable. To corroborate the adsorption energies, the binding distance between the fibrin molecules' functional group atoms and Au-/Ag-nanoparticles surface atoms were computed. To probe the mobility of the atoms in the fibrin-nanoparticle complexes, the mean square displacements graphs were plotted. Such plots suggests that the H, C, N and O atoms from the fibrin molecules' functional groups, may diffuse easily into Au- and Ag-nanoparticles. The Mülliken charges were further analysed through population analysis to trace the distribution of charges from the fibrin molecules onto the Au/Ag-nanoparticles.

1. Introduction

Biophysics based research suggests that introduction of metal nanoparticles into blood plasma may provide sustainable treatment of ailments based on nanotechnology and nanomedicine. A close focus on the ongoing research on gold (Au) and silver (Ag) nanoparticles reveals an aligned interest on their clinical applications in areas of biological sensing, biomedical imaging as well as drug delivery aspects [1, 2]. Other studies reveal that nanomaterials, which include gold (Au) and silver (Au) nanoparticles, serve as excellent candidates for the immobilisation of certain proteins [3, 4]. In such scenarios, proteins are usually covalently bound to the surface of the nanoparticles through accessible functional groups of exposed amino acids, resulting in a nanoparticle-protein complex. Existing data [5, 6] indicates that the acknowledged kinds of nanoparticles for clinical practice is quite low, which implies that more investigations aligned with biomedical applications need to be carried out. This is attributed to the limited understanding and control of the nanoparticles interactions with biological systems [7, 8].

Whenever a nanoparticle enters the physiological fluid environment, a spontaneous interaction between the surface atoms of the nanoparticle and the protein molecules functional group atoms ensues [9]. Eventually, a protein molecule wraps itself around the nanoparticle forming a nanoparticle-protein corona complex. Added interaction of a nanoparticle-protein corona with the surroundings will depend on the properties of the generated corona complex [9]. Such properties relate to the type and number of protein species involved, as well as the type, shape, and size of nanoparticle. In this study, nanoparticles

are referred to as Au- and Ag-nanospheres. Explicitly, 92 atoms Au- and Ag-nanospheres, as well as 1-, 2- and 3-fibrin protein molecules were modelled. A choice of 92 atoms nanosphere falls in the structural magic numbers of generating stable, improved surface and yet reactive nanoparticles [10], whereas a fibrin protein molecule which is a derivative of fibrinogen has been identified as an emerging and promising species suitable for biomedical applications [11]. Thereafter, interactions of Au- and Ag-nanospheres with 1-, 2- and 3-fibrin protein molecules were also sequentially modelled. The objectives of the study is to understand how the fibrin protein molecules bind with Au/Ag-nanoparticles. A combination of computational techniques; specifically, Monte Carlo absorption, density functional theory and molecular dynamics are used to obtain the results reported herein.

2. Methodology

The model consists of Au- and Ag-nanospheres' interaction with 1-, 2- and 3-fibrin protein molecules. 92 atom-containing Au- and Ag-nanospheres including fibrin molecules were modelled using $Dmol^3$ code [12] for all the density functional theory-based geometry optimisations. A simple fibrin molecule has a molecular formula $C_5H_{11}N_3O_2$, classified as a peptide [13], meaning its chains or concentration can be replicated into more chains or specifically 2- and 3-fibrin molecule chains in this instance. The polar nature of fibrin molecules functional group atoms bondings on CH_3, CH_2, CO, NH_2, and NH [13] will encourage further covalency sharing of electrons with the Au/Ag-nanosphere's surface electron density. Regarding Au- and Ag-nanospheres, it is noted that not any whole integer number of atoms can form a well-defined spherical or circular shape, but certain structural magic numbers do conform to this choice [10]. Hence Au-/Ag-nanospheres with 92 atoms obey such choice which generates a convincing and acceptable sphere or circular shape. The structures of the Au- and Ag- nanospheres as well as fibrin molecules were firstly optimised individually using $DMol^3$ code [12] in order to reach their most stable energy geometry structures. Thereafter, the Monte Carlo method [14] was used to model the adsorption of fibrin molecules onto Au/Ag-nanospheres. In the adsorption process, fibrin molecules were considered as adsorbates and Au-/Ag-nanospheres as substrates. Upon adsorption, $DMol^3$ code was applied to optimise the Au-/Ag-nanosphere + fibrin molecule complexes in order to reach their most stable conformations. Lastly, molecular dynamics simulations [15] were performed on Au-/Ag-nanosphere + fibrin molecules complexes to achieve final energy-optimised systems. The NVT ensemble at room temperature (298 K) was used where N represents the number of particles, V is volume of the system at temperature T. A time step of 1.0 fs and a total dynamic time of 100 ps were also adopted. The simulations were carried out at a room temperature (298 K) to mimic the body temperature (310 K), as the two temperatures do not show major significant differences. The effect of time on interatomic interaction movement may not show significant differences as long as the temperature is kept close to the body temperature. But a longer time is recommended to allow a satisfactory interaction of the atoms.

3. Results and discussion

3.1. Adsorption energies and bond lengths of Au- and Ag-nanospheres interactions with fibrin molecules

Adsorption energies and bond lengths involved when gold (Au) and silver (Ag) nanospheres interact with 1, 2, and 3-fibrin protein molecules respectively are discussed sequentially. To further clarify the situation, the following model interactions were considered for the Au-nanospheres and 1-, 2- and 3-fibrin molecules: Au(92) + 1-fibrin, Au(92) + 2-fibrin, Au(92) + 3-fibrin, as shown in Figure 1. Similar model interactions were considered for the Ag-nanospheres and 1, 2 and 3-fibrin molecules: Ag(92) + 1-fibrin, Ag(92) + 2-fibrin, Ag(92) + 3-fibrin, as shown in Figure 2.

In order to predict possible formation of the nanosphere-fibrin corona complexes, adsorption energies of the individual corona complexes were calculated. Adopting the method used by Carr *et al*. [16], the adsorption energy is given by:

$$E_{Ad} = E_{(ns+fm)} - nE_{fm} - E_{ns}. \quad (1)$$

In this case, E_{Ad} is the adsorption energy of the fibrin protein molecules on the nanosphere surface. $E_{(ns+fm)}$ is the total energy of the nanosphere + fibrin molecule (*ns*+*fm*) corona complex. E_{fm} and E_{ns} are total energies of the fibrin molecule and nanosphere respectively, whilst *n* = 1, 2, and 3 respectively for 1-, 2- and 3-fibrin molecules present in the nanosphere binding. Based on this equation, a negative adsorption energy indicates that the fibrin molecule(s)' adsorption on the nanoparticle is exothermic, resulting in an energetically stable nanosphere-fibrin corona system [16-18]. Adsorption energies of 1-, 2- and 3-fibrin molecules onto Au(92) and Ag(92) nanospheres were ultimately calculated and given in Table 1. The negative adsorption energies were attained for all the fibrin molecules binding onto the Au/Ag-nanosphere configurations. As per equation (1), this suggest that all the nanosphere-fibrin corona complexes presented in Table 1 are energetically stable. To quantify this, the energy values further indicate the strength of the nanosphere – fibrin molecules binding. So, the lowest negative adsorption energy is associated with the most stable nanosphere-fibrin corona complex. Stable and exothermic interactions suggest that nanospheres can successfully transport fibrin protein molecules species into cellular environment.

To corroborate the adsorption energies, the bond lengths were also explored for possible Au-H, Au-O, Au-N and Au-C as well as Ag-H, Ag-O, Ag-N and Ag-C atomic bonding as listed in Table 2. The arrangement is for 1-, 2- and 3-fibrin molecules possible bonding permutations with Au(92) and Ag(92) nanospheres. The permutations of the possible bonding assist in recognising which of the bond lengths have the smallest values. Such bond lengths would then be associated with the most probable bonding between the Au/Ag-nanospheres and the fibrin molecules. Excessively large bond lengths would also suggest unlikely occurrence of the bonding. As the number of the fibrin chain molecules increases, some of the bond lengths tend to decrease as indicated in Table 2.

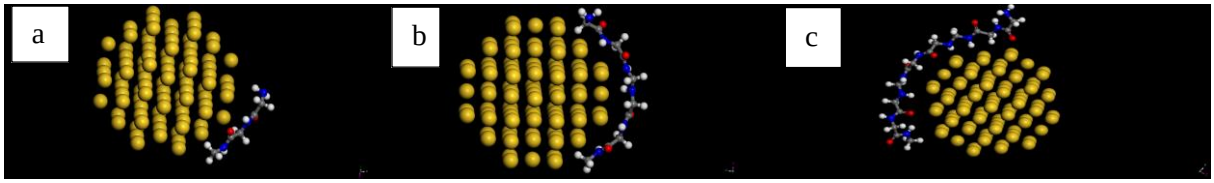


Figure 1. Different fibrin molecule chains adsorption onto Au(92) nanosphere. (a) 1-fribin molecule + Au(92) nanosphere, (b) 2-fibrin molecule + Au(92) nanosphere, (c) 3-fibrin molecule + Au(92) nanosphere. The yellow spheres represent the Au atoms. The grey, blue, red and white atoms represent the C, N, O and H atoms in the 1-, 2- and 3-fibrin molecules.

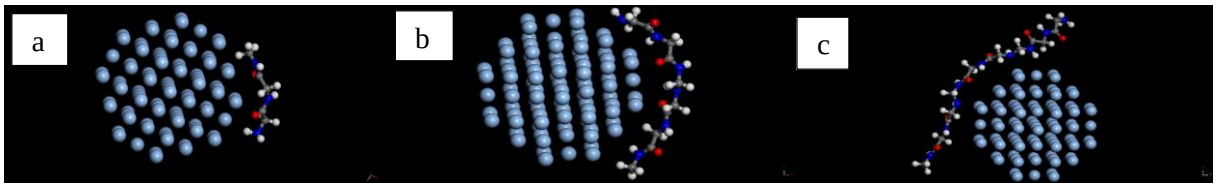


Figure 2. Different fibrin molecule chains adsorption onto Ag(92) nanosphere. (a) 1-fribin molecule + Ag(92) nanosphere, (b) 2-fibrin molecule + Ag(92) nanosphere, (c) 3-fibrin molecule + Ag(92) nanosphere. The purple spheres represent the Ag atoms. The grey, blue, red and white atoms represent the C, N, O and H atoms in the 1-, 2- and 3-fibrin molecules.

Table 1. Adsorption energies (E_{Ad}) of 1-, 2- and 3-fibrin molecules onto Au(92) and Ag(92) nanospheres.

	Au(92) nanosphere +			Ag(92) nanosphere +		
	1-fibrin	2-fibrin	3-fibrin	1-fibrin	2-fibrin	3-fibrin
Adsorption energy (eV)	-0.54	-0.55	-0.54	-0.81	-0.82	-0.81

Table 2. Predicted bond lengths between Au-/Ag-atoms and specified *n*-fibrin atoms in Angstroms (Å). All possible permutations of Au(92) or Ag(92) and 1-, 2- and 3-fibrin molecules have been considered.

		Au(92) nanosphere +				Ag(92) nanosphere +		
		1-fibrin	2-fibrin	3-fibrin		1-fibrin	2-fibrin	3-fibrin
Bond lengths (Å)	Au-O	3.291	3.045	2.548	Ag-O	2.878	2.454	3.974
	Au-C	3.724	3.058	3.240	Ag-C	3.232	3.368	3.260
	Au-H	2.694	2.431	2.828	Ag-H	2.745	2.805	2.929
	Au-N	3.497	3.060	2.900	Ag-N	3.178	2.494	2.461

3.2. Mean square displacements and diffusion coefficients

To probe the possible mobility of the Au, Ag, O, H, N and C atoms, the mean square displacement (MSD) graphs were plotted for a 1-fibrin molecule adsorbed onto Au(92), and Ag(92) nanospheres, as elaborated in Figure 3. The mobility is investigated in order to check the possible diffusion of the H, C, N and O atoms into the Au/Ag-nanospheres matrices. Based on the results from the study by Gao and Qu [19], on the calculated values of the diffusion coefficients, an ion with a high value of the diffusion coefficient diffuses better. In this work, the MSD graphs in Figure 3 suggests that all the H, C, N and O atoms will diffuse easily into the Au/Ag-nanospheres. The diffusion coefficients constants (*D*) were also calculated from the slopes of the MSD graphs and are presented in Table 3. To quantify the results; looking at the Ag-nanosphere, all the H, C, N and O atoms will diffuse readily into the nanosphere. In the case of Au-nanosphere, a delayed response to diffusion could be expected. The diffusion coefficients in Table 3 further suggests that all the hydrogen (H) atoms will diffuse easily into their respective nanospheres as compared to other fibrin molecules atoms i.e., C, N, O. All this explains the mechanism of fibrin protein species binding into the Au- and Ag-nanospheres.

*3.3. Carbon (C) atom Mulliken charges (in the Units of **e**) for 1-, 2- and 3-fibrin molecules before and after adsorption onto Au- and Ag-nanospheres*

Beside energetics and diffusivities, electronic structure of the bonding modes can also assist in understanding the interface interactions. The electronic structure on the bonding sites can describe the charge transfer which occurs between Au/Ag-nanospheres and 1-, 2- and 3-fibrins. Mülliken charges were also computed with reference to 1-, 2- and 3-fibrin molecules before and after adsorption onto Au- and Ag-nanospheres. In this process, the atomic charge values can be attained from Mülliken population analysis.

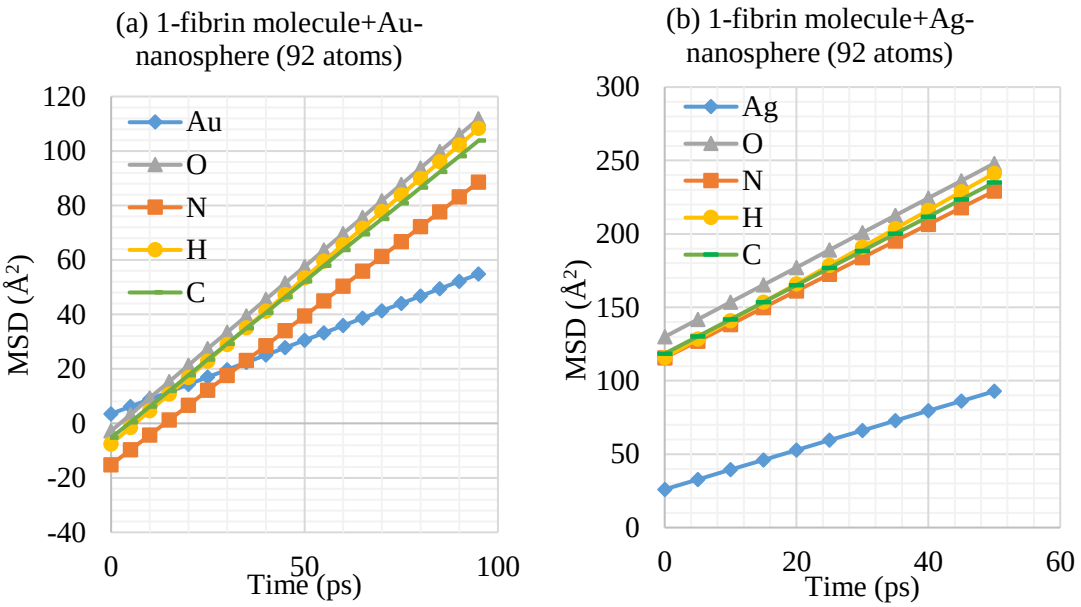


Figure 3. MSD on 1-fibrin molecule adsorbed onto (a) Au(92) and (b) Ag(92) nanospheres for Au, Ag, O, H, N and C atoms.

Table 3: Diffusion coefficients constants (*D*) for 1-fibrin molecule adsorbed onto Au(92) and Ag(92) nanospheres.

		D of 1-fibrin +	
		Au(92) (Å²/ps)	Ag(92) (Å²/ps)
Atoms	H	0.2033	0.4180
	C	0.1914	0.3890
	N	0.1821	0.3782
	O	0.2012	0.3937
	Au/Ag	0.09001	0.2227

Mülliken charge populations are useful when characterising the nature of atomic charge distributions in molecules through bonding of atomic orbitals in a neighborhood of atomic clusters [20]. In this instance, the Mülliken charges were computed for carbon atoms in fibrin molecules before and after adsorption as the carbon atoms outline the spine of these fibrin proteins. Mülliken charge distribution of each carbon atom with reference to Au- and Ag-nanospheres on a particular *n*-fibrin molecule are listed in Table 4. As observed, the negative charges for carbon atoms with reference to Au-nanosphere before adsorption becomes more negative after adsorption due to their attachment or association with acceptor atoms such as H in the CH_2 or CH_3 group. Similarly, as can be observed, some of the positive charges before adsorption becomes less positive after adsorption with a few becoming more positive due to their attachment or association with donor atoms such as O and N in the C=O and NH_2 or NH_3 group, respectively.

Regarding Ag-nanosphere, the same behaviour of charges was observed, where the negative charges for carbon atoms before adsorption also becomes more negative after adsorption due to their attachment or association with acceptor atoms such as H in the CH_2 or CH_3 group. Likewise, some of the positive charges before adsorption also becomes less positive after adsorption with some becoming more positive

due to their attachment or association with donor atoms such as O and N in the C=O and NH_2 or NH_3 groups.

During the adsorption process, the negative charge of either Au- or Ag-nanospheres is transferred by the fibrin bonding atom and charge conservation to the first carbon (C1) of the fibrin molecule enhancing or increasing the negativity value of the C1 atom. Considering the Coulomb's law of charges, this leads to second (C2) and third (C3) carbon atoms of the fibrin also acquiring extra positive and negative charges sequentially. The same behaviour of charge transfer through the carbon and sulphur atoms was observed in the findings of Carr *et al*. [16]. So, it can be concluded that the electrons are gained through the Au-H/Ag-H bond lengths, where H is associated with the CH_2 or CH_3 group as accumulated in Table 2. Conversely, some electrons get lost through the Au-O/Au-N and Ag-O/Ag-N bondings, where O and N atoms belong to C=O and NH_2 or NH_3 functional groups respectively.

Table 4. Mülliken charges for carbon atoms (charge units are e) for 1-, 2- and 3-fibrins before and after adsorption onto Au(92) and Ag(92) nanospheres.

		Au(92) nanosphere +						Ag(92) nanosphere +					
		1-fibrin		2-fibrin		3-fibrin		1-fibrin		2-fibrin		3-fibrin	
		before	after	before	after	before	after	before	after	before	after	before	After
Carbon atoms	C1	-0.173	-0.230	-0.174	-0.224	-0.169	-0.239	-0.173	-0.218	-0.174	-0.236	-0.169	-0.231
	C2	0.401	0.402	0.428	0.426	0.426	0.385	0.401	0.412	0.428	0.427	0.426	0.420
	C3	-0.153	-0.208	-0.160	-0.222	-0.170	-0.208	-0.153	-0.243	-0.160	-0.214	-0.170	-0.208

4. Conclusion

The binding nature of fibrin protein molecules onto Au- and Ag-nanospheres was successfully investigated through density functional theory, Monte Carlo adsorption study, and molecular dynamics simulations. The negative adsorption energies of -0.54 (Au(92) + 1-fibrin and Au(92) + 3-fibrin), -0.55 (Au(92) + 2-fibrin), -0.81 (Ag(92) + 1-fibrin and Ag(92) + 3-fibrin) and -0.82 eV (Ag(92) + 2-fibrin) were attained implying energetically stable modelled systems. Moreover, 2-fibrins became the most stable adsorptions onto both Au- and Ag-nanospheres. When the fibrin molecule chains are extended, some bond lengths between Au- and Ag-nanospheres and the fibrin molecules bonding functional group atoms (i.e H, C, N and O) are compromised, others enhanced. The MSD suggest that the functional group H, C, N and O atoms may diffuse easily into the Au- and Ag-nanospheres. In addition, the functional group atoms have high probability to diffuse into the 92 atom Ag-nanosphere than the Au-nanosphere. The enhanced negative and positive charges experienced through the Mülliken charge distributions before and after adsorption is attributed to the charge transfers which originate from the Au- and Ag-nanospheres onto the fibrin molecules. Such charge transfers may alter the original molecules functionality or be the source of immobilisation. But the overall results of this study suggest Au and Ag-nanospheres as potential transporters of fibrin protein species into cellular media.

References

[1] Elahi N and Kamali M, *a review Talanta*, **184** (2018), pp. 537-556.
[2] Gurunathan S, Park J H, Han J W and Kim J H, *Int. J. Nanomed*. **10** (2015), pp. 4203-4222.
[3] Pudlarz A.M, Ranoszek-Soliwoda K, Czechowska E, Tomaszewska E, Celichowski G, Grobelny J, Szemraj J, *Appl. Biochem. Biotechnol*., **187** (2019), pp. 1551-1568.
[4] Czechowska E, Ranoszek-Soliwoda K, Tomaszewska E, Pudlarz A, Celichowski G, Gralak-Zwolenik D, Szemraj J, Grobelny J, *Colloids Surf. B Biointerfaces*, **171** (2018), pp. 707-714.

[5] Farjadian F, Ghasemi A, Gohari O, Roointan A, Karimi M, MR Hamblin, *Nanomedicine (Lond.)* **14** (2019), pp. 93-126
[6] Xiaoting S, Xiang G, Jie L, Jingyuan W, Yaping L, Zhiwen Z, *Acta Pharma. Sin. B*, **12** (2022), pp. 3028-3048.
[7] Vilanova O, Mittag J.J, Kelly P.M, Milani S, Dawson K.A, Radler J.O, Franzese G, *ACS Nano* **12** (2016), pp. 10842-10850.
[8] Zanganeh S, Spitler R, Erfanzadeh M, Alkilany A.M, Mahmoudi M, *Int. J. Biochem. Cell Biol*. **75** (2016), pp. 143-147.
[9] Xu J.X and Fitzkee N.C, *Front. Physiol*. **12** (2021), 715419.
[10] Charles P. Poole Jr, Frank J. Owens, Introduction to Nanotechnology, john wiley & sons (2003).
[11] Dominik H, Anne B, Fabian K, Marcel H, Alejandro Or, Adrian K, Guido G, Klaus H, and Oliver I. S, *Biomacromolecules*, 22 (2021), pp. 4084−4094.
[12] Delley B, *J. Chem. Phys*. **113** (2000), pp. 7756-7764.
[13] Pubchem.ncbi.nlm.nih.gov/compound/Fibrin
[14] Murti BT. Putri A.D, Kanchi S, Sabela M.I, Bisetty K, Inamuddin, Asiri A.M., *J. Photochem. & Photobiol, B*: Biology **188** (2018), pp. 159-176.
[15] Frenkel D, Smit B, Understanding Molecular Simulation, 2nd ed., Academic, San Diego, (2002).
[16] Carr J.A, Wang H, Abraham A, Gullion T, Lewis J.P, *J. Phys. Chem. C* **116** (2012), pp. 25816-25823.
[17] Tsuda A and Venkata N.K, *Nano Impact*, **2 (**2016**),** pp 38-44.
[18] Zeng W, Liu T, Wang Z, Tsukimoto S, Saito M, Ikuhara Y, *Mater. Trans*., **51** (2010), pp. 171- 175.
[19] Gao F and Qu J, *Mater. Lett.* **73** (2012), pp. 92-94.
[20] Nikolova V, Cheshmedzhieva D., Ilieva S., Galabov B., *J. Org. Chem*. **84** (2019) pp.1908-1915.

Computational studies of pressure dependence of Monazite $LnPO_4$ (Ln = Pr, Nd)

L Motsomone*[1], R G Diale[2], P E Ngoepe[1], H. Bissett[3] and H R Chauke[1]

[1] Materials Modelling Centre, University of Limpopo, Private Bag X 1106, Sovenga, 0727, South Africa
[2] Advanced Materials Division, MINTEK, Private bag X 3015, Randburg, 2125, South Africa
[3] The South African Nuclear Energy Corporation SOC Ltd. (Necsa), P.O. Box 582, Pretoria, 0001, South Africa

E-mail: **lmotsomone@gmail.com**

Abstract. Monazite is an orthophosphate mineral with the formula A[PO_4] [where A = light rare earth elements, thorium, and uranium]. Studying the structural state of monazite is of vital importance to earth and materials sciences. A first-principle calculation based on density functional theory was used to study the structural and mechanical properties of $NdPO_4$ and $PrPO_4$ structures under compression. The calculated structural parameters of monazite systems decrease with an increase in pressure. Furthermore, bulk, shear, and Young's moduli increase with an increase in pressure. Pugh and Poisson ratios show that monazite systems are ductile, and their ductility can be improved with increasing pressure. The insights gained from high-pressure studies are of interest for the applications of monazite.

1. Introduction

The mineral monazite is composed of radioactive phosphates and is found to contain light rare earth metals such as cerium, lanthanum, neodymium, praseodymium, and yttrium, as well as thorium and uranium [1]. This mineral is generated through diverse geological processes and under varying environmental conditions. The creation of monazite configurations takes place through diverse pressure and temperature circumstances within sedimentary, metamorphic, and igneous rock formations [2]. Monazite has truly captured the scientific community due to its distinctive characteristics and potential applications, especially in modern technology and industry. This includes fields like medicine, aerospace, generation of renewable energy, catalysis and optical devices [3]. Additionally, it can serve as geochronometers for geological dating and aid in the identification of the distribution of rare earth elements [4]. Furthermore, monazite is considered as a promising option for nuclear waste disposal because of their capacity to confine radioactive elements.

The primary objective of the study is in developing better methods for extracting monazite constituents (REEs, Th and U) from the ore. Unfortunately, using convectional processing as a solution has proven to be both time-consuming and expensive [1]. Additionally, there are concerns about the safe disposal of the radioactive waste generated during monazite processing. Despite these challenges, the demand for rare earth elements and thorium continues to drive interest in monazite mining and processing. Studying the process, requires one to also understand how monazite molecules behave under extreme conditions (such as high pressure and temperature). Hence, the current study aims to analyse the behaviour of monazite type $NdPO_4$ and $PrPO_4$ when subjected to high pressure (0-60 GPa) which is crucial for their potential use in various technological applications. Density functional theory (DFT) calculations were employed to investigate the structural and mechanical properties of monazite systems

at different pressures. The results provide insights into the pressure-induced changes in the physical properties of $NdPO_4$ and $PrPO_4$, which can be useful for designing new materials with improved properties.

2. Methodology

In this work, we use ab initio calculations to investigate how changing the pressure affects the crystal structure and mechanical properties of the monazite types $NdPO_4$ and $PrPO_4$. The calculations are carried out with the help of density functional theory (DFT) and the Vienna ab initio simulation package (VASP), an electronic structure calculation code based on plane waves [5]. In order to achieve accurate results, the study conducted a series of test calculations to investigate the impact of varying k-point samplings and plane-wave cut-off energy. The electronic properties were ultimately calculated using a 12x13x11 k-point sampling and a cut-off energy of 600 eV, with a system size of 24 atoms for each structure. The geometry optimization is accomplished by employing the generalized gradient approximation (GGA) [6] of Perdew-Burke-Ernzerhof (PBE) [7]. Atomic forces are reduced to less than 0.01 eV/ Å by thoroughly optimizing the position of atoms in the simulation cell with the lattice parameters, and the associated angles using the conjugate gradient approach. To keep any resulting computation mistakes within acceptable bounds, the precision was set to ″accurate’’. Elastic characteristics were determined using a strain of 0.006 for the lattice deformation.

3. Results and Discussion

3.1. Structural properties

The structure of monazite systems has been optimized and is presented in Figure 1. The composition of monazite crystals is $LnPO_4$ (Ln = Nd, Pr), which exhibits monoclinic symmetry with a $P2_1/n$ space group and contains four formula units. The axes in question have different lengths and orientations. Specifically, two of the axes are perpendicular to each other, while the third axis has an angle of 103° and is not perpendicular to the other two. Furthermore, the structure described results in a unit cell that is asymmetrical in the shape of a rectangular prism.

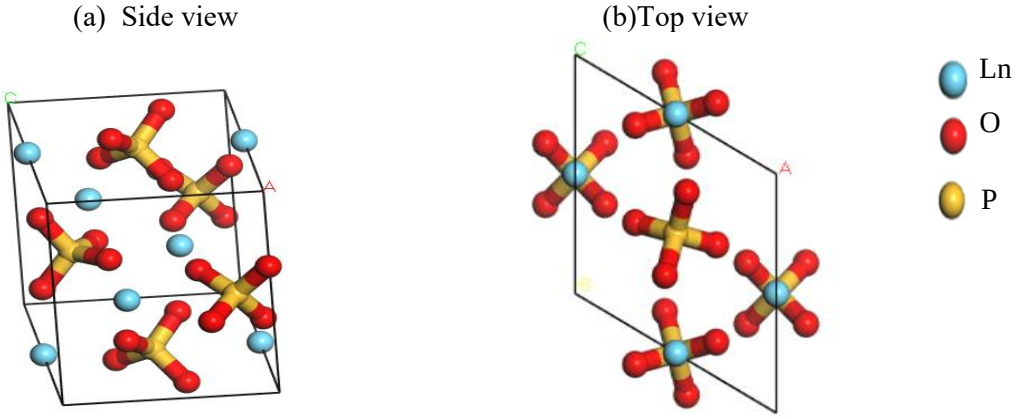


Figure 1. Atomic arrangement of monazite type $NdPO_4$ and $PrPO_4$ structures, (a) represents the side view, whereas (b) represents top view. The blue, mustard, and red atoms represent Ln (Ln= Nd or Pr), P, and O elements, respectively.

Table 1 lists the calculated and experimental lattice parameters for the optimised structures of various monazite systems. The calculated equilibrium lattice parameters were in good agreement with available experimental values (to within 5 %) at zero pressure.

Table 1. The lattice parameters for monazite type $LnPO_4$ (Ln = Nd , Pr) structures.

Structure		Lattice parameters (Å)		
		a	b	c
$NdPO_4$	This study	6.735	6.950	6.405
	Exp. [8]	6.744	6.956	6.408
$PrPO_4$	This study	6.760	6.981	6.416
	Exp. [8]	6.741	6.961	6.416

3.2. Relative lattice parameters and volume under pressure

The variation of the relative lattice parameters a/a_o, b/b_o, and c/c_o, as well as the β (beta) angle, with respect to pressure for monazite structures of $NdPO_4$ and $PrPO_4$ are shown in Figure 2. In addition, a_o, b_o, and c_o represent the equilibrium lattice parameters at zero pressure. The results from Figure 2 (a and b) indicates that when pressure is increased, the normalized or relative lattice parameters decreases. However, the normalized lattice parameter c/c_o exhibits a more rapid decrease compared to the other lattice parameters in both structures. This suggests that the c-axis exhibits greater compressibility in comparison to the a and b-axes. Furthermore, these findings suggest that the atomic bonding in the a and b directions with the nearest neighbours is comparatively stronger than in the c direction. The figure also shows that the β angle decrease with an increase with pressure.

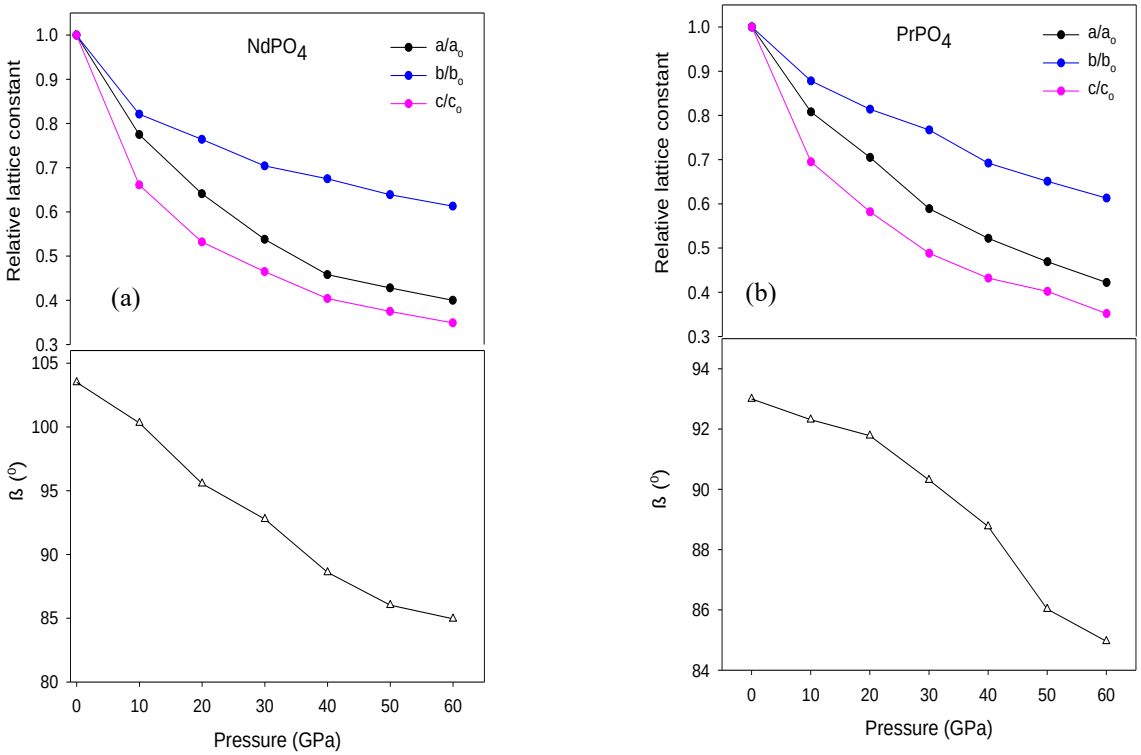


Figure 2. Relative lattice parameters and β of monazite type (a) $NdPO_4$ and (b) $PrPO_4$ at high pressure.

Figure 3 shows the relative volume against pressure ranging from 0 to 60 GPa. Volume is one of the components that are important to determine the physical properties of a material [9]. The results revealed that the relative unit cell volume decreases with an increase in pressure. However, it is noted that the rate of reduction in $PrPO_4$ is higher than that of $NdPO_4$. So, this means that the resistance of $NdPO_4$ is higher than of $PrPO_4$ against applied pressure.

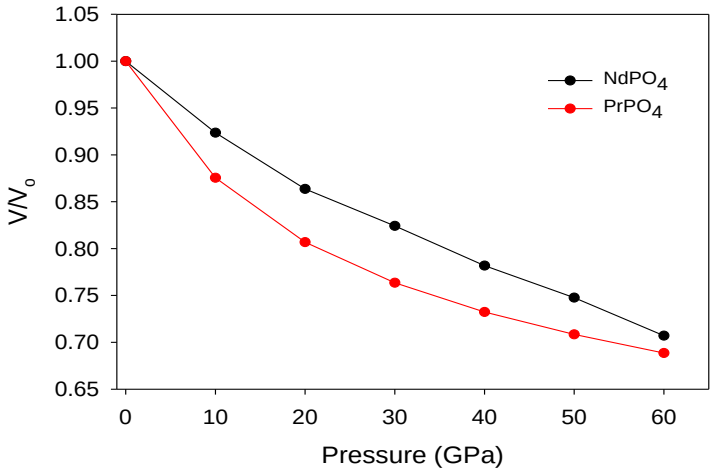


Figure 3. The variation of relative unit cell volume of monazite-$NdPO_4$ and $PrPO_4$ type with pressure.

3.3. Elastic moduli

Bulk (B), shear (G), and Young 's modulus (E) of monazite systems were determined and shown in Figure 4. The bulk modulus is a measure of resistance to volume change under pressure, it is used to measure the compressibility of the material. Shear modulus describes the resistance to shape change caused by shearing force, in this case it is used to check the material rigidity. Lastly, Young's modulus describes the relative stiffness of a material.

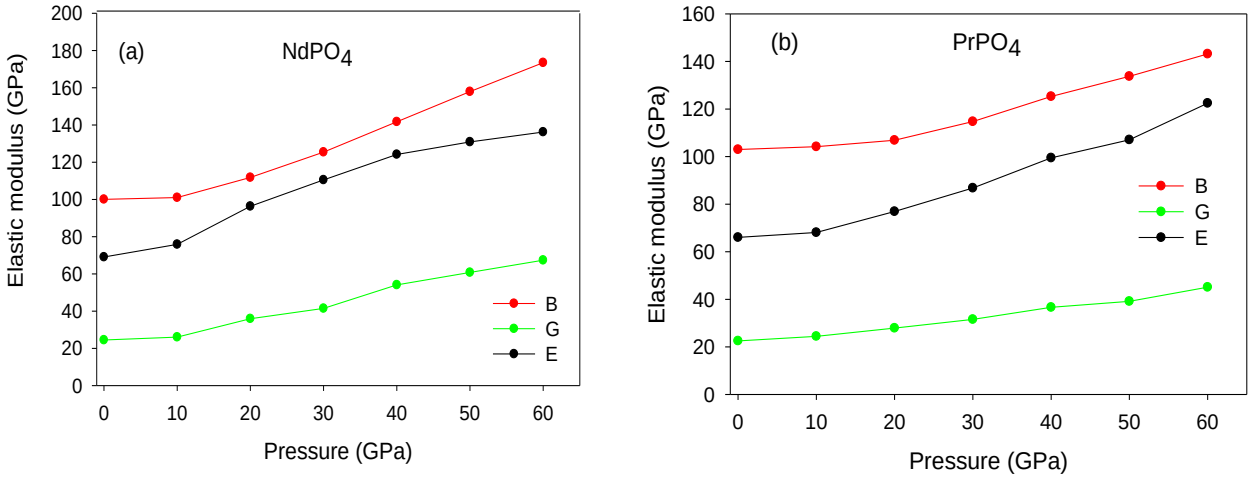


Figure 4. Calculated Bulk (B), Shear (G) and Young's (E) modulus of (a) $NdPO_4$ and (b) $PrPO_4$ at high pressure.

The results show that when the pressure is increased, the bulk modulus for $NdPO_4$ and $PrPO_4$ also increases. The discovery leads one to believe that the capacity of monazite structures to withstand uniform compression, rises. Moreover, we were able to draw a conclusion about the average atomic bond strength based on the behaviour of the bulk, and in this particular instance, the atomic bond strength of $NdPO_4$ and $PrPO_4$ rises with pressure. In addition to this, the shear modulus rises with pressure for all systems regardless of the type. This suggests that it is very difficult to generate the shear deformation when the pressure is increased. It is widely recognized that a higher shear modulus indicates stronger directionally oriented interatomic bonding. As a result, a rise in pressure causes the bonding within

monazite systems to become more directionally focused. Lastly, the Young 's modulus was also determined. When the pressure is raised, the Young 's modulus is shown to behave in a manner that is comparable, implying that it increases as pressure increases. This suggest that the monazite structures become more difficult to stretch at high pressures. The Pugh (B/G) and Poisson (σ) ratios were determined and presented in Figure 5. This was to analyse the ductility properties of monazite systems at high pressures. Pugh proposed that there is a connection between the value of B/G and the material's malleability and brittleness. Specifically, a high value of B/G is proposed to indicate malleability, while a low value is suggested to indicate brittleness. Furthermore, the value of 1.75 [10] is a critical point that distinguishes between ductile and brittle materials.

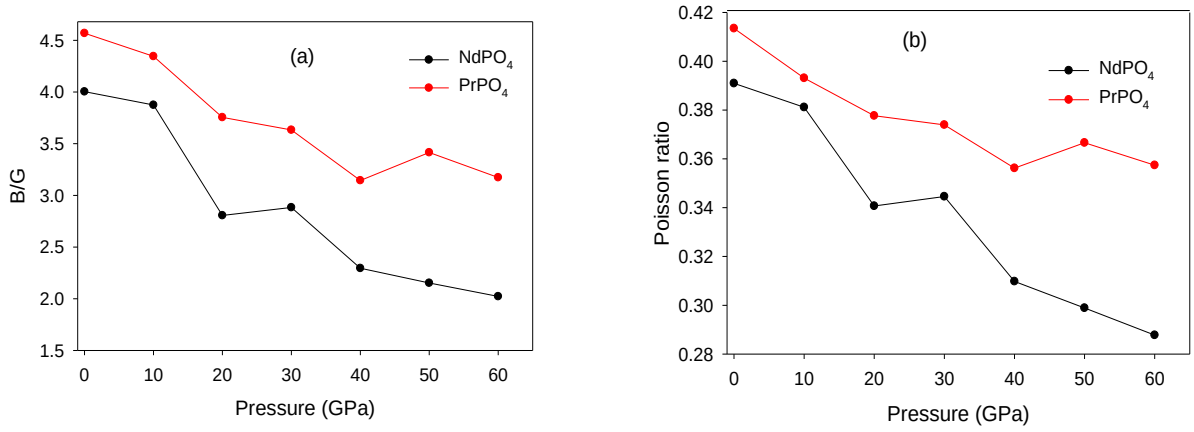


Figure 5. Calculated (a) Pugh (B/G) and (b) Poisson ratios of $NdPO_4$ and $PrPO_4$ structures at high pressure.

The results indicate that both systems are ductile as B/G is greater than 1.75. The σ also shows the decreasing behaviour in response to an increase in pressure. To achieve ductility, its value must be more than 0.26 [11], and the data obtained indicates that the structures demonstrate ductility as their values are greater than 0.26 for both systems. However, the results indicate that there is an inverse relationship between pressure and the observed values of ratios, where an increase in pressure results in a decrease in the B/G and σ values for both systems. The observation implies that the ductility of the monazite systems decreases with an increase in pressure.

4. Conclusion

The study investigated the structural, mechanical, and vibrational characteristics of $NdPO_4$ and $PrPO_4$ structures at high pressure using first-principles calculations. The research discovered that increasing pressure caused a decline in lattice parameters. The study also discovered that the structures of $NdPO_4$ and $PrPO_4$ display significant resistance to deformation with increasing pressure, as seen by higher B, G, and E values at high pressure levels. It was also found that the structures are ductile. The pressure-induced changes in the structural mechanical and vibrational properties of $NdPO_4$ and $PrPO_4$ have practical implications for the creation of new materials with improved properties for magnets and glass applications.

5. Acknowledgement

The work was carried at the Materials Modelling Centre, University of Limpopo. The authors would like to acknowledge the Centre of High-Performance Computing (CHPC) and the financial support of the Department of Science and Innovation (DSI), through the Nuclear Materials Development Network (at Necsa) of the Advanced Materials Initiative (AMI). The support of South African Research Chair initiative of Department of Science and Innovations is highly appreciated.

Reference

[1] Kemp D and Cilliers A C 2016 *J. South. Afr. Inst. Min. Metall.* **116** 901-6
[2] Hoshino M, Sanematsu K and Watanabe Y 2016 *Phys. Chem.* **49** 129-291
[3] Xaba S M, Nete M and Purcell W 2018 *Mater. Sci. Eng.* **430** 012006
[4] Oelkers E H and Poitrasson F 2002 *Chem. Geol.* **191** 73-87
[5] Kresse G and Hafner J 1993 *Phy.Rev. B* **47** 558
[6] Perdew J P and Yue W 1991 *Phy. Rev. B* **24** 13298
[7] Perdew J P, Burke K K and Ernzerhof M 1996 *Phy. Rev. Lett.* **77** 3865
[8] Mullica D F, Grossie D A and Boatner L A 1985 *J. Solid state Chem.* **58** 71-77
[9] Yaws C L 1977 *Gulf Professional Publishing.* **1** 1-26
[10] Pugh S F 1954 *J. Sci.* **367** 823-43
[11] Greaves G N, Greer A L, Lakes R S and Rouxel T 2011 *Nat. Mater.* **10** 823-37

Computer simulation of Silver (Ag) and Nickel (Ni) Nanomaterials

MG Matshaba[1], NS Makhathini[1], MA Mahladisa[1] and TE Mosuang[1].
[1]Department of Physics, University of Limpopo, Private Bag x 1106, Sovenga, 0727, South Africa.

E-mail: **malili.matshaba@ul.ac.za**

Abstract. The study of silver (Ag) and nickel (Ni) nanoparticles is of paramount importance in research and technological advancement. Energetics and structural changes in Ag and Ni nanoparticles were studied using classical molecular dynamics approach. The formulation chosen make use of the NVT ensemble under hoover thermostat. Sutton-Chen potentials have been preferred to explain the silver-silver as well as nickel-nickel interactions in the respective nanoparticles systems. Total energy against temperature have been varied in order to observe possible phase transformations. Radial distribution functions further provide information on the nanoparticles at different temperatures. Results obtained agree well with related studies.

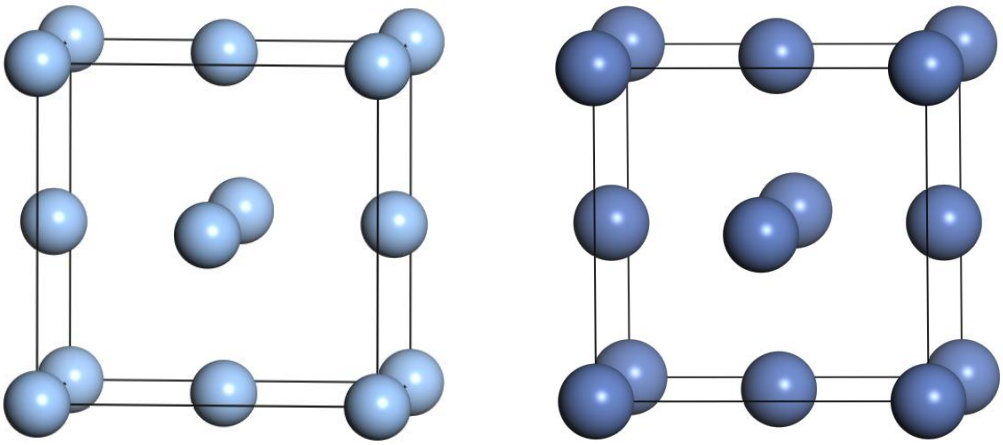
1. Introduction

Silver (Ag) and nickel (Ni) are well known metals which have face centred cubic (fcc) crystal structure [1, 2, 3]. Their distinctive properties which include electrical, optical and thermal give them a wide range of applications [4]. For instance, Ag and Ni high melting points of 1235 and 1728 K, respectively make them ideal for high temperature usages [5, 6]. In the health industry, Ag has been utilised exceptionally as an anti-bacterial agent [7]. At nanoscale, silver (AgNPs) and nickel nanoparticles (NiNPs), also have a wide range of applications due to their resistance to corrosion. Nanoscale research interest on Ag and Ni is growing widely because of their size, shape, and the ability for their chemical properties to be easily altered. Most of their applications are for magnetic sensor, biomolecular separation, high electrical conductivity, automotive catalytic converters, catalyst, energy storage and medical diagnostic [8, 9, 10 and 11]. AgNPs production is of a high scale though it has a low capacity [12].

In certain instances, conducting experiments can be expensive and time-consuming, yet not being able to acquire the expected results. Such anomalies can be caused by experimental processes not having the capability to capture the actual change of atomic states with times. Incorporating computational methods such as molecular dynamics (MD) into such materials' based studies can improve the expected outcomes. MD is an effective instrument in modelling since it provides freedom to comprehend the mobility of molecules and the structures at various thermodynamic conditions. With the help of MD simulations, sources of Ag^{1+} ions toxicity associated with skin argyria, as well as kidney and liver functioning in humans maybe traced [13]. In this paper, MD simulations are employed to study energetics and structural evolution of Ag and Ni nanoparticles with changing temperature. We hope the acquired results will assist in curbing unnecessary anxiety whenever metal nanoparticles – human cells interactions occur.

2. Methodology

Appropriate crystal structures were generated making use of lattice parameters, space groups and atomic positions of Ag and Ni respectively. Such crystal structures are presented in figure 1(a) and (b). Both have space group of Fm3m 225 with lattice parameters Ag: a = 4.090 Å and Ni: a = 3.524 Å.



(a) (b)

Figure 1. Typical face centred cubic Ag (a) and Ni (b) conventional unit cells.

Molecular dynamics DL_POLY code was employed to perform all the classical calculations [14, 15]. Calculations were carried out using the many body Sutton-Chen potential in the NVT ensemble. The many body Sutton-Chen potentials is defined by the equation:

$$U = \epsilon \left[\sum_{ij} \left(\frac{a}{r_{ij}} \right)^{n} - C \sum_{i} \sqrt{\sum_{j} \left(\frac{a}{r_{ij}} \right)^{m}} \right]. \quad (1)$$

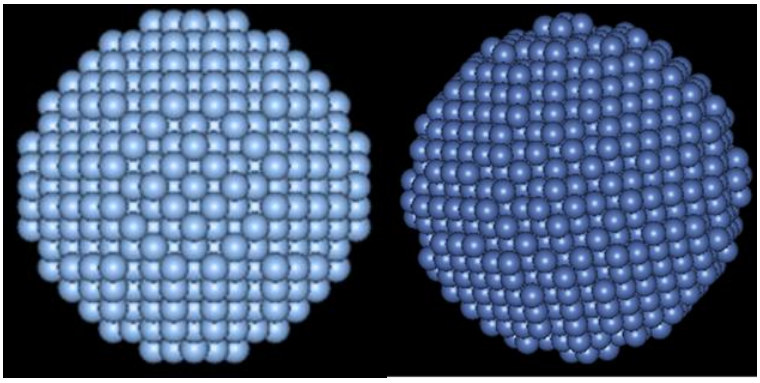
In this equation (1), ϵ is the parameter with dimensions of energy, *a* the lattice parameter and *C* is the positive dimensionless parameter. Also, r_{ij} is the separation between atoms *i* and *j*. *n* and *m* are positive integers, where *n* > *m*. All parameters utilised for Ag and Ni nanoparticles are listed in table 1.

Table 1. Parameters used in the Sutton-Chen potentials [16].

	ϵ (eV)	*a* (Å)	*n* and *m*	*C*
Ag	0.025	4.090	12 and 6	144.4
Ni	0.157	3.520	9 and 6	39.43

3. Results and discussion

Structures of the nanoparticles were generated from conventional unit cells (depicted in figure 1) using METADISE code [17]. Generated silver and nickel nanoparticles are depicted in figure 2, where AgNP has 1205 and NiNP has 1553 atoms.



(a) (b)

Figure 2. Nanoparticles of (a) silver with 1205 atoms and (b) nickel with 1553 atoms.

3.1. Relationship between energy and temperature

In order to observe the total energy of the Ag and Ni nanoparticles, their temperature was varied from 0 to 1000 K. Obtained data from such comparisons is depicted in figures 3 (AgNP) and 4 (NiNP). Both graphs indicate a linear correlated increase between total energy and temperature. Solid state physics reveals; an increasing temperature of the nanoparticles can be associated with the change in entropy which may lead to a possible phase change [18]. Thus as the temperature escalates, atoms of the nanoparticles acquire enhanced vibrational frequencies, which lead to enhanced atomic collisions. However, calculations objectives was not to determine the exact transition temperatures. AgNP energy – temperature variations does not provide possible prediction of the possible transition temperature. However, the NiNP variations suggest a possible transition temperature around ±900 K.

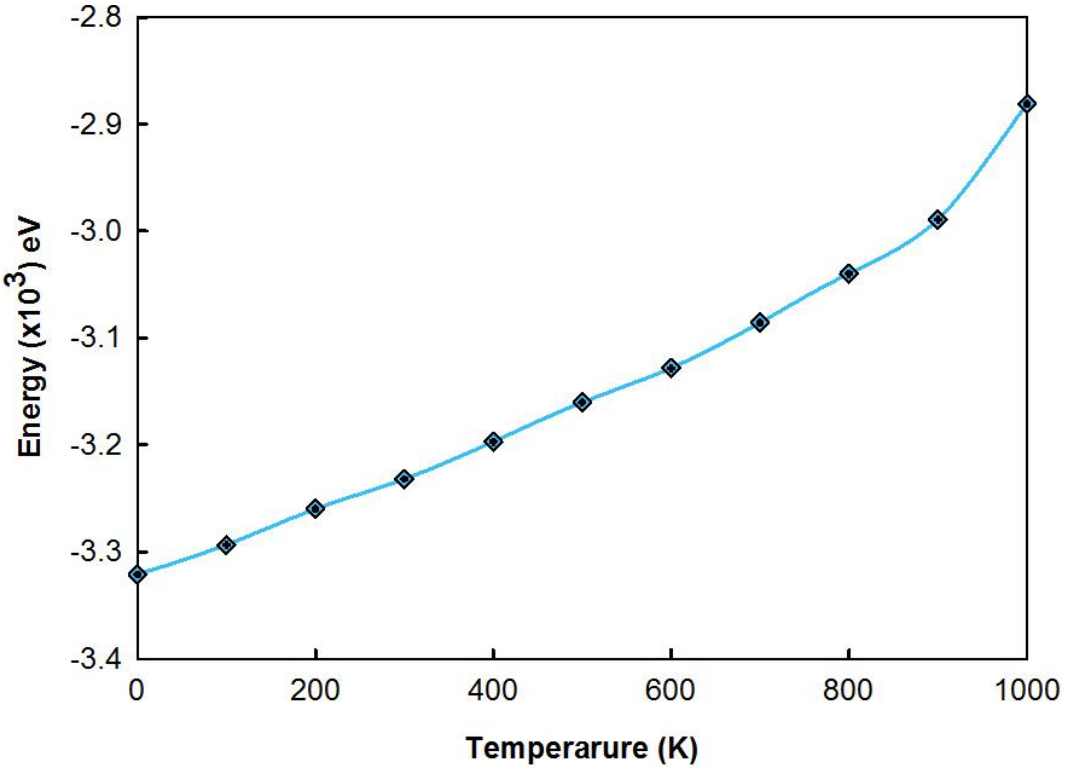


Figure 3. The graph of total energy as a function of temperature for silver nanoparticles.

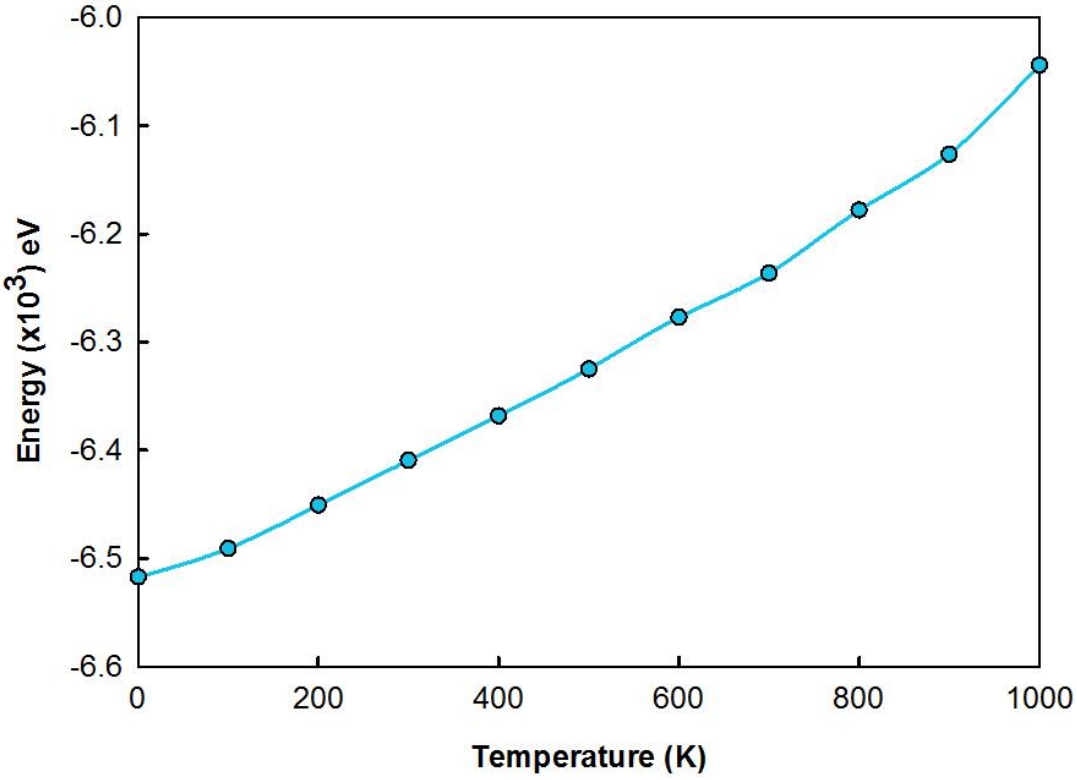


Figure 4. The graphs of total energy as a function of temperature for nickel nanoparticles.

3.2. Nanoclusters at different temperatures

The structural configurations of the nanoparticles of Ag and Ni at different temperatures are displayed in figures 5 and 6. The structures maintain original atomic arrangements up to 900 K, indicating that the nanoparticles have not reached the possible transition temperature. An upward bend in the curve at ±900 K, is noticeable in AgNP and NiNP curves in figures 3 and 4 respectively, which suggest a possible phase transformation. As the temperature increases, the nanostructures of both AgNP and NiNP adopt a highly disordered arrangement. In view of AgNP, transformation to the amorphous phase can be experienced at lower temperature of 600 K as compared to Ni NP which experiences such transformation around 500 K. Visualisation of figures 5 and 6 suggest that the initial transition temperature of AgNP is lower than that of NiNP. Findings by Gould et al. [18], Guo et al. [19] and Kryshtal et al. [20], further confirm that AgNP transform to amorphous phase earlier than NiNP. Both nanoparticles figures 5 and 6 demonstrate the change in shape and spacing with temperatures. Literature studies [21], further report that amorphous metal nanoparticles can regain their original form through recrystallisation. Studies by Goud et al. [18], further account on the AgNPs' transformation to an amorphous phase at lower temperatures, which may also create some voids within the structure.

3.3 Radial distribution functions for Ag and Ni NP's

Radial distribution functions (RDFs) of both Ag and Ni nanoparticles are depicted in figure 7. At 300 K, plots display distinct periodic peaks at higher values of g(r) than those at higher temperatures (i.e. 600, 800, and 1000 K). Such distinct and prominent peaks in the RDFs alludes to crystallinity of the material, while lesser peaks alludes to amorphous arrangement. Both the RDFs of Ag and Ni nanoparticles at 1000 K show little to no peaks indicating that the nanoparticles are now transforming the amorphous phase.

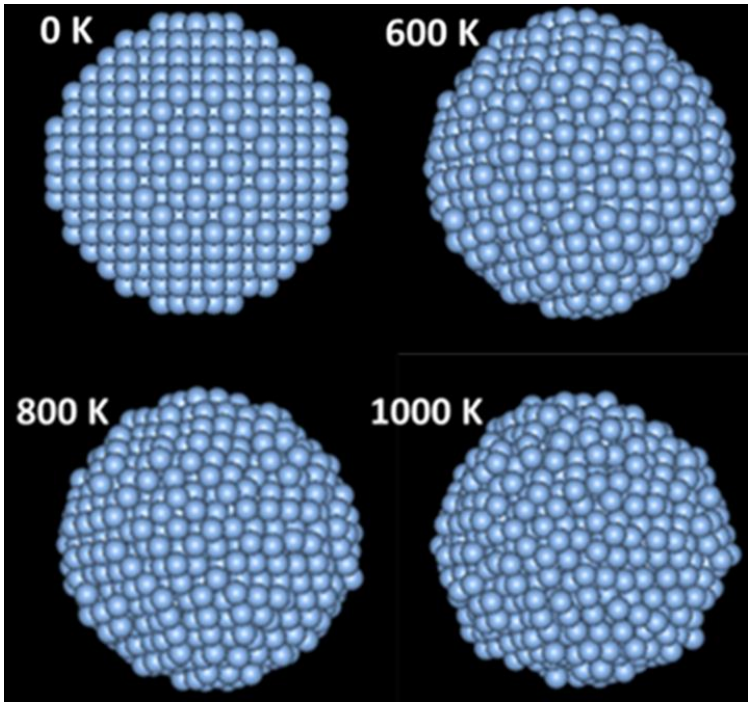


Figure 5. Visualisation of 1205 atoms of AgNP at different temperatures.

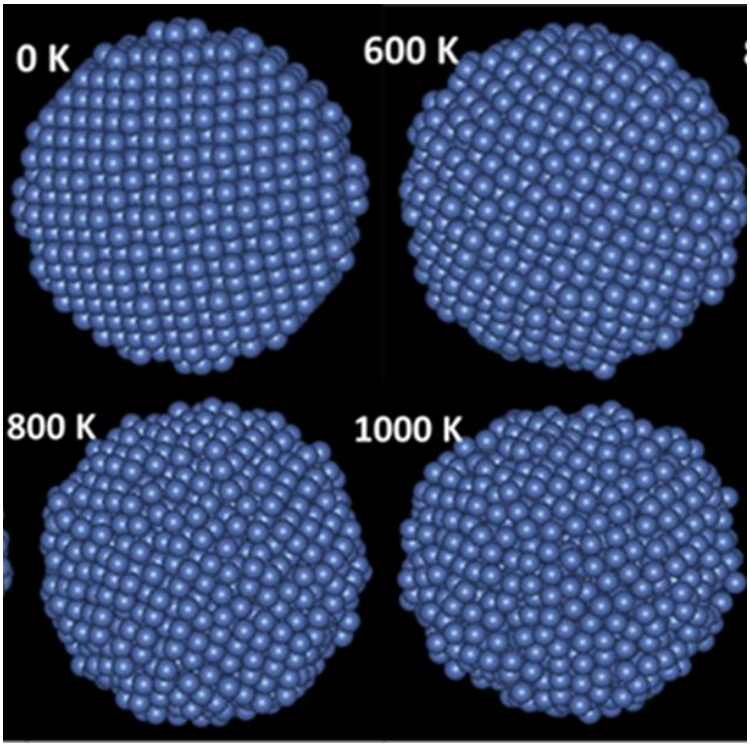


Figure 6. Visualisation of 1553 atoms of NiNP at different temperatures.

Additional information on RDFs and how Ag and Ni atoms are distributed in the nanoparticles appears in tables 2 and 3. Table 2 itemise statistical information on 1st, 2nd, and 3rd nearest neighbour distances extracted figure 7 at 300 K. Likewise, table 3 further list such statistical RDF information, now at 1000 K. It must be noted that at 1000 K, only the 1st nearest neighbour data can be extracted as extra other peaks are no longer distinct. Our results are comparable with Cui et al. [20] from the literature.

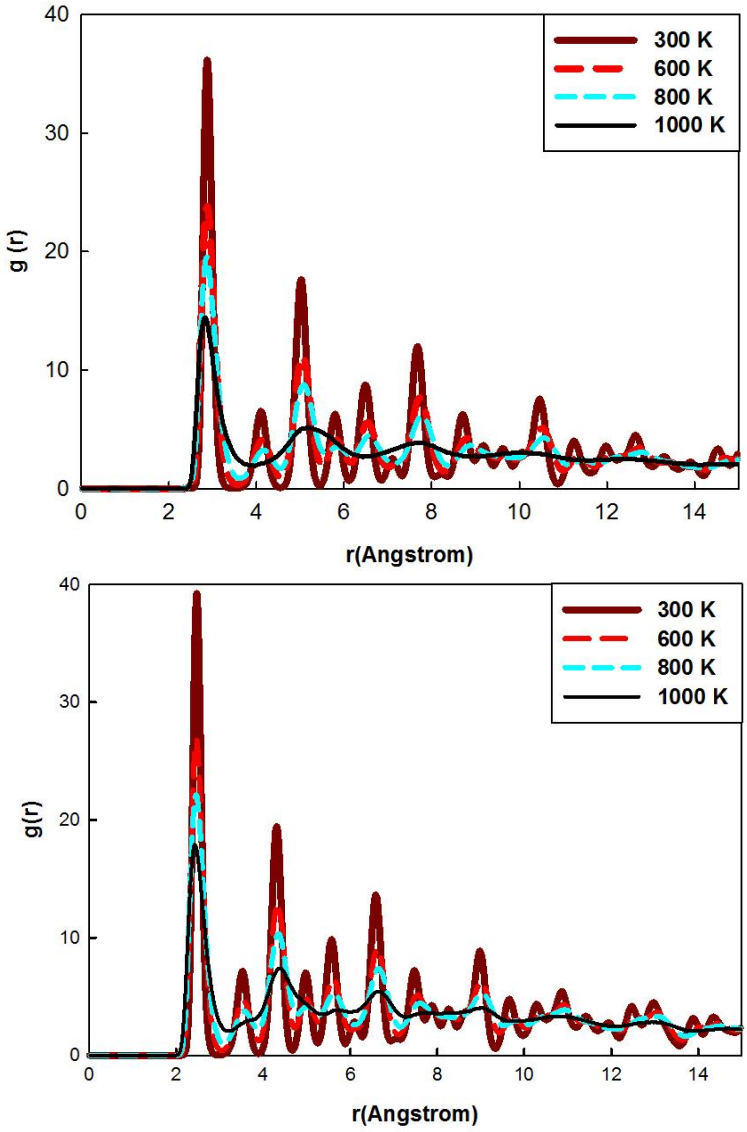


Figure 7. Radial distribution functions (RDFs) of Ag (top) and Ni NPs (bottom).

Table 2. First, second, and third nearest neighbour atom distribution in AgNP and NiNP at 300 K. r_1, r_2, and r_3 are positions of the 1st, 2nd, and 3rd nearest neighbour atoms. n_1, n_2, and n_3 are the numbers of the 1st, 2nd, and 3rd nearest neighbour atoms.

	r_1	n_1	r_2	n_2	r_3	n_3
AgNP	3.0	14	4.0	3	5.0	9
NiNP	2.5	12	3.6	2	4.3	8

Table 3. First nearest neighbour atoms distribution in AgNP and NiNP at 1000 K. r_1 is position of the 1st nearest neighbour atoms. n_1 is the number of the 1st nearest neighbour atoms.

	r_1	n_1
AgNP	3.0	9
NiNP	2.5	9

4. Conclusion

AgNPs and NiNPs were successfully generated using a METADISE code. Direct proportionality is observed on total energy against temperature up to ±900 K in both Ag and Ni nanoparticles. This signifies how the associated increase in entropy change may ultimately leads to phase transformation. RDFs results suggest that both AgNPs and NiNPs maintain only the first nearest neighbour ordered structures at a higher temperature of 1000 K. This also indicates that at nano-level, the structures may be utilised for some high temperature applications. However, under normal conditions (300 K), the well-defined crystallinity can be traced using the first, second, and third nearest neighbour atomic order distribution.

Acknowledgement

We would like to acknowledge centre for high computing (CHPC) for allowing us to utilise their systems and the University of Limpopo for giving us opportunity to do our research in their institution.

References

[1] Hu F., Luyang H.W., He R.L., Guan Z.J., Yuan S.F. and Wang Q.M. Face-Centered Cubic Silver Nanoclusters Consolidated with Tetradentate Formamidinate Ligands. J. Am. Chem. Soc. 2022, 144 (42), 19365 - 19371.
[2] Karpuz A., Kockar H., Alper M., Karaagac O. and M. Haciismailoglu M. Electrodeposited Ni–Co films from electrolytes with different Co contents Appl. Surf. Sci. 2012, 258, 4005 – 4010.
[3] Wang Z.K., Kuok M.H., Ng S.C., Lockwood D.J., Cottam M.G., Nielsch K., Wehrspohn R.B. and Gösele U. Spin-Wave Quantization in Ferromagnetic Nickel Nanowires. Phys. Rev. Lett. 2002, 89, 027201, 1 - 3.
[4] Fabara A., Cuesta S., Pilaquinga F. and Meneses L. Computational modeling of the interaction of Silver nanoparticles with the lipid layer of the skin. J. Nanotechnol. 2018, 4927017.
[5] Higashino M., Murai S. and Tanaka K. Improving the Plasmonic Response of Silver Nanoparticle Arrays via Atomic Layer Deposition Coating and Annealing above the Melting Point. J. Phys. Chem. C 2020, 124, 27687 – 27693.
[6] Teijlingen A., Davis S.A. and Hall S.R. Size-dependent melting point depression of nickel nanoparticles. Nanoscale Adv. 2020, 2, 2347 – 2351.
[7] Kholoud M.M, Abou El-Nour, Ala'a Eftaiha, Abdulrhman Al-Warthan and Reda A.A. Ammar. "Synthesis and applications of silver nanoparticles Arabian Journal of Chemistry". 2010, 3, 135 - 140.
[8] Lee K.-B., Park S. and Mirkin C.A. Multicomponent Magnetic Nanorods for Biomolecular Separations. Angew. Chem. Int. Ed. 2004, 43, 3048 - 3050.
[9] Li W.R., Xie X.B., Shi Q.S., Zeng H.Y., Ou-Yang Y.S. and Chen Y.B. Antibacterial activity and mechanism of silver nanoparticles on Escherichia coli. Appl. Microbiol. Biotechnol. 2010, 8, 1115 - 1122.
[10] Dolez P.I. Chapter 1.1—Nanomaterials Definitions, Classifications, and Applications. In Nanoengineering; Dolez, P.I., Ed.; Elsevier: Amsterdam, The Netherlands, 2015, 3 – 40.
[11] Jasni A.H., Ali A.A., Sagadevan S. and Wahid Z. Silver nanoparticles in various new applications. Silver Micro-Nanoparticles. 2021, 10, 5772, 1 – 19.
[12] Temizel-Sekeryan S. and Hicks A.L. Global environmental impacts of silver nanoparticle production methods supported by life cycle assessment. Resour. Conserv. Recycl. 2020, 156, 104676 – 104686.
[13] Lewinski N., Colvin V. and Drezek R. Cytotoxicity of nanoparticles, Small. 2008, 4(1), 26 – 49.
[14] Todorov I. T., Smith W. and Trachenko K. DL POLY 3: New Dimensions in Molecular Dynamics Simulations via Massive Parallelism. J. Mater. Chem. 2006, 16, 1911 – 1918.

[15] Smith W., Forester T. R. and Todorov I. T. The DL POLY User Manual (STFC Daresbury Laboratory, Version 2.18. 2007.
[16] Sutton A.P. and Chen J. Long-range Finnis-Sinclair potentials. Philos. Mag. Lett. 1990, 139 – 146.
[17] Watson G., Kelsey E.T., de Leeuw N.H., Harris D.J. and Parker S.C. Atomistic simulation of dislocations, surfaces and interfaces in mgo. J. Chem. Soc., Faraday Trans. 1996, 92, 433 – 438.
[18] Gould A.L., Kadkhodazadeh S., Wagner J.B., Catlow C.R., Logsdail A.J. and Vece M.D. Understanding the Thermal Stability of Silver Nanoparticles Embedded in a-Si. J. Phys. Chem. C. 2015, 119, 41, 23767 – 23773.
[19] Guo Y.Q., Yu J.K., Yang H.B., Qiao Q., Li Q.Y. and Zhao X.C. Improved thermal stability and oxidation resistance of electrodeposited NiCrP amorphous alloy coatings. Bull. Mater. Sci. 2020, 43, 98, 1 – 6.
[20] Gui J., Yang L. and Wang Y. Molecular dynamics simulation study of the melting of silver nanoparticles, Integr. Ferroelectr. 2013, 145 (1), 1 – 9.
[21] Sayle D. C. and Watson G. W. Strutural exploration of thin-film oxide interfaces via "simulated amorphisation and recrystallisation. Surf. Sci. 2001, 437, 97 - 107.

Structural investigation of Pd/Zr/Pd/Ti/Pd multi-layered stack systems implanted with 150 keV argon ions for hydrogen storage applications

CT Thethwayo[1*], CB Mtshali[2], ZM Khumalo[2], CL Ndlangamandla[1], PN Biyela[1], NF Vilakazi[1]

[1]Department of Physics, University of Zululand, Private Bag X1001, KwaDlangezwa 3886, South Africa
[2]iThemba LABS, National Research Foundation, PO Box 722, Somerset West 7129, South Africa

E-mail: **ThethwayoC@unizulu.ac.za**

Abstract. The thermal stability of a multi-layered stack system of Pd/Zr/Pd/Ti/Pd implanted with 150 keV argon ions was studied using X-ray diffraction (XRD) and Rutherford backscattering spectroscopy (RBS). XRD reveals that there was a diffusion of atoms between the layers at room temperature (RT), resulting in the formation of $TiPd_3$ and $Pd_{1.12}Zr_{1.88}$ alloys before ion implantation and annealing under a hydrogen environment. Also, this intermixing was confirmed by RBS. In the samples implanted with 150 keV argon ions at the fluencies of 1×10^{12} $\mathrm{ions.cm^{-2}}$, 1×10^{14} $\mathrm{ions.cm^{-2}}$, and 1×10^{16} $\mathrm{ions.cm^{-2}}$ respectively, the decrease in O concentration was observed, but no noticeable alteration in the crystal structures. Both as-prepared and 150 keV argon ion implanted samples showed oxidation between the two layers, resulting from the residual gases. The secondary reaction of the layers was observed under the H environment, including the diffusion of Pd toward the surface forming a new layer

1. Introduction

Thermal structural stability is one of the critical aspects of hydrogen storing materials because hydrogenation is currently performed at higher temperatures [1, 2]. This includes thermal stability, resistance to corrosion, irradiation, time, and other environmental factors that a material may be exposed to during practical applications [1]. The properties of any material are determined by the arrangement of the atoms in the material's crystal structure, the type of atoms, and the electron configuration. Hence optimizing structural stability and material properties to improve materials' performance is important in the hydrogen economy. One of the methods used to improve material properties is ion implantation whereby foreign atoms are injected into a target material. Ion implantation alters the material's physical, chemical, and electrical properties by distorting the lattice crystal structure [2]. For example, in electronic applications, ion implantation can be used to introduce foreign species into the bulk sample, changing the band gap of a semiconductor.

This research work presents a multi-layered system of titanium (Ti), palladium (Pd), and zirconium (Zr) films deposited on a pure commercial titanium substrate (CP-Ti) to form a stack of Pd/Zr/Pd/Ti/Pd/CP-Ti. These materials were selected based on their properties, performance, and catalytic effect on hydrogen absorption as reported in the literature [2-5]. Since some layers of the multi-layered stacks belong to reactive transition elements, which are Ti and Zr, atomic interlayer mixing normally takes place at relatively low temperatures, and solid solutions or amorphous alloys are formed in the vicinity of the interfaces [4, 6]. Khumalo *et al*. [1] studied the interfacial reaction and phase

formation in Pd/ZrO/Pd/TiO/Pd multi-layered stack on silicon substrates, the results showed that the stack was stable up to ~370 °C. At temperatures above 480 °C, the reaction between Pd and silicon substrate forms a new layer of Pd_2Si.

In the current study, the multi-layered stack system was implanted with 150 keV argon ions at fluences of 1×10^{12} ions. cm^{-2}, 1×10^{14} ions. $cm^{-2}$and 1×10^{16} ions. cm^{-2}, respectively, to study the effect of ion implantation fluences. Furthermore, the effect of hydrogenation on structural thermal stability was investigated. The structural transformation of layers at ambient temperatures and high temperatures of 550 °C was investigated using RBS and XRD. Several researchers investigated the impact of temperature on hydrogenation and found that at 550 °C, there was a significant volumetric amount of hydrogen absorbed [7, 8].

2. Experimental method

The multi-layered stack system was prepared following the previously published experimental work, using e-beam evaporator [1, 2, 4]. The annealing process was performed using a DC magnetron sputtering chamber, which was evacuated to a pressure of 2.8×10^{-3} Torr. The temperature of the chamber was raised from RT up to 550 °C, in the pure hydrogen environment. The hydrogen flow rate was kept at 20 Standard cubic centimetres per minute (SCCM). The annealing was performed for 1 hour at 550 °C and then cooled to RT in a hydrogen environment.

Both RBS and XRD give a detailed analysis of phase transformation and interlayer mixing. The D8 Advance Bruker XRD diffractometer, equipped with copper X-ray source (Cu-Kα radiation (λ ≈ 1.54 Å)), was used and the results were analysed with X'pert highscore plus search and match program to identify phases, with their corresponding crystal structure type. RBS was employed using a near oxygen resonance energy of 3.00 MeV He^{++} ions, emanating from a 3 MV Tandetron accelerator at iThemba LABS, and the spectra were simulated using SIMNRA software, version 7.03.

3. Results and discussion

3.1. Rutherford backscattering spectroscopy.

SRIM2013 software was used to simulate the penetration range of the 150 keV argon ion implantation in the Pd/ZrO/Pd/TiO/Pd multi-layered stack system and was found to be 88.5 nm. Despite the presence of the Pd as a protective layer, a high concentration of oxygen (O) was observed between the Ti and Zr layers, which was attributed to the oxidation during deposition, see Figure 1. The ratio of Ti and O, as well as Zr and O, was found to be 1:1 and 1:3 respectively for as-prepared, Figure 1 (a), and 1:1 and 1:2 for argon ion implanted, see Figure 1(b-d). The oxidation of Ti and Zr during deposition was also reported by Khumalo *et al.* [1] and credited to the residual gases that were present in the deposition chamber. Since multi-layered stack systems were deposited using the same e-beam evaporator, which had a maximum pressure of roughly $\sim 10^{-6}$Torr, oxidation was anticipated.

Besides the decrease in the concentration of O in 150 keV argon ions implanted samples, there was no change in the crystal structure. Based on the observed RBS results, this could be evidence that 150 keV argon ion implantation can assist in reducing oxygen in highly reactive metals like Zr and Ti, which reduces hydrogen absorption. The interdiffusion between the layers was observed at RT in both as-prepared and ion implanted samples, forming a solid solution.

Figure 2 shows RBS results for both as-prepared and 150 keV argon ion implanted samples after annealing in an H environment. The annealing chamber size (~20 liters) and pressure ($\sim 3 \times 10^{-3}$ Torr) should be noted as they significantly impact how the gas is distributed, resulting in different gas concentrations in the chamber. In theory, the layers of the multi-layered stack system should be preserved under a hydrogen environment. The concentration of the residual gases present in the chamber depends on the system pressure, high concentration of residual gases will cause a secondary reaction between H and residual gases, which will prevent the absorption of H. Furthermore, it should be noted

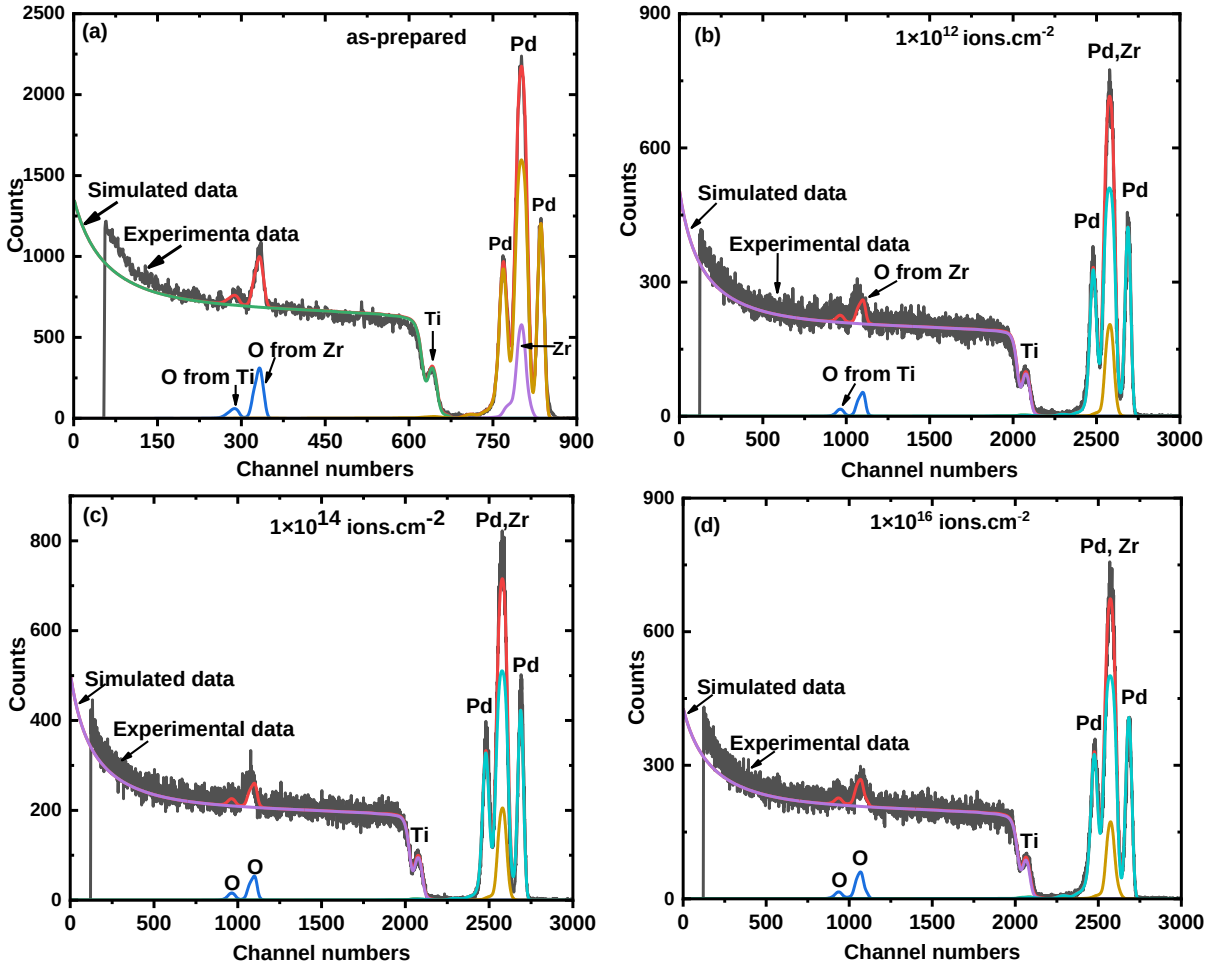


Figure 1. Rutherford backscattering (RBS) results for as-prepared (a) and 150 keV argon ion implanted samples (b-d). (b) sample implanted with the fluence of 1×10^{12} ions. cm^{-2}, (c) implanted with 1×10^{14} ions. cm^{-2}fluence and (d) the sample implanted with the fluence of 1×10^{16} ions. cm^{-2}. The experimental data is labeled in the diagrams, the solid lines indicate the simulated data, and the elements present in each sample.

that the thermal stability in this context depends greatly on the hydrogen flow rate and the concentration each sample received during the hydrogenation process.

RBS analysis demonstrates the evidence of reaction in all the samples as-prepared and 150 keV argon ion implanted, see Figure 2 (a-d). The diffusion of Pd (the Pd layer that was in contact with the substrate) toward the substrate forming a new layer of Ti and Pd, with a ratio of 1:3, was observed in all samples (as-prepared and implanted). Interdiffusion between the layers was observed, resulting in irreversible compounds (solid solutions or amorphous alloys). Zr was found to diffuse toward the Pd layers on the surface and the Pd layers in the bulk. The higher Ti concentration was found to diffuse toward the surface layers, which is thought to be the result of substrate diffusion toward the surface. Lastly, the interdiffusion between the Pd layers was observed; the surface layer diffused toward the Zr layer, and the interlayers toward the bulk and the surface.

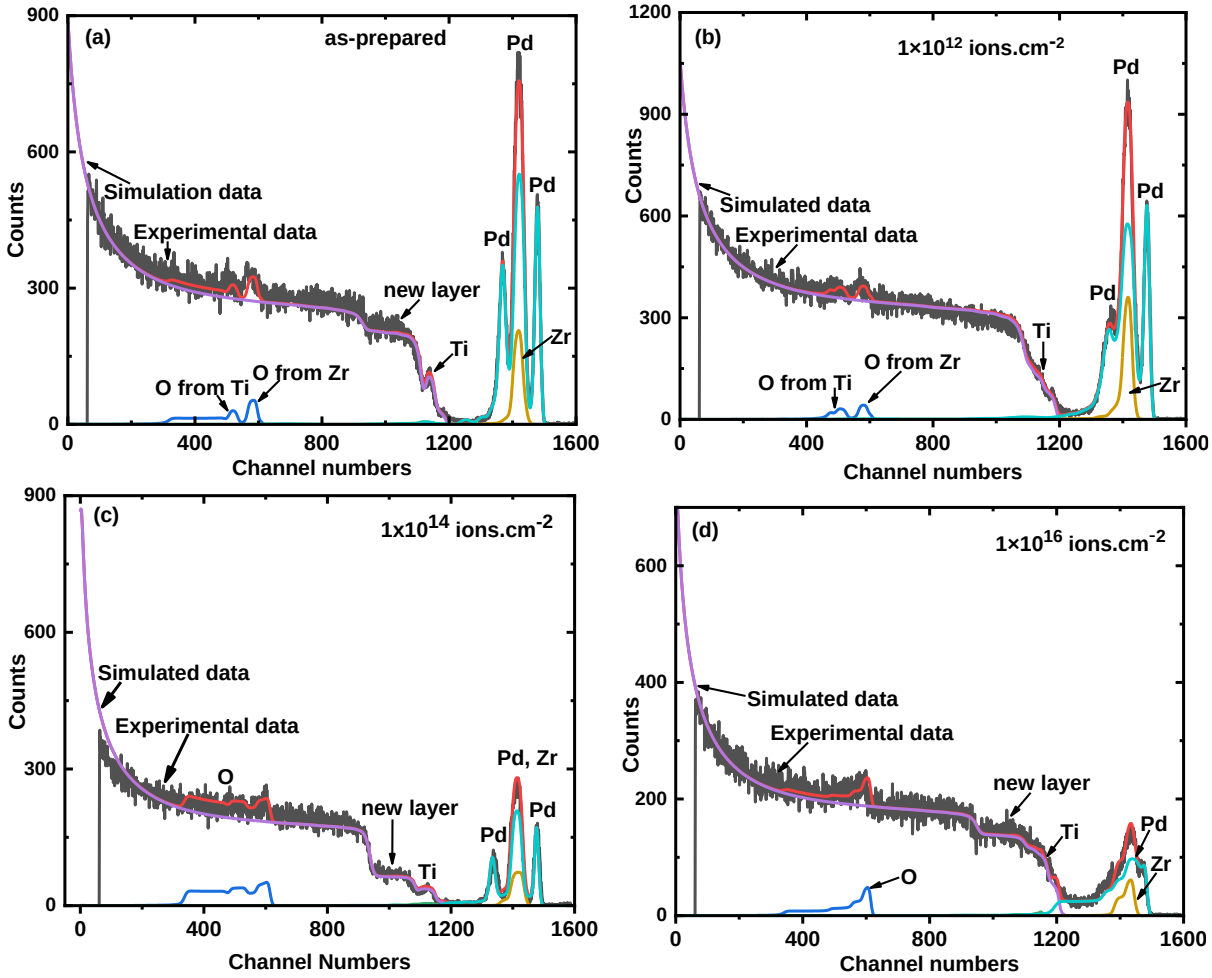


Figure 2. Rutherford backscattering results after annealing in hydrogen environment at 550 °C at a pressure of 3×10^{-3} Torr, for the sample used as-prepared and 150 keV argon ion implanted samples (a-d). (a) as prepared sample, (b) sample implanted with a fluence of 1×10^{12} ions. cm^{-2}, (c) implanted with 1×10^{14} ions. cm^{-2} fluence, and (d) implanted with the fluence of 1×10^{16} ions. cm^{-2}.

3.2. X-ray diffraction

To better understand the crystal structures and phase transformation, XRD analysis was conducted on both as-prepared and ion implanted samples, before and after annealing. There were no noticeable effects of 150 keV argon ion implantation in the multi-layer stack system, (see Figure 3 (a)), which was found to be in agreement with RBS results. XRD results, before hydrogenation, support the RBS results that showed the intermixing between the layers at ambient temperature. The reaction between Ti, Zr, and Pd was observed, showing a hexagonal crystal structure (ref. no: 00-035-1361) of $TiPd_3$ and $Zr_{1.12}Pd_{1.88}$ which corresponded to a tetragonal crystal structure (03-065-7877). There were two phases of Ti observed, α-Ti and β-Ti, corresponding to the hexagonal crystal structure (ref. no: 01-089-3073) and cubic structure (ref. no: 01-089-3726), respectively. Also, the Zr showed two phases belonging to α-Zr with the hexagonal crystal structure (ref. no: 00-005-0665), and β-Zr, which is categorized as the cubic crystal structure (ref. no: 01-089-4916). Lastly, the peaks corresponding to the cubic crystal structure of Pd were observed (ref. no: 01-089-4897).

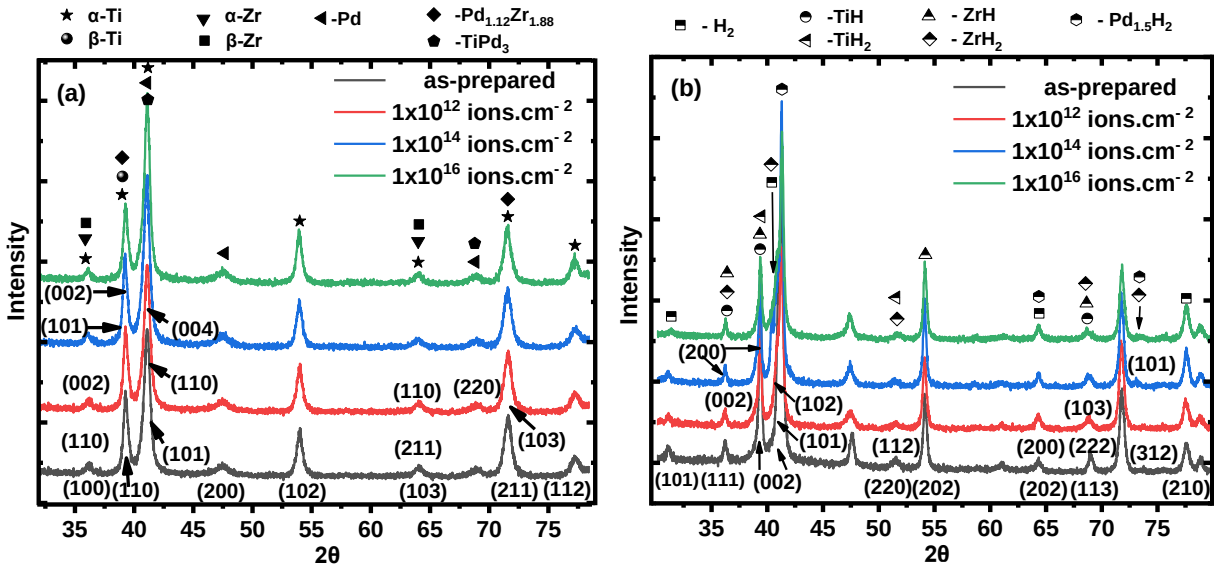


Figure 3. X-ray diffraction (XRD) analysis of Pd/Zr/Pd/Ti/Pd multi-layered stack systems (a) for the sample used as-prepared and 150 keV argon ion implanted samples at different fluences before annealing (b) after annealing in a hydrogen environment at 550 °C, for the sample used as-prepared and 150 KeV ion implanted with fluences from $1 \times 10^{12} - 1 \times 10^{16} \text{ions.cm}^{-2}$. Starting from the bottom line is as-prepared, the second line is for the sample with $1 \times 10^{12} \text{ions.cm}^{-2}$ fluence, the third is 1×10^{14} ions.cm^{-2}, and the fourth is for the sample with 1×10^{16} ions.cm^{-2} fluence, respectively.

On annealed samples, it was clear that all the crystal structures seen before annealing were also retained after annealing, with new crystal structures due to the formation of metal hydrides. The presence of hydrogen molecules (H_2), which may be on the surface or trapped on the crakes from 150 keV argon ion implantation, as well as various metal hydrides (TiH, TiH_2, ZrH, ZrH_2, and $Pd_{1.5}H_2$, were also observed, (see Figure 3 (b)). The tetragonal crystal structure was found to be dominant since it corresponded to TiH, TiH_2, ZrH, ZrH_2, and $Pd_{1.5}H_2$ (ref. no:00-065-6949, 03-065-6949, 00-020-1465, and 01-073-0004, respectively). The hexagonal crystal structure was found to correspond to H_2 (ref. no: 01-085-1243). The results analysis for both RBS and XRD demonstrated annealing under the H environment does not prevent the interdiffusion of atoms and secondary reactions, hence the thermal stability of the multi-layered stack system was compromised.

4. Conclusion

Crystal thermal stability was successfully investigated using RBS and XRD characterisation techniques. As a result of the layer-to-layer reaction, which produced solid solutions, the results showed that annealing under H environment did not improve thermal stability. The diffusion of the Pd layer toward the substrate was observed from RBS, forming a new layer of $TiPd_3$ in good agreement with XRD analysis. It was discovered that annealing caused the formation of metal hydrides, and the newly formed crystal structures were examined using XRD analysis software. The formation of the metal hydride is proof that the multi-layered system can be used in hydrogen absorption applications. The oxidation of Ti and Zr was observed at RT because Ti and Zr are highly reactive elements. Metals like Pd are imported to prevent oxidation layer in highly reactive elements. A further benefit of 150 keV argon ion implantation was its ability to lower O concentration, particularly in the Zr layer.

Acknowledgments

This work is based on the research supported by iThemba LABS NRF and the University of Zululand research office. The authors would like to thank Dr SC Masikane and Mr S Khoza,chemistry department, University of Zululand for the XRD analysis and Dr N Kheswa,iThemba LABS for e-beam deposition.

References

[1] Khumalo Z, Thethwayo C, Mtshali C, Msimanga M, Madito M, Numan N, Mongwaketsi N, Kotsedi C and Kheswa N, 2023, Interfacial reaction and phase formation in Pd/ZrO/Pd/TiO/Pd multilayer stack on silicon substrate: Investigated by ion beam techniques, *Vacuum.*, Vol. **214**, no. 112204.

[2] Maebela N, Mtshali C, Khumalo Z, Madjoe R, Arendse C, Cummings C, Madito M and Mongwaketsi N, 2022, Influence of implanted 150 keV Fe ion and 100 keV Mg ions on hydrogen absorption by Pd/Ti/V/Pd/Ti multilayer films on Ti substrate, *Nucl. Instr. Meth. B.,* Vol. **523**, no. 1-7.

[3] Liu H, Tan L, Guo Y, Chen P, Gao S, Su Z, Liu W and Zhao J, 2019, Fabrication and electrochemical hydrogen storage performance of Ti49Zr26Ni25 alloy covered with Cd/Pd core/shell particles, *Int. J. Hydrog. Energy.*, Vol. **44**, no. 24800-24809.

[4] Magogodi S, Mtshali C, Halindintwali S, Khumalo Z, Mongwaketsi N, Cloete K, Madito M, Pieters C, Sook A and McDonald E, 2019, Effect of the annealing atmosphere on the layer interdiffusion in Pd/Ti/Pd multilayer stacks deposited on pure Ti and Ti-alloy substrates, *Nucl. Instr. Meth. B.*, Vol. **461**, no. 37-43.

[5] Lai Q, Sun Y, Wang T, Modi P, Cazorla C, Demirci U, Ares Fernandez J, Leardini F, and Aguey-Zinsou K, 2019, How to design hydrogen storage materials? Fundamentals, synthesis, and storage tanks, *Adv. Sustain. Syst.*, Vol. **3**, no. 1900043.

[6] Fujimoto Y, Tadayyon S, Yoshinari O and Tanaka K 1997 Effect of hydriding on structural stability of Ni/Ti and Pd/Ti multilayers, *Mater. Trans*., Vol. **38**, no. 780-786.

[7] Lopez-Suarez A, Richards J and Trejo-Luna R 2003 Analysis of hydrogen absorption by Ti and Ti–6Al–4V using the ERDA technique, *Int. J. Hydrog. Energy*., Vol. **28**, no. 1107-1113.

[8] Topić M, Halindintwali S, Mtshali C, Nsengiyumva S and Khumalo Z 2018 Hydrogen storage in Ti-based metal hydrides investigated by elastic recoil detection analysis (ERDA), *Nucl. Instr. Meth. B.,* Vol. **450**, no. 239-243.

Electronic, elastic and thermoelectric properties of hexagonal CuSe phase

M. Ramoshaba, T.E. Mosuang
Department of Physics, University of Limpopo, Private Bag x1106, Sovenga, 0727, South Africa

E-mail: **moshibudi.ramoshaba@ul.ac.za,thuto.mosuang@ul.ac.za**

Abstract. Density functional theory using full-potential all-electrons linearised augmented plane waves was implemented to study electronic, elastic, and thermoelectric properties of hexagonal CuSe phase. Electronic band structure suggest a metallic compound of zero energy band gap. Density of states further expose the electron density responsible for such metallic identity. Elastic properties reveal mechanical stability and the possibility of being synthesisable. The compound is less compressible through positive calculated Cauchy pressure. High values of the power factor and the Seebeck coefficient allow consideration of the phase for thermoelectric applications

1. Introduction

Copper selenides (CuSe) are a component of the metal chalcogenides family with appealing yet unusual physical and chemical properties which may form part of the global energy crisis solution. Such properties make CuSe an upcoming material for applications in diverse fields such as photo-voltaics, photo-catalysis, gas sensing, thermo-electrics, and solar cells [1, 2]. Successful fabrication of some ordered 2- and 3-dimensional CuSe nanosheets and nanoplates is crucial in the design of thermo-electric converters and solar cells [2]. Peiris *et al*. [3] and Nozaki *et al* [4] indicated that hexagonal CuSe transforms to the orthorhombic phase at 323 K. Additional temperature increase induces further transition from orthorhombic to another hexagonal phase at 393 K. This high temperature hexagonal CuSe has no super-lattice reflections due to repositioning of some Cu^{2+} ions in the trigonal planar geometry. Also, Xu *et al.* [5] reports CuSe as an excellent precursor in the fabricating of ternary chalcogenides - $CuInSe_2$, and high order chalcogenides based nanostructures. Various methods which include but not limited to solvo-thermal, hydro-thermal, hot injection, and chemical aqueous have been used and reported on how CuSe polymorphs can be synthesised [1 – 7]. Li *et al*. [7] have studied the evolution and emergence of irregular growing yet ordered $Cu_{2-x}Se$ and CuSe nanocrystals. The study revealed that different stoichiometric compositions and crystal phases occur normally when free Cu^{2+} ions diffuse into a Cu-Se reaction setting under selenium rich conditions. Some of the notable phases accounted for in the literature are cubic berzeliante Cu_2Se, tetragonal umangite Cu_3Se_2, hexagonal klockmannite CuSe, cubic pyritic krutaite CuSe, and orthorhombic marcasitic krutaite $CuSe_2$ [2].

This paper is focused on the electronic, elastic, and transport properties of the hexagonal CuSe. Despite the ongoing research on various stoichiometric compositions and crystal structures of CuSe, an experimentally observed hexagonal CuSe, is not well reported. In this paper, acquired first principle density functional theory calculations seek to provide feasible parameters for intensive analysis on the nature of hexagonal CuSe phase.

2. Computational Method

Total energy first principle calculations based on linearized augmented plane waves + local orbitals were used. Generalized gradient approximation (GGA) proposed by Perdew, Burke, and Ernzerhof (PBE) revised for solids was used for the exchange and correlation functional energy [2]. The all electron full-potential computational software package; *Exciting-code* was used [7]. Brillouin zone integrals were approximated using the method of Monkhorst and Pack, [8]. The energies were allowed to converge with respect to the k points' density (4 × 4 × 2) for the density of states, band structure and elastic calculations. The transport coefficients with respect to the chemical potential at a constant temperature (300 K) were calculated using the full-potential all electron ab-initio calculations. These transport properties are based on the Boltzmann transport theory [9]. In the calculations, only k-points with Eigen energies near the Fermi level were considered to contribute to the coefficients. In order to achieve a sufficient representation of the energy bands in this area, a fine extended k-points grid (30 x 30 x 20) was adopted. Throughout the calculations, the following Wyckoff crystal positions were used: Cu_1^{2+} - 2d (0.667, 0.333, 0.250) and Cu_2^{2+} - 4f (0.333, 0.667, 0.393), as well Se_1^{2-} - 2c (0.667, 0.333, 0.750) and Se_2^{2-} - 4e (0.000, 0.000, 0.435) [13] for a complete hexagonal symmetry.

3. Results and Discussion

The electronic, elastic and transport properties of hexagonal CuSe were investigated under room temperature conditions. The electronic properties can be used to fully define the state and behaviour of electrons in the material. For example, the electronic band structure and the density of states, which describes the state of the electrons in terms of their energy. Elastic properties determine the mechanical properties of the materials. The transport properties assist with the understanding of various interactions in electronic systems such as Seebeck coefficient, electrical conductivity, thermal conductivity and power factor. Figure 1 is a typical hexagonal CuSe crystal lattice obtained using the XCrysDen programme [10], which is used to visualise different crystalline and molecular structures. The brown spheres represent the copper atoms whilst the green ones represent selenium atoms.

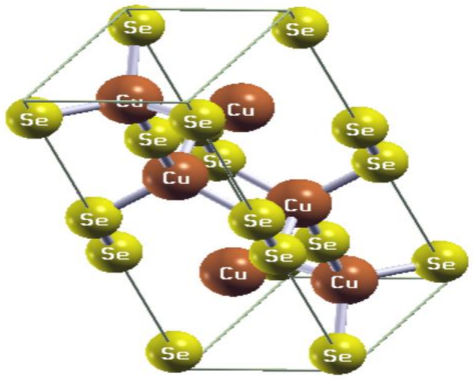


Figure 1. Visualisation of the crystal structure of hexagonal CuSe. Brown sphere – copper ions; green spheres – selenium ions.

3.1. Electronic properties

On setting up the electronic properties calculations, a reasonable k points' density of 4 × 4 × 2 was used to predict both the density of states and band structure. On the structural optimization, all atoms were allowed to relax, whilst maintaining an energy convergence limit of 1.0×10^{-5} eV. The computed band structure along the high symmetry directions in the Brillouin zone is shown in figure 2. Some of the

upper most valence band states and some of the lower most conduction band states overlap at the Femi level, which suggest a metallic character of the hexagonal CuSe [8, 11]. In order to confirm this idea, Lv *et al.* [12], enlightened that, if the energy bands have curves which produce a maxima, the effective mass of the charge carriers is negative and associated with p-type conductivity. If the curves produce a minima, the effective mass of the charge carriers is positive which suggest an n-type material. Figure 2 shows that the curvature of the bands is downward, meaning that the charge carriers have a negative effective mass and hexagonal CuSe suggests a p-type material. In addition, figure 2 shows some flat bands along symmetry routes: Γ-A, H-K and L-M, which point out that the mass of charge carriers along these directions are very hefty, so hexagonal CuSe suggests an anisotropic electronic transport behaviour [13]. The calculated total density of states (TDOS) and partial density of states (PDOS) are presented in figure 3. The PDOS will help in observing the bonding properties of hexagonal CuSe. It can be seen that the top of the valence band and the bottom of the conduction band influenced by the Cu 3d and Se 3p. Cu 3d contribute more than Se 3p in the conduction band. DOS and band structure show that, there is a gap between -4 and -3.8 eV at the valence band. The energy bands between -15 to -10 eV represent the Se 3s. The calculated band structures and TDOS of hexagonal CuSe indicate the metallic behaviour with zero band gap. This agrees well with the results of Milman *et al.* [8] and that of Shojae *et al.* [11].

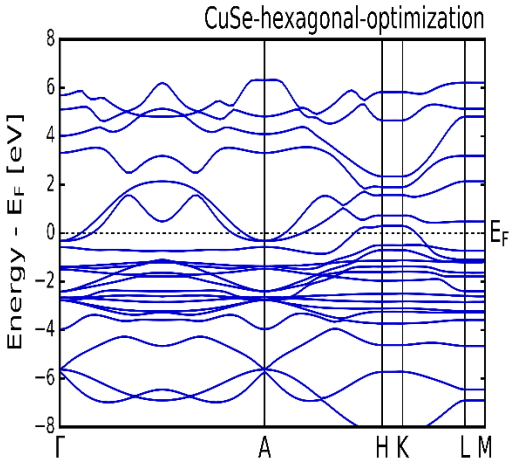


Figure 2. Band structure of CuSe.

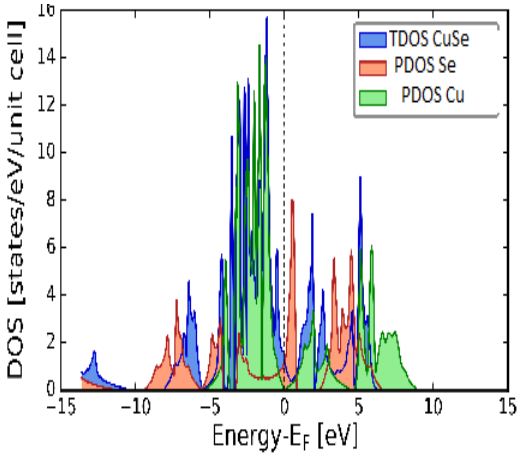


Figure 3. Density of states of CuSe.

3.2. Elastic properties

Systematically, the elastic properties of hexagonal CuSe were also computed. In order to study the mechanical stability and behaviour of CuSe, elastic constants (C_{ij}), were computed. The hexagonal crystal consists of five independent elastic constants, which are C_{11}, C_{12}, C_{13}, C_{33} and C_{44}. A stable hexagonal symmetry must satisfy the following elastic constants relations: $C_{11} > 0, C_{44} > 0, C_{11} > |C_{12}|$ and $(C_{11} + C_{12})C_{33} - 2C_{13}^2 > 0$ [10-11]. In this article, elastic constants were computed using the ElaStic@exciting [14] interface, which can be used to obtain full elastic constants of any crystal structures. From the hexagonal elastic constants C_{ij} (ij = 11, 12, 13, 33, and 44); mechanical properties which include bulk modulus (B_0), shear modulus (G), Young's modulus (Y) and Poisson's ratio (v) were computed using the Voigt approximations [11-12]. In table 1, all the criteria for structural stability of a hexagonal system listed above are satisfied. C_{11} and C_{33} describe how the responds to linear compression, whilst C_{12}, C_{13}, and C_{44} express how the material resists a change in shape. Besides, C_{11} and C_{33} can also describe the stiffness felt by the material along the crystallographic *a*- and *c*-axis, respectively. C_{12} and C_{13}, which are shear elastic constants, lead to a functional stress component felt along the crystallographic *a*-axis with a uniaxial strain along the *b*- and *c*-axis, respectively. If $C_{11} < C_{33}$, this suggest that the crystal is more susceptible to lattice compressions [12].

Table 2 displays the bulk modulus (B_0), Young's modulus (Y), shear modulus (G), Cauchy pressure (C_{12}-C_{44}), and Poisson's ratio (v), Pugh's index (G/B_0) and Zener's anisotropy (A) of the hexagonal CuSe. The high bulk modulus value (71.32 GPa) of the material suggest that it is hard to compress. Moreover, the values of the bulk and shear moduli were used to calculate the Poisson's ratio. According to Arora *et al*. [13], materials having the Pugh's index (G/B_0) and Poisson's ratio (v) less than 0.500 and greater than 0.330 in that order are characterised as ductile with dominant ionic bonds. As such, a positive value of 0.403 was obtained for the Poisson's ratio, which further suggest incompressibility and ductility with major ionic bonding's [13]. To quantify this, a Cauchy pressure value (63.83 GPa) is also high and positive, further verifying ionic metallic character of the compound. Looking at the CuSe's Pugh's index, a value of 0.135, which is a quotient from G and B_0 values listed in table 2, signifies a positive, ductile, ionic, and metallic bonding property. Such properties are crucial in solar cell device fabrication [13]. Another property of interest is a Zener's anisotropy factor (A), which is a measure of the degree of anisotropy in the bonding between atoms in different planes. If a crystal possess a Zener's anisotropy factor $A \leq 1$, it suggests an elastic anisotropy [15]. Consequently, the value of A (0.174) listed in table 2, which is less than 1, signifies an anisotropic behaviour on the hexagonal CuSe.

Table 1. Calculated elastic constants C_{11}, C_{12}, C_{13}, C_{33}, C_{44} and C_{66} all in GPa.

CuSe	C_{11}	C_{12}	C_{13}	C_{33}	C_{44}	C_{66}
This work (GPa)	80.87	70.62	57.01	110.9	6.78	5.12

Table 2. Calculated bulk modulus (B_0), Young's modulus (Y), shear modulus (G), Cauchy pressure (C_{12}-C_{44}) all in GPa, Poisson's ratio (ν), Pugh's index (G/B_0) and Zener's anisotropy factor (A).

CuSe	B_0	Y	G	C_{12}-C_{44}	v	G/B_0	A
This work (GPa)	71.32	27.55	9.59	63.83	0.403	0.135	0.174

3.3. Transport properties

The transport properties were computed using the electronic structure procedure as per sub-section *3.1*. But here, the calculations considered only the k-points and Eigen energies near the Fermi level for a contribution to the transport coefficients. In order to accomplish an acceptable representation of the energy bands during the calculations, a good extended k-points grid (30 x 30 x 20) was adopted. The Seebeck coefficient (S), electrical conductivity per relaxation time (σ/τ), thermal conductivity relaxation time (κ/τ) and power factor (σS^2) all as functions of chemical potential (μ) were firstly studied at 300 K. The positive and negative values of the chemical potential stand for the electrons (n-type) and holes (p-type) doping, respectively. The doping level in a material is described by the chemical potentials, which is very crucial for accepting the thermoelectric description of a material for practical recognition [16-17]. Figure 4 presents the Seebeck coefficient as a function of chemical potential at 300 K. The maximum S value at the p-type doping is 110 μVK^{-1} and at the n-type is -90 μVK^{-1}. Given the fact that the S value at the p-type doping is greater than that at the n-type, this signify that hexagonal CuSe favours holes conductivity. Proceeding, figure 5 portrays the σ/τ against the μ. The maximum σ/τ at the p-type doping is 60000 Scm^{-1} and at the n-type doping is 48000 Scm^{-1}. The results suggest that the p-type dopants possess greater electrical conductivity compared to the n-type dopants. Results of the thermal conductivity per relaxation time against the chemical potential are shown in figure 6. The curve depict a significant decrease in thermal conductivity with the increasing chemical potential. The highest values of κ/τ appear at the p-type dopant region around 0.38 $Wcm^{-1}K^{-1}$ and 0.31 $Wcm^{-1}K^{-1}$. The

behaviour of the Power factor (σS^2) against the chemical potential (μ) is illustrated in figure 7. At the p-type dopant region, the maximum value of σS^2 is 325 μWcm^{-1}K^{-2} whilst at n-type dopant region is 220 μWcm^{-1}K^{-2}. Just like the Seebeck coefficient and the electrical conductivity, the power factor behaviour favours p-type doping which conform to holes conductivity. In unison, high positive values of the Seebeck coefficient, electrical conductivity and power factor suggest the hexagonal CuSe as a good thermoelectric material.

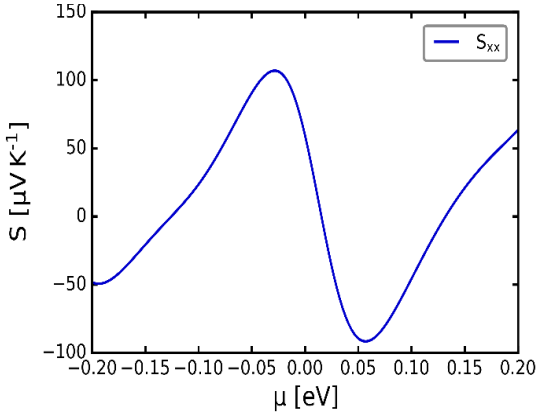


Figure 4. Variation of the Seebeck coefficient vs chemical potential.

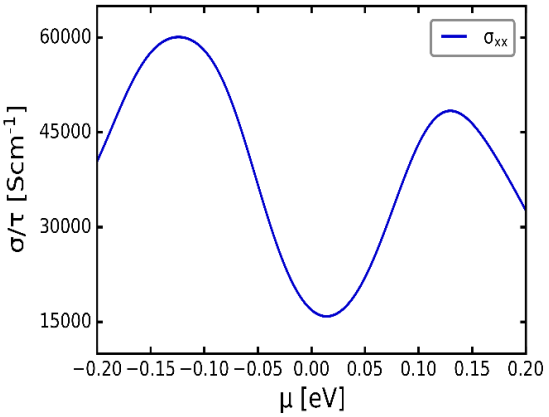


Figure 5. Variation of the electrical conductivity per relaxation time vs chemical potential.

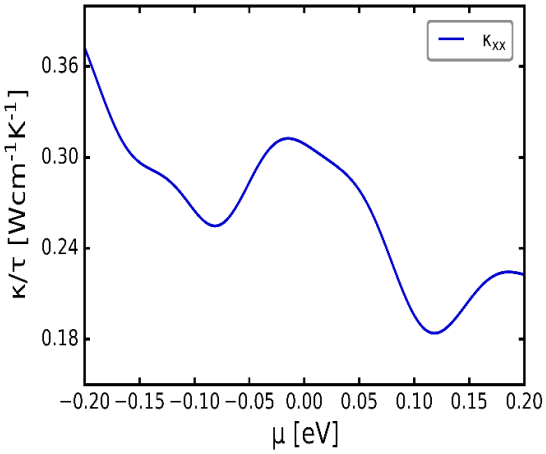


Figure 6. Variation of the thermal conductivity per relaxation time vs chemical potential.

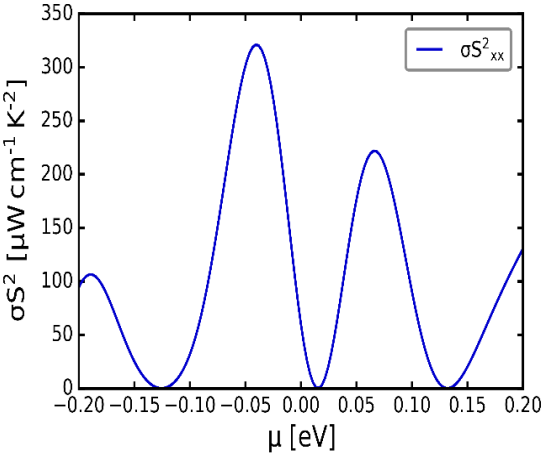


Figure 7. Variation of the Power factor vs chemical potential.

Systematically, Figure 8 displays the Seebeck coefficient vs temperature for μ = -0.136 and 0.136 eV. It must be noted that according to literature, positive chemical potential (0.136 eV) conforms to n-type doping and a negative chemical potential (-0.136 eV) conforms to p-type doping [17, 18]. In the temperature region, 300 – 700 K, both p- and n-type conductivities behave similarly. Starting from 700 – 1200 K, the p-type is elevated compared to the n-type conductivity. This signifies how the μ = -0.136 eV, p-type conductivity dominates the hexagonal CuSe at high temperatures in accordance with the electronic structure data of sub-section *3.1*.

Variation of the electrical conductivity per relaxation time (σ/τ) against temperature can be seen in figure 9. In both μ = -0.136 and 0.136 eV curves, σ/τ decreases exponentially with the increasing temperature. Furthermore, the μ = -0.136 eV curve, across the entire temperature range (300 – 1200 K), experiences the electrical conductivity per relaxation time which is greater than that of μ = 0.136 eV

curve by a uniform ±13 x 10^3 Scm^{-1} margin. Another confirmation of the p-type conductivity behaviour of hexagonal CuSe. Figure 10 show graphs of thermal conductivity per relaxation time (κ/τ) against temperature for μ = -0.136 eV and 0.136 eV. At both chemical potentials, κ/τ increases with the temperature, in the manner in which metallic materials do. But for a change, the μ = 0.136 eV n-type curve, throughout the temperature range (300 – 1200 K) is greater than the μ = -0.136 eV p-type conductivity curve. A clarification of this anomaly came from Narjis *et al*. [19]; materials with high electrical conductivity must have a very low thermal conductivity in order to possess good thermoelectric effect. As reflected in figure 10, a good thermoelectric material with p-type conductivity is supposed to have quite a low thermal conductivity. Lastly, figure 11 provides the variation of the Power factor (σS^2) against temperature. For μ = -0.136 eV curve, the Power factor starts by decreasing to zero at ±450 K, thereafter it increases almost linearly to 1200 K. Similarly, the μ = 0.136 eV curve, starts by decreasing to zero at ±400 K, from then on it increases to exceed that of μ = -0.136 eV at 1200 K. This behaviour could be accounted for by the decreasing Seebeck coefficient with temperature in region 300 – 450K for both μ = -0.136 and 0.136 eV curves (Figure 8). Hasan *et al* [20] explained that if the maximum σS^2 was obtained at the p-type doping region (μ <0), then it could be concluded that hexagonal CuSe perform better as a p-type thermoelectric material. Similarly, when maximum power factor was obtained at the n-type doping region (μ >0), then the material works well as an n-type thermoelectric material [20]. Figure 11 further shows that at a low temperature (300 K), the p-type doping power factor is higher than that of the n-type doping. This further suggest that, hexagonal CuSe works better as a p-type thermoelectric material at low temperatures, it can also perform better as an n-type thermoelectric material at high temperatures.

4. Conclusion

The computed electronic band structure and total-partial density of states of hexagonal CuSe show a metallic character with a zero energy gap. The curvature of the band-lines in the band structure suggest a p-type material. Somehow flat band-lines at the valence band region, suggest an anisotropic behaviour of hexagonal CuSe.

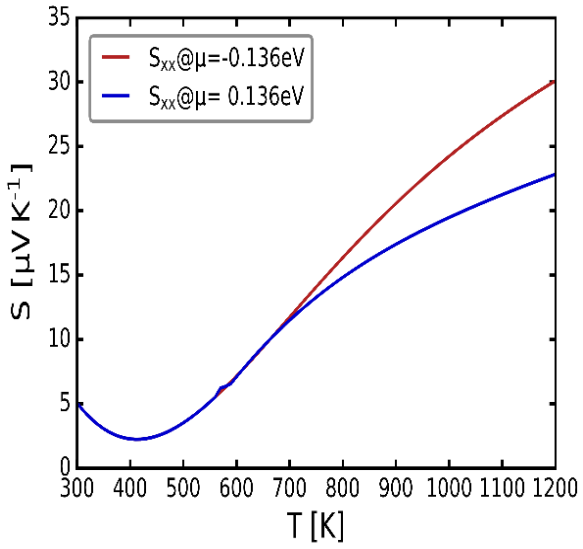


Figure 8. Variation of the Seebeck coefficient vs temperature.

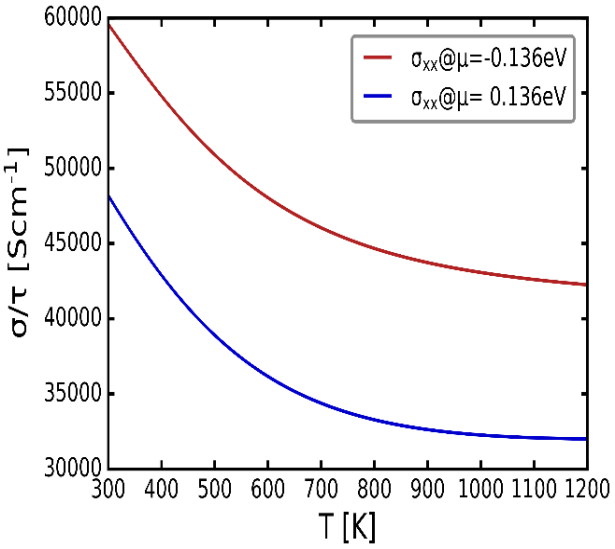


Figure 9. Variation of the electrical conductivity per relaxation time vs temperature.

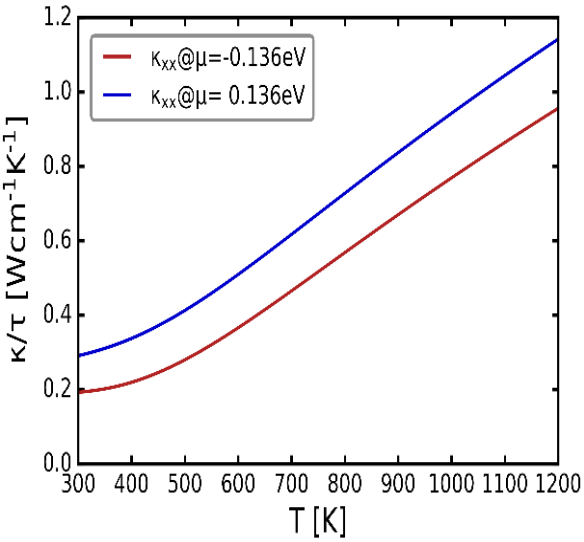


Figure 10. Variation of the thermal conductivity per relaxation time vs temperature.

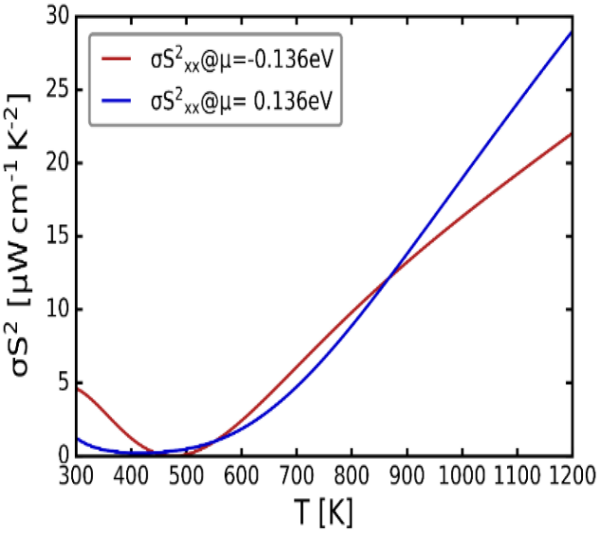


Figure 11. Variation of the Power factor vs temperature.

The TDOS further imply that Cu^{2+} ions d electrons contribute more to the electronic structure than the Se^{2-} ions. The values of the elastic constants obtained indicate that hexagonal CuSe is mechanically stable.

Further on, the properties suggest that the CuSe system is ductile and susceptible to compression. The Cauchy pressure is positive, which suggests the metallic characteristics with ionic bonding of the compound. The Pugh's index and Poisson's ratio values show the ductility behaviour with the possibility of being applicable to solar cells device fabrication. In addition to that, the value of C_{44} which is the smallest elastic constant, suggest that if a shear stress is applied in the horizontal plane along the vertical direction, the shear deformation can easily occur.

At μ = -0.136 eV, which aligns with p-type doping, hexagonal CuSe has high and low electrical conductivities which suggest good thermoelectric effect. The thermal conductivity increases with temperature at both μ = -0.136 and 0.136 eV. At high temperatures, the p-type Seebeck coefficient is greater than that of n-type doping. This also emphasise the Seebeck coefficient effect on the p-type doping. In all transport coefficients against the chemical potential curves; the high values of the Seebeck coefficient, electrical conductivity, Power factor, and the low value of thermal conductivity, all confirm that hexagonal CuSe is a good thermoelectric material.

5. Acknowledgments

Special thanks to my supervisor, Prof TE Mosuang, University of Limpopo for funding and the Centre for High Performance Computing for computational resources.

References

[1] Xiao, L., Wu, J., Ran, J., Liu, Y., Qiu, W., Lu, F., Shao, F., Tang, D. and Peng, P., 2016. Near-infrared radiation absorption properties of covellite (CuS) using first-principles calculations. *AIP Advances*, *6*(8).

[2] Willamson E.M., Sun Z., Tappan B.A., Brutchery R.L., Predictive synthesis of copper selenidesusing a multidimensional phase map constructed with a data driven classifier. *J. Am. Chem. Soc.* 2023, 145, 17954-19964

[3] Peiris, S. M.; Pearson, T. T.; Heinz, D. L. Compression of Klockmannite, CuSe. *J. Chem. Phys*. 1998, 109, 634−636

[4] Nozaki, H., Shibata, K., Onoda, M., Yukino, K. and Ishii, M., 1994. Phase transition of copper selenide studied by powder x-ray diffractometry. *Materials research bulletin*, *29*(2), pp.203-211.

[5] Xu, J., Lee, C.S., Tang, Y.B., Chen, X., Chen, Z.H., Zhang, W.J., Lee, S.T., Zhang, W. and Yang, Z., 2010.

Large-scale synthesis and phase transformation of CuSe, CuInSe2, and CuInSe2/CuInS2 core/shell nanowire bundles. *ACS nano*, *4*(4), pp.1845-1850.

[6] Liew, J.Y.C., Talib, Z.A., Zainal, Z., Kamarudin, M.A., Osman, N.H. and Lee, H.K., 2019. Structural and transport mechanism studies of copper selenide nanoparticles. *Semiconductor Science and Technology*, *34*(12), p.125017.

[7] Li, D., Zheng, Z., Lei, Y., Ge, S., Zhang, Y., Zhang, Y., Wong, K.W., Yang, F. and Lau, W.M., 2010. Design and growth of dendritic Cu2− xSe and bunchy CuSe hierarchical crystalline aggregations. *CrystEngComm*, *12*(6), pp.1856-1861

[8] Milman, V., 2002. Klockmannite, CuSe: structure, properties and phase stability from ab initio modeling. *Acta Crystallographica Section B: Structural Science*, *58*(3), pp.437-447.

[9] Stedman, T. and Woods, L.M., 2020. Transport theory within a generalized Boltzmann equation for multiband wave packets. *Physical Review Research*, 2(3), p.033086.

[10] Kokalj, A., 1999. XCrySDen—a new program for displaying crystalline structures and electron densities. *Journal of Molecular Graphics and Modelling*, 17(3-4), pp.176-179.

[11] Shojaei, A.R., Nourbakhsh, Z., Vaez, A. and Dehghani, M., 2013. Ab initio investigation of the structural and unusual electronic properties of α-CuSe (klockmannite). *Chinese Physics B*, *22*(12), p.127102.

[12] Lv, Z., Cui, H., Huang, H., Li, X., Wang, H. and Ji, G., 2017. Study of the electronic, bonding, elastic and acoustic properties of covellite via first principles. *Journal of Alloys and Compounds*, *692*, pp.440-447.

[13] Arora, S., Kabra, K., Joshi, K.B., Sharma, B.K. and Sharma, G., 2020. Structural, elastic, thermodynamic and electronic properties of covellite, CuS. *Physica B: Condensed Matter*, *582*, p.311142.

[14] Gulans, A., Kontur, S., Meisenbichler, C., Nabok, D., Pavone, P., Rigamonti, S., Sagmeister, S., Werner, U. and Draxl, C., 2014. Exciting: a full-potential all-electron package implementing density-functional theory and many-body perturbation theory. *Journal of Physics: Condensed Matter*, 26(36), p.363202.

[15] Al-Qaisi, S., Abu-Jafar, M.S., Gopir, G.K., Ahmed, R., Omran, S.B., Jaradat, R., Dahliah, D. and Khenata, R., 2017. Structural, elastic, mechanical and thermodynamic properties of Terbium oxide: First-principles investigations. *Results in physics*, 7, pp.709-714.

[16] Khan, B., Yazdani-Kachoei, M., Aliabad, H.R., Khan, I., Jalali-Asadabadi, S. and Ahmad, I., 2017. Effects of chemical potential on the thermoelectric performance of alkaline-earth based skutterudites (AFe4Sb12, ACa, Sr and Ba). *Journal of Alloys and Compounds*, 694, pp.253-260.

[17] Scheidemantel, T.J., Ambrosch-Draxl, C., Thonhauser, T., Badding, J.V. and Sofo, J.O., 2003. Transport coefficients from first-principles calculations. *Physical Review B*, 68(12), p.125210.

[18] Zhao, L.D., Berardan, D., Pei, Y.L., Byl, C., Pinsard-Gaudart, L. and Dragoe, N., 2010. $Bi_{1-x}Sr_xCuSeO$ oxyselenides as promising thermoelectric materials. *Applied Physics Letters*, 97(9).

[19] Narjis, A., Outzourhit, A., Aberkouks, A., El Hasnaoui, M. and Nkhaili, L., 2018. Structural and thermoelectric properties of copper sulphide powders. *Journal of Semiconductors*, *39*(12), p.122001

[20] Hasan, S., San, S., Baral, K., Li, N., Rulis, P. and Ching, W.Y., 2022. First-Principles Calculations of Thermoelectric Transport Properties of Quaternary and Ternary Bulk Chalcogenide Crystals. *Materials*, 15(8), p.2843.

Study of femtosecond laser annealed Sn/Ti bimetallic films using RBS as a probing tool.

K. Lebesana[1,2], I. K. Segola[1,2], G. Magchiels[3], A. Guga[2], C. Mtshali[2], Z. M. Khumalo[2], L. Kotsedi[2], A. Vantomme[3], M. Msimanga[1]

[1] Department of Physics, Tshwane University of Technology, Private Bag X680, Pretoria, 0001, South Africa

[2] iThemba LABS-NRF, Tandetron laboratory, Old Faure Road, Cape Town, Eerste river,7121, South Africa

[3] Quantum Solid-State Physics, KU Leuven, Celestijnenlaan 200D, 3001 Leuven, Belgium

E-mail: keletso77lebesana@gmail.com

Abstract. The experimental study of femtosecond laser-induced structural effects on epitaxially deposited Sn/Ti bimetallic thin films is reported. Thin films composed of varying thicknesses of 15 nm bilayers and 25 nm bilayers were deposited by Molecular Beam Epitaxy on SiO_2/Si virtual substrate. The as-deposited bimetallic thin films were annealed under non-isothermal conditions using a $\lambda = 1030$ nm fs laser, with a fixed repetition rate of 500 kHz under dynamic scanning mode at an area of 1.5 mm $\times$ 1.5 mm and varied laser powers. The surface morphology and microstructural features of the thin films were investigated by scanning electron microscopy. The structural and compositional changes due to non-isothermal annealing were analysed using Rutherford Backscattering Spectrometry, progressed intermixing between the bilayers in the films was observed and more pronounced a function of laser power. Particularly, pronounced intermixing was observed at 250 mW for the 15 nm bilayer films and at 500 mW for the 25 nm bilayers. Apart from intermixing, changes in the elemental composition in the films are attributed to heat accumulation, resulting in nanopatterning of the subsurface and laser-induced oxidation producing well-defined oxide top layers. The thickness of the top oxide layers increases with increasing laser power and grows at the expense of the subsurface layer.

1. Introduction

Since their development, thin films have continued to be at the forefront of nanotechnology due to their unique physical, chemical, mechanical, and optical properties. In particular, bimetallic thin films are superior to their monometallic and bulk counterparts because of the synergistic effects between the two metals resulting in enhanced properties [1, 2]. These types of materials have since allowed material researchers to design, explore, and develop novel nanoscale devices with precise control of their zero-dimensional to 3-dimensional metallic nanostructures [3]. The recent increased demand for miniaturization in nanotechnology has led to the need to create novel nanomaterials with certain matrix nanostructures with specific properties, high performance, and long lifetime using surface modification techniques [4, 5]. Recently, non-isothermal annealing using a fs laser has emerged as an advanced surface modification technique because of its ability to generate novel and complex nanostructures across a diverse range of materials including metals, glass, and semiconductors [6]. Femtosecond laser modification makes use of the primary benefits of localized heating within an extremely rapid timescale, thereby resulting in minimized

heat-affected zones which is usually a challenge with continuous nanosecond and longer laser pulses [5]. However, in contrast to conventional thermal annealing, there is still an increased and critical need to expand the scope of knowledge and investigate the structural effects and modifications induced by fs laser annealing. This is especially important for bimetallic thin films since they are susceptible to fs laser surface modifications on the nano and micro-scale [7]. A practical and effective approach to investigate the properties of such thin films is to perform accurate characterization which also enables material researchers to develop new materials with desired chemical composition and thickness [8]. Therefore, in this paper, the analysis of epitaxial Sn/Ti bimetallic thin films annealed by a fs laser was performed. The results reported here develop thin film processing technology and methods for manufacturing novel bimetallic thin films.

2. Experimental method

Sn and Ti epitaxial layers were grown using Molecular beam epitaxy (MBE) on a clean 200 nm SiO_2 on a Si substrate under ultra-high vacuum (UHV) base pressure of 2×10^{-10} Torr. Two sets of thin films consisting of two 15 nm layers of Sn and Ti, and 25 nm layers of Sn and Ti were evaporated from Knudsen cells. Sn with a low melting point was first deposited atop the substrate and Ti atop Sn layer. To prevent oxidation of the grown bilayers, a 5 nm Si capping layer was deposited on the 25 nm thin films. The epitaxially grown thin film bilayers were annealed under non-isothermal annealing conditions. The samples were mounted on a precise computer-controlled X-Y-Z (3D) translation stage shown in Figure 1, where raster scans were performed in ambient conditions. Non-isothermal annealing was performed using a $\lambda = 1030$ nm fs laser with a pulse duration of 300 fs and a beam spot size of 40 μm. The repetition rate, scan speed, spacing between scans, and spacing between areas were fixed to 500 kHz, v = 1 mm.s^{-1}, 1.5 mm × 1.5 mm, 40 μm, 0.7 mm resulting in approximately N = 20 000 effective number of pulses. Detailed surface morphology after fs laser annealing was examined firstly by scanning electron microscope (SEM), while Rutherford backscattering spectroscopy (RBS) was used to analyse the extent of atomic depth profile and compositional changes in the bimetallic films.

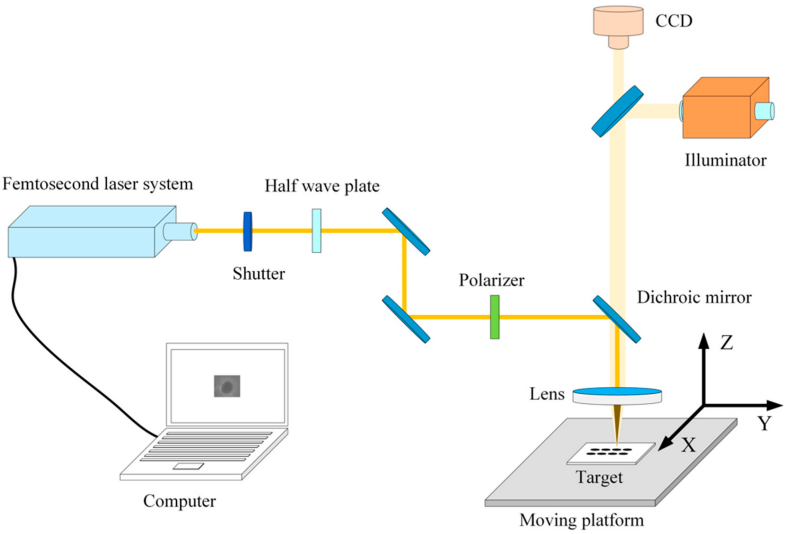


Figure 1. 3D computerized femtosecond laser stage [9]

3. Results and discussions

The laser-induced morphological changes were investigated using SEM and the images are shown in Figure 2 (a)-(c) and Figure 3 (a)-(c). The as-deposited 15 nm uncapped sample shows circular nanodroplet-like structures, and clear boundaries shown in in Figure 2(a), with an average particle size of 93 nm calculated with image J. The 25 nm as-deposited capped sample in Figure 3(a) shows nanoisland circular-like structures observed in the uncapped sample which are larger in size as confirmed by their average particle size of 268 nm. Figure 2 (b)-(c) shows the SEM micrographs of the uncapped sample annealed at 250 mW, and 500 mW respectively. At 250 mW, the micrograph in Figure 2(b) exhibits sparsely and randomly distributed nanostructures with an exposure of the (sub)surface in the form of random voids. Increasing the laser power to 500 mW led to more (sub)surface exposure with randomly distributed nanostructures. The observed laser-induced surface morphology especially the subsurface voids can be attributed to the heat accumulation in the surface often resulting in either thermo-elastic stresses, evaporation leading to delamination and dewetting in the heat-affected zone [10]. The 25 nm capped sample, at 250 mW in Figure 3(b), shows a melted surface that can result from laser-induced thermo-elastic stresses [10]. At 500 mW, randomly distributed nanostructures and more pronounced (sub)surface exposure in the form of voids with small spherical nanostructures are seen in Figure 3(c).

RBS analysis of the as-deposited and laser-annealed samples was performed to investigate the effect of laser heating on the thickness and compound compositions of the films. SIMNRA software was used to reproduce the RBS experimental data as shown in Figure 2 (b), (d) and (f), and Figure 3 (b), (d) and (f). The respective Ti, Sn, Si, and O elements contained in the thin films are also indicated. Despite Ti being at the top layer, it appears at lower channels (between 750 and 800) due to its smaller atomic number of (22) compared to Sn which has a higher atomic number of (50) and appears at higher channels (between 950 to 960). Additionally, the high oxygen peak contents observed in the spectra are attributed to the oxygen resonance energy, of 3.039 MeV which was used during the analysis [11]. The expected layer thicknesses in the as-deposited samples were approximately 15 nm layers of Sn and Ti, and 25 nm layers of Sn and Ti with a Si capping layer of 5 nm. However, the RBS analysis showed that the as-deposited 15 nm uncapped sample consisted of a TiO_2 top impurity layer in the 15 nm layer thin films. On the other hand, the 25 nm capped sample with a Si capping layer also indicated the existence of a tiny $SiTiO_2$ top layer. The existence of the oxide layers can be attributed to the fact that Ti and Si as top layers both have high oxygen affinity [12, 13]. To effectively show the effects of laser annealing on the thin films, the changes in the elemental concentrations and thicknesses are shown in Table 1 and Table 2 corresponding to RBS spectra in Figures 4(a) and 4(b) respectively.

Table 1. Elemental concentrations of the intermixed layer thicknesses in units of ($\times 10^{15}$ at.cm^2) after fs laser annealing of 15 nm uncapped sample.

Laser power (mW)	Layer 1				Layer 2			Layer 3	
	Ti	Sn	O	thickness	Ti	Sn	thickness	Sn	thickness
As-dep	0.11		0,89	58,46	1,00		30.17	1,00	50,53
250	0,08	0,24	0,69	63,72	0,51	0,49	63,34		
350	0,15	0,18	0,68	81,20	0,48	0,52	30,21		
500	0,17	0,16	0,62	102,28	0,49	0,51	20,20		

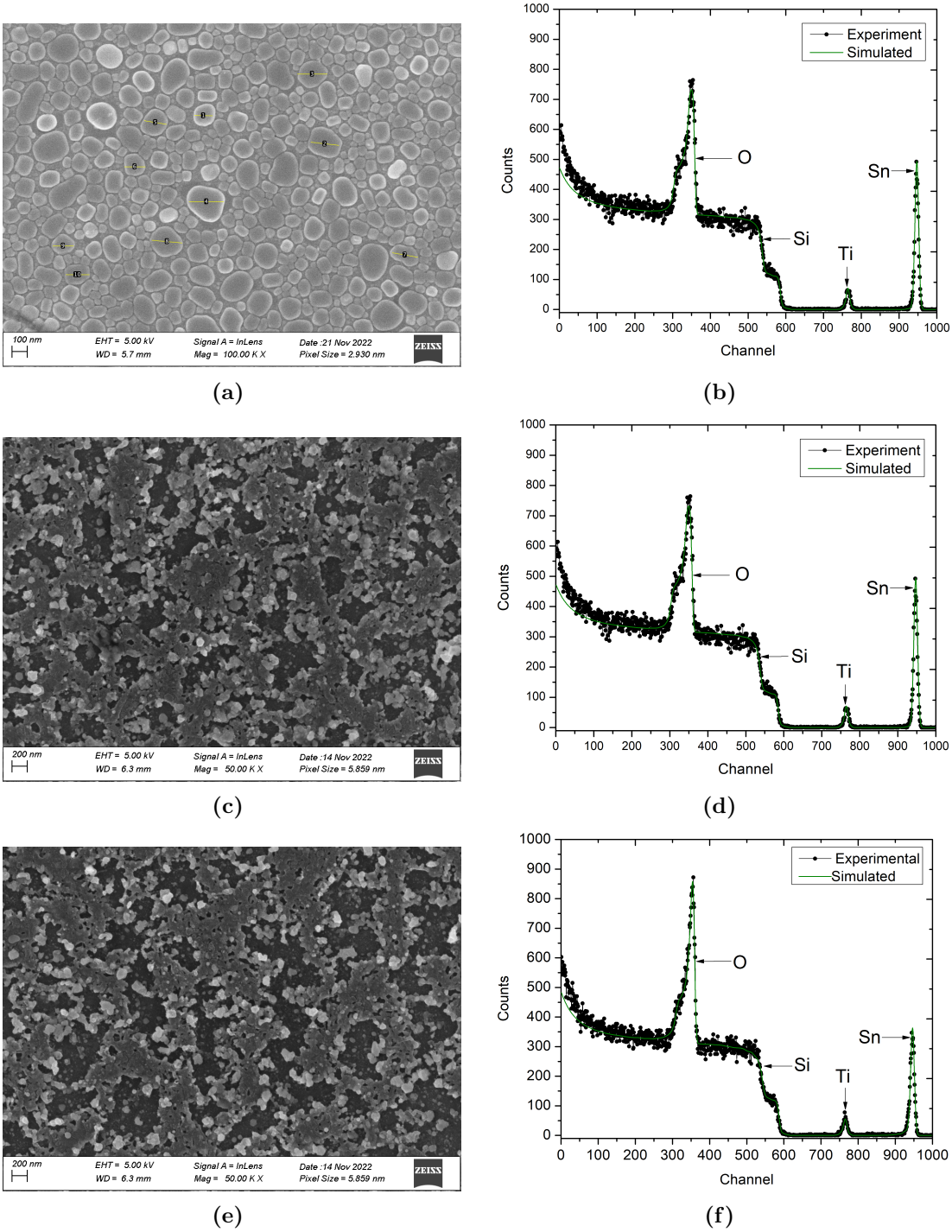


Figure 2. Represents SEM images and the corresponding RBS results for 15 nm samples where: (a), (b) represents the as-deposited, (c), (d) 250 mW, and (e), (f) 500 mW.

In the 15 nm uncapped sample, the concentrations of Ti and Sn in layer 2 remain relatively the same while the corresponding thickness decreases with increasing laser power, shown in Table 1. However, the thickness of layer 1 increases with increasing laser power. Additionally,

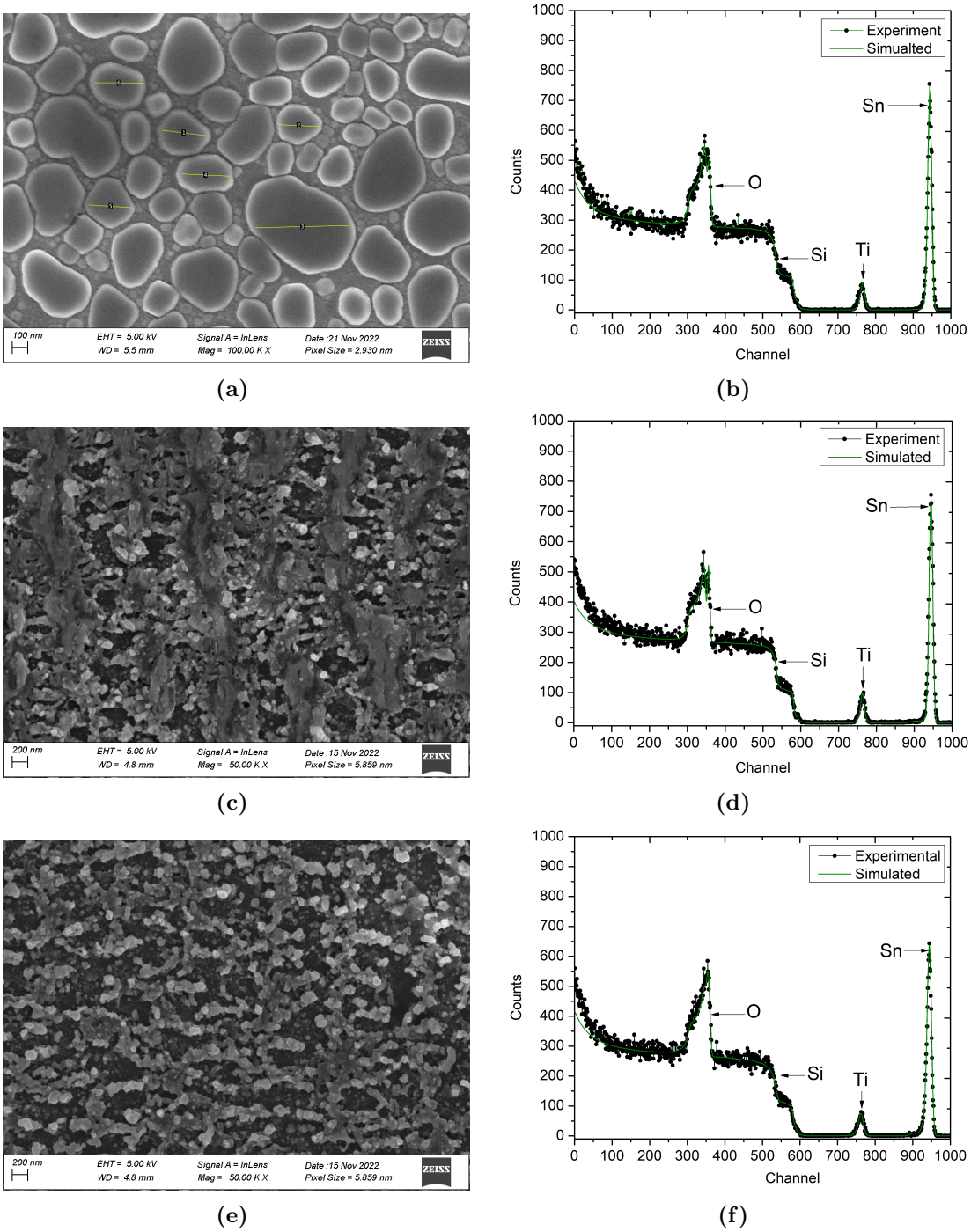


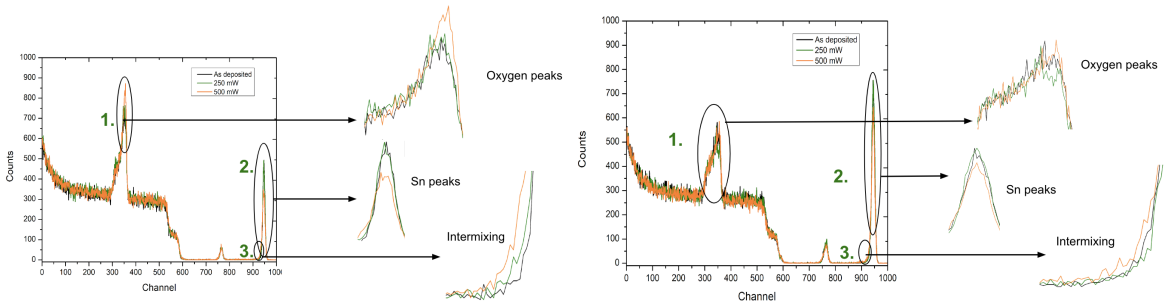
Figure 3. Represents SEM images and the corresponding RBS results for 25 nm samples where: (a), and (b) represent the as-deposited, (c), (d) 250 mW, and (e), (f) 500 mW.

the concentration of Ti increases with increasing laser power. In contrast, the Sn concentration decreases with increasing laser power which mostly indicates preferential intermixing and or ablation of Sn because of its lower melting point compared to Ti, as observed in Figure 4(a), region marked 3. Additionally, the SIMNRA simulation shows that the Sn peaks decrease

Table 2. Elemental concentrations of the intermixed layer thicknesses in units of ($\times 10^{15}$ at.cm^2) after fs laser annealing of 25 nm capped sample.

Laser power (mW)	Layer 1					Layer 2			Layer 3	
	Ti	Sn	Si	O	thickness	Ti	Sn	thickness	Sn	thickness
As-dep	0,20		0,30	0,50	93,04	1,00		30.77	1,00	66,02
250	0,15	0,13	0,19	0,53	68,11	0,42	0,58	127,72		
500	0,17	0,32	0,13	0,39	157,44	0,42	0,58	68,99		
750	0,18	0,27	0,16	0,40	206,52	0,41	0,59	40,64		

from approximately 550 counts in the as-deposited sample to 450 and 350 counts at 250 mW, and 500 mW, respectively. In the 25 nm capped sample, layer 1 on Table 2 shows a gradual increase of Ti concentration with increasing laser power while the concentration of the Si capping layer decreases with increasing laser power. On the other hand, Sn shows a similar increase in concentration at 250 mW, and 500 mW, however at 750 mW, the concentration of Sn decreases. The decrease in the concentration of Sn at 750 mW indicates that intermixing and interdiffusion in the layers was more pronounced at this laser power as can be seen in region marked 3 in Figure 4(b) Furthermore, Sn counts decreased from 850 counts in the as-deposited sample to 650 counts at 750 mW in Figure 4(b). Overall, at 250 mW and 500 mW, the spectra look relatively the same in all regions 1, 2, and 3 as indicated in Figure 4(b). For layer 2:(TiSn) shown in Table 2 which is closer to the substrate, the concentration of Ti and Sn does not change but remains relatively the same, while the thickness of the layer decreases with increasing laser power which is similar to the observations from the 15 nm bilayer films in Table 1. One clear similarity between both the 15 nm and 25 nm layers is that the amount of oxygen concentration in layer 1 decreases with increasing laser power, seen in Table 1 and Table 2. The observed increase in O peaks in the spectra appeared to be more pronounced in the 15 nm layers than in the 25 nm layers which can be attributed to the energy loss of the backscattered alpha particles which is higher in the 25 nm sample than in the 15 nm sample. The difference between the



(a) 15 nm uncapped sample **(b)** 25 nm capped sample

Figure 4. RBS spectra of the samples before and after laser heating at 250 mW and 500 mW.

concentrations observed in Table 1 and 2 and the respective oxygen peaks in the spectra shown in Figure 4(a) and (b) leads to the conclusion that the oxygen in the layers does not necessarily increase with increasing laser power. However, the laser-induced delamination and dewetting of the thin films as a function of laser power resulted in the exposure of the SiO_2 substrate surface

(voids) hence less energy loss on the film but the increase in the oxygen resonance peaks. This observation allows a correlation with the observed surface voids in the SEM images shown in Figure 2 (a)- (c) and Figure 3 (a) - (c). Additionally, the oxygen concentration on top (layer 1) for both samples decreases with increasing laser power, leading to the conclusion that layer 1 grows and increases at the expense of the (sub)surface layer (layer 2) in both samples.

4. Conclusion

Epitaxial Sn/Ti thin films were successfully grown and annealed under non-isothermal conditions using an ultrafast $\lambda = 1030$ nm fs laser. The surface morphology was remarkably changed after dynamic fs laser annealing. Generally, fs laser heating at 250 mW resulted in melting, delamination and resolidification which induced voids in the subsurface layer 2. Secondly, the irradiation was performed in ambient air, which resulted in oxidation producing a well-defined protective oxide layer (layer 1). Sn concentration decreases in the top layer (layer 1) while the concentration of Ti increases which can be attributed to the fact that heat accumulation resulted in the preferential diffusion and/or ablation and segregation of Sn with high mobility, and low melting temperature compared to Ti.

Acknowledgments

The authors would like to acknowledge the Tshwane University of Technology and the iThemba LABS (NRF), Cape Town for the financial support. The authors also would like to acknowledge KULeuven, iThemba LABS (NRF), and Dipartimento di Fisica, Politecnico di Milano for providing the necessary resources, facilities, and access to equipment for the success of this project.

References

[1] Mazhar T, Shrivastava V and Tomar R S 2017 *Journal of Pharmaceutical Sciences and Research* **9** 102
[2] Akman E, Oztoprak B G, Gunes M, Kacar E and Demir A 2011 *Photonics and nanostructures-fundamentals and applications* **9** 276–286
[3] Paras, Yadav K, Kumar P, Teja D R, Chakraborty S, Chakraborty M, Mohapatra S S, Sahoo A, Chou M M, Liang C T *et al.* 2022 *Nanomaterials* **13** 160
[4] Neri G 2015 *Chemosensors* **3** 1–20
[5] Ma Z C, Zhang Y L, Han B, Chen Q D and Sun H B 2018 *Small Methods* **2** 1700413
[6] Kucherik A, Arakelyan S, Kutrovskaya S, Osipov A, Istratov A, Vartanyan T, Itina T *et al.* 2017 *Journal of Nanomaterials* **2017**
[7] De Damborenea J 1998 *Surface and Coatings Technology* **100** 377–382
[8] Arunkumar P, Kuanr S K and Babu K S 2015 *Thin Film Structures in Energy Applications* 1–49
[9] Wu X, Yin H and Li Q 2019 *Applied Sciences* **9** 3045
[10] Talbi A, Coddet P, Tabbal M, Thomann A L, Millon E, Stolz A, Boulmer-Leborgne C, O'Connor G and Semmar N 2019 *Applied Surface Science* **476** 303–307
[11] Ager F, Justicia I, Gerbasi R, Battiston G, McSporran N and Figueras A 2006 *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **249** 490–492
[12] Fang Z Z, Middlemas S, Guo J and Fan P 2013 *Journal of the American Chemical Society* **135** 18248–18251
[13] Ritala M, Kukli K, Rahtu A, Raisanen P I, Leskela M, Sajavaara T and Keinonen J 2000 *Science* **288** 319–321

Structural and magnetic properties of $Mn_{50}Pt_{50-x}Ni_x$ alloys: A first principles study

R G Diale[1], P E Ngoepe[2], H R Chauke[2], S J Moema[1], and M J Phasha[1]
[1] Advanced Materials Discipline (AMD), Mintek, Private Bag X 3015, Randburg 2125, South Africa
[2] Materials Modelling Centre, University of Limpopo, Private Bag X 1106, Sovenga, 0727, South Africa

E-mail: **ram@mintek.co.za**

Abstract. $L1_0$ MnPt alloy has attracted much attention recently since it can be easily produced and provides magnetic properties of interest to spintronic applications. According to previous studies, this alloy retains its antiferromagnetism at room temperature. It is estimated that only a few studies have been conducted on a ferromagnetic (FM) MnPt alloy, despite its high magnetic moment compared with FePt. In this study, first-principles calculations were performed in the framework of density functional theory to identify new spintronic material compositions in Mn-based ternary alloys. By using this method, the structural, thermodynamic, elastic and magnetic properties were determined to track the stability of $Mn_{50}Pt_{50-x}Ni_x$ (x=0, 6.25, 12.5, 18.75, 25) alloys. According to heats of formation calculations, B2 and L10 $Mn_{50}Pt_{50-x}Ni_x$ alloys are thermodynamically stable when c/a is less than 1.10, indicating the possibility for these alloys to be experimentally explored. Furthermore, it was found that the magnetic moment for B2 phase decreased with increase in Ni composition, except for composition with 18.75 atomic percent Ni where the magnetic moment increased. As a result of these predictions, guided experimental work can be undertaken on promising $Mn_{50}Pt_{50-x}Ni_x$ alloys for potential use in future spintronic devices.

1. Introduction

The Mn-based alloys have recently attracted much attention in the information technology industry due to their interesting magnetic and structural properties [1]. Intermetallic alloys containing Mn, such as MnPt, have antiferromagnetic exchange interactions [2]. At low temperatures, this alloy undergoes a diffusionless transformation from the cubic phase B2 (CsCl) to the tetragonal phase $L1_0$ (CuAu-I). It has been reported that MnPt $L1_0$ phase has a very high Néel temperature (T_N) of 970 K [2]. Furthermore, this alloy phase has excellent chemical and thermal stability and a large magnetocrystalline anisotropy (MAE), making it more suitable for spintronic applications [3, 4]. To further enhance these attractive properties for use in data storage devices, it is vital to conduct investigations that will provide insights on the fundamentals governing these magnetic properties. The research studies carried out so far on MnPt focused mainly on B2 paramagnetic (PM) phase which martensitically transforms to $L1_0$ antiferromagnetic (AFM) phase at low temperature [2]. Only very few studies have been conducted on ferromagnetic (FM) MnPt alloy [5, 6], despite its high magnetic moment compared to FePt, rendering it suitable candidate material for use in electronic devices. This lack of motivation to pursue FM MnPt alloy could be attributed to the difficulty to experimentally stabilize this alloy at room temperature.

This remains so despite the fact that the above mentioned material processing challenge can be overcame through alloying. Subsequently, in attempt to revive interest in this material, theoretical investigation on the effects of Cr and Fe elements on MnPt alloy were studied by the same authors using

first-principles calculations [5]. It was revealed that $Pt_{50}Mn_{43.75}Cr_{6.25}$ shows improved magnetism when c/a is 1.10, whereas Fe is found to reduce the magnetic moment below $Pt_{50}Mn_{50}$. Furthermore, in a similar study, the effect of Co on the magnetic properties of $Mn_{50}Pt_{50}$ was also studied [6]. According to the findings, the addition of Co favours the formation of $L1_0$ compared with B2, as a result of a negative heat of formation. The study revealed that the addition of Co on both B2 and $L1_0$ $Mn_{50}Pt_{50}$ results in improved magnetic strength and mechanical stability on Pt-site. Based on the findings above, it is clear that magnetic strength can be improved on Pt-site as compared to Mn-site. Consequently, in order to have a complete view on the alloying effect of AFM and FM transition metals, the current study investigates the effect of Ni on magnetic properties of MnPt alloys on Pt-site using DFT-based first-principles calculations.

2. Methodology

A density functional theory framework, implemented in the Vienna ab initio simulation Package (VASP) [7, 8], was used for all calculations in this study. The atomic potentials used in these calculations were described by the projector-augmented wave (PAW) pseudopotentials [9]. The exchange correlation function was handled by the generalized gradient approximation (GGA) [10] with spin polarization enabled proposed by Perdew, Burke and Ernzerhof (PBE) [11]. The size of the plane-waves basis set and the number of k-points were used to assess if the total energy was converging. To increase the valence electronic wave functions, a plane-wave basis set with a 500 eV energy cutoff is utilized in each case. A mesh-grid of 14x14x14 for the B2 phase and 15x15x12 for the $L1_0$ phase is used in the construction of the k-point sample of the Brillouin zone according to the Monkhorst-Pack scheme [12]. On the MnPt system, a 2x2x2 supercell with a cubic crystal made up of 16 atoms and a tetragonal crystal made up of 32 atoms was built. A substitutional search engine within the Medea software platform was used to create the ternary $Mn_{50}Pt_{50-x}Ni_x$ alloys, and Pt was swapped out for Ni. The search engine offered the most stable compositions, such as 6.25, 12.50, 18.75, and 25 at. % at the necessary symmetry. The setting of "accurate" was applied throughout the calculation procedure to reduce errors and avoid wrap-arounds. The structure was fully relaxed with regard to volume, shape, and internal atomic locations prior to the determination of the elastic constants until the Hellman-Feynman forces were less than 0.01 eV/Å for the unit cell. The difference in total energy within 1×10^{-6} meV/atom was established as the tolerance for geometry optimization. The maximum strain value was set to 0.005 % while calculating elastic parameters like elastic constants.

3. Results and discussion

3.1. Thermodynamic and magnetic properties

The Wyckoff atomic locations of the B2 MnPt are Mn: 1a (0, 0, 0) and Pt: 1b (1/2, 1/2, 1/2), whereas those of the $L1_0$ MnPt are Mn: 1a (0, 0, 0) and Pt: 1d (0, 1/2, 1/2), which belong to space group P4/mmm. To obtain the ground state structures, spin-polarized calculations are carried out to obtain the total energy of MnPt intermetallic compounds using GGA approximation. Figure 1 (a) and (b) show the calculated heats of formation and the magnetic moments of the $Mn_{50}Pt_{50-x}Ni_x$ alloys along the Bain path. Through the projected heats of formation (ΔH_f), the stability of $Mn_{50}Pt_{50-x}Ni_x$ is explored and can be approximated by:

$$\Delta H_f = E_C - \sum_i x_i E_i \quad (1)$$

where E_C is the determined total energy of the system and E_i is the determined total energy of elements in the system. Generally, the more negative heats of formation is, the better stability is [13]. The calculated heats of formation of the compounds are presented in Figure 1 (a). From Figure 1 (a), it can be noted that the heats of formation of both B2 and $L1_0$ $Mn_{50}Pt_{50}$ alloy are negative, which indicates the possibility for these compounds to form when c/a is less than 1.10. At 6.25 at.% Ni, $L1_0$ phase (c/a=1.10) is thermodynamically more favourable as compared to B2 phase as shown by the more negative heats

of formation, suggesting a higher stability at 0 K. Even though their stability is reduced, the $L1_0$ phase is still thermodynamically favourable compared to B2 phase at 12.50 and 18.75 at. % Ni compositions. Similar behaviour was observed at 25 at. % Ni. The $L1_0$ phase is thermodynamically unstable when c/a is above 1.10 as indicated by positive heat of formation, thus unlikely to form.

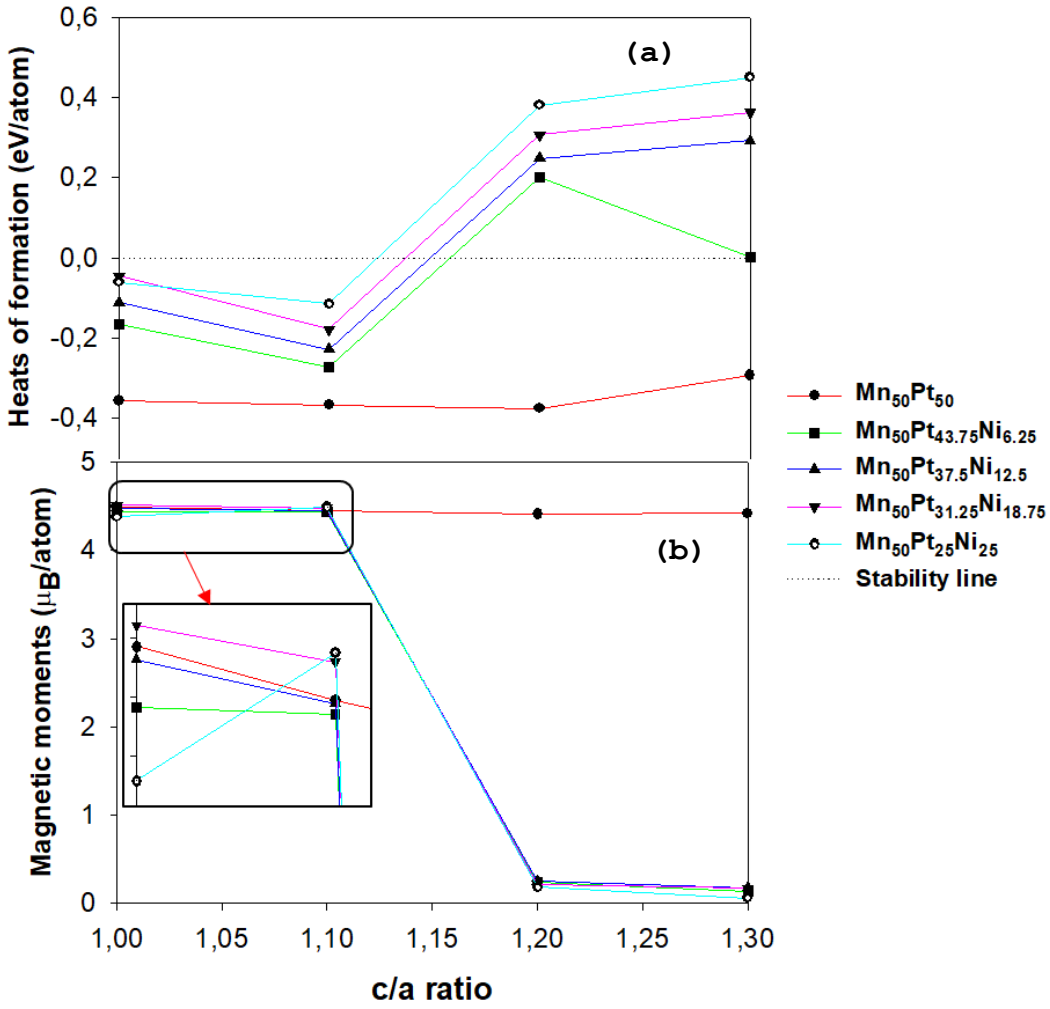


Figure 1. (a) Heats of formation, ΔH_f and (b) magnetic moments versus c/a ratio for $Mn_{50}Pt_{50-x}Ni_x$ alloys (x= 6.25, 12.50, 18.75 and 25).

To find out how strong the structures' magnetic fields are, the magnetic moments were calculated. Figure 1(b) shows that the magnetic moments of the B2 and $L1_0$ $Mn_{50}Pt_{50}$ structures remain constant. When c/a=1.20, the most stable structure is discovered, suggesting that $L1_0$ is more thermodynamically stable than B2 while still keeping the same magnetic moments (See Figure 1(b)). It is noted that the B2 $Mn_{50}Pt_{50-x}Ni_x$ structures shows higher magnetic moments at 18.75 at. % Ni as compared to $Mn_{50}Pt_{50}$. This behavior implies that the B2 phase at 18.75 at. % Ni can benefit from the addition of Ni to increase its magnetic strength. Figure 1 (b) shows that the magnetic moments of the $L1_0$ phase increase above 18.75 at. % Ni when c/a is 1.10. This suggests that the ferromagnetic behaviour is improved at these compositions. Furthermore, it is observed that B2 phase has slightly higher magnetic moments at 18.75 at. % Ni compared to $L1_0$ phase. However, when c/a is above 1.10 the magnetic moment drops gradually and show paramagnetic behaviour for the desired composition range ($0 \leq x \leq 25$). According to the findings, Ni can raise the magnetic moment of $Mn_{50}Pt_{50}$ by 18.75 at. % in both the B2 and $L1_0$ phases.

3.2. Structural properties

The calculated equilibrium lattice parameters for $Mn_{50}Pt_{50-x}Ni_x$ alloys ($0 \leq x \leq 25$) are shown in Figure 2 (a, b). Full geometry optimization was carried out on the structures that are still likely to form when Ni added, on the basis of calculated heats of formation presented in section 3.1. The estimated lattice parameters were compared to the theoretical values to validate a binary model ($Mn_{50}Pt_{50}$). Binary B2-paramagnetic (PM) $Mn_{50}Pt_{50}$ is anticipated to have an equilibrium lattice parameter of 3.00 Å (3.01 Å) [14], which is well in agreement with theoretical data (in parentheses) to within 2%. The ferromagnetic (FM)-B2 and $L1_0$ phase's lattice parameters (c/a ratio=1.10-1.30) have no known theoretical or experimental values. It can be seen that the equilibrium lattice parameters of the B2 $Mn_{50}Pt_{50-x}Ni_x$ reduce as Ni content is increased as shown in Figure 2 (a). Since Pt's atomic radius (1.83 Å) is greater than Ni's (1.49 Å), this is to be expected. With a *c/a* ratio of 1.10, the lattice parameter *a* increases minimally while the lattice parameter *c* decreases for the $L1_0$ phases with Ni addition, for the desired composition range ($0 \leq x \leq 25$). According to Bain path transformation [15], the lattice contracts in direction *a* and expands in direction *c* with Ni introduction.

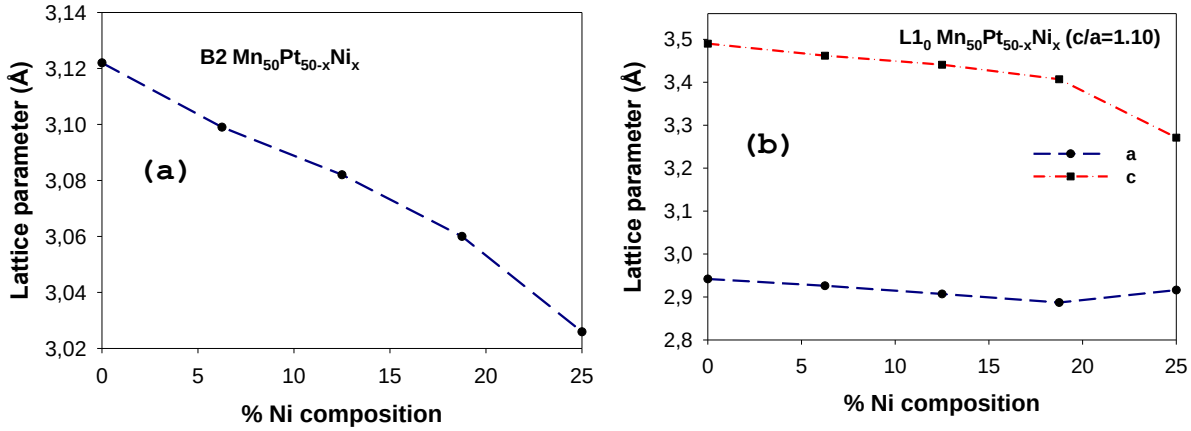


Figure 2. Lattice parameters (a) B2 and (b) $L1_0$ (c/a=1.10) against at. % Ni for $Mn_{50}Pt_{50-x}Ni_x$ alloys (*x*= 6.25, 12.50, 18.75 and 25).

3.3. Elastic properties and mechanical stability

The elastic properties offer useful knowledge about the mechanical behaviour of crystals, the forces at work in solids, and they also offer crucial information for creating interatomic potentials. They also give significant details regarding bonding properties close to equilibrium. Therefore, in order to comprehend the mechanical behaviour of $Mn_{50}Pt_{50-x}Ni_x$ compounds at equilibrium, it is crucial to analyse the elastic constants. Each lattice has its own elastic constants, for example, cubic crystals have three (c_{11}, c_{12}, c_{44}) and tetragonal crystals have six (c_{11}, c_{12}, c_{13}, c_{33}, c_{44}, c_{66}) independently. Additionally, the constraints on the elastic constants are as follows [16] because of the related requirements of mechanical stability in a cubic crystal:

$$c_{44} > 0; c_{11} > c_{12} \text{ and } c_{11} + 2c_{12} > 0, \quad (2)$$

The following formula can be used to compute the shear modulus (C'):

$$C' = \frac{1}{2}(c_{11} - c_{12}) \quad (3)$$

As for tetragonal crystal stability condition is as follows:

$$c_{44} > 0,\ c_{66} > 0,\ c_{11} > |c_{12}| \text{ and } c_{11}+c_{12} - \frac{2c_{13}^2}{c_{33}} > 0, \quad (4)$$

The calculated elastic constants of the B2 and $L1_0$ $Mn_{50}Pt_{50-x}Ni_x$ alloys ($0 \leq x \leq 50$) are shown in Figure 3 (a, b). The stability condition for the elastic constants must be satisfied in order for the structure to be stable. The crystal is mechanically stable if C' is positive ($(1/2(c_{11}-c_{12})>0$); otherwise, it is unstable. The binary B2 $Mn_{50}Pt_{50}$ alloy is shown to have positive C', which contributes to its mechanical stability. As Ni is added, the elastic constant c_{11} becomes less than c_{12}, causing C' to be negative and indicating instability below 12 at. % Ni. All elastic constants for B2 phase met the stability criterion only at 18.75 at. % Ni indicating mechanical stability. Furthermore, the elastic constants of $L1_0$ $Mn_{50}Pt_{50-x}Ni_x$ when the c/a = 1.10 are shown in Figure 3 (b). As the composition of Ni is added, the elastic constants (c_{11}, c_{12}, c_{13}, c_{33}, c_{44}, c_{66}) are all positive, satisfying the mechanical stability condition. It is noted that the C' is positive for the necessary composition range and also satisfies the requirements for tetragonal stability, indicating that the structures are mechanically stable.

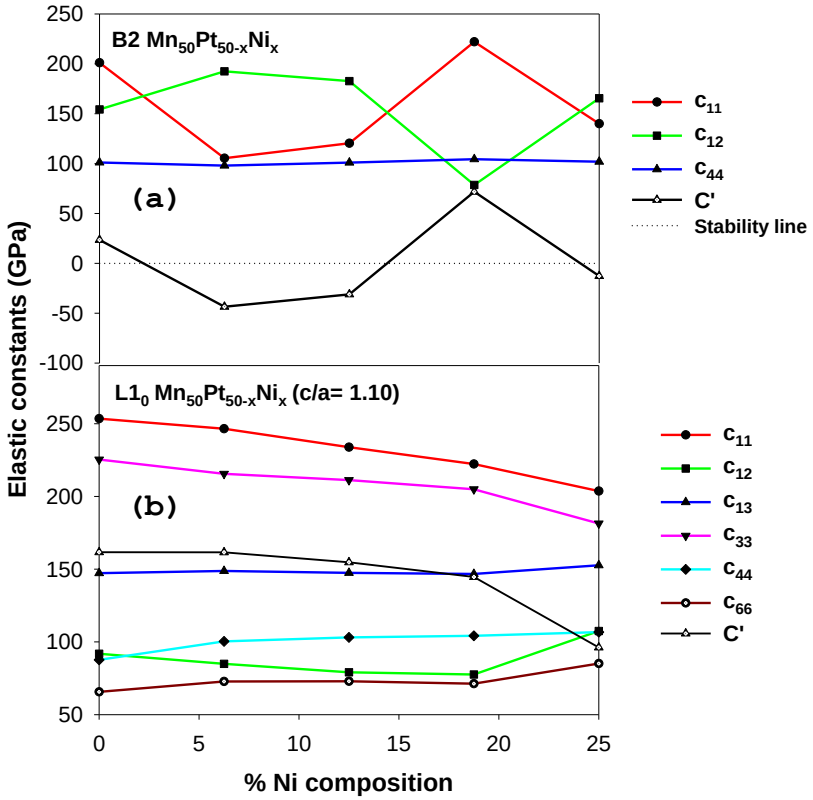


Figure 3. Elastic constants (GPa) of (a) B2 and (b) $L1_0$ (c/a=1.10) for the $Mn_{50}Pt_{50-x}Ni_x$ ($0 \leq x \leq 25$) ternary alloys.

4. Conclusion

First-principles calculations based on DFT are used to examine various $Mn_{50}Pt_{50-x}Ni_x$ ($0 \leq x \leq 25$) composition features for potential spintronic applications. Our findings regarding the lattice parameters for PM B2 $Mn_{50}Pt_{50}$ were discovered to be in excellent agreement with the known theoretical values to within 2%. At $Mn_{50}Pt_{50}$, we found the $L1_0$ phase to be the most thermodynamically stable structure than B2 maintaining the same magnetic moments when *c/a*=1.20. The results suggested that thermodynamic stability can be achieved for both B2 and $L1_0$ $Mn_{50}Pt_{50-x}Ni_x$ alloys (x=0, 6.25, 12.5, 18.75, 25) when *c/a* is less than 1.10 due to negative heat of formation. Furthermore, it was found that the magnetism is improved for B2 $Mn_{50}Pt_{50-x}Ni_x$ (x = 18.75) and $L1_0$ $Mn_{50}Pt_{50-x}Ni_x$ (x = 18.75 and 25) when c/a=1.10. In

addition, B2 $Mn_{50}Pt_{50-x}Ni_x$ at 18.75 at. % Ni satisfies the stability criteria for cubic crystals. It was revealed that increasing Ni could effectively improve the magnetism of the compound when c/a is less than 1.10.

5. Acknowledgments

Computer resources from the Centre for High Performance Computing (CHPC) in Cape Town and the Materials Modelling Centre (MMC) at the University of Limpopo were used to complete the task. The Department of Science and Innovation (DSI), through Mintek, provided financial assistance for the authors' work through the Advanced Metals Initiative (AMI). The Department of Science and Technology's support of the South African Research Chair Initiative is likewise greatly valued.

References

[1] Yang J, Yang W, Shao Z, and Liang D 2018 *Chin. Phys. B* **27** 117503.
[2] Baltz V, Manchon A, Tsoi M, Moriyama T, Ono T, and Tserkovnyak Y 2018 *Rev. Mod. Phys.* **90** 0150.
[3] Farrow R F C, Marks R F, Gider S, Marley A C, Parkin S S P, and Mauri D 1997 *J. Appl. Phys.* **81** 4986.
[4] Lu Z, Chepulskii R V, and Butler W H 2010 *Phys. Rev. B* **81** 094437.
[5] Diale R, Ngoepe P. Moema J. Phasha M, and Chauke H. 2022 *MATEC Web Conf.* **370** 02006.
[6] Diale R G, Ngoepe P E, Moema J S, Phasha M J, Moller H, and Chauke H R 2023 *MRS Adv.* **8** 651-655.
[7] Kresse G and Hafner J 1993 *Phys. Rev.* B **47** 58-561.
[8] Kresse G and Furthmüller J 1996 *Phys. Rev.* B **54** 11169-11186.
[9] Blöchl P E 1994 *Phys. Rev. B* **50** 17953-17979.
[10] Perdew J P, Burke K, and Ernzerhof M 1996 *Phys. Rev. Lett*. **77** 3865.
[11] Kohn W and Sham L J 1965 *Phys. Rev.* **140** 1133-1138.
[12] Monkhorst H J and Pack J D 1996 *Phys. Rev. B* **13** 5188-1592.
[13] Liu Q J, Zhang N C, Liu F S and Liu Z T 2014 *J. Alloy. Compd.* **589** 278.
[14] Wang J, Gao A, Chen W, Zhang X D, Zhou B and Jiang Z 2013 *J. Magn. Magn. Mater.* **333** 93-99.
[15] Bain E C 1924 *Trans. Am. Inst. Min. Metall. Engrs.* **70** 25.
[16] Mehl M J and Klein B M 1994 *Intermetallic Compd.* **1** 1-26.

Numerical Simulations of Defect-Decorated (5,5) Carbon Nanotubes under Pressure

R Warmbier[1] and X Maphisa[2]

[1] School of Physics, University of the Witwatersrand, 1 Jan Smuts Avenue, Johannesburg, South Africa

[2] Department of Physics, University of Johannesburg, Cnr Kingsway and University Road, Auckland Park, South Africa

E-mail: robert.warmbier@wits.ac.za

Abstract. Carbon nanotubes (CNTs) are amongst the strongest materials known with very high tensile strength and Young's modulus. Unfortunately the ideal parameters primarily refer to single nanotubes. In amorphous materials and non-twisted bundles the macroscopic strength is limited by the relatively weak van der Waals interaction between the tubes.

In recent years researchers have tried to modify CNT-based materials in various ways, including irradiation to create defect-decorated CNTs. In this paper we present Density Functional Theory and Molecular Dynamics simulations of defect-decorated CNTs. We show that even small fractions of extra defects in the material can lower the collapse pressures significantly. The extra defects also facilitate deformations into oval and race-track shapes at lower pressures. The interplay between defect-decorations and pressures further allows for local morphological changes, including CNT-interlinking, which do not lessen the strength of CNTs further, but might improve other material properties.

1. Introduction

Carbon nanotubes (CNTs) have received a great deal of interest given their outstanding combination of properties and prospective applications [1]. In particular the combination of high strength and flexibility as well as excellent thermal and electrical conductivity come to mind. Individual single wall carbon nanotubes have shown a Young's modulus of about 1000 GPa and tensile strength above 100 GPa [2, 3].

Exposure to high pressures can change the properties of nanotubes considerably, both reversibly and irreversibly. For example high pressure in CNTs can be used to promote interlinks at the high curvature points resulting from the deformation. It has been observed that sp^3 interlinks induced at low pressure are often reversible when the structure is recovered to ambient conditions [4]. In contrast going to high pressure has been shown to cause irreversible structural damage [4, 5, 6, 7, 8, 9]. While experimental results vary [4], several authors report persistent deformation [7] and broadening of Raman peaks [5, 6], which were associated with structural change, see Loa [4] for a more complete discussion.

In this paper we investigate changes in the electronic structure of (5,5) CNT bundles under isotropic pressure and the possible interlinking associated with these changes.

2. Methodology

The unit cell used in this work consisted of two parallel CNTs in a close pack configuration, which is the smallest possible configuration to observe adequate deformation under pressure. (5,5) single wall CNTs with diameter of about 7 Å were chosen as they are small enough to contain few atoms around the circumference, but are within, or close to, the lower range of diameters found in real samples [10, 11, 12]. With a length of about 20 Å each unit cell had a total of 320 carbon atoms except for cases where vacancy type defects were created. Periodic boundary conditions were used in all directions. Simulation structures were created and modified using the Atomic Simulation Environment (ASE) software [13].

Molecular dynamics simulations (MD) were performed using the open source Large-scale Atomic/Molecular Massively Parallel simulator (LAMMPS) software [14]. The $\text{ReaxFF}_{\text{C-2013}}$ force field was used, as it is known to describe sp^2 as well as sp^3 bonds accurately [15]. MD simulations were performed using a very conservative time step of 0.2 fs. The Nosé-Hoover thermostat and barostat [16, 17, 18], as implemented in LAMMPS, were used for temperature and pressure control. A constant temperature of 100 K was set unless stated otherwise and isotropic pressure was applied. Applying equal pressure in all directions allows the nanotubes to radially collapse as well as to buckle axially. The pressure was increased in intervals of 5 GPa at a rate of 20 ps/GPa until total collapse was observed. At the end of each interval the system was allowed to equilibrate for 10 ps before proceeding.

The DFT calculations were performed using the GPAW software [19, 20], using the default double-ζ polarised (DZP) numerical atomic orbital (NAO) basis. The PBE xc-functional [21, 22] was used with Γ-point only to describe the Brillouin zone. Visualizations of atomic structures were created using ASE [13], Ovito [23] and VESTA [24].

3. Results

3.1. Collapse Pressure

In order to obtain a baseline for the different pressure regimes for the system at hand a pressure series was recorded in steps of 1 GPa. As pressure increases a system undergoes consecutive geometrical transitions, which are marked by an abrupt change in volume. As CNTs are low density structures which collapse under pressure, these changes are quite prominent, as shown in Figure 1. The insets a) to c) highlight the corresponding structural changes at selected points. In addition to the pristine cell, a system with 2 single vacancy defects is shown as well. The defects were placed randomly, and simulations were repeated for several random placements. In the presented case both defects appear on the central CNT, on opposite sides and about 2.5 Å separated along the axis.

The pristine nanotube bundle behaves differently from defect-decorated nanotubes under pressure despite the relatively low defect content of 0.6 % of the defect-decorated nanotubes. As can be expected, the volume of the pristine and defect systems are almost identical before pressurization. Small differences arise as the constant pressure ensemble leads to fluctuations of the cell volume during the MD simulation. The nanotubes change from circular to elliptical shape in the first phase transition, labeled a) in Figure 1, which occurs around 10 GPa for the defect-decorated system and 15 GPa for the pristine system. The pristine system thereafter shows again a linear compression characterized by a reduction in CNT–CNT distance and increase in CNT eccentricity, see b) in Figure 1, before a rapid total collapse at about 25 GPa, c) in Figure 1. The defect-decorated CNT system shows a qualitatively different behaviour in the second phase, b). In addition to the effects discussed, it also shows signs of gradual collapse, with the final collapse, c), at 21 GPa being much less pronounced as a result. The total collapse occurs at different pressures but identical volume in both cases. After the collapse the systems behave similarly. In both cases the CNTs collapse to sheet-like structures, which are barely compressible independent of defect-content. As the pressure in the simulations was also applied

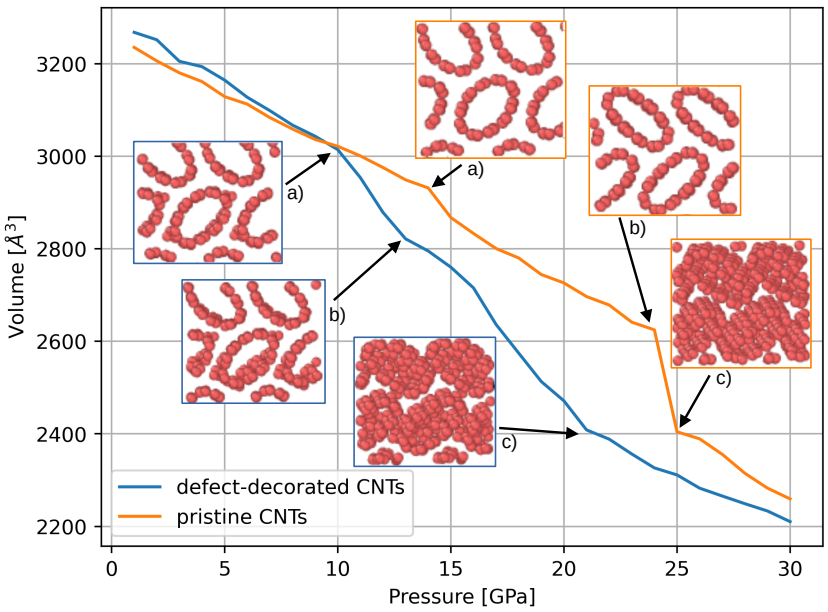


Figure 1. Relationship between volume and pressure applied for (5,5) pristine (light orange) and single vacancy defect (darker blue) CNT bundles. Insets show the structural changes at selected points.

axially, warping along that direction occurs as well. Residual reduction is volume stems primarily from reduction of the separation of the CNT sheets. C–C bond length barely decrease from about 1.44 Å at ambient conditions to 1.39 − 1.40 Å at 20 GPa to no less than 1.38 Å at 30 GPa.

3.2. Band Structures

To gain more insight into the changes in the electronic structure under pressure we selected geometries of the pristine CNT sample from the MD runs at 0 GPa and 25 GPa for band structure calculations, as shown in Figure 2. These structures were first relaxed in DFT, at constant volume, into their equilibria.

At both pressures the central $k_z = 0$ cut, $\Gamma \to X \to S \to Y \to \Gamma$, is showing a gaped behaviour while the corresponding cut through the zone boundary $k_z = \pm\pi$, $Z \to U \to R \to T \to Z$, has bands crossing the Fermi level, which makes these nanotubes metallic.

The bands in the $k_z = 0$ plane show changes related to the deformation of the CNTs, but are qualitatively similar. In the $k_z = \pm\pi$ plane however, the bands are pushed away from the Fermi level under pressure. The total number of bands crossing the Fermi level in this band path is reduced from 6 to 2 as a consequence. While this is not nearly enough to open a band gap in this case, such band repulsion could in principle be used to force a metal to semiconductor transition in selected CNTs under pressure as was predicted by Lammert *et al.* [25]. Such transitions have been observed experimentally using doped nanotubes and without pressure [26].

For completeness and comparison we also show a similar figure for the band structure of a system of defect-decorated CNTs. In this case a bundle with 4 vacancies was chosen. As can be seen in Figure 3 even at ambient pressure the introduction of defects leads to significant changes in the band structure. In general the bands are flatter, pointing at a overall more localized electronic structure. This is not unexpected as the defects disrupt the periodicity of the system. In this case the $k_z = 0$ plane now also shows a metallic characteristic. Under 15 GPa of pressure the band structure flattened further.

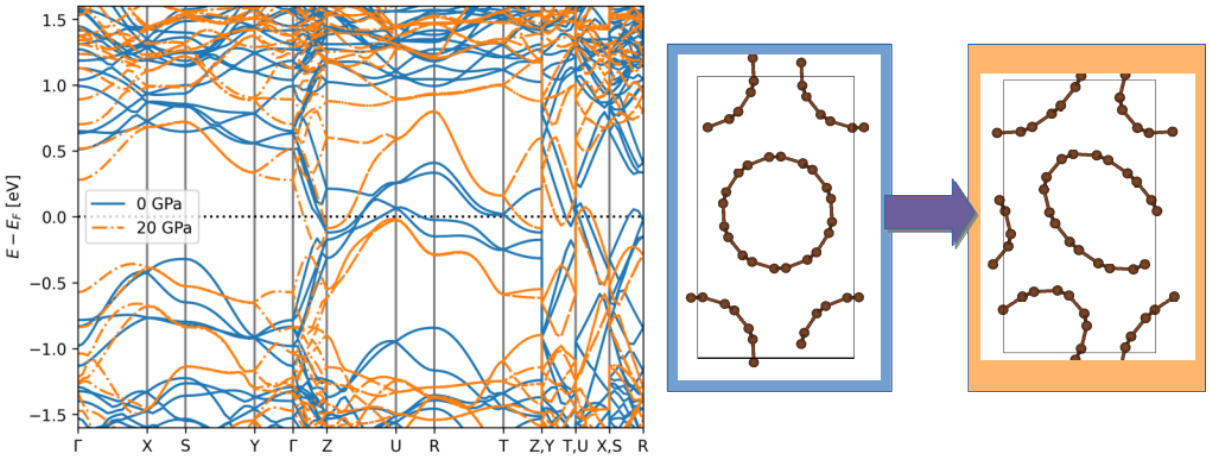


Figure 2. DFT band structure of pristine (5,5) CNT bundles at 0 GPa (blue) and 20 GPa (orange) pressure. The top views of the corresponding geometries are shown for illustration.

3.3. Interaction between Nanotubes

A closer look at the atomic structure shown in Figure 3 reveals points of close proximity between neighboring CNTs under high pressure, especially near the high curvature points, as indicated by the bonds drawn between the nanotubes. These could indicate bonding between these nanotubes, which is of particular interest for the modification of CNT materials to improve their overall strength and thermal conductivity [12]. If the system is recovered to ambient pressure, the CNT–CNT bonds seem to persist as shown in Figure 4. As before, we relaxed the recovered structure in DFT to lower the risk of unphysical results, which can appear in MD simulations as those use semi-empirical force fields of difficult to ascertain reliability. Another example of potential CNT–CNT interlinking is shown in Figure 5. Here a pristine bundle was simulated at 500 K in MD and pressurized up to 30 GPa. The bundle was recovered to ambient pressure after it suffered irreversible changes. A number of CNT–CNT interlinks are retained,

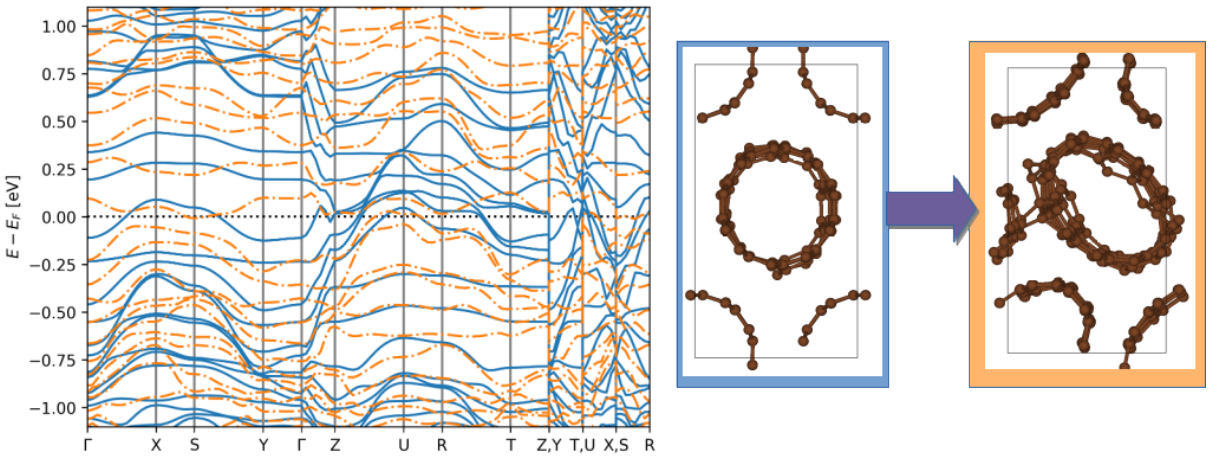


Figure 3. DFT band structure of defect decorated (5,5) CNT bundles at 0 GPa (blue) and 15 GPa (orange) pressure. The top views of the corresponding geometries are shown for illustration.

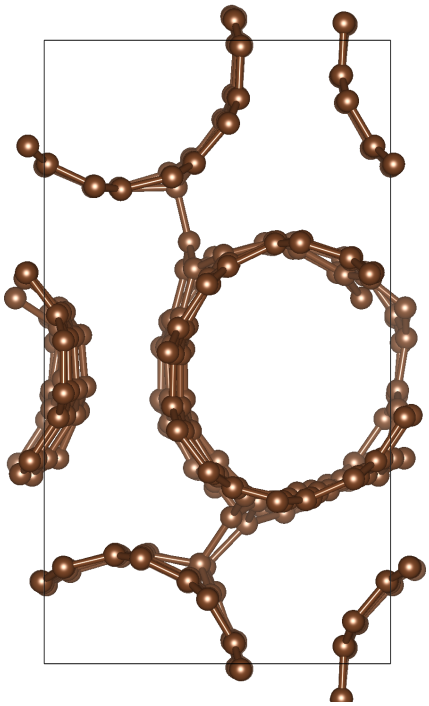


Figure 4. Structure of defect-decorated CNT from Figure 3 recovered to ambient pressure in MD and then relaxed in DFT at constant volume.

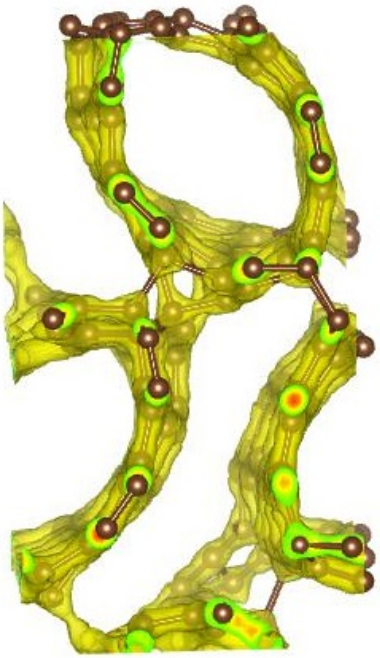


Figure 5. Electron density iso-surface plot of pristine CNT bundle obtained after pressure is removed from 30 GPa at 500 K.

but the mechanical properties of this sample might be compromised by the damage.

4. Conclusions

In this paper we presented a collapse pressure study of a closed packed (5,5) CNT bundle, which follows the known circle, ellipse, race-track, collapse sequence. We also show that the introduction of even small vacancy defects reduces the collapse pressure significantly, as they act as structural weak points. In accordance with these findings, the introduction of vacancies leads a partial localisation of many bands around the Fermi level. A combination of pressure and defect-decoration might also lead to a substantial amount of nanotube interlinking.

To further understand the exact physical processes in these nanotubes system the reconstruction of vacancy defects in different temperature and pressure regimes need to be investigated.

Acknowledgments

The authors would like to thank the DSI-NRF Centre of Excellence in Strong Materials (CoE-SM) for support. RW would like to thank the Mandelstam Institute for Theoretical Physics for support.

References

[1] Ajayan P M and Zhou O Z 2001 *Carbon Nanotubes* 391–425
[2] Salvetat J P, Briggs G A D, Bonard J M, Bacsa R R, Kulik A J, Stöckli T, Burnham N A and Forró L 1999 *Physical Review Letters* **82** 944
[3] Lu J P 1997 *Physical Review Letters* **79** 1297
[4] Loa I 2003 *Journal of Raman Spectroscopy* **34** 611–627

[5] Sood A K, Teresdesai P V, Muthu D V S, Sen R, Govindaraj A and Rao C N R 1999 *physica status solidi (b)* **215** 393–401
[6] Teredesai P, Sood A, Sharma S, Karmakar S, Sikka S, Govindaraj A and Rao C 2001 *physica status solidi (b)* **223** 479–487
[7] Obraztsova E 2001 *In situ Raman Investigations of Single-Wall Carbon Nanotubes Pressurized in Diamond Anvil Cell* (Dordrecht: Springer Netherlands) pp 473–482 ISBN 978-94-010-0520-3
[8] Jensen B D, Kim J W, Sauti G, Wise K E, Dong L, Wadley H N, Park J G, Liang R and Siochi E J 2020 *Carbon* **156** 538–548
[9] Sakurai M and Saito S 2011 *Physica E* **43** 673–676
[10] Adam T and Hashim U 2014 *Current nanoscience* **10** 695–699
[11] Yonezawa S, Chiba T, Seki Y and Takashiri M 2021 *Scientific Reports* **11** 5758
[12] Hearne G, Kapesi L, Erasmus R, Naidoo S and Warmbier R 2021 *Physical Review Materials* **5** 033607
[13] Larsen A H, Mortensen J J, Blomqvist J, Castelli I E, Christensen R, Dułak M, Friis J, Groves M N, Hammer B, Hargus C *et al.* 2017 *Journal of Physics: Condensed Matter* **29** 273002
[14] Plimpton S 1995 *Journal of Computational Physics* **117** 1–19
[15] Srinivasan S G, Van Duin A C and Ganesh P 2015 *The Journal of Physical Chemistry A* **119** 571–580
[16] Nosé S 1984 *Molecular Physics* **52** 255–268
[17] Hoover W G 1985 *Physical Review A* **31** 1695
[18] Hünenberger P H 2005 *Thermostat Algorithms for Molecular Dynamics Simulations* (Berlin, Heidelberg: Springer Berlin Heidelberg) pp 105–149 ISBN 978-3-540-31558-2
[19] Mortensen J J, Hansen L B and Jacobsen K W 2005 *Physical Review B* **71** 035109
[20] Enkovaara J, Rostgaard C, Mortensen J J, Chen J, Dułak M, Ferrighi L, Gavnholt J, Glinsvad C, Haikola V, Hansen H *et al.* 2010 *Journal of Physics: Condensed Matter* **22** 253202
[21] Perdew J P, Burke K and Ernzerhof M 1996 *Physical Review Letters* **77** 3865–3868
[22] Perdew J P, Burke K and Ernzerhof M 1997 *Physical Review Letters* **78** 1396–1396
[23] Stukowski A 2009 *Modelling and Simulation in Materials Science and Engineering* **18** 015012
[24] Momma K and Izumi F 2011 *Journal of Applied Crystallography* **44** 1272–1276
[25] Lammert P E, Zhang P and Crespi V H 2000 *Phys. Rev. Lett.* **84**(11) 2453–2456
[26] Vavro J, Kikkawa J M and Fischer J E 2005 *Phys. Rev. B* **71**(15) 155410

Optimization of silver decorated functionalized multi-walled carbon nanotubes nanofluids for improved heat transfer.

HB Masolo[1], SS Nkosi[1], DE Motaung[3] and M Diale[2]

[1]Department physics, University of Limpopo, Private Bag X1106, Sovenga 0727, South Africa
[2]Department of Physics, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa
[3]Department of physics, University of Free state, PO Box 339, Bloemfontein 9300, South Africa

E-mail: **beautymasolo1@gmail.com**

Abstract. Carbon nanotubes (CNTs) are excellent for heat transfer in nanofluids due to their outstanding material properties, thermal conductivity of (9932,43 W/mK) and specific heat capacity of (1329,83 W/m^{2}K). However, the fundamental chemical structure of the material tends to aggregate, limiting their applications. In this study, surface modifications were made to overcome the agglomeration and increase the dispersibility of CNTs, leading to enhanced interactions of the functionalized and silver decorated CNTs with polymer matrix materials. X-ray diffraction (XRD), high-resolution transmission electron microscopy (HR-TEM), Raman spectroscopy and Fourier-transform infrared spectroscopy (FTIR) were used to characterize the synthesized silver nanoparticle (Ag-NP), decorated functionalized multi-walled carbon nanotubes (f-MWCNTs) and hybrid nanocomposite (Ag/f-MWCNTs). Raman spectroscopy analysis showed an increase in the defects to Graphitic ratio of Ag/f-MWCNTs nanocomposites when compared to that of f-MWCNTs and pristine multi-walled carbon nanotubes (p-MWCNTs). It also showed the maintenance of the graphitic structure of the material after functionalization as confirmed by XRD analysis, revealing MWCNT-specific Bragg diffraction peaks, graphite crystalline phases, and Ag NPs in f-MWCNT. The improvement in dispersibility and thermal conductivity was achieved by Ag/f-MWCNTs nanocomposite in the synthesis of nanofluids for heat transfer.

1. Introduction

Carbon nanotubes, also known as buck tubes were discovered by Iijima in 1991[1]. These are cylindrical in shape and made from a single layer of carbon atoms producing single-walled carbon nanotubes (SWCNTs) with a diameter of less than 1 nanometer-sized(nm) and multi-walled carbon nanotubes (MWCNTs) with up to 100 nm made up of multi concentrically interconnected nanotubes. They are of particular interest because they can be utilized as advancement for heat transfer [2].

MWCNTs are among the most exciting materials due to their remarkable mechanical, electrical, chemical and optical properties such as high elastic modulus, high charge carrier mobility absorption and photoluminescence [3]. The material is therefore promising for numerous applications such as heat exchangers, boilers, and solar collectors [4]. However, several aspects such as solubility, dispersibility, processability and agglomeration due to their long-tangled structure still limit their applications. Kuila el at. [5] conducted surface modification enhanced by chemical functionalization increasing surface dispersibility and chemical compatibility. Chemical functionalization generates functional groups on MWCNT sidewalls. The most often used reagents for oxidation treatment are HNO_3, H_2O_2, H_2SO_4, and $KMnO_4$, or a combination of them [6, 7]. As a result, the surface of MWCNTs will acquire several functional groups, such as carboxylic acid (-COOH) and amine (-NH2), while the raw MWCNTs are simultaneously debundled and shortened [8]. Additionally, functionalized MWCNTs have more surface area and defects, which enables the coupling of other materials with them [9].

Decoration of f-MWCNTs with attractive noble metal NPs such as palladium (Pd), gold (Au), platinum (Pt) and silver (Ag) are gaining interest due to their potential use in advanced nanocomposites [10, 11].

Enhancing heat transmission and improving performance are two primary requirements in industrial equipment. Heat exchangers, boilers and solar collectors are among the applications that might result in energy savings, speed up processing, increase the thermal system rating, and extend equipment life [12]. The additions of NPs into based fluids form nanofluids, which can improve the heat transfer performance of the base fluids considerably affecting the thermal conductivity and viscosity of the base fluid. NPs can interact with liquids at the molecular level and transfer heat more effectively than base fluids because of their small size (often less than 100 nm) in a liquid mixture [13]. Carbon-based nanofluids have been observed to exhibit improved thermophysical properties such as heat transfer coefficients, thermal conductivity, viscosity and thermal diffusivity compared to those of base fluids [14].

In this study, surface modifications were made to overcome the agglomeration and increase the use of Ag/f-MWCNTs in heat transfer due to their remarkably high thermal conductivity of more than 3000 W/mK at room temperature [15]. Both deionized and ethylene glycol were utilized as base fluids in this investigation to explore the impact of base fluids. The results show a great improvement in the dispersibility and thermal conductivity of Ag/f-MWCNTs.

2. Experimental procedure

2.1. Synthesis of f-MCNTs and Ag/f-MWCNTs 414

All the utilized chemicals were of analytical grade and employed directly out of the package. Sulfuric acid (H_2SO_4), Nitric acid (HNO_3), Ethylene glycol ($CH_2OH)_2$, Polyvinylpyrrolidone (PVP) and MWCNTs were obtained from Sigma Aldrich. Glassware used in the tests were carefully cleaned, rinsed with deionized water (di-H_2O), and dried. Synthesis involves dispersing NPs in a 3:1 mixture of ethylene glycol and di-H_2O.

2.1.1. f-MWCNTs by acid cutting method.

Firstly, 10 mL of a 3:1 solution of concentrated H_2SO_4 and HNO_3 was used to treat 0.01 g of raw MWCNTs. The mixture was subjected to stirring for 8 hours at room temperature and then sonicated for 3 hours in an ultrasonic bath. Carboxyl acid groups were added to MWCNTs. When the pH of the mixture reached 7, the resulting solution was filtered, rinsed, and diluted with di-H_2O. The produced f-MWCNTs were vacuum-dried for 6 hours at 40 °C.

2.1.2. Wet impregnation.

Wet impregnation and ultrasonication were used to decorate f-MWCNTs with Ag NPs. 100 mL of ethanol and 1.0 g of f-MWCNTs were mixed, treated for 30 minutes, and gradually added 0.05 M AgNO3 to the ethanol. When the desired ratio of Ag and di-H_20 was achieved, ethanol evaporated by drying the precursor mixture at 40 °C for 6 hours.

2.1.3. f-MWCNTs and Ag/f-MWCNTs nanofluids synthesis by a two-step method.

The synthesis of f-MWCNTs and Ag/f-MWCNTs was achieved by a two-step method where NPs were dissolved in base fluid with surfactants added to enhance stability, modify viscosity, and increase thermal conductivity.

2.2. Characterization.

The structural properties of the prepared MWCNTs, f-MWCNTs and Ag/f-MWCNTs NPs were investigated using XRD measurements. The Raman spectra of the NPs were analyzed with a laser of 532 nm at a laser power of 1,45 mW. The element composition and morphology of the NPs were investigated by energy-dispersive X-ray spectroscopy (EDS) and field emission-scanning electron microscopy (FE-SEM) systems, which were coupled using the Zeiss Ultrafast 540 instrument. MWCNTs, Ag-NPs, and Ag/f-MWCNTs nanocomposites shape and size distribution were investigated by high-resolution transmission electron microscopy (TEM, JEOL-JEM 1010) at a magnification of 1.0

M. Fourie-transform Infrared spectrometer was employed to study the optical properties of the MWCNTs.

3. Results and discussion

3.1. Structural properties

XRD patterns shown in Figure 1, validated the chemical and structural integrity of the produced f-MWCNTs and Ag/f- MWCNTs. MWCNT samples showed a specific Bragg diffraction peak at 30° and 50° assigned to the (002) and (200) of the graphite crystalline phase. The peak exhibited an upshift as compared to the typical graphite peak at 2Θ = 26.5°, explained by a reduction in the spacing between the sp^2, C = C layers [16]. The presence of Ag in f-MWCNT showed the face-centered cubic (fcc) phase represented by the other four observed peaks at 2Θ = 32,8°, 37,6°, 53,4°, 58,2°, and 62,9°, respectively. Ag is indexed as (002), (111), (200), (220), and (311), respectively.

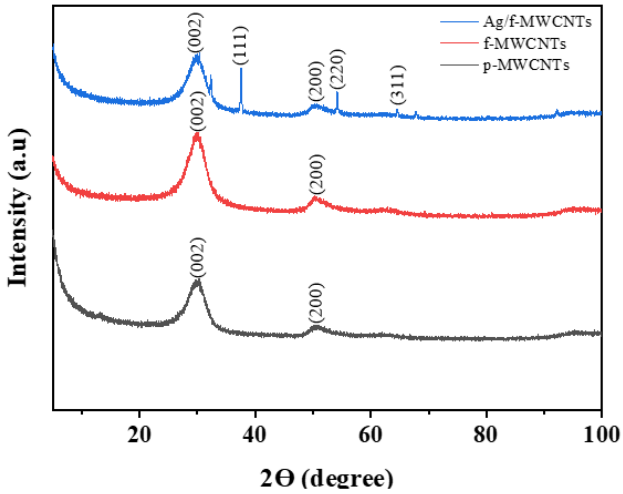


Figure 1. XRD for p-MWCNTs, f-MWCNTs, and Ag/f-MWCNTs NPs

3.2. Surface morphology properties

Figure 2(a) shows an HR-TEM image of p-MWCNTs indicating large particle length and smooth surface with visible walls of the material. HR-TEM image of f-MWCNTs nanocomposites shown in Figure 2(b) have the roughness on the surface of the material and reduced size of the particle. Ag/f-MWCNTs studied in Figure 2(c) show an HR-TEM image of Ag/f-MWCNTs, the walls of the material are no longer visible due to the presence of Ag NPs grafted on the surface.

Figure 2(d) shows SEM and EDS images of MWCNT structure defects, carbon, and oxygen content. Figure 2(e) shows EDS for f-MWCNTs with sulfur, carbon, and oxygen due to particle compactness. Furthermore, Figure 2(f) EDS spectra confirms Ag NPs on the material surface with shortened particle lengths. SEM images show the breaking of large particles into nanotubes, shortening MWCNTs, and Ag NPs, while EDS shows chemical composition before and after oxidation.

3.3. Optical properties

The Raman spectra of p-MWCNTs, f-MWCNTs, and Ag/f-MWCNTs are displayed in Figure 3(a) showing two bands: the graphitic band (G-band) and the structural defects band (D-band). All samples were observed to have the D and G bands assigned to the peak at 1340 cm^{-1} and 1597 cm^{-1} respectively. Ag/f-MWCNTs have a higher D-band intensity peak, indicating high defects in their structure. This increases the surface area of Ag/f-MWCNTs, enabling them to form couples with other materials.

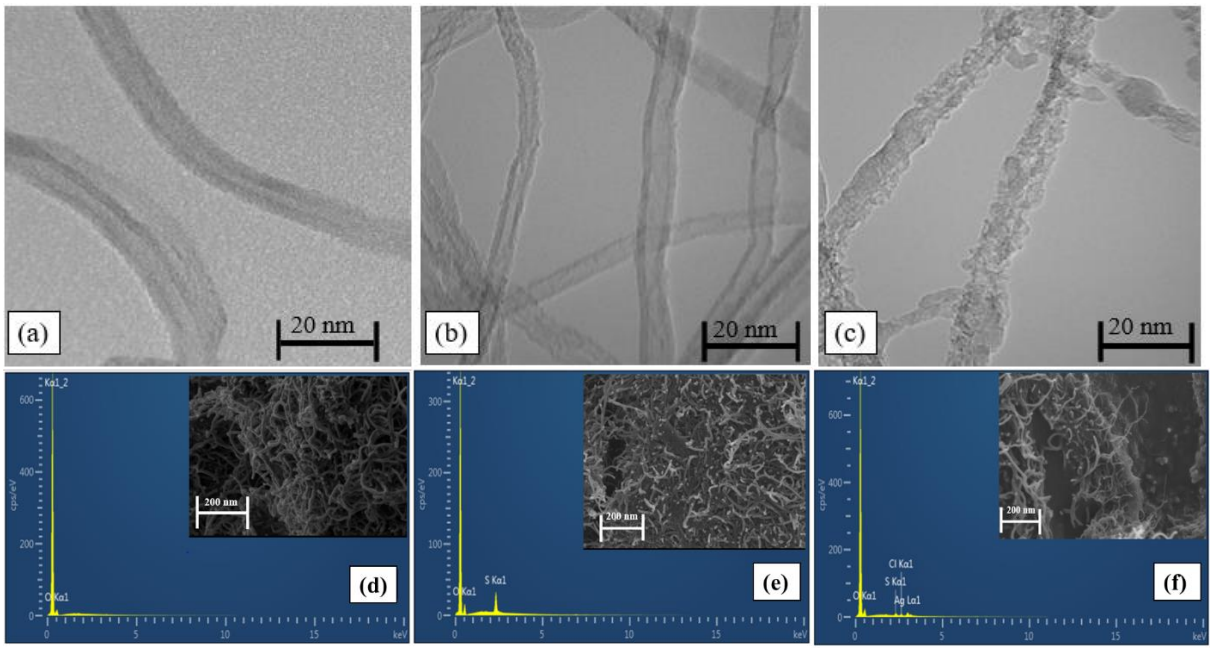


Figure 2. HR-TEM images (a) p-MWCNTs, (b) f-MWCNTs and (c) Ag/f-MWCNTs. SEM images of (d) p-MWCNTs, (e) f-MWCNTs, and f) Ag/f-MWCNTs.

The FTIR spectra shown in Figure 3(b) provide information about the functionalization of MWCNTs and their interaction with metal cations. Herein, p-MWCNT lacks functional groups, reducing complex formation, and MWCNT functionalization allows complex formation with different metal cations. This is reflected in the spectrum, where no distinctive peaks related to functional groups are observed. The peak observed at 1110 cm-1 corresponds to the carbon nanotube's stretching vibration of C-C bonds. In addition, both f-MWCNTs and Ag/f-MWCNT exhibited characteristic absorption peaks at 1640 cm-1 and 1220 cm-1 assigned to the C=O stretching vibration. Furthermore, the presence of free hydroxyl groups was indicated by the absorption peak at 3435 cm-1 in the spectrum of Ag/f-MWCNTs. The existence of a small peak at 612 cm-1 can be assigned to the deformations in the metal-oxygen bonds within Ag-O in the Ag/f-MWCNTs spectra[17]. The FTIR spectrum confirmed the successful functionalization of MWCNTs and demonstrates the presence of functional groups that enable complex formation with different metal cations.

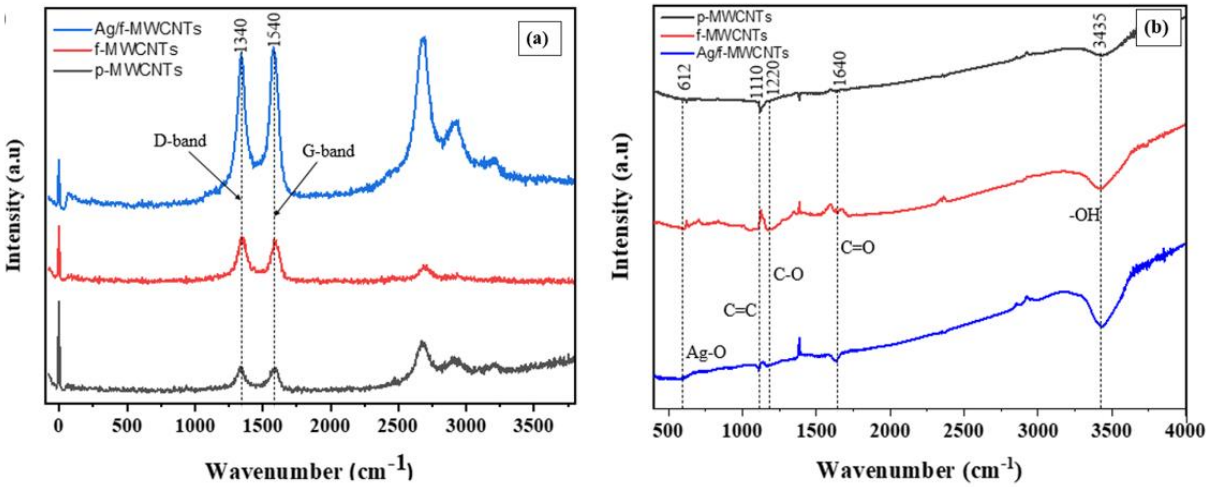


Figure 3. (a) Raman and (b) FTIR spectra of f-MWCNTs, p-MWCNTs and Ag/f-MWCNTs.

3.4. Thermophysical properties

The specific heat capacity of the listed materials in Table 1 was calculated using the gradient from the graph shown in Figure 4(a). The specific heat capacity (C) can be determined using the equation [18]:

$$C_{sample}=\frac{1}{gradient}\times\frac{1}{mass} \quad (1)$$

Based on the calculations, it is observed that the specific heat capacity of carbon-based nanofluids is lower compared to that of base fluids. This implies that a smaller amount of energy is required to transfer heat in carbon-based nanofluids. In particular, the specific heat capacity of Ag/f-MWCNTs nanofluid was found to be 1329.89 W/m^{2}K. This indicates that the energy required to transfer heat in this nanofluid is lower compared to the other fluids. As a result, using Ag/f-MWCNTs as a nanofluid can be beneficial for heat transfer applications due to their lower specific heat capacity. This characteristic can be advantageous for various heat transfer processes and applications.

The thermal conductivity measurements presented in Table 1 provide information about the heat transfer properties of different nanofluids, including p-MWCNTs, f-MWCNTs and Ag/f-MWCNTs nanofluids. The thermal conductivity of a nanofluid describes its ability to transfer heat, the higher the thermal conductivity value, the more effectively the material can transfer heat. The thermal conductivity of the nanofluid is inversely proportional to the gradient of the temperature-time response shown in Figure 4(b). Based on the measurements, Ag/f-MWCNT nanofluids exhibit the highest thermal conductivity of 9932.43 W/mK. This indicates that these nanofluids are superior in terms of heat transfer efficiency compared to the other fluids (f-MWCNT fluid and p-MWCNT fluid) in the study.

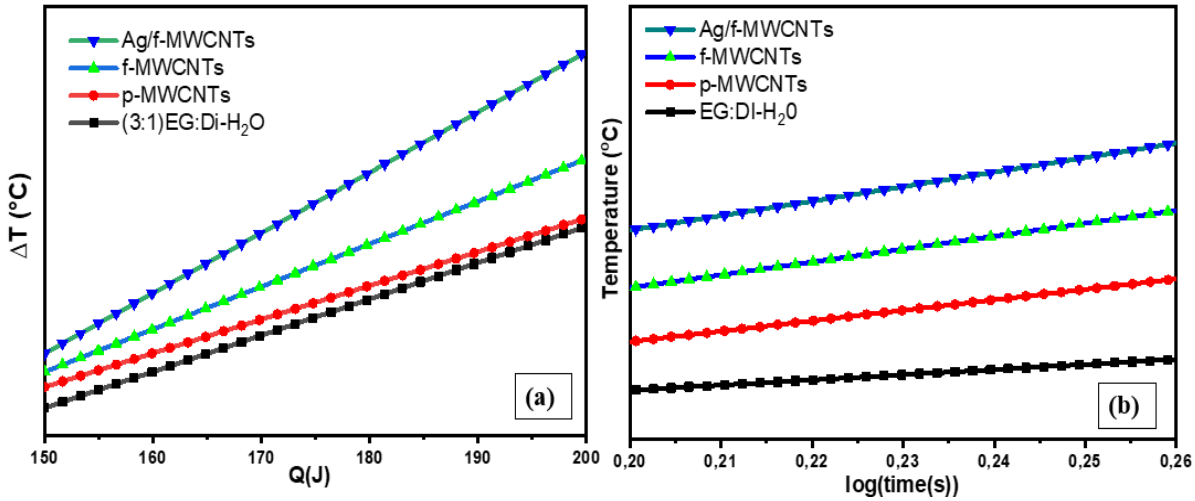


Figure 4. (a) Graph of ΔT as the material has been heated in relation to time interval and (b) Plot of temperature in relation to the natural log of time

Table 1. Heat measurements

Material	Specific heat capacity (W/m^{2}K)	Thermal conductivity(W/mK).
EG: Di-H_2O	3803,94	1428,57
p-MWCNTs	3432,61	3486,75
f-MWVNTs	1938,36	6423,76
Ag/f-MWCNTs	1329,83	9932,43

4. Conclusion and future work

Ag/f-MWCNTs nanofluids were successfully synthesized via acid cutting, wet impregnation, and a two-step method without degradation of the MWCNTs. The method involved the use of strong acids for oxidation and surfactants, which are suitable for large-scale production. The surface chemistry examined by Raman and FTIR spectroscopy analyses confirmed the formation of Ag/f-MWCNTs NPs.

The FTIR spectrum confirmed the attachment of oxygenated functional groups by appearing essential bands, such as C=O (from the carboxyl group), O-H, C-O, and C=C. The XRD, HR-TEM, SEM, and EDS confirmed the crystalline character of the Ag/f-MWCNTs and the presence of Ag NPs. The thermal conductivity measurements for Ag/f-MWCNTs nanofluids showed an enhancement in heat transfer with a low specific heat capacity of 1329,89 W/m^2K and a thermal conductivity of 9932,43 W/mK both contributed by enhanced stability and uniform dispersion of NPs.

5. Acknowledgement

The authors thank the Department of Science and Innovation (DSI) and Leap-Re for financial support, the University of Limpopo, University of Pretoria and University of the Free state for the program and SACHI (UID: 115463) are also gratefully acknowledged.

References

[1] Iijima S 1991 Helical microtubules of graphitic carbon. *Nature.* **354** 56
[2] Dresselhaus M, et al 2004 Electronic, thermal and mechanical properties of carbon nanotubes. *Philos. Trans. Royal Soc. A.* **362** 2065
[3] Venkataraman A, Amadi E V, Chen Y and Papadopoulos C 2019 Carbon nanotube assembly and integration for applications. *Nanoscale research letters*, **14** 1
[4] Kalogirou S A 2004 Solar thermal collectors and applications. *Progress in energy and combustion science* **30(3)** 231
[5] Kuila T, Bose S, Mishra A K, Khanra P, Kim N H and Lee J H 2012. Chemical functionalization of graphene and its applications. *Progress in Materials Science*, **57(7)** 1061
[6] Hou P X, Liu C and Cheng H M 2008 Purification of carbon nanotubes. *carbon*. **46** 2003
[7] Gojny F H, et al 2003 Surface modified multi-walled carbon nanotubes in CNT/epoxy-composites. *Chem. Phys. lett*. **370** 820
[8] Saeb M R, et al 2013 Cure kinetics of epoxy nanocomposites affected by MWCNTs functionalization: a review. *Sci. World J.* **2013**
[9] Likodimos V, et al 2014 Controlled surface functionalization of multiwall carbon nanotubes by HNO3 hydrothermal oxidation. *J. Carbon.* **69** 311
[10] Wildgoose G G, Banks C E, and Compton R G 2006 Metal nanoparticles and related materials supported on carbon nanotubes: methods and applications. *Small*. **2** 182
[11] Ahmadpoor F, Zebarjad S M and Janghorban K 2013 Decoration of multi-walled carbon nanotubes with Ag nanoparticles and investigation on its colloid stability. *Materials Chemistry and Physics* **139** 113
[12] Kumar L, Hasanuzzaman M, and Rahim N A 2019. Global advancement of solar thermal energy technologies for industrial process heat and its future prospects: A review. *Energy Conversion and Management.* **195** 885
[13] Said Z, Sundar L S, Tiwari A K, Ali H M, Sheikholeslami M, Bellos E and Babar H 2022 Recent advances on the fundamental physical phenomena behind stability, dynamic motion, thermophysical properties, heat transport, applications, and challenges of nanofluids. *Physics Reports.* **946** 1
[14] Banisharif A, Estelle P, Rashidi A, Van Vaerenbergh S and Aghajani M 2021 Heat transfer properties of metal, metal oxides, and carbon water-based nanofluids in the ethanol condensation process. *Colloids and Surf. A Physicochem. Eng. Asp.* **622** 126720
[15] Mingzheng Z, et al 2012 Analysis of factors influencing thermal conductivity and viscosity in different kinds of surfactant solutions. *Experimental Thermal and Fluid Science*. **36** 22
[16] Stobinski L, et al 2014 Graphene oxide and reduced graphene oxide studied by the XRD, TEM and electron spectroscopy methods. *J. Electron Spectrosc. Relat. Phenom*. **195** 145
[17] Markova-Deneva I 2010 Infrared spectroscopy investigation of metallic nanoparticles based on copper, cobalt, and nickel synthesized through borohydride reduction method. *J. Chem. Technol. Metall*. **45** 351
[18] Greenbowe T and Meltzer D 2003 Student learning of thermochemical concepts in the context of solution calorimetry. *Int. J. Sci. Educ*. **25** 779

Structural, Thermodynamic and Electronic Properties of Calcium Manganese Oxide ($CaMn_2O_4$) Polymorphs: A First-Principles Study

Oswald Roberts and Ndanduleni Lethole
University of Fort Hare P Bag X 1314 Alice, 5700 South Africa

E-mail: **jessicaoswaldxs15@gmail.com, nlethole@ufh.ac.za**

Abstract. This study presents first principles studies on the structural, thermodynamics, and electronic properties of Pbca-$CaMn_2O_4$ and Pbcm-$CaMn_2O_4$ polymorphs for potential application as cathode materials in Ca-ion rechargeable batteries. All calculations were performed using the CASTEP code as embedded in Material Studio. We employed the density functional theory (DFT) and the generalized gradient approximation (GGA) of Perdew-Burke-Ernzerhof (PBE) for the exchange correlation functional. Since we were dealing with transition metal oxides, we employed the Hubbard U-correction to take care of the highly correlated electrons in the *d*-orbital of the Mn atoms. The calculated lattice parameters were more than 90 % in agreement with available experimental results, thus validating the approach employed. The heats of formation were negative, suggesting that the structures can be synthesized experimentally. During the analysis of the density of states, the considered $CaMn_2O_4$ polymorphs were characterized by a shallow pseudo-gap and hence can be classified as semiconductors.

1. Introduction

For years, we have relied heavily on fossil fuel for energy supply and gasoline-powered models for transportation [1, 2], which are two of the leading causes of global warming [3, 4]. Hence, alternative energy sources are inevitable to avert global warming and environmental challenges. Renewable energy sources such as Photovoltaic solar systems and wind energy are some of the currently most promising energy supplies [5, 6]. However, energy storage devices are essential for the current renewables to be sustainable and reliable [7]. Energy storage devices characterized by inherent high energy density, appropriate operating voltage, environmental benignity, abundance, and cost-effectiveness will increase the implementation of renewables as an alternative to the current resources [6, 8, 9]. Rechargeable lithium-ion batteries (LIBs) have been used as energy storage devices for years due to their high energy density and other desirable properties [7, 10]. However, its single charge (Li^+) brings about minimal energy density, raising concerns about the application of these batteries in large-scale industries [6]. Moreover, the projected high demand and extensive use of LIBs may lead to the exhaustion of resources and, consequently, increased costs and jeopardize sustainability [9]. Multivalent (MV) battery chemistries such as Mg^{2+}, Zn^{2+}, and Ca^{2+} ions present a more significant potential for future applications in battery storage [11]. Multivalent battery chemistries are of great interest since the multivalent ion insertion/extraction is associated with double/triple electron transfer per ion in the intercalation reaction, leading to higher specific energy density and volumetric power than monovalent ions [12]. Calcium ion battery (CIB) chemistry is particularly promising due to the high abundance of calcium on the earth's

crest, relatively lower Ca^{2+} charge-to-radius ratio, and high capacity (1340 $mAh.g^{-1}$) of calcium anodes [13]. Finding a suitable cathode to intercalate a multivalent ion at high voltage and high capacity has been a significant challenge for the development of CIBs [8, 12, 14]. Materials used in batteries are complex and require deep investigation to understand their complexity. The charge and discharge processes include electrical and chemical processes; hence, it is essential to extensively investigate the properties of the suitable cathode [15].

Various polymorphs of $CaMn_2O_4$ (Marokite, spinel, and $CaFe_2O_4$-type) have shown promise to initiate the quest for cathode materials for Ca ion batteries [13]. They are characterized by the edge-sharing chains of MnO_6 octahedra forming double-rutile chains connected by edge-sharing and corner-sharing oxygens, which cause tunnels in the framework, providing 8-coordinated sites for large calcium atoms to occupy and diffuse as seen in Figure 1 [13, 16]. The current study presents the preliminary first principles studies on the structural, thermodynamic, and electronic properties of Pbca-$CaMn_2O_4$ and Pbcm-$CaMn_2O_4$ polymorphs at ambient conditions to mimic stability (Pbca and Pbcm refers to the space group of each polymorph). We have particularly calculated the lattice parameters, heats of formation, and density of states.

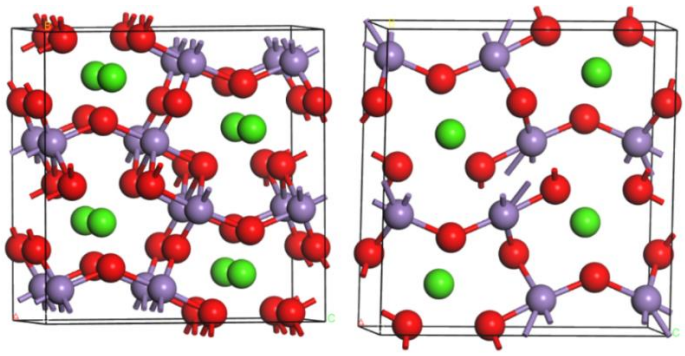


Figure 1. Schematic representation of atomic coordination in the orthorhombic (a) Pbca-$CaMn_2O_4$ and (b) Pbcm-$CaMn_2O_4$ systems.

2. Methodology

The study performed the spin-polarized density functional theory (DFT) calculations using the CAmbridge Serial Total Energy Package (CASTEP) code embedded in the Materials Studio software platform [17]. Generalized gradient approximation (GGA) of Perdew−Burke−Ernzerhof (PBE) were chosen for the exchange-correlation functional [18]. The on-the-fly generated (OTFG) ultrasoft pseudopotentials were adopted for all calculations. The spin polarization was set at collinear, and for the relativistic treatment, the Koelling-Harmon scheme was employed [19]. A single point numerical energy convergence test calculations on both Pbca-$CaMn_2O_4$ and Pbcm-$CaMn_2O_4$ polymorphs were performed to determine the plane wave cut-off energies for the electronic wave functions and appropriate Monkhorst-Pack [20] gamma-centered k-point mesh suitable for Brillouin zone sampling. The plane wave cut-off energy of 500eV for both structures and k-mesh points of $5 \times 4 \times 4$ and $9 \times 3 \times 3$ for Pbca-$CaMn_2O_4$ and Pbcm-$CaMn_2O_4$ respectively were found sufficient to converge the total energy to within 0.1 meV. Since $CaMn_2O_4$ has Mn atoms, the Hubbard ($U - J = 4\ eV$) parameter was used to deal with the strongly correlated *d* orbitals of Mn ions [21, 22]. A study by Arroyo-de Dompablo, et al reported that a U value of 4 eV gives a good prediction of the electronic structure of the $CaMn_2O_4$ polymorphs [13].

3. Results and Discussions

3.1. Structural and Thermodynamics Properties

Table 1 presents the equilibrium structural lattice parameters and heats of formation for Pbca-$CaMn_2O_4$ and Pbcm-$CaMn_2O_4$ polymorphs at ambient conditions. The calculated lattice parameters overestimate the experimental data (see Table 1 for the values) [16, 23] for Pbca-$CaMn_2O_4$ and underestimate for Pbcm-$CaMn_2O_4$. More importantly, our DFT+U calculated lattice parameters are more than 90% in agreement with the experimental data, which indicates the validity of the approach employed during this study. Furthermore, the two polymorphs exhibit relatively large cell volumes. Specifically, the Pbca-$CaMn_2O_4$ polymorph displays a larger volume of 653.83 Å3, whereas the Pbca-$CaMn_2O_4$ polymorph has a smaller volume of 300.91 Å3. This suggests that the Pbca-$CaMn_2O_4$ polymorph is likely to deliver higher volumetric excellent Ca reversible rate capacities during operation [24].

The heats of formation (ΔH_f) which are essential in determining the thermodynamic stability of a material were calculated using equation 1 [25, 26]:

$$\Delta H_f = [E_{total} - E_{Ca} - 2E_{Mn} - 4E_O], \quad (1)$$

where E_{total} is the total energy of the geometrically fully relaxed $CaMn_2O_4$, while E_{Ca}, E_{Mn} and E_O are the total elemental energies of Ca, Mn, and O in their respective ground states. The heats of formation for Pbca-$CaMn_2O_4$ and Pbcm-$CaMn_2O_4$ structures are negative, indicating thermodynamic stability and that these polymorphs can be synthesized experimentally. Similar observations were reported by Arroyo-de Dompablo et al [13]. Moreover, the Pbca- $CaMn_2O_4$ is more stable than the Pbcm-$CaMn_2O_4$ since it has lower heats of formation (-6.796 eV).

Table 1. Calculated equilibrium lattice parameters and enthalpy of formation for Pbca-$CaMn_2O_4$ and Pbcm-$CaMn_2O_4$. Experimental data is also given where available.

Lattice Parameter	Pbca-$CaMn_2O_4$	Exp. [16]	Pbcm-$CaMn_2O_4$	Exp. [23]
a (Å)	6.534	6.25	3.143	3.15
b (Å)	10.222	9.90	9.947	9.95
c (Å)	9.789	9.63	9.600	9.67
V ($Å^3$)	653.83	596.07	300.091	302.92
ΔH_f (eV)	-6.796	-	-3.395	-

3.2. Electronic Properties

To determine the electronic properties of the Pbca-$CaMn_2O_4$ and Pbcm-$CaMn_2O_4$ polymorphs, we have calculated the spin-up total densities of states (DOS) as presented in Figure 2.

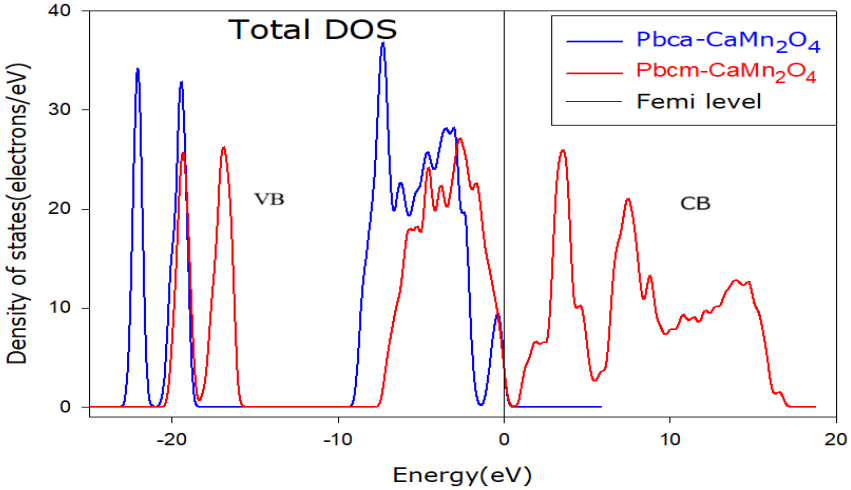


Figure 2. Total density of states for Pbca-$CaMn_2O_4$ and Pbcm-$CaMn_2O_4$. The Fermi energy is used as zero on the energy scale.

The DOS plots show similar trends, i.e. peaks at almost similar energy ranges for both polymorphs and are characterized by two distinguishable patterns: the valence band (VB) and the conduction band (CB). We note that the Fermi level is located on the edge of the upper valance band in both polymorphs, suggesting that most valence states do not overlap with the conduction band. Moreover, both systems show a shallow pseudo-gap around the Fermi level, indicating stability and semiconductor behaviour characteristics [25, 27], which agrees with the study done by Arroyo-de Dompablo et al [13]. Furthermore, Pbca-$CaMn_2O_4$ has lower states around the Fermi level as compared to Pbcm-$CaMn_2O_4$, suggesting that the Pbca-$CaMn_2O_4$ polymorph is more stable [28]. This is consistent with the calculated heats of formation.

4. Conclusion

The structural, thermodynamic, and electronic properties of Pbca-$CaMn_2O_4$ and Pbcm-$CaMn_2O_4$ polymorphs at ambient conditions were successfully determined by DFT-based first-principles calculations to mimic their stability. The calculated equilibrium lattice parameters agreed with experimental results, with a percentage difference of less than 10 % for both structures. Moreover, Pbca-$CaMn_2O_4$ was predicted to have more volumetric power and better Ca diffusion than Pbcm-$CaMn_2O_4$ due to the large cell volume. Both polymorphs were predicted to be thermodynamically stable due to the negative heats of formation, with the Pbca-$CaMn_2O_4$ being the most stable. The total density of states showed both polymorphs are characterized by the presence of the pseudo-gap around the Fermi, indicating stability. Moreover, both polymorphs are predicted to have semiconductor behaviour characteristics. Furthermore, Pbca-$CaMn_2O_4$ has lower states around the Fermi level than Pbcm-$CaMn_2O_4$, suggesting that the Pbca-$CaMn_2O_4$ polymorph is the more stable. The predicted large cell volumes, thermodynamic stability and semiconductor behaviour show that these polymorphs, in particular the Pbca-$CaMn_2O_4$ polymorph are potential candidates for application as cathode material in Ca-ion batteries. However, further studies on structural tuning to enhance electronic conductivity are needed to fully explore their application in this field.

References

[1] Harjanne A and korhonen JM 2019 Abandoning the concept of renewable energy *En. Pol.* **127** 330
[2] Gieelen D, Boshell F, Saygin D, Bazilian MD, Wagner N and Gorini R 2019 The role of renewable energy in the global energy transformation *En. St. Rev.* **24** 38
[3] Houghton J 2005 Global warming *Rep. Prog. Phys.* **68** 1343
[4] Zandalinas SI, Fritschi FB and Mittler R 2021 Global warming, climate change, and environmental pollution: recipe for a multifactorial stress combination disaster *T. Plants.* **26** 588
[5] Qazi A, Hussain F, Rahim N, Hardaker G, Algazzawi D, Shaban K and Harvna K 2019 Towards sustainable energy: A systematic review of renewable energy sources, technology and public opinions *IEEE Access.* **7** 63837
[6] Garcia-Olivares A, Sole J and Osychenk O 2018 Transportation in a 100% renewable energy system *En. Con. Man.* **158** 266
[7] Zubi G, Dufo-Lopez R, Carvalho M and Pasaoglu G 2018 The lithium-ion battery: state of the art and future perspectives *J. Ren. Sus. En. Rev.* **89** 292
[8] Lipson AL, Pan B, Lapidus SH, Liao C, Vaughey JT and Ingram BJ 2015 Rechargeable Ca-ion batteries: a new energy storage system *Chem. Mat.* **27** 8442
[9] Liang Y, Zhao C, Yuan H, Chen Y, Zhang W, Huang J, Yu D, Liu Y, Titirici M, Chueh Y, Yu H and Zhang Q 2019 A review of rechargeable batteries for portable electronic devices (PED) In*foMat* **1** 6
[10] Khan T, Garg AK, Gupta A, Madan AK and Jain PK 2020 Comprehensive review on latest advances on rechargeable batteries 2022 *J. En. Sto*. **57** 106204
[11] Noorden RV The rechargeable revolution: A better battery 2014 *Nature* **507** 26
[12] Gummow RJ, Vamvounis G, Kannan MB, and He Y 2018 Calcium-ion batteries: current state-of-the-art and future perspectives *Adv. Mat*. **30** 1801702
[13] Arroyo-de Dompablo M, Krich C, Nava-Avendano J, Biskup N, Palacin M and Barde F 2016 A Joint Computational and Experimental evaluation of CaMn2O4 polymorphs as cathode materials for Ca ion batteries *Chem. Mat.* **28** 6886

[14] Muldoon J, Bucur CB, and Gregory T 2014 Quest for nonaqueous multivalent secondary batteries: magnesium and beyond *Chem. Rev.* **2014** 11683
[15] Meng YS and Arroyo-de Dompablo ME 2009 First principles computational materials design for energy storage materials in lithium ion batteries *En. Env. Sc.* **2** 589
[16] Zouari S, Ranno L, Cheikh-Rouhou A, Isnard O, Pernet M, Wolfers P and Strobel P 2003New model for the magnetic structure of the marokite-type oxide CaMn2O4 *J. A. Comp.* **353** 5
[17] Segall MD, Lindan PJD, Probert MJ, Pickard CJ, Hasnip SJ, Clarke SJ and Payne MC 2002 First-principles simulation: ideas, illustrations and CASTEP code *J. Phys: Con. Matt.* **14** 2717
[18] Perdew JP, Burke K and Ernzerhof M Generalized Gradient Approximation Made Simple 1996 *Phys. Rev. Lett.* **77** 3865
[19] Koelling DD and Harmon BN 1977 A technique for relativistic spin-polarised calculations *J. Phys C: Sol. St. Phys.* **10** 3107
[20] Monkhorst HJ and Pack JD 1976 Special points for Brillouin-zone integrations *Phys. Rev. B.* **13** 5188
[21] Hubbard J 1963 Electron correlations in narrow energy bands *Roy. Soc* **276** 1365
[22] Wang Z, Brock C, Matt A and Bevan KH 2017 Implementation of DFT+U and its use in simulating polarons in oxide- and battery material *Phys. Rev. B.* **96** 125150
[23] Ling CD, Neumeier JJ and Argyriou DN 2001 Observation of antiferromagnetism in marokite CaMn2O4 *J. Sol. St. Chem.* **160** 167
[24] Wei, X, Zhang M, Miao B and Liu R 2021 Volumetric capacity enhancement in $LiFePO_4$ cathodes by hot isostatic pressing Scripta Materialia, **194** 113638
[25] Lethole NL and Ngoepe PE 2020 Structural, thermodynamic, electronic and mechanical properties of spinel and phonon-harvested AMn2O4 (A:Li, Na, Mg) system: a first-principles study *Mat. T. Com.* **22** 1
[26] Rogers DW, Zavitsas AA and Matsunaga N 2012 Determination of enthalphies ('Heats') of formation *W. Int. Rev: Comp. Mol. Sc.* **3** 21
[27] Lethole NL, Ngoepe PE and Chauke HR 2020 First-principles study: effect of lithium and sodium intercalation in transition metal phosphates, MPO4 (M: Mn, Fe, Co) *Comp. Con. Matt.* **22** e00437
[28] Zhao W, Tan PH, Liu J and Ferrari AC 2011 Intercalation of few-layer graphite flakes with FeCl3: Raman determination of fermi level, layer by layer decoupling and stability *J. Amer. Chem. Soc.* **133** 5941

Effect of Ruthenium dopant on the sensitivity of alpha iron oxide (α-Fe_2O_3) to Flammable and Hazardous Gases ype the title of your paper here

N G Cebekhulu[a], S A Ogundipe[a], C L Ndlangamadla[a], S S Nkosi[b] and H Swart[c]

[a]Department of Physics, University of Zululand, Private Bag X1001, KwaDlangezwa ZA3886, South Africa

[b]Department of Physics, University of Limpopo, Private Bag X1106, Sovenga, ZA0727, South Africa

[c]Department of Physics, University of the Free State, Box 339, Bloemfontein, ZA9300, South Africa

E-mail: **ntokozongc200@gmail.com**

Abstract. The increase in the number of manufacturing industries in recent decades has both positive and negative impacts on human health. Release of flammable and hazardous gases by manufacturing companies, industries, and fire into the atmosphere requires monitoring. Extensive research over the past decades has been conducted on the detection and monitoring of these gases using metal oxide semiconductor materials. This paper seeks to present the gas sensing performance of ruthenium-doped alpha iron oxide results towards the flammable and hazardous gases. The alpha iron oxide doped with a different weight percentage of ruthenium was synthesized by the co-precipitation method. Various characterization techniques such as X-ray diffractometry, thermogravimetric analysis, x-ray photoelectron spectroscopy, Brunauer-Emmett-Teller surface area analysis, scanning electron microscopy, high-resolution transmission electron microscopy, and X-ray photoelectron spectroscopy we used to investigate the properties of the materials. Sensors were fabricated by using the drop-casting method, and the sensors were tested for gas sensing performance at the operating temperature of 225 ºC, toward Liquefied petroleum gas, Ethanol, Propanol, Ammonium, and hydrogen sulfide respectively. The pure sample was found to be sensitive to the target gases with a response of 26.01 toward ammonia, while the response decreased upon the addition of ruthenium. The samples exhibited a shift toward the Liquefied petroleum gas with a response of 24.40 as the ruthenium content increased. Due to the reproduction of the oxygen vacancy sites the addition of Ru in the samples results in unstable RuO_2 as compared to Fe_2O_3

1. Introduction

A study conducted by the World Health Organization (WHO) found that air pollution alone caused 4.2 million deaths in 2016, while household air pollution caused 3.6 million. According to 2019 statistics, air pollution is among the biggest environmental health risks [1, 2]. It is equally important to maintain indoor air quality since people spend most of their time in indoors, as well as outdoor air pollution must be controlled [1, 3]. MOS-based gas sensors have been considered a promising candidate for gas detection in recent years. It is still necessary to improve the sensing properties of MOS-based gas sensors to satisfy the higher requirements for a wide range of applications, including medical diagnosis based on breath, gas detection in harsh conditions, and others. It is essential for MOS-based gas sensors to exhibit excellent selectivity, low power consumption, fast response/recovery, low humidity dependence, and a

low limit of detection concentration simultaneously in these fields. More and more researchers have been devoted to improving the sensing performance of MOS-based gas sensors in recent years [4]. Detecting and monitoring toxic, hazardous gases is crucial to protecting the environment and human health. Liquefied petroleum gas (LPG) is regarded as one of the most hazardous gases among other gases [5]. The presence of various toxic gases in the air that are harmful to human health has led to an increased interest in gas sensors in both research and industry [6]. The effects of toxic, hazardous, and flammable gases can include asthma, skin burning, dizziness, cancer, lung issues, and even weight loss [7].

2. Experimental procedure

2.1. Synthesis

To prepare samples of α-Fe_2O_3 and Ru-Fe_2O_3 with varying percentages, co-precipitation has been identified as the most preferred route. In this process, iron chloride hexahydrate ($FeCl_3 .6H_2O$) was used as the precursor while ammonia solution (NH_2OH) was used as the precipitant. The deoxygenated distilled water of 100 ml was mixed with Iron (III) chloride hexahydrate ($FeCl_3.6H_2O$) in a beaker. To obtain 0.01 M, the solution was stirred for 30 minutes at 80 ºC with a magnetic stirrer. A similar technique was applied to get (0.1 M) for the other solutions of Ru- α-Fe_2O_3 to investigate the impact of concentration on the size of the material. As a precipitating agent, ammonia hydroxide (NH_4OH) was added gradually and drop by drop until pH 11 was attained. The solution was heated for 3 hours while being stirred magnetically continuously. A centrifugation machine (6000 rpm) was used to collect the precipitated product. The product underwent multiple ethanol and distilled water washings. The product was then dried in an oven for three hours at 80 °C. FeOOH was transformed into α-Fe_2O_3. By calcining the material for 4 hours at 700 °C in the open air.

2.2. Characterization of Hematite and Ru dope hematite

X-ray diffraction (XRD) was used to determine the crystal structure of those powders using German Bruker D8 equipment with Cu/K as a radiation source and a wavelength of (λ = 1.5418 Å). Scanning electron microscopy (SEM, ZEISS Sigma VP-03-07) was used to capture an image of the powder's surface morphology and particle size. HR-TEM of the samples was done using JEOL 1400 instrument. The thermal behavior was evaluated by Thermo gravimetric (TGA) and differential thermal analysis (DTA) in the air by using TGA Q500 TA systems or instruments. The surface area and pore size were determined by a nitrogen adsorption /desorption isotherms analyzer. Nitrogen adsorption/desorption isotherms of the adsorbent samples were carried out on a 77 k Micrometrics TRISTAR II 3020 volumetric adsorption analyzers.

2.3. Sensor fabrication and measurement

The powder samples of α-Fe_2O_3 and Ru-Fe_2O_3 were sonicated for two hours in the ultrasonicate bath, and then dropped onto an alumina substrate that had a gold electrode screen-printed on top using the drop-casting technique. The gas sensing measurement was conducted by using a KS026K16 (KENOSISTES model of the 2016 year). Sensors were placed inside the gas testing chamber equipped with electrical and gas feeds for sensing measurement. Operating temperatures for all the measurement was set to 225 ºC. The flow of the dry air (79% N_2 and 21% O_2) inside the chamber controlled the concentration of the target gases. The target gases included liquefied petroleum gas, ammonia, ethanol, propanol, and hydrogen sulfide. The chamber was first filled with dry air for 30 minutes to allow the sensor to reach equilibrium, and then 10 minutes of each concentration of a target gas were added. The chamber was then filled with dry air once more for ten minutes to allow the sensor to recover. For all gas-sensing measures, this was done.

3. Result and Discussion

3.1. X-ray diffraction and Surface morphology

The XRD patterns of samples of α-Fe_2O_3 and Ru-Fe_2O_3 are depicted in Fig. 1(a). On the XRD pattern, only nine peaks are visible, but only five of them have high intensity, and two of them, the ones at (113) and (204), have lesser intensity. It indicates the existence of hematite with a rhombohedral structure with lattice parameters of a = 0.5036 nm and c = 1.3749 nm. The peak at the plane (104) in ruthenium-doped samples (Ru-α-Fe_2O_3) started decreasing in intensity when ruthenium was added, whereas the peak at the plane (110) increased in intensity. The secondary phase (RuO_2) was generated when the Ru concentration increased and was identified at 2θ = 28. 267. SEM image of sample 0.05Ru in Fig. 1 (b) shows well-distributed particles. While the HR -TEM shows the crystallinity of the material. The bright rings in the SAED patterns (Fig.2 (a)) confirm the crystallinity of alpha iron oxide (Ru-α-Fe_2O_3) samples. This is consistent with the data analysis of XRD (see Fig.1 (a)).

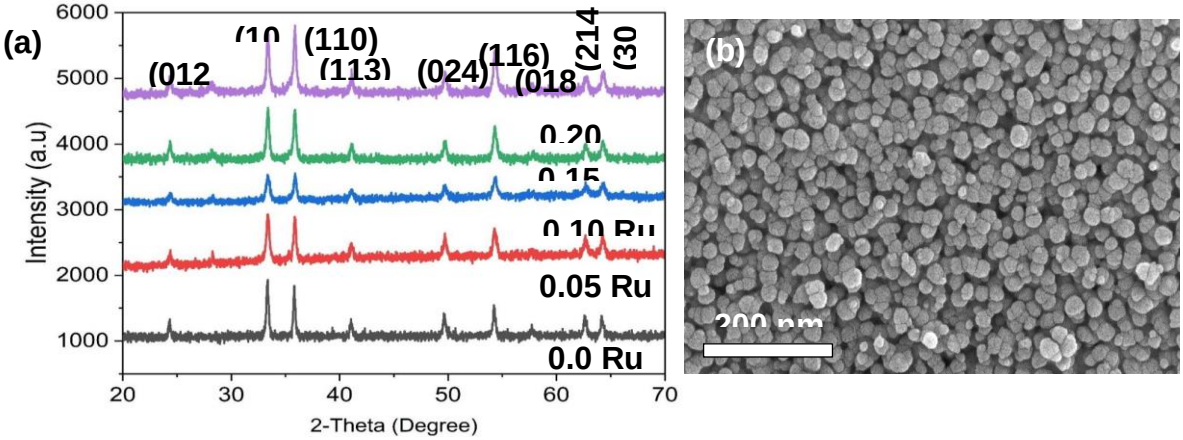


Figure 1. (a) The XRD patterns of pristine α-Fe_2O_3 and Ru-Fe_2O_3 samples (b) SEM images for sample 0.05Ru

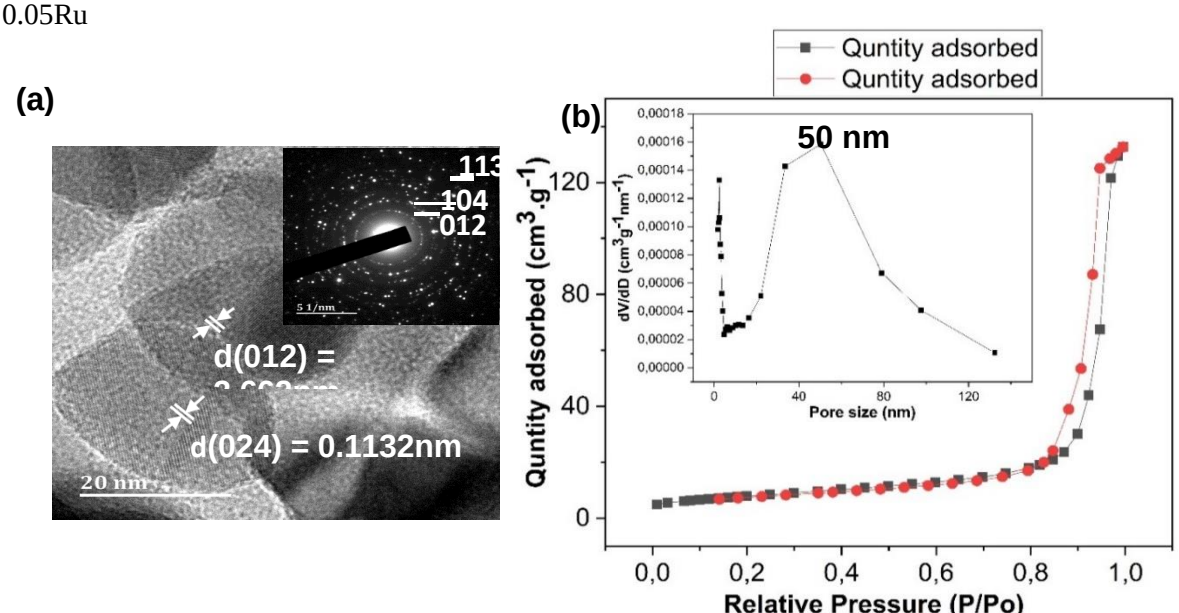


Figure 2. (a) HR-TEM image for sample 0.05Ru and (b) N_2 adsorption-desorption isotherm: The **inset** corresponds to the pore size distribution.

Nitrogen adsorption-desorption experiments were done to determine if sample surfaces are macro- or meso-porous. Fig.2 (b) shows nitrogen adsorption-desorption isotherm curves of the pristine α-Fe_2O_3. It was found the pose size of the pristine was 50 nm as shown on the pore size distribution curve.

The surface and chemical composition of the materials were studied by using X-ray photoelectron spectroscopy. The samples, α-Fe_2O_3 and Ru-Fe_2O_3 underwent survey scans and XPS measurements. The survey demonstrates the element present in the materials shown in Fig.3 (a) in particular. But most importantly it confirmed the presence of Fe, O, and Ru in the materials. As shown in Fig.3 (b-d), the survey scan reveals that the sample of α-Fe_2O_3 contains three elements: O1s, Fe2p, and Ru3d. The binding energy of each element was then confirmed by further XPS analysis. Only three peaks can be identified in the Fe2p scan in Fig.3 (b), and they correspond to the binding energies of 710.1 eV, 711.5 eV, and 725.1 eV, respectively. When the binding energy is lower, at 710.1 eV, the high peaks appear. The scan for O1s and Ru3d shown in Fig.3(c-d) shows only two peaks, but for the O1s high peaks occur at the lower binding energy of 530.0 eV. For the Ru3d the high peak occurs when the binding energy is high, at 285.1 eV.

3.1. Gas Sensing Properties

Semiconducting metal oxides have gas-sensing properties when exposed to hazardous and flammable gases at an operating temperature of (225 ºC**).** Fig.4 shows the 3D selectivity plot of the different flammable gases. Ammonia (1000 ppm) was substantially detected by the pristine α-Fe_2O_3-based sensor, with a response of 26.01 at 225 °C operating temperature. However, the addition of ruthenium decreases the gas sensor's response and the selectivity shifted towards the LPG.

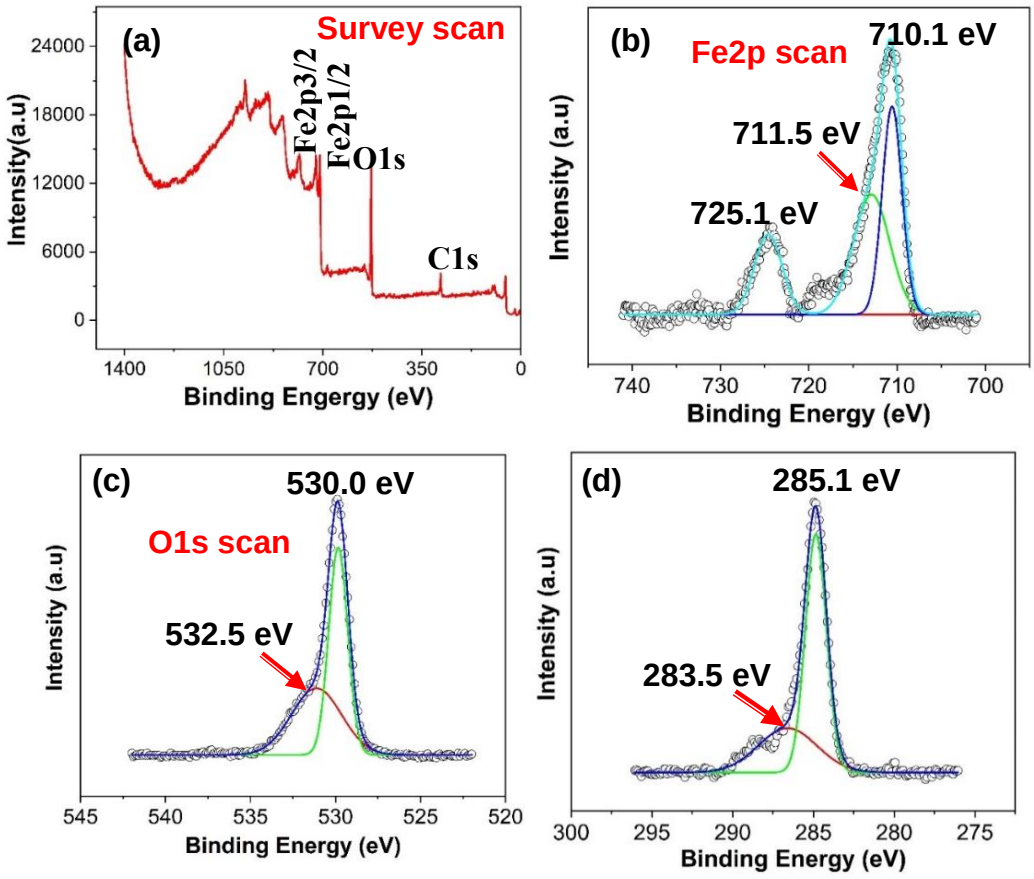


Figure 3. X-ray photoelectron spectroscopy results for pure alpha iron oxide (α-Fe_2O_3).

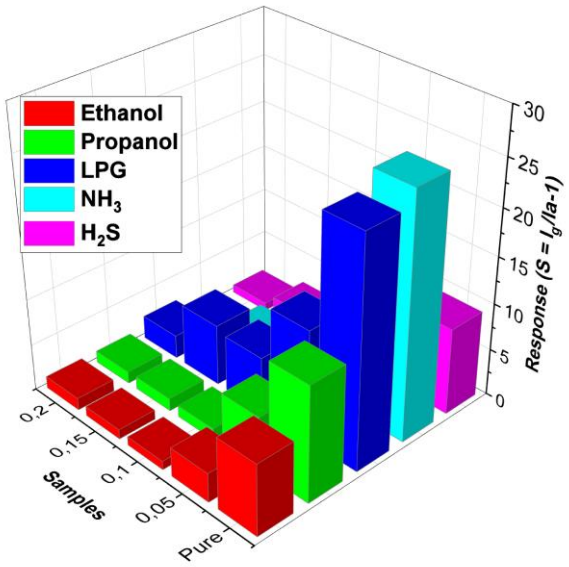


Figure 4. Selectivity 3D plot of n-type α-Fe_2O_3 and Ru-Fe_2O_3 toward various flammable gases, ethanol, propanol, ammonia, hydrogen sulfide, and liquefied petroleum gas at 225° C operating temperature.

Bar graphs in Fig. 5 (a) show the result of each sample and it also gives more information about the selectivity. The pure sample was found to be more selective of the ammonia gas since it showed a high response compared to the other samples. In addition to the ruthenium to host material, they selectively shifted towards Liquefier petroleum gas. Gas response decreases upon the addition of the ruthenium this is shown in Fig. 5 (b).

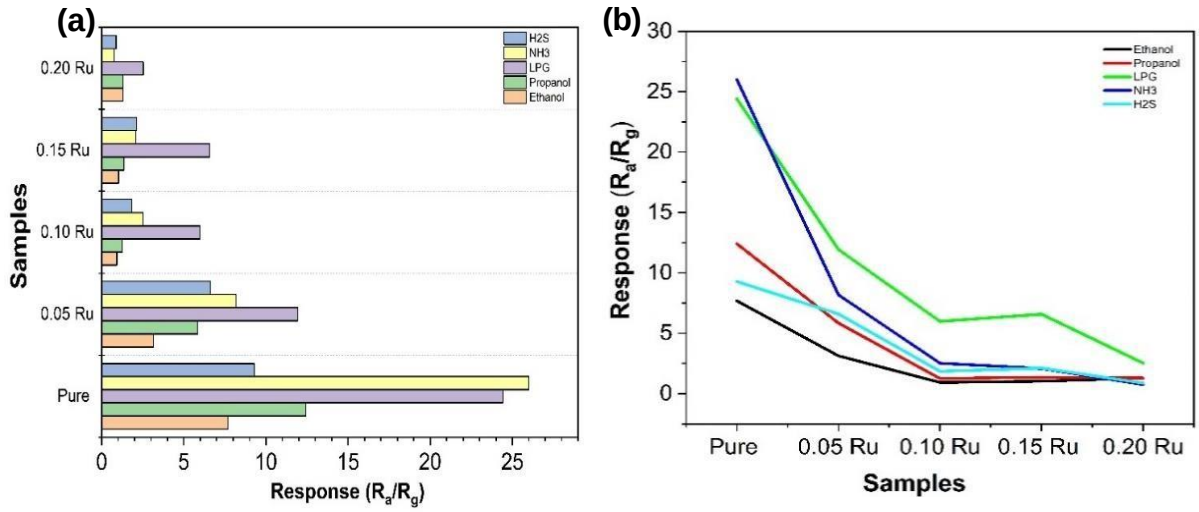


Figure 5. (a) Bar graphs show the gas sensing result of all samples. (b) Shows the response of the samples toward various gases at the operating temperature is 225° C.

4. The sensing mechanism of alpha iron oxide doped with ruthenium.

In general, gas sensing only occurs at the semiconducting material's surface where oxygen is chemically adsorbed during the reaction. However, the free electron from the conduction band is often taken by the chemisorbed oxygen on the surface of α-Fe_2O_3 to ionize it. This process increases the resistance of air

by forming depletion layers [9]. Depending on the operating temperature, the ambient oxygen catches the free electrons on the sensor material to produce $O2^-$, O^-, or O^{-2} [10]:

$$O_2\ (gas) \rightarrow O2\ (ads). \quad (2.1)$$

$$O_2\ (ads) + e^- \rightarrow O^-\ (ads)\ (T< 100^oC), \quad (2.2)$$

$$O_2\ (ads) + e^- \rightarrow 2O^-\ (ads)\ (100^oC < T<300^oC), \quad (2.3)$$

$$O^-(ads) + e^- \rightarrow O^{2-}\ (ads)\ (T> 300^oC). \quad (2.4)$$

Whenever n-type semiconductors like hematite and graphite are exposed to a target gas, the previously created electron-depletion layers when the sensor is in the air become thin, and electrons return to the conduction band. In the beginning, the oxygen molecules from the air pass through the surface of the hematite and are adsorbed to the surface. Following this, the oxygen-free electrons form oxygen ions on the hematite surface with trapped electrons, as indicated in the following chemical reactions [10-13].

5. Conclusion

Finally, it was found that ruthenium is not suitable to be used as the dopant material, since the addition of the ruthenium to the alpha iron oxide decreases the gas response. Pure alpha iron oxide was found to be more selective towards ammonia. The selectivity shifted to liquefied petroleum gas while the ruthenium concentration increased, this is due to the switch in Fe_2O_3 to RuO_2

6. Acknowledgement

The authors wish to thank the research office of the University of Zululand for the financial support during the cause of this work. Also, to thank NRF for financial support. Finally, I would acknowledge the University of the Free State for XPS techniques in this paper.

Reference

[1] Organization, W.H., *Ten threats to global health in 2019. 2019*. 2019.
[2] Sovacool, B.K., et al., *Climate change and industrial F-gases: A critical and systematic review of developments, sociotechnical systems and policy options for reducing synthetic greenhouse gas emissions.* Renewable and sustainable energy reviews, 2021. **141**: p. 110759.
[3] Chmielewski, A., *Monitoring, Control and Effects of Air Pollution*. 2011: BoD–Books on Demand.
[4] Zhang, J., et al., *Metal-oxide-semiconductor based gas sensors: screening, preparation, and integration.* Physical Chemistry Chemical Physics, 2017. **19**(9): p. 6313-6329.
[5] Nkosi, S., et al., *The effect of stabilized ZnO nanostructures green luminescence towards LPG sensing capabilities.* Materials Chemistry and Physics, 2020. **242**: p. 122452.
[6] Akram, R., et al., *Integrated Capacitive-and Resistive-Type Bimodal Relative Humidity Sensor Based on 5, 10, 15, 20-Tetraphenylporphyrinatonickel (II)(TPPNi) and Zinc Oxide (ZnO) Nanocomposite.* ACS omega, 2022. **7**(34): p. 30590-30600.
[7] Dhall, S., et al., *A review on environmental gas sensors: Materials and technologies.* Sensors International, 2021. **2**: p. 100116.
[8] Akram, R., et al., *Integrated Capacitive-and Resistive-Type Bimodal Relative Humidity Sensor Based on 5, 10, 15, 20-Tetraphenylporphyrinatonickel (II)(TPPNi) and Zinc Oxide (ZnO) Nanocomposite.* ACS Omega, 2022.
[9] Abbasi, Y.F. and H. Bera, *Konjac glucomannan-based nanomaterials in drug delivery and biomedical applications*, in *Biopolymer-Based Nanomaterials in Drug Delivery and Biomedical Applications*. 2021, Elsevier. p. 119-141.
[10] Nkosi, S.S., et al., *Abnormal p-type to n-type switching during nitric oxide gas sensing: Ni (OH) 2 nanoplatelets on amorphous NiO seed layers.* Vacuum, 2022. **200**: p. 111032.
[11] Mkwae, P.S., et al., *The heat rate kinetics on the liquefied hydrocarbon gases sensing and food quality*

control detecting strategy. Materials Chemistry and Physics, 2022. **277**: p. 125550.
[12] Shozi, N.N., et al., *Extremely sensitive and selective flammable liquefied hydrocarbon gas sensing and inter-dependence of fluctuating operating temperature and resistance: Perspective of rare-earth doped cobalt nanoferrites.* Journal of Alloys and Compounds, 2021. **859**: p. 157846.
[13] Li, C., et al., *Enhancement of gas-sensing abilities in p-type ZnWO4 by local modification of Pt nanoparticles.* Analytica Chimica Acta, 2016. **927**: p. 107-116

A machine learning approach to prediction of bandgap and optimum efficiency of $MASn_xPb_{1-x}I_3$ perovskite solar cells based on SCAPS 1-D data simulation

D R Mashamba[1,2], A D Kapim[1], M A Mavuso[1], G H Mhlongo[3] and M Msimanga[1,2]

[1] PV Nanocomposites R&D Platform, Department of Physics, Tshwane University of Technology, Private Bag X680, Pretoria, 0001, South Africa

[2] iThemba LABS TAMS, National Research Foundation, Private Bag 11, WITS, 2050, Johannesburg, South Africa

[3] DST/CSIR National Centre for Nano-Structured Materials, Council for Scientific and Industrial Research, Pretoria, 0001, South Africa

E-mail: **dakalorollet@gmail.com**

Abstract. Recent advances in perovskite solar cells (PSCs) have yielded power conversion efficiencies (PCEs) exceeding 25% within a decade from inception. Despite the rapid growth in PCE, one of the major drawbacks hindering large scale commercialization is the toxicity of lead (Pb) component in their composition. Partially or wholly substituting Pb with environmentally friendly metals offers a potential solution. However, composition exploration for mixed perovskites still largely necessitates labour-intensive trial and error experiments. A machine learning (ML) approach to predict the bandgap and PCE based on the composition of methylammonium tin/lead iodide ($MASn_xPb_{1-x}I_3$) using solar cell capacitance simulator (SCAPS) datasets is reported. The adopted ML models show good prediction capabilities with high R^2 scores of >0.98 and >0.99 for bandgap and solar cell performance parameter predictions respectively. Our results suggest that ML can innovatively accelerate discovery of efficient and less toxic PSCs by narrowing down possible perovskite combinations to a few permutations.

1. Introduction

Perovskite solar cells (PSCs) have recently emerged as the most progressive technology in the realm of photovoltaic research owing to their composition diversity and bandgap tunability [1-3].

Perovskite is any material with a crystal structure represented by ABX_3, where A is an organic or inorganic cation e.g., methylammonium (MA) or formamidinium (FA), B denotes a metal cation, typically lead (Pb) or tin (Sn) and X is an anion e.g., iodine (I) or bromine (Br) [2,3]. Miyasaka et al. [1] synthesized the first PSC of $MAPbI_3$ in 2009 resulting in power conversion efficiency (PCE) of 3.8%. Research in this domain continued to make thundering progress yielding Pb-based PSCs device PCE approaching 26% in 2022, making them great contenders for the most commercialized silicon solar cells [1]. Despite the high PCE, one of the factors impeding PSCs market adaptability is the toxicity of Pb in the B-site of perovskite crystal structure. Moreover, this typical perovskite material exhibits a larger

bandgap value of 1.55 – 1.6 eV, which does not fall within the optimal range in the Shockely-Queisser (S-Q) limit for single junction solar cells [1, 10]. The preferable bandgap for single junction solar cell is between 1.1 – 1.4 eV [2, 3], therefore even higher PCE can be achieved if the bandgap can be reduced to lower values as in $MASnI_3$ PSCs with a bandgap value of about 1.3 eV. Nevertheless, $MASnI_3$ is susceptible to degradation due to the ready oxidation of Sn^{2+} to Sn^{4+} under ambient conditions [1-3].

Current PSCs research trends are moving towards mixed perovskites to realise their maximum potential [1-5]. The unparalleled advancement of Pb-based PSC efficiency is attributed to the device architecture and layer properties including transport layers, bandgap and absorber layer composition. In the present work, we envisage that partially substituting Pb in the absorber layer with non-toxic Sn may be a pathway towards reducing Pb toxicity while maintaining PCE. The commonly adopted PSC fabrication procedure largely involves trial and error methods such as continuous synthesis and property optimization. However, due to the vast chemical landscape of perovskite materials, realization of optimum properties through continuous experiments can be tedious and time consuming. In this perspective, we illustrate an approach that combines numerical simulation with machine learning (ML) supported by experimental results from literature [7] to accelerate discovery of efficient and less toxic PSCs. We probe the relative impact of partially substituting Pb with Sn cation on the bandgap and performance parameters of $MASn_xPb_{1-x}I_3$ (where, $0 \leq x \leq 1$) PSC. This approach can provide a robust pathway and a much-needed bridge between artificial intelligence and data driven materials research.

2. Materials and methods

2.1. SCAPS 1-D simulation

Solar cell capacitance simulator (SCAPS) 1-D software, developed by Burgelman et al. [8] was utilized to simulate the performance parameter datasets. SCAPS is a versatile software designed based on basic semiconductor equations (i.e., Poisson, hole and electron continuity equations) for modelling solar devices under steady conditions. In this work, $MASn_xPb_{1-x}I_3$ solar cell performance parameters were generated through a combination of 9 bandgap values and 14 thickness values as input parameters.

The designed PSC device architecture in the present work consists of a transparent conducting oxide (TCO) electrode, electron transport layer (ETL), perovskite absorber layer, hole transport layer (HTL) and metal contact electrode arranged the sequence as: (*FTO/TiO_2/$MASn_xPb_{1-x}I_3$/Spiro-OMeTAD/Au).*

Table 1. Layer input parameters for ETL, absorber layer and HTL

Layer properties	TiO_2 (ETL)	$MASn_xPb_{1-x}I_3$ (Absorber)	Spiro (HTL)
Thickness (nm)	25	200-1500 (Δ: 100)	20
Bandgap (eV)	3.2	1.18-1.6	2.9
Electron affinity (eV)	4.0	4.0	2.2
Dielectric permittivity (ε_r)	10.0	10.0	3.0
Conduction band density of states (N_c) (cm^{-3})	1×10^{21}	1.2×10^{18}	2×10^{18}
Valence band density of states (N_v) (cm^{-3})	2×10^{20}	2.8×10^{18}	2×10^{18}
Donor concentration (N_D) (cm^{-3})	1×10^{18}	1.21×10^{9}	0
Acceptor concentration (N_A) (cm^{-3})	0	1.21×10^{9}	1.3×10^{15}
Electron mobility μ_n (cm^2/V.s)	1×10^{-4}	2.0	1×10^{-4}
Hole mobility μ_h (cm^2/V.s)	1×10^{-4}	2.0	$1 \times 1\ 0^{-4}$
Total defect density (cm^{-3})	1×10^{15}	2.5×10^{13}	1×10^{15}

While overall PSC performance mostly depends on the transport and absorber layer properties [6], HTL and ETL properties were fixed along with some of the absorber layer properties. Material properties for each layer in our SCAPS simulation are listed in Table 1 [4, 6]. The work function for electrodes in our simulation were 5.0 and 5.47 eV for FTO and Au respectively [4]. The experimental bandgaps with their corresponding compositions for $MASn_xPb_{1-x}I_3$ were extracted from a single peer reviewed publication [7] measured under the same laboratory conditions, which can arguably improve the reliability and accuracy of the prediction models. Solar cell performance parameters of open circuit voltage (V_{oc}), fill factor (FF), current density (J_{sc}) and PCE were obtained through variation of 9 bandgaps and 14 thickness values of the absorber layer. SCAPS simulation settings were configured to A.M. 1.5G (1000 $W.m^{-2}$) spectrum and temperature of 300 K. All possible combinations of the varied input parameters were simulated to generate 126 PSCs with unique performance parameters for each simulated device.

2.2. Machine learning

Scikit-learn ML tools [9] were utilized for bandgap and solar cell performance parameter prediction models. Polynomial regression (PR) and Decision tree regressor (DT) models were used during the investigation. First, polynomial regression model predicted the bandgap from $MASn_xPb_{1-x}I_3$ composition as shown in Figure 1(a). Secondly, using SCAPS simulated data, polynomial regression and decision tree models predicted the performance parameters using input parameters of thickness and experimental bandgap as shown in Figure 1(b). The models were validated using cross validation scores variation with different polynomial degrees and maximum depths for polynomial regression and decision tree, respectively. The dataset was split into 80-20% to train and test the models. Model evaluation employed traditional statistical error metrics of Root Mean Square Error (RMSE) and R-squared (R^2) scores, where R^2 score of 1.0 and RMSE of 0.0 denote a perfect prediction with no error. Reliability of the models was further evaluated by predicting performance parameters using the mutually exclusive predicted bandgaps and thickness as shown in Figure 1(c) of the ML workflow.

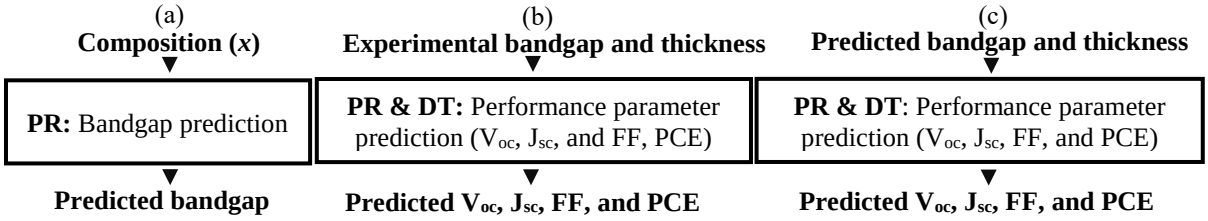


Figure 1. Illustration of the utilized machine learning workflow

3. Results and discussion

Figure 2(a-d) presents R^2 and RMSE scores for different polynomial degrees and maximum depths. R^2 scores increase with polynomial degrees, and peak at degree 4 with a score of 0.9983. Similarly, in Figure 2(b), the RMSE scores decrease with increase in polynomial degrees reaching the lowest score of 0.1578 at degree 4. The scores indicate optimal model performance at degree 4, with degrees beyond 4 causing overfitting as observable in Figure 2(a), where R^2 test scores are lower than train scores. Figure 2(c) shows an initial increase in R^2 score for train and test data with increasing maximum depths which signifies continuous improvement in performance of the model. The decrease in RMSE scores in Figure 2(d) as maximum depth increase indicates reduction in prediction error towards higher maximum depths. However, the optimal maximum depth of the decision tree model is unclear, prompting cross-validation to gain insight of cross-validation score variation with maximum depths. Similarly, cross-validation was conducted for polynomial regression to confirm the initially observed optimal degree.

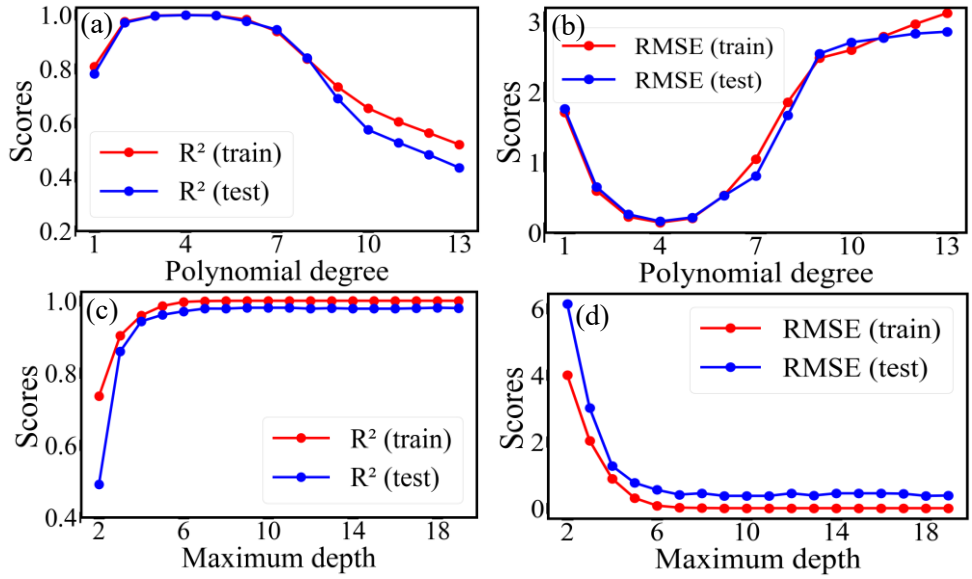


Figure 2. R^2 and RMSE scores for train and test data as a function of polynomial degree and maximum depth

Cross-validation scores for different polynomial degrees and maximum depths are presented in Figure 3 (a-b). Notably, cross-validation scores peak around degree 4 and depth 10 for polynomial regression and decision tree models respectively. Beyond degree 4 and depth 10, the scores start declining suggesting that these values present optimal performance of our models. In predicting the performance parameters using the predicted bandgap, polynomial regression achieved RMSE and R^2 scores of 0.3003 and 0.9960 while decision tree achieved 0.5397 and 0.9884 respectively.

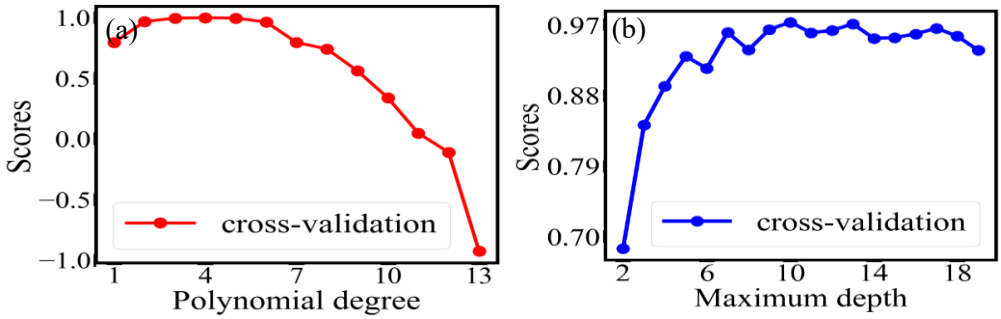


Figure 3. (a) Polynomial regression and (b) decision tree cross-validation scores as a function of polynomial degree and maximum depth

Correlation between the SCAPS simulated and predicted performance parameters is presented in Figure 4. For all four solar cell performance parameters, decision tree predicted values slightly deviate from the reference line, whereas polynomial regression predicted values relatively align well with the reference line. This supports RMSE and R^2 scores as illustrated in Figure 2, where both models show good R^2 for train and test scores. The slight deviation in decision tree predicted values can be associated with the relatively high RMSE (above 50%) in predicting performance parameters based on the unseen (predicted) bandgap. The results demonstrate that polynomial regression is a best performing model for our dataset. For bandgap prediction, polynomial regression was utilized by interpolating a polynomial

fit line as illustrated in Figure 5(a). The model demonstrated good predictive capabilities achieving low RMSE of 0.0138 and R^2 score of 0.9895 indicating a good fit between the predicted and experimental bandgaps as shown in Figure 5(b). The model built in this manner allows prediction of bandgaps corresponding to compositions not present in our original dataset providing guidance on synthesizing optimized perovskite materials. Figure 5(a) illustrates dependence of the bandgap on composition. With increase in *x,* the bandgap decreases almost linearly at lower compositions and non-linearly at intermediate compositions showing parabolic dependence on the composition between end compounds.

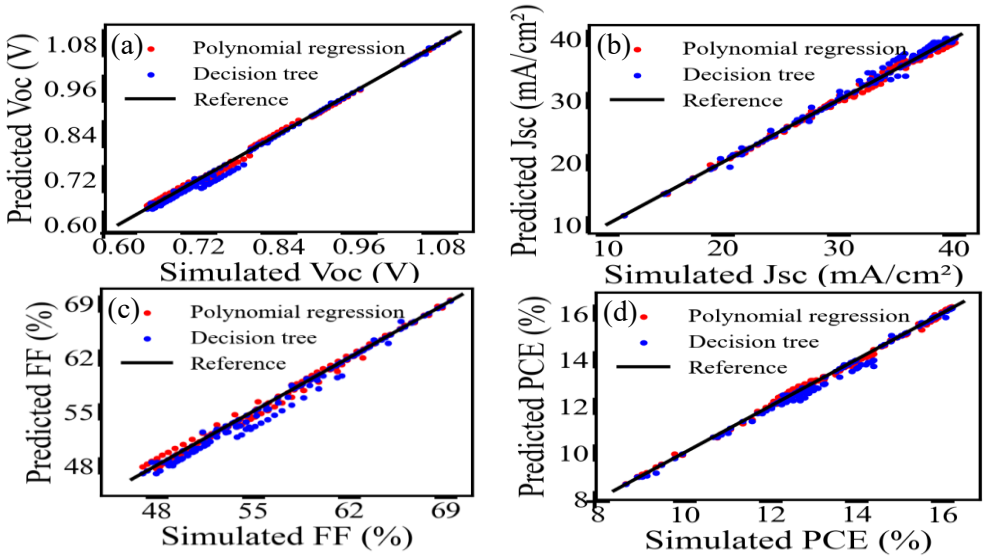


Figure 4. Comparison of simulated and predicted solar cell performance parameters

When x=0.5, the bandgap displays lowest value (1.18 eV), lower than the lowest $MASnI_3$ end compound. This non-linearity is referred to as bandgap bowing effect and is consistent in $ASn_xPb_{1-x}I_3$ for all possible A-site cation studied so far [2, 3].

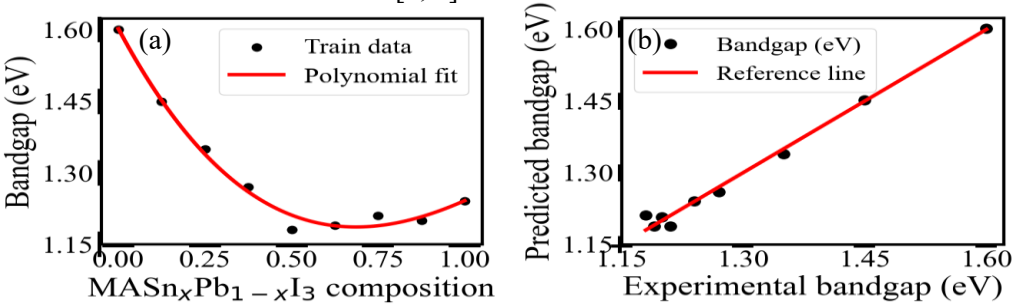


Figure 5. (a) Composition-dependent bandgap variation (b) Comparison of predicted and experimental bandgap

The implication is that this effect can also be attained through utilization of other A-site cations including formamidinium and cesium or multiple A-sites. Rajagopal et al. [3] reported that $MAPbI_3$ and $MASnI_3$ crystallize in the tetragonal and (pseudo)cubic phases respectively at room temperature and with increase in x, a gradual tetragonal to (pseudo)cubic phase transition as x approaches 0.5 is observed. This transition not only offers an advantage of reduced bandgap but also slower oxidation kinetics of $MASnI_3$, leading to improved stability of $MASn_xPb_{1-x}I_3$ compared to pure $MASnI_3$ [1, 5, 10]. Therefore, depending on the composition ratio and temperature, $MASnI_3$ crystal structure may transform from

(pseudo)cubic to tetragonal or a combination of both influencing its stability and optoelectronic properties. For example, perovskite compounds with 50% Sn or less are reportedly more stable than with pure Sn compounds [5]. While mixed compounds enhance degradation lifetime compared to pure $MASnI_3$, overall stability remains inadequate for long term exposure to ambient conditions. Current research efforts to address this issue include additive engineering and encapsulation layers such as protective polymer coating [1]. Figure 6 shows correlation between SCAPS simulated and ML predicted optimum solar cell performance parameters. SCAPS simulated performance parameters: V_{oc}= 0.91 V, FF = 62.78 %, J_{sc} = 28.47 mA/cm^2 and PCE = 16.22% are in good correlation with ML predicted performance parameters as also evident in the RMSE and R^2 scores of the models in Figure 2.

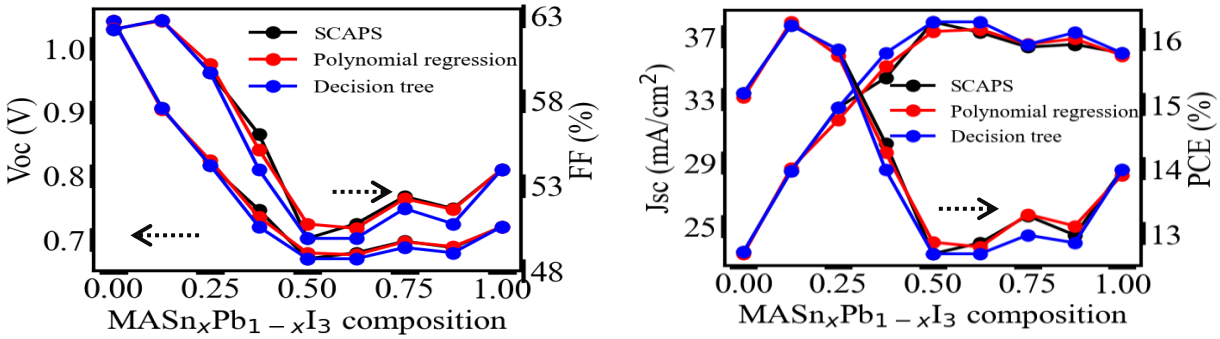


Figure 6. Comparison between SCAPS simulated, polynomial regression and decision tree predicted optimum solar cell performance parameters.

These parameters correspond to the bandgap of 1.45 eV and Sn composition of 0.125. Notably, PCE decreases with increase in Sn composition, suggesting a trade-off between toxicity reduction with PCE and stability. The simulated device PCE obtained is less than the S-Q maximum efficiency limit estimated at 30% by Lim et al. [10]. Therefore, optimization of the simulated device such as varying the HTL, ETL and contact electrode properties can be pursued to further improve PCE of the device towards the optimal theoretical S-Q limit.

4. Conclusion

The ML workflow is expected to guide experimentalists on minimizing trial and error experiments through rapid composition dependent bandgap and performance parameter prediction before fabrication. Our results suggest that polynomial regression exhibit remarkable accuracy as it achieves high R^2 score with minimal error rates for both seen and unseen bandgaps. The versatility of our workflow is that it can be applied independently to PSCs with analogous structures, expanding its usability beyond SCAPS simulation framework. The PSC structure herein achieves PCE of 16.22% for the bandgap of 1.45 eV and Sn composition of 0.125 and it offers multi-advantages such as bandgap reduction to single junction solar cell range. Moreover, toxicity can be reduced through a reasonable trade off with PCE and stability. Although mixed Sn/Pb PSCs improves stability of the oxidation prone $MASnI_3$, it is not yet adequate for prolonged exposure to ambient conditions. Ongoing research have shown that the lifetime and PCE of $MASn_xPb_{1-x}I_3$ can further improve through additive engineering and/or protective coating layer.

References

1. Wang H, He J, Xiang H, Ran R, Zhou W, Wang W, and Shao Z 2023. *Energy & Fuels* **37**(9): 6401–6423
2. Cai, X., Liu, F., Yu, A., Qin, J., Hatamvand, M., Ahmed, I., Luo, J., Zhang, Y., Zhang, H., and Zhan, Y. 2022 *Light: Sci. & Appl.* **11**(1): 234.

3. Rajagopal, A., Stoddard, R.J., Hillhouse, H.W., and Jen, A.K.Y. 2019. *J. Mater Chem* A, **7**(27): 16285-16293.
4. Hasanzadeh Azar, M., Aynehband, S., Abdollahi, H., Alimohammadi, H., Rajabi, N., Angizi, S., Kamraninejad, V., Teimouri, R., Mohammadpour, R., and Simchi, A. 2023. *Photonics*,**10**(3), 271.
5. Leijtens, T., Prasanna, R., Gold-Parker, A., Toney, M.F., and McGehee, M.D. 2017. *ACS Energy Lett*, **2**(9), 2159-2165.
6. Islam, M.S., et al. 2022. *ACS Omega*,**7**(26), 22263-22278.
7. Guan, L., Xu, X., Liang, Y., Han, S., Guo, J., Wang, J., and Li, X. 2020. *Phys. Lett.* A,**384**(8), 126173.
8. Burgelman, M., Nollet, P., and Degrave, S., 2000. *Thin Solid Films*, **361**, 527-532.
9. Pedregosa, F., et al. 2011. *JMLR*, **12**, 2825-2830
10. Lim, E.L., Hagfeldt, A., and Bi, D., 2021. *Energy Environ. Sci.*, **14**(6), 3256-3300

First-principle study of structural, thermodynamic and mechanical stability of ternary $NaVSe_2$

DM Tshwane[1,2*], RS Dima[1], L Mogakane[1], T Ngcobo[1], PM Maleka[1] and RR Maphanga[1,2]

[1] Next Generation Enterprises and Institution Cluster, Council for Scientific and Industrial Research, P.O. Box 395, Pretoria 0001, South Africa
[2] National Institute for Theoretical and Computational Sciences (NITheCS), Gauteng 2000, South Africa

E-mail: **DTshwane@csir.co.za**

Abstract. Ternary metal dichalcogenide compounds have piqued the interest of the industrial and scientific communities because of their numerous applications in catalysis, electronics, aerospace, and electrode materials. Among them, the $NaVSe_2$ dichalcogenide compound promises excellent intercalation and high conductivity. However, details on the structural and electronic properties of $NaVSe_2$ remain limited, especially at the atomistic scale. The present study used a density functional theory (DFT) approach using different generalized gradient approximation functionals such as Perdew-Burke-Ernzerhof (GGA-PBE), Perdew-Burke-Ernzerhof revised for solids (PBEsol), and local density approximation (LDA) to investigate the structural, thermal, and electronic stability of $NaVSe_2$. The results showed that the heats of formation (ΔH_f) less than 0 for $NaVSe_2$, which implies that the phase is thermodynamically stable. Furthermore, the GGA-PBE functional was found to be the most suitable function to predict the $NaVSe_2$ properties compared to the GGA-PBEsol and LDA functional. Furthermore, the recent findings reveal that $NaVSe_2$ meets all mechanical stability requirements across the board. It was also discovered that LDA overestimates elastic constants while GGA-PBE underestimates them.

1. Introduction

Transitional metal dichalcogenides (TMDs) with the usual sandwich structure are among the most promising materials with unique electrical, magnetic and mechanical properties and have been the subject of extensive research in recent decades [1]. Due to its multiple uses in catalysis, electronics, aerospace, and electrode materials, this family of compounds has garnered a great lot of interest from the technological and scientific sectors. This is because of its intriguing properties, such as its low toxicity, high chemical and mechanical stability. The localised behaviour of the transition metal d bands and the mixing of the transition metal state between the atoms of the transition metal and the surroundings of their chalcogen ligands play a significant role in determining the electronic characteristics of TMDs [2]. TMDCs have a variety of electrical characteristics, from metallic to semi-conductor. Researchers have been studying the extensive variety of electrical, optical, mechanical, chemical, and thermal properties of TMDCs for decades [3, 4]. Due to recent developments in sample preparation, there is currently a renaissance of scientific and engineering interest in TMDCs in their atomically forms [5, 6].

Various atomic species of M (metal) and X (chalcogen) have already been used to synthesise and characterise compounds of the TMD family with chemical composition, MX_2. These comprise metallic, and superconductor electronic phases, making them potential candidates for a variety of devices and processes, including catalysts [7], Li-ion batteries (LIB) [8], and lubricants [9]. Popov *et al.* [10], used density functional theory (DFT) and DFT+U approaches to investigate the atomic geometry, electronic structure, and magnetic properties of VS_2 and VSe_2.

Recently, there has been a lot of interest in vanadium dichalcogenides (VX_2) such as VSe_2, VTe_2 and VS_2 because of their exceptional optical, electrical, and chemical properties [11]. Salavati *et al.* [12] investigated the use of highly stretchable and conductive 2D 1T VS_2 and VSe_2 as anode materials for Li-, Na-, and Ca-ion storage using DFT and are reported to have a promising performance. Dichalcogenides of group IV and V can form a non-distorted phase, which is sometimes the thermodynamically stable structure in compounds such as TiS_2, $TiSe_2$, or VSe_2. The study of the activity of the basal planes of metallic polymorphs should greatly benefit from the thermodynamic stability [13]. Due to high surface area, stability, and great electrical conductivity, TMDs are a valuable material for electrochemical energy storage devices, such as supercapacitors and batteries. TMDs can be employed as cathode materials in batteries due to their high specific capacity and ability to store and release a significant number of Li-ions [14]. However, the electronic properties of $NaVSe_2$ are yet to be determined. Therefore, it is worthwhile to predict their electronic properties using different functionals to determine more accurate properties of $NaVSe_2$.

In this work, DFT was employed to investigate the structural stability of ternary $NaVSe_2$ crystal using different functionals such as generalized gradient approximation (GGA) by Perdew-Burke-Ernzerhof (PBE), PBEsol and local density approximation (LDA). DFT has proven to be an effective theoretical technique for electronic structure calculations of material properties and simulation material process including details on chemical bonding, hybridization, and atomic interaction [15].

2. Computational Method

The work used density functional theory to characterise and investigate the structural, electronic, and mechanical properties of the $NaVSe_2$ structure using the plane wave technique implemented in the Cambridge Serial Total Energy Package (CASTEP) [16] code of BOIVIA Materials Studio software. Different functionals, namely, LDA [17] and PBE [18] and PBEsol [19] were employed to approximate the electron exchange and correlation. All calculations used a cutoff energy of 710 eV for the set of plane wave basis and a Monkhorst-Pack k-point of 6x6x1 [20]. During geometry optimisation, the Broyden-Fetcher-Goldforb-Shanno minimisation scheme algorithm was used. The convergence tolerance and the absolute maximum force were set at $1x10^{-5}$ eV/atom and 0.03 eV/atom, respectively. The heats of formation (ΔH_f) of the $NaVSe_2$ structure were calculated as follows: $E_{tot.}$ and E_i refer to the system's and individual's total energies, respectively, where *i* denotes for the atomic configuration and *n* for the number of atoms.

$$E_{HF} = E_{tot.} - \sum_i n_i E_i \quad (1)$$

3. Results and Discussion

3.1. Structural parameters and heats of formation

Figure 1 presents the optimised atomic structure of the $NaVSe_2$ crystal structure considered as an ordered atomic configuration with the space group R-3m. The structure is based on a trigonal phase with Na atoms bounded to six Se that form $NaSe_6$ octahedra that share equivalent VSe_6 octahedra. Using various functionals, that is, PBE, PBEsol, and LDA, structural optimisation was achieved by minimising force and stress tension. The computed lattice parameters a=b≠c, were a= 3.55 Å and c =21.36 Å for PBE functional. When comparing the computed lattice parameters for the various functionals, it was discovered that the lattice parameters of the PBE functional were larger than those of PBEsol and LDA as presented in Table 1. Comparing different GGA functionals, PBEsol underestimates all lattice parameters while PBE overestimates the lattice parameter *a* [21].

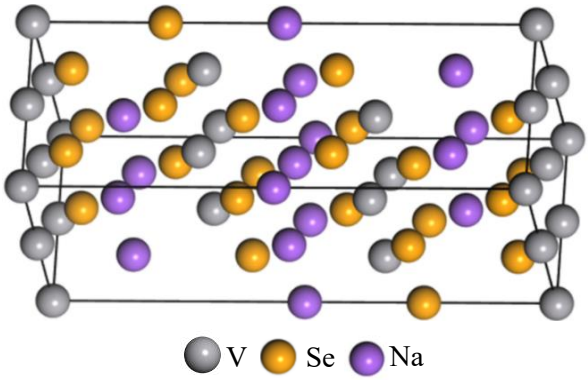


Figure 1. The optimised atomic structure of $NaVSe_2$ compound with aspace group R-3m.

Furthermore, the computed ΔH_f for $NaVSe_2$ in relation to the different functionals is shown in Table 1. According to the results, the calculated value of ΔH_f for the functional PBE (-1.173 eV) functional is largely negative (suggesting more stability) compared to the functionals, PBEsol (-1.085 eV) and LDA (-1.071 eV) as presented in Table 1. It was discovered that for all the functionals, the ΔH_f value is negative, suggesting the exothermic production phase. The negative number of ΔH_f implies that these materials are expected to have a solid solution [22]. In this current study, the ΔH_f difference is -0.088 eV for PBE and PBEsol and -0.102 eV for the functional PBE and LDA. This implies that the functional PBE is energetically favourable for the formation of $NaVSe_2$.

Table 1. Equilibrium lattice parameters (Å) and heats of formation (ΔH_f) of $NaVSe_2$ structures at different functionals: PBE, PBEsol and LDA.

Structure	Parameter(s)	PBE	PBEsol	LDA
$NaVSe_2$	a	3,55	3,48	3,44
	c	21,36	21,37	23,23
	a/c	0,166	0,163	0,148
	ΔH_f	-1,173	-1,085	-1,071

3.2. Electronic density difference

The electron density difference (EDD) was calculated to define the distribution of electronic charge in $NaVSe_2$ structure and to identify the type of interatomic bonding between a specific pair of atoms. The distribution of charges and electrons inside the crystal structure was assessed using the EDD. Figure 2 shows a 2D plot of the atoms in the electron density distribution of the system. The magnitude of electrons and charges are indicated by different coloured regions. The red area denotes positively charged electrons, whereas the blue region denotes negatively charged electrons. It is evident that the V atoms are mostly surround by the blue region, whereas the Se atoms dominate the red zone. The V atom has a significant impact on the -d states, while the Na and Se atoms do not contribute to the d states. Further observations revealed that both atoms (V and Se) show blue and red region charged this is due to atomic interaction and bonding.

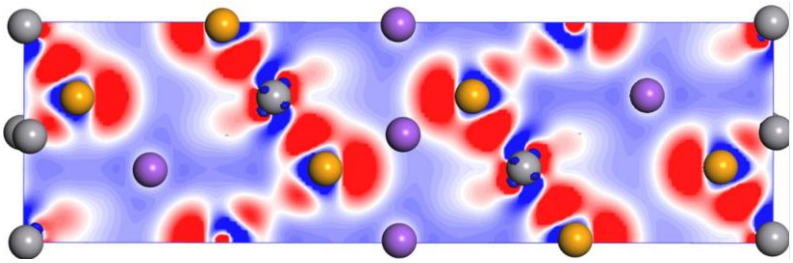


Figure 2. EDD for the $NaVSe_2$ crystal structure with the blue and red regions indicating positive and negative charge, respectively.

3.3. Elastic properties

Understanding the elastic constants is essential to comprehend the anisotropic character of the bonding, the structural stability, and the bonding characteristics between the adjacent atomic planes. The elastic characteristics of the crystal are critical to understanding the microscopic and macroscopic behaviour of the crystal and creating new materials [23]. Six deformation modes can be used to estimate C_{11}, C_{12}, C_{13}, C_{33}, C_{44}, and C_{66} for the tetragonal phase and the criteria of the tetragonal system are $C_{11} > 0$, $C_{33} > 0$, $C_{44} > 0$, $C_{66} > 0$, C_{11}-$C_{12} > 0$, C_{11}+C_{33}-2$C_{13} > 0$, 2C_{11}+C_{33}+2C_{12}+4$C_{13} > 0$ [24]. Calculations were performed using strain energy theory to investigate how the atomic configuration and different functionals affect the mechanical characteristics of the $NaVSe_2$ phase.

Table 2. Elastics constants C_{ij} and moduli for $NaVSe_2$ structures predicted by three different functionals, PBE, PBEsol, and LDA.

C_{ij} (Gpa)	Functionals		
	PBE	PBEsol.	LDA
C_{11}	106	110	115
C_{12}	23	26	17
C_{13}	41	53	49
C_{33}	101	108	99
C_{44}	35	35	41
C_{66}	41	39	49
B	58	66	62
G	36	34	40
B/G	1.6	1.94	1.55
Youngs modulus	90	89	100
Poison's ratio	0.24	0.27	0.23
Anisotropy	0.84	0.88	1.4

The strain energy theory and the elastic constant of the matrix should both be positive and symmetric, according to the mechanical stability criteria. Furthermore, the current findings show that $NaVSe_2$ satisfies all the mechanical stability requirements in all functionals. The elastic constants C_{11} and C_{33} are closely related to uniaxial, while C_{12}, C_{13}, C_{44}, and C_{66} are largely associated with non-axial sound propagation [25]. Therefore, for all functionals, C_{11} and C_{33} have the highest values among the constants evaluated, which implies that $NaVSe_2$ cannot be easily compressed under uniaxial stress. In addition, the Voigt and Reuss methods could be used to compute the mechanical characteristics of polycrystalline elastic materials, such as the shear modulus (G), bulk modulus (B), Young's modulus (Y), and Poisson's ratio (ν). According to the hypothesis of Pugh, brittleness (> 1.75) and ductility (<1.75) are predicted by its bulk-to-shear modulus ratio. PBE and LDA predicted that the $NaVSe_2$ structure is brittle, while PBEsol predicted possible ductility. The Possion ratio, on the other hand, relates to how brittle and ductile a material is, with a compound being called brittle if the value is less than 0.26 and ductile otherwise. It is important to note that the PBE and LDA functionals in the current results also indicated that $NaVSe_2$ is brittle in nature, while PBEsol exhibits ductility properties with a lower value of 0.27. Generally, LDA tends to be overestimate elastic constant while PBE form to underestimate them, although some elastic constants calculated by LDA usually are better. The creation of PBEsol was developed to correct PBE underbonding in solids and could account for this precision, also the overall trends for PBE are more accurate than LDA [26].

4. Conclusion

In summary, DFT was employed successfully to investigate the structural, thermodynamic, and mechanical stability of crystal $NaVSe_2$ utilising three distinct functionals, PBE, PBEsol, and LDA. The findings revealed that $NaVSe_2$ compound is thermodynamically stable with ΔH_f<0. It was observed that the PBE functional predicts the stable heats of formation compared to those of PBEsol and LDA.

Furthermore, the computed lattice parameters for the LDA functional are shorter (3.44 Å) than PBEsol, (3.48 Å) and PBE (3.55 Å). It was also noted that LDA tends to overestimate elastic constants while PBE form to underestimate them. The fact that PBEsol was developed especially to correct PBE underbonding in solids may account for this accuracy. Additionally, both Na and Se have robust atomic contacts and bonding in both their negatively and positively charged regions. The mechanical characteristics and comparative structural study of the $NaVSe_2$ compound maybe applied as potential anode and cathode TMDs materials.

Acknowledgements
The authors acknowledge the Department of Science and Innovation (DSI) Foundational Digital Capability Research and Council for Scientific and Industrial Research for financial support. Center for High-Performance Computing (CHPC) for computing resources.

References

[1] Huang C, Xie L, Zhang H, Wang H, Hu J, Liang Z, Jiang Z, Feasible S F 2022 *Nanomater.* **12** 2518.
[2] Kutana A, Penev E S and Yakobson B 2014 *Nanoscale* **4** 5820.
[3] Wang Q H, Kalantar-Zadeh K, Kis A, Coleman J N and Strano M S 2012 *Nat. Nanotechnol.* **7** 712.
[4] Arora A, Koperski M, Nogajewski K, Marcus J, Faugeras C, Petemski M 2015 *Nanoscale* **7** 10421.
[5] Yang R, Fan Y, Zhang Y, Mei L, Zhu R, Qin J, Hu J, Chen Z, Ng Y H, Voiry D, Li S, Lu Q, Wang Q, Yu J C, Zeng Z 2023 *Angew. Chem. Int. Ed. Engl.* **62** 202218016.
[6] Chowdhury T, Sadler E C and Kempa T J 2020 *Chem. Rev.* **1220** 12591.
[7] Drescher T, Niefind F, Bensch W and Grünert W 2012 *J. Am. Chem. Soc.* **134** 18896 2012.
[8] Wang Z, Chen T, Chen W, Chang K, Ma L, Huang G, Chen D and Lee J Y 2013 *J Mater. Chem. A* **1** 2210.
[9] Mosleh M, Atnafu N D, Belk J H and Nobles O M 2009 *Wear* **267** 1225.
[10] Popov Z I, Mikhaleva N S, Visotin M A, Kuzubov A A, Entani S, Naramoto H, Sakai S, Sorokin P and Avramov P V 2016, *Phys. Chem. Chem. Phys.* **48** 33052.
[11] Sanyal G, Lakshmy S, Vaidyanathan A, Kalarikkal N and Chakraborty B 2022 *Surf. Interfaces* **29** 101816.
[12] Salavati M and Rabczuk T 2019 *Compt. Mater. Sci.* **160** 367.
[13] Masaityte L 2023 PhD Thesis.
[14] Shi Z, Huang H, Wang C, Huo M, Ho S H and Tsai H S 2022 *Chem. Eng. J.* **447** 137469.
[15] Hafner J 2008 *J. Comput. Chem.* **29** 2045.
[16] Segall M D, Philip J D L, Probert M J, Pickard C J, Hasnip P J, Clark S J and Payne M C 2002 *J.Phys: Condens. Matter.* **14** 2744.
[17] Kohn W and Sham L J 1965 *Phys. Rev.* **140** A1133.
[18] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865.
[19] Perdew J P, Ruzsinszky A, Csonka G I, Vydrov O A, Scuseria G E, Constantin L A, Zhou X, and Burke K 2009 *Phys. Rev. Lett.* **102** 039902.
[20] Monkhorst H J and Pack J D 1976 *Phys. Rev. B* **13** 5188.
[21] Dima R S, Maleka P M, Maluta E N and Maphanga R R 2022 *Mater. Tod.: Proc.* **62** S11.
[22] Dima R S, Maleka P M, Maluta N E and Maphanga R R 2022 *Mater.* **15** 5280.
[23] Mehl M J, Klein B M and Papaconstantopoulos D A 1994 New York: Wiley 1994.
[24] Watt J P and Peselnick L 1980 *J. Appl. Phys.* **51** 1531.
[25] Hu Z, Xu W, Chen C, Wen Y and Liu L 2018 *Adv. Mater. Sci. Eng.* 9.
[26] Xing W D, Meng F Y, Ning J L, Sun J W and Yu R 2021 *Sci. China Technol. Sci.* **64** 2761.

Deposition time-dependent properties of electrodeposited CdSe thin films from a cadmium nitrate source for energy harvesting applications.

AU. Yimamu[1, 3], BF. Dejene[4], OK. Echendu[5, 6], JJ. Terblans[2], HC. Swart[2] and SJ. Motloung[1]

[1]Department of Physics, University of the Free State, 9866, Phuthaditjhaba, South Africa.
[2]Department of Physics, University of the Free State, P.O. Box 339, Bloemfontein, 9300, South Africa
[3]Department of Physics, Adama Science and Technology University, Adama, Ethiopia
[4]Department of Physics, Walter Sisulu University (Mthatha), Mthatha, South Africa
[5]Department of Physics, Federal University of Technology, P. M. B, 1526, Owerri, Nigeria
[6]Africa Center of Excellence in Future Energies and Electrochemical, Systems (ACE-FUELS), Federal University of Technology, P, M, B, 1526, Owerri, Nigeria.

E-mail: **ahemed.dilla2010@gmail.com and tlousj@gmail.com**

Abstract Electrodeposition of CdSe thin films was done using a two-electrode configuration for different deposition periods of 5, 10, 15, 20, and 30 min. The structural, and optical properties, surface morphology, surface roughness, and elemental composition of the as-deposited CdSe thin film samples were investigated by using X-ray powder diffraction, ultraviolet- visible (UV-VIS) spectroscopy, scanning electron microscopy, scanning probe microscopy and energy-dispersive X-ray spectroscopy, respectively. The CdSe thin films were cubic in nature. The optical properties analysis revealed that the absorbance of the CdSe thin films increased with deposition time. The energy band gap varied from 1.69 eV to 1.85 eV for the as-deposited samples. The glass substrates were covered uniformly, and the shape of the grains, grain size, and morphology changed with deposition time. The average surface roughness was recorded as 35.7, 45.45, and 53.6 nm for deposition times of 5, 20, and 30 minutes, respectively. Both Cd and Se were present in the CdSe thin films, and their percentage composition varied with deposition time. In thin film solar cell devices, material layer thickness plays a significant role and is controlled by the deposition time. The CdSe film deposited at a short time of 5 min has potential application as a window (buffer) layer. For a longer time, 30 min, the CdSe film can be used as a solar cell absorber layer.

1. Introduction

Energy is one of the primary research areas around the globe. Utilizing energy sources that are not renewable such as coal, gas, and oil, accelerates global warming [1]. Sustainable, renewable, and environmentally friendly energy sources are the way forward. Solar energy is the best source in all aspects, such as being free from transportation and labour force. Solar energy conversion materials play a significant role in converting this large amount of energy to electricity. Based on these solar energy materials, energy usage is controlled by quantum efficiency, production cost and stability, and lifetime [2]. Traditional crystalline Si-based solar cells are the leading ones in the present market. However, the major issue is the high cost of production of Si and its indirect band transition, as well as its low absorption coefficient. For these reasons, Si wafer needs material thicknesses of up to 300 μm to absorb more light, and this adds to the cost of production [3]. Thin film CdTe-based solar cells are alternative low-cost solar cells with moderate efficiency. A thin thickness of nearly 2 μm is required to absorb 99% of visible light. There is literature mentioning CdTe-based solar cell current status [4]. First, solar achieved a new benchmark for CdTe thin film conversion efficiency in 2016 of 22.1% [5]. One of the main issues of CdTe-based solar cells is a suitable window layer for *p-n* junction formation. Cadmium selenide (CdSe) is a group III-V compound semiconductor material with a high absorption coefficient, direct energy band gap, and a high photoconductive nature. It can be used for different solid-state devices, such as sensors [6], supercapacitors [7], and photovoltaic solar cells [8]. CdSe thin films were

used as buffer or absorber layers based on the preparation method and synthesis conditions. CdSe thin films were synthesized with the aid of a cheap two-electrode electrodeposition method to further reduce the cost of production. The advantages of this method were discussed in previous reports. These include scalability, extended bath life, little waste generation, and self-purification of electrolytes. Deposition parameters and processing steps, such as electrolyte pH, deposition temperature, deposition period, doping, and post-deposition treatment, play significant roles [9]. CdSe thin film preparation using two and three-electrode electrodeposition methods and with cadmium precursors such as cadmium sulphate [10], cadmium chloride [11], and cadmium acetate have been reported. There is no report on the preparation of CdSe thin films from cadmium nitrate as a Cd precursor by using the two-electrode electrodeposition configuration in the potentiostatic mode. Specifically, the effect of deposition time plays a great role in controlling the thickness of the film for device applications. CdSe thin films, with thickness of up to 1 µm, can be used as absorber material with very thin thicknesses of less than 100 nm, it can be used as window material.

In this study, the structure and optical properties, surface and elemental composition of CdSe thin films grown by the two-electrode electrodeposition method from cadmium nitrate as a Cd source at different preparation times are reported.

2. Materials and Method

Cadmium selenide (CdSe) thin film coatings were electrochemically synthesised on conductive fluorine doped tin oxide (FTO) glass substrates using a potentiostat in the two-electrode mode. The sheet resistivity of the FTO glass substrate was 8Ω/square. The electrolytic solution contained 0.03 M SeO_2 and 0.3 M Cd $(NO_3)_2.4H_2O$ as precursors of selenium and cadmium, respectively. The solution was prepared in 400 ml deionized water, and deposition voltage, temperature, and pH were fixed at 1950 mV, 75 ^oC, and 2.3, respectively. A magnetic stirrer was used to gently agitate the mixture throughout the deposition. The deposition was done without further purification of the electrolyte solution. Five FTO glass substrates were washed using an ultrasonic bath for 30 min and then washed in ethanol, acetone, and methanol, respectively. Finally, the substrates were washed with deionized water and then dried in air. A computerized Gill AC potentiostat (ACM instrument, United Kingdom) was used as a source of electrical power for the two-electrode setup. Since the cathodic deposition technique was employed, a high-purity graphite rod that functioned as the working electrode (cathode) was bonded to the FTO glass substrate using an insulating polytetrafluoroethylene (PTFE) thread seal tape. Another high-purity graphite rod was used as the counter electrode (anode). The films were prepared at different deposition times of 5, 10, 15, 20, and 30 min. The structural and optical properties, surface morphology, surface roughness, and elemental composition of the as-prepared samples were analysed and reported.

3. Results and discussion

3.1 Structural properties

A Bruker D8 Advance X-ray powder diffractometer was used for X-ray powder diffraction (XRPD) measurements in order to examine the structural characteristics of the CdSe thin films. To determine the degree of crystallinity and the CdSe layers crystal structure the incidence angle was adjusted from 20 to 70^o. Figure 1 (a) shows the XRPD patterns of the as-deposited CdSe thin films. The results confirmed that the CdSe thin films were formed in the cubic structure, and the (111) peak intensity of the phase increased with deposition time. The CdSe diffraction peaks can be detected at the angles with 2θ values of 25.50, 42.51, and 49.81^o corresponding to the lattice planes (111), (220), and (311), respectively. The outcome is consistent with the JCPDS card 190191. In addition, the intense peaks corresponding to the FTO glass substrates were detected in the XRPD patterns. The crystallite size was estimated by utilized the Scherrer equation.

$$D = \frac{0.94\lambda}{\beta \cos\theta} \quad (1)$$

where θ is the Bragg diffraction angle and β is the peak's full width at half maximum (FWHM) in radians. The 111 peak was used to estimate the crystallite sizes. Figure 1(b) shows the graph of the FWHM of the (111) plane and the crystallite size in relation to the period of deposition. It is evident that when the Coating duration enhanced from 5 to 30 min, the crystallite size of the CdSe thin films decreased. The crystallite size obtained for the 5 min deposition time was the maximum.

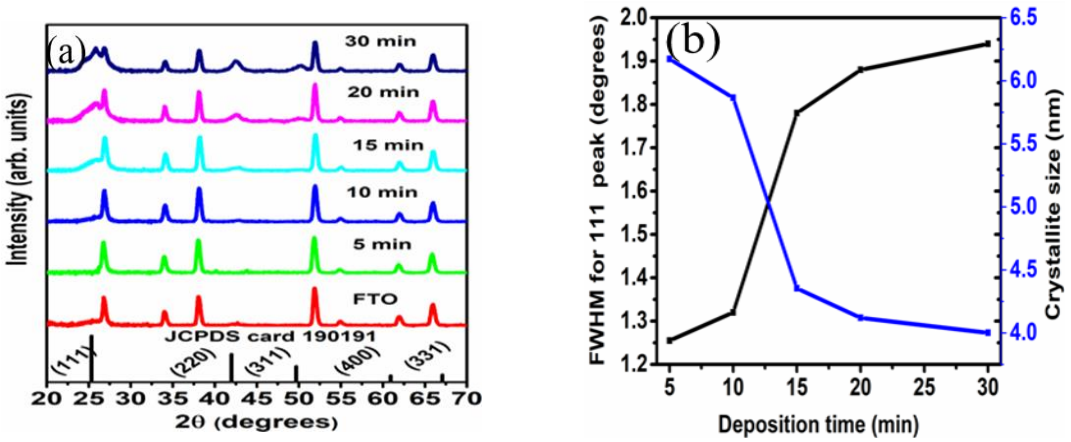


Figure 1. (a) CdSe thin films' XRPD plot for the different deposition times. (b) Graph of FWHM of the (111) peak and the crystallite size with the period of deposition.

3.2 Optical properties

To determine the films band gap, optical absorption investigations of the CdSe layers were made using a ultraviolet- visible (UV-VIS) spectroscopy (Win lab-scan lambda 950). Figure 2 (a) shows the absorbance spectra of the CdSe thin films with respect to wavelength.

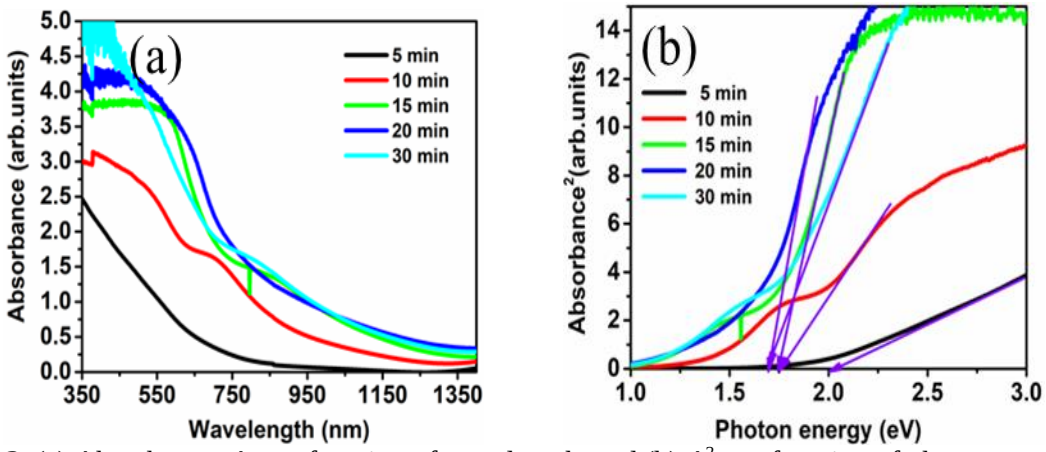


Figure 2. (a) Absorbance, A as a function of wavelength, and (b) A^2 as a function of photon energy, for the CdSe films deposited for different durations.

The results confirmed that the absorbance of the film increased as the deposition time increased in the visible part of solar spectra. This situation is not same beyond 750 nm. This is due to the increase in the thickness of the film. Figure 2(b) demonstrates how the square of absorbance changes with photon energy (eV). Extrapolating the tangent of the graph's straight-line segment to the photon energy axis allowed for estimating the energy band gap. The results indicate that the band gap changed as the duration of deposition increased. The energy band gaps were recorded as 2.11, 1.79, 1.77, 1.73 and 1.71 eV with deposition times of 5, 10, 15, 20, and 30 min, respectively.

3.3 Morphological property

To examine the surface morphology, grain size, and homogeneous of the film covering the substrate, scanning electron microscope (SEM) images were collected using a Jeol JSM7800F FE-SEM (field emission scanning electron microscope). Figure 3 shows the SEM images of the CdSe thin films grown for different deposition times. The results demonstrated that a thin coating of CdSe had been applied to the FTO glass substrate.

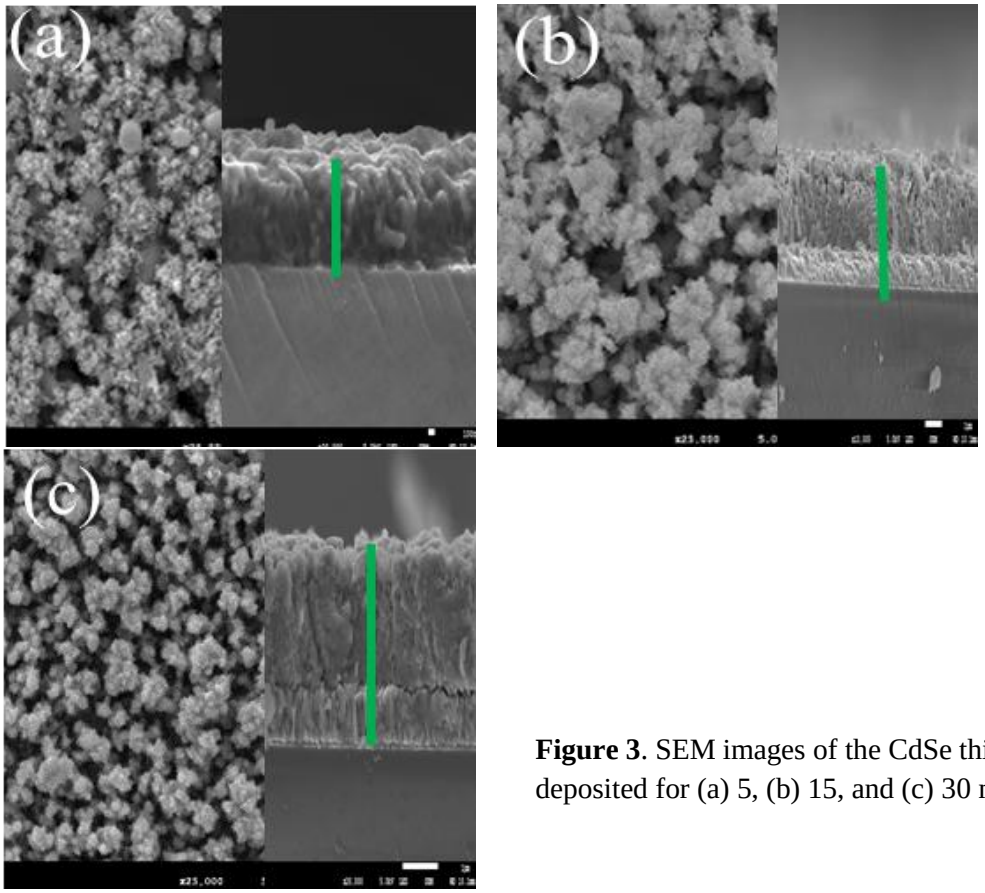


Figure 3. SEM images of the CdSe thin films deposited for (a) 5, (b) 15, and (c) 30 min.

The result confirmed that the morphology of the film varied with deposition time. At deposition time of 5 min, agglomerated grains with spherical shape were observed. When the deposition time is 15 min more agglomerated and pin holes are observed. When the time is 30 min independent grain are visible which is suitable for absorber layer. The thickness of the film increased with the length of the deposition time. The experimental thickness of the CdSe thin films was estimated from the SEM-cross-section images and the theoretical thickness was calculated by using Faraday's law of electrolysis as given in equation 2 [2].

$$T = \frac{JMt}{Fn\rho} \tag{2}$$

where T is the film thickness, t is the deposition period, M is the molar mass of CdSe (191.37 g.mol^{-1}), J is the average current density (A.cm^{-2}) during deposition, F is the Faraday constant 96,485 C.mol^{-1}, ρ is the density of CdSe (5.82 g.cm^{-3}), and n is the number of electrons that are transferred during the deposition of one CdSe molecule (n = 6). The experimental thicknesses of the films were recorded as 0.31, 0.94, and 1.9 µm for deposition times of 5, 15, and 20 min, respectively. For the theoretical thickness calculation, the thickness was recorded as 0.43, 1.27, and 2.5 µm for deposition times of 5, 15, and 30 min, respectively. It should be noted that the theoretical thicknesses derived from Faraday's law of electrolysis exhibit larger values than the observed thicknesses from SEM. Faraday's law of

electrolysis, considers that all the electronic charges contributed to the formation of CdSe layers. In electrodeposition however, all electrons may not contribute to the thin film formation due to the simultaneous electrolysis of water taking place during the deposition.

3.4. Surface roughness

The surface roughness properties of the electrodeposited CdSe thin films were examined using scanning probe microscopy (SPM). The surface properties of materials have a significant role in device fabrication. The roughness of the surface indicates its degree of smoothness. Figure 4 shows the SPM images of the CdSe thin films that were deposited for different growth times. The results confirmed that the average surface roughness (R_a)of the CdSe thin films changed with the deposition period. The R_a values recorded for the as-deposited films were 35.7, 45.45 and 53.6 nm for deposition time of 5, 20, and 30 min, respectively. It is clear from the topography images that deposition period has an impact on the surface roughness.

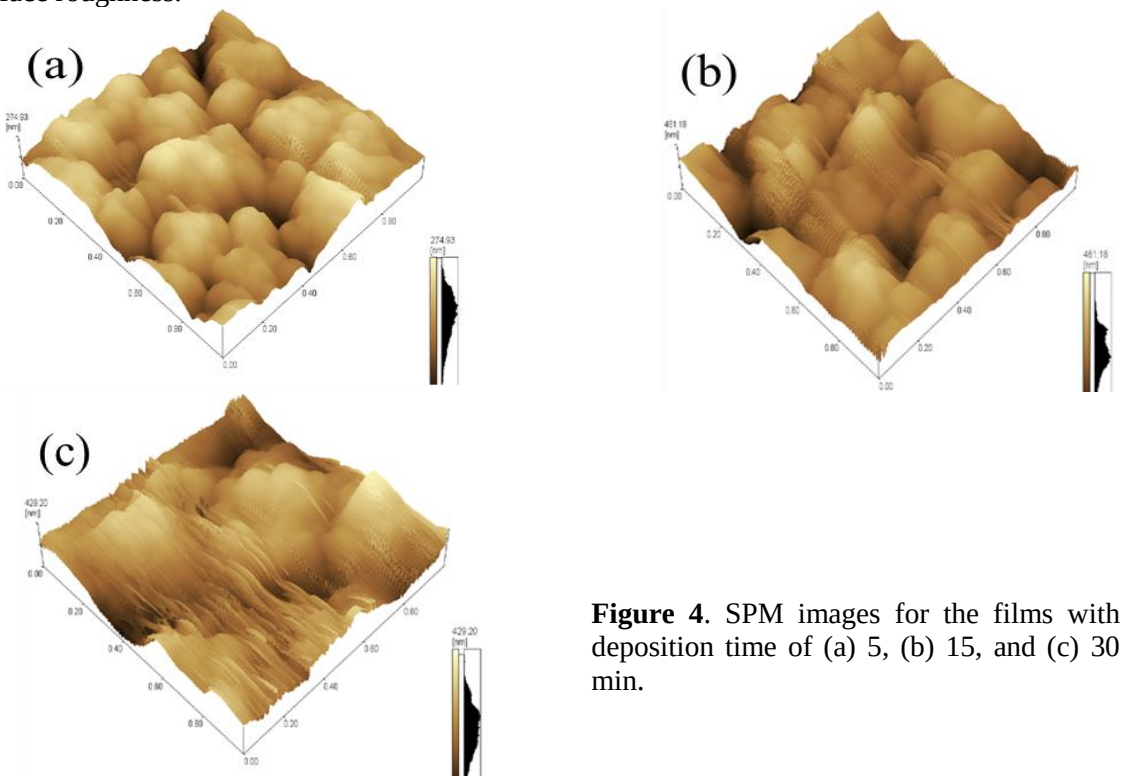


Figure 4. SPM images for the films with deposition time of (a) 5, (b) 15, and (c) 30 min.

3.5 Compositional properties

CdSe thin films produced for various deposition times were examined for their elemental composition using Energy Dispersive X-Ray Spectroscopy (EDS). Figure 5 shows the CdSe thin films' EDS spectra for different growth times. The results show that Cd and Se were present in the CdSe films coated on the FTO glass substrates. The compositions of the Cd and Se varied with the deposition period. The Si, Sn and O elements are related to glass/FTO substrate while C is from the carbon tape used in sample preparation for SEM measurements.

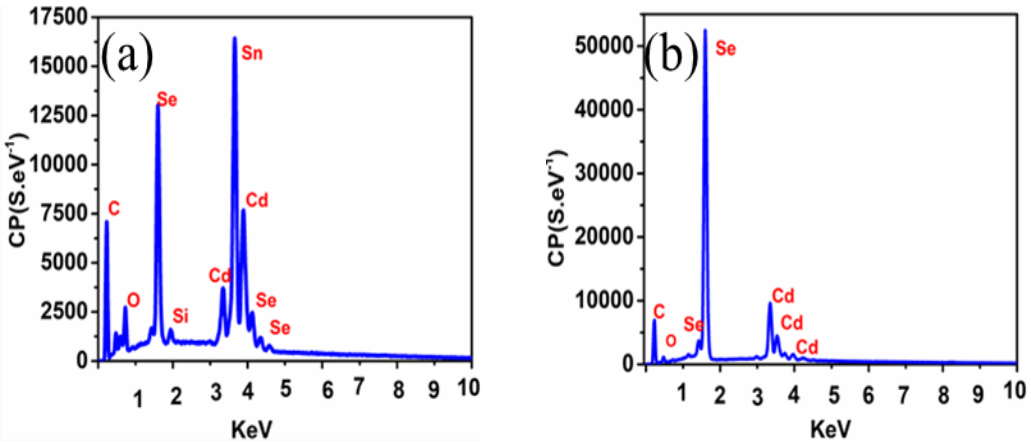


Figure 5. EDS plot of the representative CdSe thin films grown for (a) 5 and (b) 30 min.

4. Conclusion

The electrodeposition approach was used to effectively grow CdSe thin films on glass/FTO substrates utilizing cadmium nitrate and selenium dioxide as the Cd and Se sources, respectively. The thin films were electrodeposited for 5, 10, 15, 20, and 30 min. The CdSe thin films were in the cubic phase, and their crystallite sizes varied with the deposition period. The biggest size was noted at the lowest deposition time of 5 min. The absorbance and band gap of these materials varied due to their thickness being increased with increasing deposition time. The films' average surface roughness changed as their growth time increased. The compositions of the Cd and Se in the CdSe thin films also varied with deposition time. The overall analysis confirmed that in the electrodeposition method, CdSe thin film deposition time affected the material properties. Films deposited at high deposition time have the potential application as a possible absorber layer, while the one deposited at a shorter time of 5 min can be used as a window or buffer layer for thin-film solar cells.

Acknowledgments

The authors gratefully thank the National Research Fund (NRF) and the University of the Free State for financial assistance.

Reference

[1] T. Sinha, D. Lilhare, and A. Khare, "A review on the improvement in performance of CdTe/CdS thin-film solar cells through optimization of structural parameters," *J. Mater. Sci.*, vol. 54, no. 19, pp. 12189–12205, 2019, doi: 10.1007/s10853-019-03651-0.

[2] A. U. Yimamu *et al.*, "Electrodeposition of CdTe thin films using an acetate precursor for solar energy application: The effect of deposition voltage," *Mater. Today Commun.*, vol. 35, no. February, p. 105673, 2023, doi: 10.1016/j.mtcomm.2023.105673.

[3] P. M. Voroshilov, C. R. Simovski, P. A. Belov, and A. S. Shalin, "Light-trapping and antireflective coatings for amorphous Si-based thin film solar cells," *J. Appl. Phys.*, vol. 117, no. 20, 2015, doi: 10.1063/1.4921440.

[4] A. U. Yimamu *et al.*, "Influence of growth time on the properties of CdTe thin films grown by electrodeposition using acetate precursor for solar energy application," *Mater. Res. Express*, vol. 10, no. 5, pp. 0–13, 2023, doi: 10.1088/2053-1591/acd322.

[5] S. He *et al.*, "Study of CdTe/ZnTe composite absorbing layer deposited by pulsed laser deposition for CdS/CdTe solar cell," *Mater. Sci. Semicond. Process.*, vol. 67, no. April, pp. 41–45, 2017, doi: 10.1016/j.mssp.2017.05.009.

[6] A. Bafti, V. Mandić, I. Panžić, L. Pavić, and V. Špada, "CdSe QDs modified cellulose microfibrils for enhanced humidity sensing properties," *Appl. Surf. Sci.*, vol. 612, no. November 2022, 2023, doi: 10.1016/j.apsusc.2022.155894.

[7] L. Teng, L. Tian, W. Li, H. Feng, and Z. Xing, "Preparation and characterization of CdSe as electrode materials for supercapacitors," *Synth. Met.*, vol. 291, no. July, p. 117169, 2022, doi:

10.1016/j.synthmet.2022.117169.
[8] J. Xue *et al.*, "Improved Carrier Lifetimes of CdSe Thin Film via Te Doping for Photovoltaic Application," *ACS Appl. Mater. Interfaces*, vol. 15, no. 14, pp. 17858–17866, 2023, doi: 10.1021/acsami.3c00461.
[9] S. Kumari, G. Chasta, R. Sharma, N. Kumari, and M. S. Dhaka, "Phase transition correlated grain growth in CdSe thin films: Annealing evolution to cadmium chloride activation," *Phys. B Condens. Matter*, vol. 649, no. August 2022, p. 414422, 2023, doi: 10.1016/j.physb.2022.414422.
[10] R. Mariappan, V. Ponnuswamy, S. M. Mohan, P. Suresh, and R. Suresh, "The effect of potential on electrodeposited CdSe thin films," *Mater. Sci. Semicond. Process.*, vol. 15, no. 2, pp. 174–180, 2012, doi: 10.1016/j.mssp.2011.10.007.
[11] O. I. Olusola, O. K. Echendu, and I. M. Dharmadasa, "Development of CdSe thin films for application in electronic devices," *J. Mater. Sci. Mater. Electron.*, vol. 26, no. 2, pp. 1066–1076, 2015, doi: 10.1007/s10854-014-2506-x.

Structural, electronic, and thermodynamic properties of ternary $NaVS_2$: A computational study

L Mogakane[1], R S Dima[1], D M Tshwane[1,2], P M Maleka[1], and R R Maphanga[1,2]
[1]Next Generation Enterprises and Institution Cluster, Council for Scientific and Industrial Research, P.O. Box 395, Pretoria 0001, South Africa.
[2]National Institute for Theoretical and Computational Sciences (NITheCS), Gauteng 2000, South Africa

E-mail: **Lmogakane@csir.co.za**

Abstract. Sodium-ion batteries have gained a lot of attention because they are much safer and less expensive. Recently, potential candidates for SIB were reported to include transition metal chalcogenides. In this study, the first-principle approach was used to investigate ternary $NaVS_2$ using different exchange correlation functions, GGA-PBE, GGA-PBEsol and LDA. The result revealed that the $NaVS_2$ phases are thermodynamically stable with $\Delta H_f < 0$. The calculated lattice parameters using GGA-PBEsol functional were found to be in good agreement with previously reported results. Furthermore, the material's electronic properties, namely the band structure and density of states calculations were performed and revealed that the material has a zero-band gap, displaying a metallic behavior. The Mulliken population analysis were investigated to understand the atomic behaviour of $NaVS_2$ phases which reveals that the present net charges (e) of Na, and S atoms indicate a charge transfer between the two atoms.

1. Introduction

In the past two decades, a significant progress has been made in cathode materials for sodium-ion batteries that operates at an ambient temperature. Poly-anionic materials, sodium layered oxides, framework oxides, NASICONs, sulfate-fluorides, and fluorides are a few examples of these materials [1, 2, 3, 4, 5]. The cathode material, being a crucial constituent of sodium-ion batteries (SIB), exerts a direct influence on the overall battery performance and imposes an upper limit on the energy density of the cell. In order for the cathode materials to be deemed suitable for widespread use, they must meet the following set of specifications. Firstly, the cathode material must possess a significant redox potential to sustain the electrodynamic force of the SIBs and generate a substantial output voltage [6, 7]. Additionally, the cathode material must exhibit a commendable energy storage capacity to efficiently store and release energy [7]. Gaining a comprehensive comprehension of the operational mechanisms of the cathode is of utmost significance in achieving the goal of obtaining a suitable cathode material. Transition metal chalcogenides (TMCs) such as MoS_2 and WS_2 were recently reported to be promising candidates for SIB. Previous studies have provided insight into the electrochemical properties of various binary disulfides, including VS_2, TaS_2, ZrS_2, and NbS_2 [6]. Among these materials, VS_2 shows notable characteristics owing to its inherent characteristics, namely, higher surface area, higher energy density, and extensive cycling performance, which are attributed to its conductivity resulting from its metallic behaviour. These properties suggest that the material VS_2 has good electrochemical characteristics which makes it a promising candidate for SIB since it offers more Na^+ storage sites and accelerate charge transfer [6]. Furthermore, it has been reported that VS_2 as the electrode of metal ion batteries (e.g., Li^+, Na^+, K^+, Zn^{2+}, Mg^{2+}, and Al^{3+}) enhances the electrochemical performance in terms of theory and experiment. The unique structure and properties of VS_2 make it an ideal host for the insertion/extraction of alkali metal ions such as Na, Li, and K [8, 9].

Understanding the workings of the cathode is very crucial to obtaining the correct cathode material. $NaVS_2$ is a compound that has garnered interest in the context of SIBs due to its unique properties such as serving as an electrode material in SIBs, typically functioning as the cathode. The compound undergoes electrochemical reactions during the charging and discharging processes, with sodium ions ($Na^{+)}$ intercalation and deintercalation within the structure. In electrochemical reactions, $NaVS_2$'s metallic conductivity helps promoting effective charge transfer. This feature enhances the battery's overall performance and cycling stability.

Density functional theory (DFT+U) was used to investigate structural, electronic, and magnetic properties of 2D VS_2 and VSe_2 monolayer. The stability of the phase of VS_2 and VSe_2 freestanding monolayers has been demonstrated by a comparison of ground-state energy, which is produced by strong electron correlations within vanadium d-shells [10]. There are limited studies on investigating TMCs as potential materials for SIB applications. Ge *et al.* [11] investigated the ternary TMD Ti_2PX_2 (X = S, Se, Te) structure using a DFT technique, and the Ti_2PTe_2 structure was reported as a promising candidate for SIBs applications. Furthermore, the study reported that the lattice constants and thickness decrease in atomic order of Te, Se and S. The properties of mostly 2D TMC compounds have been thoroughly examined by researchers; nevertheless, there is still a dearth of knowledge regarding ternary TMCs, particularly with regard to atomic modelling. In this work, DFT was used to investigate structural, electronic, and thermodynamic properties of bulk $NaVS_2$ compound for potential application in SIB.Several DFT functionals, including the local density approximation (LDA), Perdew-Burke-Ernzerhof (PBE), and PBESol, were used to compute heats of formation, lattice constants, band structures, and density of states (DOS).

2. Computational Method

All calculations were performed using Materials Studio software employing the CASTEP code [13]. Geometry optimizations were performed utilizing various exchange correlations, including the generalized gradient approximation (GGA) in Perdew-Burke-Ernzerhof [14]and PBEsol and the local density approximation. To determine the structural, electronic, and thermodynamic stability, calculations were performed using the plane-wave pseudopotential method [15]. From the convergence tests, the kinetic cut-off energy for all the different functionals is determined to be 600 eV [9]and with the k-points sampling 19 x 19 x 3. Structural optimisation was carried out using the Broyden-Fetcher-Goldforb-Shanno (BFGS) minimizations scheme algorithm. The lattice parameters were calculated and compared with previous reported values. Heats of formation energy (ΔH_f) per unit for $NaVS_2$ structure was calculated to determine the thermodynamical stability of the structures and is calculated using the analytical equation 1. The $E_{tot.}$ and E_i present the total energies of the system and individuals, respectively, and *i* denotes the atomic configuration and *n* is the number of atoms.

$$E_{HF} = E_{tot.} - \sum_i n_i E_i \quad (1)$$

3. Results and Discussion

3.1. Structural properties and heats of formation

Figure 1 presents the optimized atomic structure of $NaVS_2$, exhibiting a crystal structure known as Caswell Silverite that crystallizes into the trigonal $\bar{R}3m$ space group. In this structure sodium cations (Na^{1+}) are coordinated with six sulfur anions (S^{2-}) in a manner that results in the formation of NaS_6 octahedra. These octahedra share their corners with six equivalent vanadium sulfide (VS_6) octahedra, their edges with six equivalent NaS_6 octahedra, and their edges with six equivalent VS_6 octahedra. The compound S^{2-} is chemically connected to three Na^{1+} and three V^{3+} atoms in a manner that results in the formation of a combination of Na_3V_3S octahedra, where the bonding occurs through both edge-sharing and corner-sharing interactions.

The octahedra that shares corners are not in a slanted configuration. The optimized lattice parameters of the $NaVS_2$ compounds were computed using the different functionals for comparison and are listed in Table 1. Upon analysis of the lattice parameters obtained by different functionals, it was observed that the PBE functional exhibited larger lattice parameters as compared to PBEsol and LDA. From our calculations it was clear that the PBEsol exchange correlation lattice parameters correlate with the reported experimental results. The heats of formation (ΔH_f) were calculated using equation 1 to determine the thermodynamic stability of the structure as predicted by the different exchange correlation functionals. The computed values of ΔH_f were found to be negative, indicating that the $NaVS_2$ is thermodynamically stable. In other words, this makes $NaVS_2$ thermodynamically advantageous due to its energy release during the formation process. It can be concluded that the $NaVS_2$ material has a solid solution and that it is experimentally feasible. Notably, the PBE exchange correlation functional exhibited greater stability with lower ΔH_f = -1.37 eV/atom. A larger negative value denotes more stable and favourable thermodynamic state. It was discovered that the order of the current thermodynamic stability of $NaVS_2$ follow $\Delta H_f^{PBE} > \Delta H_f^{PBEsol} > \Delta H_f^{LDA}$

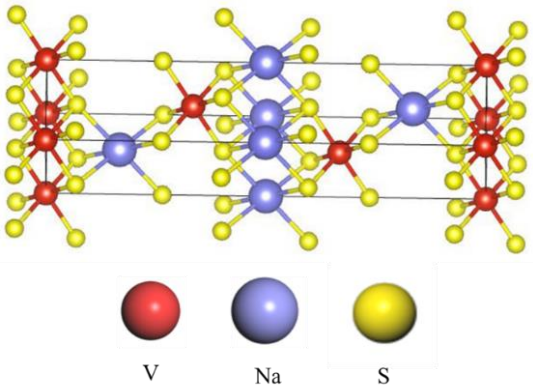


Figure 1. The optimized crystal structure of $NaVS_2$ compound.

Table 1. Computed lattice parameters (Å), heats of formation, (ΔH_f eV/atom) of $NaVS_2$ crystal structure obtained using different functionals.

Structure	Parameters(s)	Experimental [12]	PBE	PBESol	LDA
$NaVS_2$	a	3.346	3.414	3.346	3.308
	c	21.02	20.405	20.268	20.027
	a/c	0.159	0.167	0.165	0.165
	Volume	-	206.007	196.512	189.856
	ΔH_f	-	-1.37	-1.31	-1.27

3.2 Electronic properties

3.2.1 Band structures

The electronic structure of a material is one of the significant aspects in solid state to explain a variety of material features including opto-electronic properties, electronic thermal conductivity, and electrical conductivity. The electronic properties of the ternary $NaVS_2$ compounds are described by calculating the electronic band structures as well as the partial densities of states for the PBE, PBE-sol and LDA exchange correlation functionals. The electronic band structures are depicted in Figure 2. The Fermi level, also known as E_F, is denoted by the broken red line in the horizontal. The calculations in Figure 2 reveals that there is no band gap of $NaVS_2$ at the Fermi level since there is an overlapping of the valence band and conduction band. These finding were observed using all three functional, GGA-PBE, GGA-PBEsol and LDA. The material $NaVS_2$ exhibits a metallic behaviour as indicated above. $NaVS_2$ being metallic in nature is an advantage for battery applications since metallic materials typically offer better electrical conductivity and may lead to improved performance in certain battery electrode applications.

The valence and conduction bands of $NaVS_2$ would overlap meaning that the material would conduct electricity efficiently and exhibit high electron density at the Fermi level.

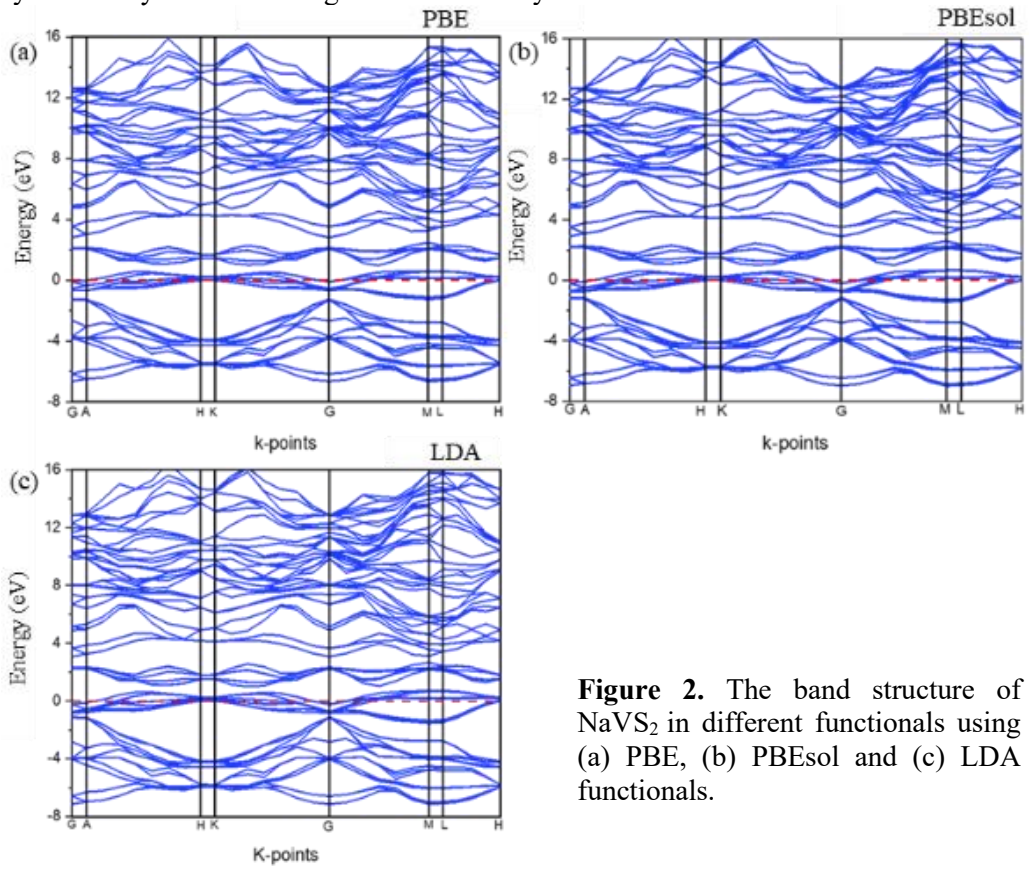


Figure 2. The band structure of $NaVS_2$ in different functionals using (a) PBE, (b) PBEsol and (c) LDA functionals.

3.2.2 Density of states

Figure 3 depicts the total and orbital projected partial density of states (PDOS) plots for the $NaVS_2$ compound. The DOS for $NaVS_2$ demonstrates a continuous distribution of electronic states throughout the energy range, characterized by a non-zero DOS at the Fermi level. The Fermi level is a fundamental concept in solid-state physics that characterizes the energy level at which the electronic states of a material are occupied when the temperature approaches absolute zero. The existence of electronic states available for conduction in $NaVS_2$ is shown by the non-zero density of states at the Fermi level. The presence of several accessible states facilitates the efficient mobility of electrons, hence contributing to the notable electrical conductivity exhibited by the material. Consequently, there is a notable increase in the abundance of electronic states available for conduction, resulting in a heightened level of electrical conductivity compared to VS_2 [6]. The valence and conduction bands exhibit an overlapping region, wherein the existence of states near the Fermi level facilitates the efficient mobility of electrons, leading to a notable enhancement in electrical conductivity. The findings indicate that V-3d exhibits contributions on both sides near the Fermi level. The overlapping and the hybridization of electronic peaks were observed at approximately of -7 to -2 eV which correspond to S-4p and V-3d. This represents mainly the atomic interaction and bonding of the atom in the system. In addition, it was found that the -d orbital state contributed significantly to the Fermi level compared to -s and -p orbitals contributed less. This suggests that V-3d states are primarily responsible for the system's metallic behaviour. There is also a contribution from Na-2p orbital in the conduction band far from the band edge around 4-5.

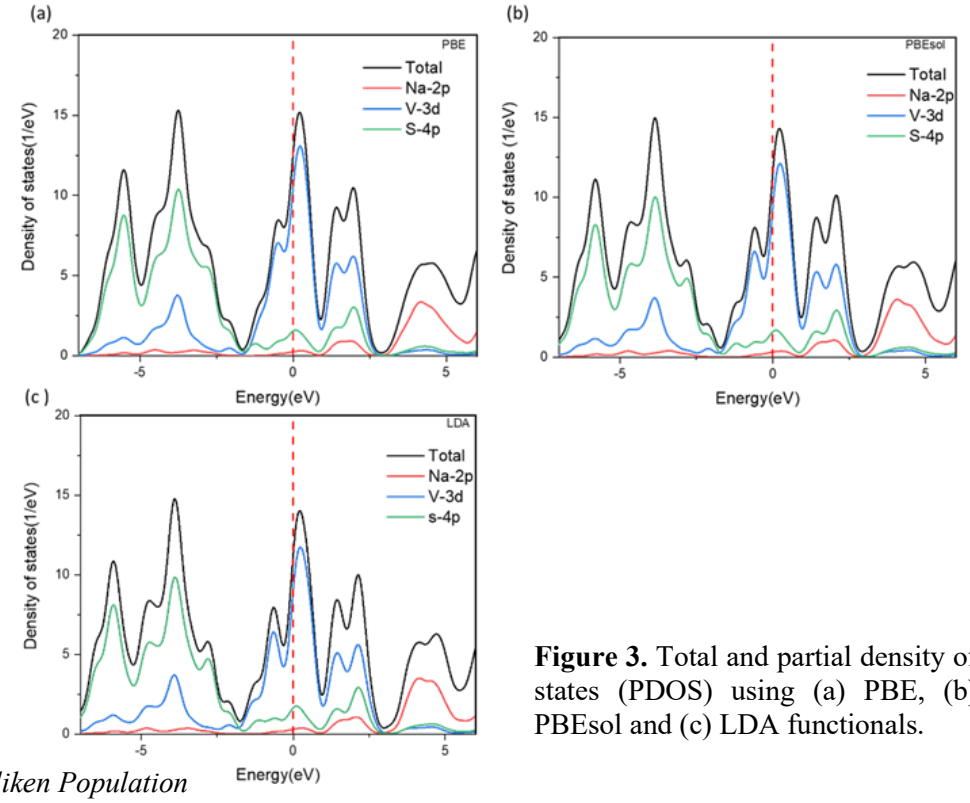


Figure 3. Total and partial density of states (PDOS) using (a) PBE, (b) PBEsol and (c) LDA functionals.

3.2.3 Mulliken Population

Further exploration of the electronic properties of $NaVS_2$ from Na, V, and S atoms can be obtained from the Mulliken population analysis. The partial atomic charges of Na, V, and S atoms are shown in Table 2 and the analysis reveals that the total charge of V atoms is greater than that of Na and S atoms, with the difference contributed from the 3d states of V atoms. Table 2 reveals that S atoms have negative charges (behave as anions) whereas Na and V atoms have positive charges (behave as cations). The results shows that the material $NAVS_2$ has a partial charge of about +0.50 to +0.60e for Na, −0.45 to −0.79e for S, and −0.29 to −0.42e for V, using these three functionals PBE, PBEsol and LDA, respectively. these charges are all less than +1 for V and S. The present net charges (e) of Na, and S atoms indicate a charge transfer between the two atoms. The effective valence is also represented in detail in Table 2 and is defined as the difference between the formal ionic charge and the Mulliken charge on the cation species. It is utilized to quantify ionicity or covalency of a compound. A covalency bond corresponding to an effective value greater than zero, while effective values of zero is an ionic bond. In the $NaVS_2$ compound, V atom has a higher effective valence compared to Na and S atoms.

Table 2. The computed orbital charge, Milliken charge and effective charge of $NaVS_2$ using three different exchange correlation functionals.

Functionals		s	p	d	Total	Charge	Effective
PBE	Na	2.13	6.28	-	8.40	+0.60	0.36
	S	1.81	4.53	-	6.34	-0.34	1.66
	V	2.41	6.71	3.80	12.91	+0.09	4.91
PBEsol	Na	2.15	6.30	-	8.50	+0.50	0.50
	S	1.80	4.48	-	6.28	-0.28	1.72
	V	2.40	6.74	3.86	13.00	0.00	5.00
LDA	Na	2.15	6.30	-	8.45	+0.55	0.45
	S	1.80	4.48	-	6.28	-0.28	1.72
	V	2.38	6.71	3.90	13.00	0.00	5.00

4. Conclusion

The structural, electronic, and thermodynamic properties of $NaVS_2$ was performed successfully using DFT. These properties were investigated using three different exchange correlation functionals, GGA-PBE, GGA-PBEsol, and LDA. The lattice parameters calculated were found to be in good agreement with previously reported results using GGA-PBEsol functional. All the functionals that were considered and analyzed have shown to have negative values of ΔH_f, which indicate the presence of an exothermic solid solution, suggesting that $NaVS_2$ structure is thermodynamically stable. Moreover, it was observed that GGA-PBE functional predicts the lowest heat of formation compared to PBEsol, and LDA. The electronic band structure and density of states demonstrated that $NaVS_2$ has a metallic behaviour making this material suitable for SIB.

5. Acknowledgements

We would like to thank the Centre for High Performance Computing (CHPC) for using their computing facilities. We would also like to acknowledge the Department of Science and Innovation (DSI) Foundational Digital Capability Research and Council for Scientific and Industrial Research for financial support.

References

[1] Takada K 2013 *Acta Mater.* **61** 759-770.

[2] Sakuda A, Takeuchi T, Shikano M, Sakaebe H, Kobayashi H 2016 Front Energy Res. **4** 25.

[3] Zheng Y, Yao Y, Ou J, Li M, Luo D, Dou H, Li Z, Amine K, Yu A and Chen Z 2020 Chem. Soc. Rev. **49** 8790.

[4] Dima R S, Maleka P M, Maluta N E an Maphanga R R 2022 *Mater.* **15** 5280.

[5] Dima R S, Maleka P M, Maluta N E and Maphanga R R 2022 *Mater. Today: Proc.* **62** 7-11.

[6] Xu S, Wu X, Li Y 2014 *Chin. Phys. B.* **23** 118202.

[7] Asl H Y, Manthiram A 2020 *Sci.* **369** 140-141.

[8] Yu D, Pang Q, Gao Y, Wei Y, Wang C, Chen G and Du F 2018 Energy Storage Materials **11** 1–7

[9] Xie X C, Shuai H L, Wu X, Huang KJ, Wang L N, Wang R M, Chen Y 2020 J. Alloys Compd. **847** 156288.

[10] Popov Z I, Mikhaleva N S, Visotin M A, . Kuzubov A A,Entani S, Naramoto H, Sakai S, Sorokin P and Avramov P V 2016 *Chem. Phys.*, DOI: 10.1039/C6CP06732H.

[11] Ge B, Chen B and Li L 2021 *Appl. Surf. Sci.* 550.

[12] Chen W, Zhang J M, Xia Q L, Nie Y Z and Guo G H 2020 *Chem. Phys.* **22** 16007-16012.

[13] Payne M C, Teter M P, Allan D C, Arias T A and Joannopoulos J D 1992 *Rev. Mod. Phys.* **64** 1045-1097.

[14] Perdew J P, Burke K, Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865.

[15] Hohenberg P, Kohn W 1964 *Phys. Rev.***136** 864-867.

[16] Abraham K M 1982 *Solid State Ionics.* **7** 199-212.

[17] Schuster J, He G, Mandlmeier B, Yim T, Lee K T, Bein T 2012 *Angew. Chem.* **51** 3591.

Optimization of cathodic deposition voltage on the CdZrS thin films

TKW Mohapi[1]*, AU Yimamu[1,2], KG Tshabalala[1], and SJ Motloung[1]

[1]Department of Physics, University of the Free State, Private Bag X13, Phuthaditjhaba, 9866, QwaQwa campus, South Africa
[2]Department of Physics, Adama Science and Technology University, Adama, Ethiopia

E-mail: **tshepomohapi54@gmail.com, tlousj@gmail.com**

Abstract. Cadmium zirconium sulfide (CdZrS) thin films were synthesized by the two-electrode electrodeposition method at different growth cathodic deposition voltages 1540-1565 mV with intervals of 5 mV on the conductive fluorine-doped tin oxide (FTO) glass substrates. The structural, surface morphology, elemental composition, surface roughness, and optical properties of the resulting films were investigated using X-ray diffraction (XRD), scanning electron microscopy (SEM), energy-dispersive X-ray spectroscopy (EDS), scanning probe microscopy (SPM), and ultraviolet-visible (UV–vis) spectrophotometry respectively. The XRD results revealed that CdZrS thin films exhibited cubic structure. Furthermore, the results indicate that the intensity of the peak at $2\theta = 32.84^{o}$ fluctuated with an increase in deposition voltage. The SEM images show that the film completely covered the glass substrate and that the grain size changed as the cathodic deposition voltage increased. EDS confirmed the presence of all elements forming the desired compound CdZrS and its analysis shows that the composition of Cd, S, and Zr varies with deposition voltage. SPM results show that average surface roughness was found in the range of 13.7-28.2 nm as deposition voltage increased. UV-Vis results indicate a decrease in absorbance and an increase in optical bandgap in the range of 2.55-2.92 eV as deposition voltage increases.

1. Introduction

Two-dimensional nanomaterials known as thin films have demonstrated potential for energy conversion [1]. The growth methods and deposition conditions, such as growth rate, electrode spacing, substrate temperature, thickness, as well as pre-and post-deposition activation treatments like doping, annealing temperature, annealing atmosphere, halide treatment, etc., all have an impact on the physical properties of thin films. Compound semiconductors with direct energy transition, such as CdTe, PbS, Cu_2S, CdS, ZnTe, and ZnS, are well suited for the design of both single-junction and multi-junction devices. Cadmium sulfide (CdS) is an II-VI semiconductor with outstanding thermal characteristics, chemical stability, potential optical absorption, and a direct bandgap of 2.42 eV at 300 K. In modern thin film, solar cells with CdTe and CIGS absorbent layers, CdS thin films are the ideal choice for window layers. CdS is routinely synthesized using a variety of techniques such as chemical bath deposition (CBD) [2], successful ionic layer adsorption and reaction (SILAR) [3], pulsed laser deposition (PLD) [4], and electrodeposition [5]. With regard to device applications, each method has its own advantages and disadvantages. Among the above-mentioned methods, the most effective method for synthesizing CdS and ZnS has been found to be low-cost electrodeposition procedures [6]. CdS electrodeposited on

several substrates, including glass, silicon, and fluorine-doped tin oxide (FTO), has been widely researched. The substrate chosen is determined by the individual application and desired features of the CdS films. The electrodeposition technique employs an electrochemical cell with a two-electrode system to deposit a CdS layer onto a substrate by providing an electric potential. The two-electrode system is simple, inexpensive, and adaptable, whereas the three-electrode system enables exact potential control and is appropriate for investigations requiring accurate potential readings or control. Continued research and development in this field are expected to improve the performance of CdS-based devices and broaden their potential applications. According to the demands of the related technology, the properties of CdS films can be changed, and the performance parameters of the device can be updated. According to numerous studies, these methods include cationic and anionic dopants that enhance the optoelectronic characteristics of CdS films. There are a number of potential effects and improvements to the material when zirconium (Zr) is added to CdS. The incorporation of Zr is expected to enhance the stability of the material, and electrical properties by adding impurity energy levels within the bandgap, which affects the conductivity of a material and carrier concentration.

It is important to note that the precise effects and improvements brought about by Zr addition to CdS might vary depending on a number of variables, including the technique of integration, and Zr concentration. The CdZrS thin films must typically be improved through additional study and testing in order to serve a particular application. A decrease in average surface roughness, and an increase in the bandgap of ZnO thin films was observed with the addition of Zr concentration [7]. A similar trend is observed when Zr is incorporated into the CdS matrix which increases the bandgap of CdZrS thin films. The average surface roughness decreased at lower voltages but increased at a higher voltage.

1. Material and Method

2.1 Deposition of CdZrS

Thin films of CdZrS were electrodeposited on fluorine-doped tin oxide (FTO) glass substrates. These substrates were cut into 2.5 cm × 2.5 cm. The cleaning of substrates was followed based on previous reported work [8]. All these starting materials were of analytical reagent-grade chemicals purchased from Sigma-Aldrich. In 400 ml of deionized water, an electrolytic solution containing 0.3 M of cadmium acetate dihydrate (Cd $(CH_3\ COO)_2).2H_2O$ as a source of cadmium (Cd), 0.03 M of sodium thiosulphate ($Na_2S_2O_3$) as a source of sulfur (S), and 0.025 M of zirconium (iv) acetate hydroxide ($Zr_3\ (C_6H_{18}O_6\ )_2$) as a source of zirconium (Zr) was prepared. The chemicals were stirred for 5 hrs to dissolve all of the substances. In order to electrodeposit CdZrS thin films, FTO glass substrates were fastened to a high−purity carbon holder using an insulating polytetrafluoroethylene (PTFE), which also serves as a working electrode. The FTO (now a working electrode) was rinsed with deionized water, dried in the air, and placed in the proper electrolyte for the electrodeposition of the necessary CdZrS thin films.

2. Result and Discussion

3.1 Structural properties

Figure 1(a) shows XRD patterns of CdZrS thin films deposited in the 1540-1565 mV voltage range. Depending on the deposition conditions, CdZrS exists in a cubic (C) crystal structure. Diffraction peaks with diffraction planes at angles of 2θ values, 27.56º (111) C, 32.82º (200) C, and 44.84º (620) C are indexed on the spectra. In figure 1(a) and (b), according to the accepted JCPDS card no. 02-0454 and 21-0829 the detected peaks support the existence of the cubic phase of CdZrS. At the cathodic deposition voltage of 1545 mV, the most intense peak is observed at 2θ = 32.84º. Larger and better-aligned crystallites contribute more to the diffraction process, resulting in higher peak intensity in figure 1(b). A stronger peak can be produced if the material has a high degree of crystallinity and contains big, well-defined crystallites. Zr^{4+} has an ionic radius of (0.74 Å) which is less than that of Cd^{2+} (0.97Å), which when added into the CdS lattice can cause strain and lattice distortions. The crystal

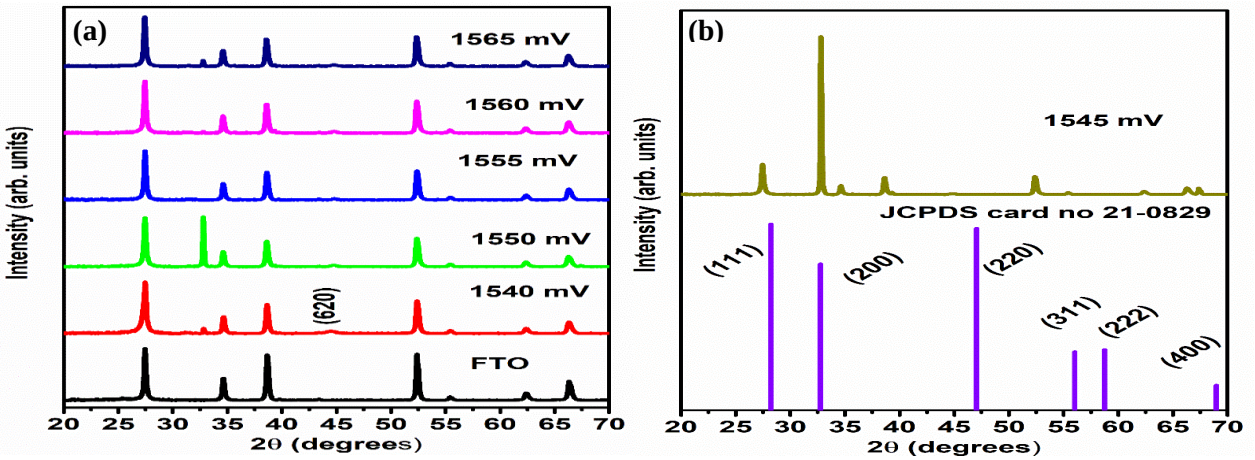


Figure 1. XRD patterns for CdZrS thin films deposited at various cathodic deposition voltages.

structure and peak intensity of CdZrS thin films can be impacted by Zr inclusion. More Zr may be absorbed into the film as the deposition voltage increases, leading to a lower peak intensity.
By enlarging the diffraction peaks or weakening the overall crystalline quality of a material, these strain effects can cause a drop in peak intensity. The CdZrS diffraction peaks can become less intense due to the presence of new phases or the segregation of Zr-rich regions.

The crystallite size, *D*, of the CdZrS thin films, was estimated at a diffraction plane of (620) C, using Scherrer's equation [9]. The crystallite size of CdZrS thin films was estimated at various deposition voltages as shown in table 1. The rate of metal ion reduction at the cathode surface increases with increasing deposition voltages. A greater flux of metal ions is deposited into the substrate as a result of the enhanced reduction rate. The crystals can grow bigger thanks to the quicker development rate, which leads to bigger crystallites as observed in table 1.

Table 1. The crystallite size for CdZrS thin films deposited at various cathodic deposition voltages.

2θ (Degrees)	Thin film	Deposition voltage (mV)	Crystallite size, *D* (nm)
44.51	CdZrS	1540	8.59
44.72	CdZrS	1545	14.53
44.86	CdZrS	1550	12.18
44.80	CdZrS	1555	16.41
44.78	CdZrS	1560	16.42
44.84	CdZrS	1565	11.11

3.2 Morphological property

Figure 2 (a-c) shows the SEM images for CdZrS thin films deposited at various cathodic deposition voltages. The nucleation process is accelerated at lower voltages (1540 mV), which causes a greater quantity of tiny grains to develop. The balanced growth in all directions gives these grains a tendency to have a more circular form. At 1550 mV decrease in transparency and an increase in light reflection occur as a result of the increased light scattering of the film. The film appears white due to this light reflection from figure 2(b). Additionally, in figure 2(c), the deposition process has a tendency to be more homogeneous across the substrate surface at higher voltages (1565 mV). As a result, there is greater uniformity in the dispersion of atoms and less growth of gleaming large grains. Shiny gigantic grains often appear when there is a localized build-up of deposition, which can happen at lower voltages or in regions with irregular current distribution.

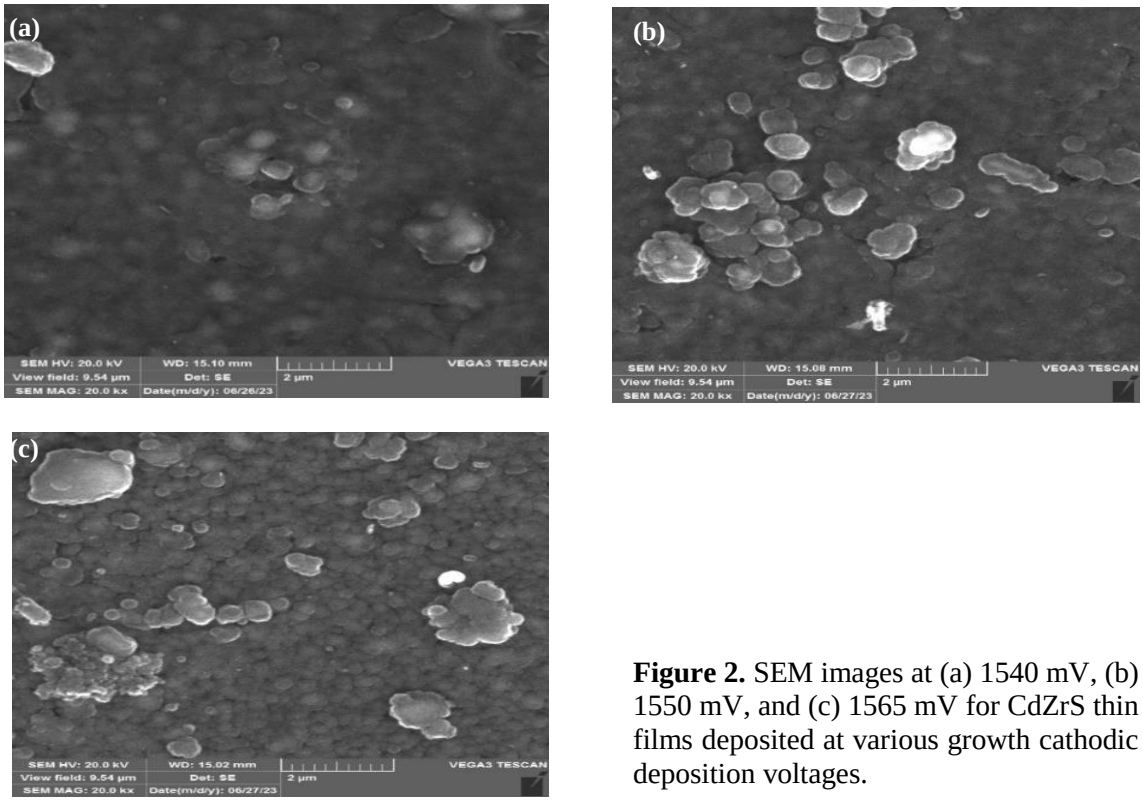


Figure 2. SEM images at (a) 1540 mV, (b) 1550 mV, and (c) 1565 mV for CdZrS thin films deposited at various growth cathodic deposition voltages.

3.3 Elemental composition

Compositional analysis of CdZrS thin films was performed using EDS to determine the elemental composition of thin films, as shown in figure 3.

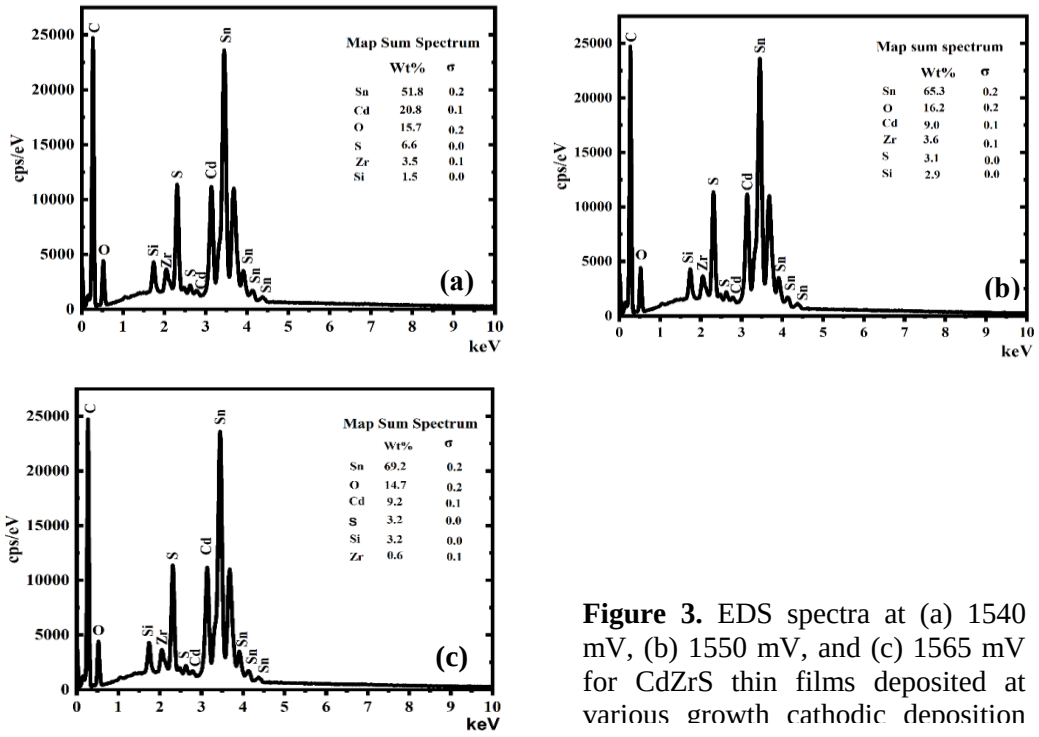
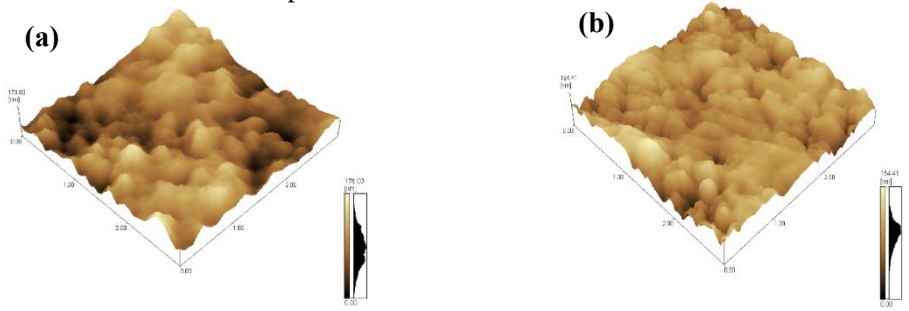


Figure 3. EDS spectra at (a) 1540 mV, (b) 1550 mV, and (c) 1565 mV for CdZrS thin films deposited at various growth cathodic deposition

The results confirm the presence of Cd, Zr, and S atoms in the films. The glass substrate contains components such as Si, O, and Sn. Figure 3 also shows how the atomic composition of Cd, Zr, and S changed when the deposition voltage varied from 1540 to 1565 mV. When the deposition voltage is at 1540 mV as shown in figure 3(a), Cd and S have higher atomic composition which makes the Cd and S film rich. The increasing electric field at higher voltages from 1540 to 1565 mV can cause accelerated migration of Cd and S ions away from the cathode surface. As a result, the atomic percentage of Cd decreases, pointing to the replacement of Cd ions by Zr ions as deposition voltage increases. When CdZrS thin films are electrodeposited, complicated interactions between the electrodeposition conditions, the composition of the deposition bath, and the electrochemical processes involved result in fluctuations in the atomic percentage of Zr as the deposition voltage rises. However, at higher voltages, a further phenomenon known as surface sputtering can occur. Surface sputtering happens when high-energy electrons attack the cathode material, causing the ejection of surface atoms. This ejection of atoms may include previously deposited Zr atoms on the surface.

3.4 Scanning probe microscopy (SPM)

Three-dimensional SPM images of CdZrS thin films are shown in figure 4(a-c) deposited in the range of 1540-1565 mV. The grains were evenly spaced over the surface and also demonstrated a high degree of film compactness. The average roughness (R_a) value in the current study is 23.3 nm when the CdZrS thin film was deposited using the ED technique at a deposition voltage of 1540 mV, as shown in figure 4(a). However, the R_a value decreased to 13.7 nm at 1550 mV in figure 4(b). Furthermore, the growth voltage of 1565 mV in figure 4(c), shows that the R_a value suddenly increased again to 18.2 nm. The reason might be an increase in grain size. A lower average roughness is often preferred for window materials since it denotes a smoother surface. Better optical transparency, less light scattering, and enhanced window material performance can all be attributed to a smoother surface. An average surface roughness of 13.7 nm as compared to 23.3 nm might be preferable, however, if the window material is used for situations where good optical clarity and less light scattering are essential. The optical qualities of smoother surfaces are often superior.



3.00×3.00 [um] Z 0.00 – 176.03[nm]

3.00×3.00 [um] Z 0.00 – 154.41[nm]

3.00×3.00 [um] Z 0.00 – 245.11[nm]

Figure 4. SPM images at (a) 1540 mV, (b) 1550 mV, and (c) 1565 mV for CdZrS thin films deposited at various cathodic deposition voltages.

3.5 Optical properties

Figure 5 shows spectra of absorbance as a function of wavelength, and absorbance square as a function of photon energy for CdZrS thin films deposited at various cathodic deposition voltages. The absorbance value drops as the deposition voltage increases, as shown in the spectra from figure 5(a). This is because

a higher deposition voltage causes more Zr to be incorporated into the CdS matrix. The incorporation of Zr into the CdS matrix, however, increases when the deposition voltage increases to 1545 mV–1565 mV. This is seen as the absorbance value drops with an increase in deposition voltages. The larger crystallite sizes also contribute to this decrease in the absorbance as deposition voltage increases from 1545-1565 mV. Lower absorbance occurs from the lower surface area since there are fewer sites accessible for light absorption. Less scattering suggests that bigger crystallites scatter light less frequently than smaller ones. Calculating the band gap involves extrapolating the graph's straight-line region to the photon energy axis ($(Abs)^2 = 0$) in figure 5(b). The incorporation of more Zr into the matrix, which influences the intensity of the characteristic peaks and the absorbance of films, results in an increase in the energy band gap of the CdZrS thin films as deposition voltage increases. The energy band gap of CdZrS increased from 2.55, 2.70, 2.75, 2.78, 2.88, and 2.92 eV respectively, while the deposition potential increased. The energy band gap of the substance grows in direct proportion to the amount of Zr integrated into the CdS matrix as the deposition voltage increases.

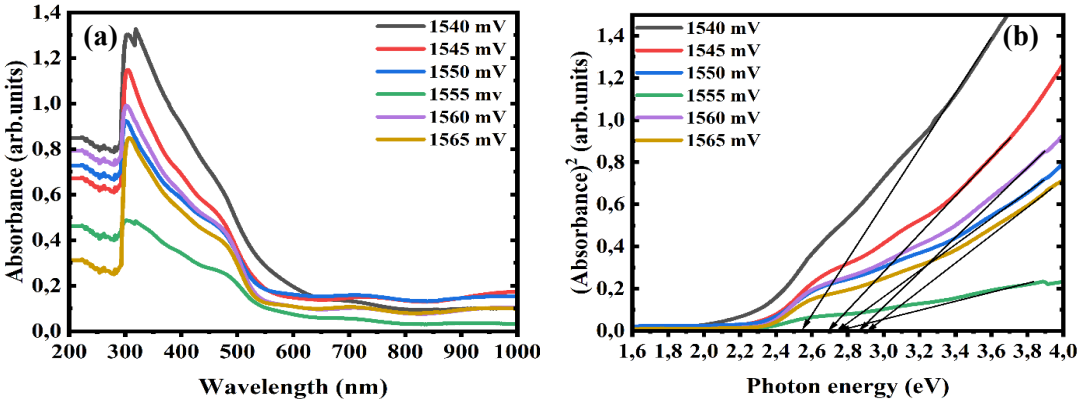


Figure 5. UV-vis spectra for (a) absorbance as a function of wavelength, and (b) absorbance square as a function of photon energy for CdZrS thin films deposited at various cathodic deposition voltages.

4. Conclusion

CdZrS thin films have been successfully deposited on FTO glass substrates. Increasing cathodic deposition voltage resulted in an increase in crystallite sizes. The morphology changed with variation of deposition voltages. EDS spectra confirmed all the expected elements. The average roughness of the material has the lowest value at 1550 mV. For the window material, a smooth surface is desired for thin films in solar cells. It is important for absorbance to decrease in a window material and 1555 mV has a lower absorbance compared to other voltages. An increase in bandgap as cathodic deposition voltage increases is an important factor in achieving high-quality window materials that have higher transparency and minimum light absorption. Overall, these results imply that CdZrS thin films can be employed as windows for optoelectronic applications when produced at higher cathodic deposition voltages.

References

[1] Faremi A A *et al*. Dec.2022 *Results in Engineering* vol. **16**
[2] Abbas K, Aamir M, Mehmood R F, Sher M and Akhtar J 2021 *Materials Today: Proceedings Elsevier Ltd* pp 339–344
[3] Deshmukh S G, Kheraj V and Panchal A K 2018 *www.sciencedirect.comwww.materialstoday.com/proceedings* [Online]. Available: 2214-7853
[4] Martínez-Landeros V H, Hernandez-Como N, Gutierrez-Heredia G, Quevedo-Lopez M A and Aguirre-Tostado F S Jul. 2019 *Thin Solid Films* vol. **682** pp 24–28
[5] Maricheva J, Bereznev S, Naidu R, Maticiuc N, V. Mikli V and Kois J Nov. 2016 *Mater Sci Semicond Process* vol. **54** pp 14–19

[6] Echendu O K and Dharmadasa I M 2015 *Energies (Basel)* vol. **8** pp 4416–4435
[7] Tsay C Y and Fan K S August 2008 *Mater Trans*, vol. **49**, no. 8, pp. 1900–1904
[8] Yimamu A U *et al.* Jun. 2023 *Mater Today Commun*, vol. **35**
[9] Lim D J, Marks N A and Rowles M R Jun. 2020 *Carbon N Y* vol. **162** pp 475–480

A comparative study on Mn-doped $CsPbI_3$ and $CsPbBr_3$ using first principle-based DFT

PM Maleka[1,2], DT Tshwane[1,3], RS Dima[1], L Mogakalane[1], OM Ntwaeaborwa[2] and RR Maphanga[1,3]

[1]Next Generation Enterprises and Institution Cluster, Council for Scientific and Industrial Research, P.O. Box 395, Pretoria 0001, South Africa

[2]School of Physics, University of the Witwatersrand, Private Bag x 3, Wits, 2050, South Africa

[3]National Institute for Theoretical and Computational Sciences (NITheCS), Gauteng 2000, South Africa

E-mail: **pmaleka@csir.co.za**

Abstract. Doping of $CsPbX_3$ (where X = I and Br) perovskite with transition metals modifies and stabilizes the material. Mn has become important in these transition metal doping halides perovskites because of its ability to control fundamental properties and enhance structural stability and luminescence of materials. In this study, the first principle-based DFT was employed to investigate the manganese (Mn) incorporation effect on the structural, optoelectronic, and mechanical properties of $CsPbX_3$ (where X = I and Br), materials for solar cell application. For both $CsPbBr_3$ and $CsPbI_3$, the results indicate that the substitution Pb atom with the Mn atom reduces the bond length and lattice constants due to Mn having a smaller ionic radius compared to the Pb atom. Both these undoped $CsPbX_3$ (X=I and Br) have a direct band of 1.488 eV and 1.688 eV, whereas Mn-doped $CsPbI_3$ and $CsPbBr_3$ materials have indirect bandgap values of 1.274 and 1.350 eV, respectively. The optical properties of these calculated compounds indicate that these materials can be used in solar cell and optoelectronics applications. The mechanical properties were also calculated, and the results show that the materials are mechanically stable and ductile.

1. Introduction

The halide perovskites have garnered significant interest due to their extensive utilization in various kinds of photovoltaic and optoelectronic devices, such as photodetectors, solar cells, and spintronic devices [1-2]. These perovskites have outstanding optoelectronic and photovoltaic properties including direct bandgap, enhanced electron mobility, significant light absorption, and long diffusion length [3, 4]. B-site doping may improve the absorption and distance of B-X bonding, thereby increasing the environmental endurance and phase stability. However, partial substitution with other ions can be difficult due to the larger formation energy of the B-site. Manganese (Mn^{2+}), cobalt (Co^{2+}), zinc (Zn^{2+}), germanium (Ge^{2+}), strontium (Sr^{2+}), and nickel (Ni^{2+}) cations, known for their relatively smaller ionic radii, have been introduced into the substitutional or interstitial sites within cesium lead halide perovskite structures [5]. These cations serve different purposes, including improving optoelectronic properties, phase stability, passivating defects, suppressing ion migration, manipulating the crystallization process, and optimizing energy band alignment. Transition metal doping is regarded as a technique that is used to introduce transition metal ions into the $CsPbX_3$ (X=Cl, Br, and I) perovskite material to modify the opto-electronic properties of the material, which may enhance the performance of solar cells. Each transition metal ion has different properties and effects on the material, and the choice of the transition metal ion will depend on the specific application. One of the key advantages of

transition metal doping is that it can introduce new energy levels into the material, which can broaden the absorption spectrum of the perovskite [6]. This can lead to an increase in the absorption of light, which may enhance the solar cell's performance. Additionally, transition metal doping can introduce new recombination pathways, which can improve the efficiency, and decrease charge carriers' recombination of solar cells.

Doping can improve the stability of the $CsPbX_3$ (X=Cl, Br, and I) perovskite material by introducing new defects and impurities into the material, which can act as recombination centres and reduce the stability of the material. However, by carefully controlling the amount and type of transition metal ions that are introduced into the material, it is possible to decrease the formation of these defects and impurities and improve the stability of the material. In this study, first-principles density functional theory (DFT) was used to investigate the structural, electronic, optical, and mechanical properties of Mn-doped $CsPbX_3$ (X=I and Br) perovskites.

2. Computational Method

The density function theory calculations were performed using the Vienna Ab initio Simulation Package (VASP) with the generalized gradient approximation (GGA) functional [7-8]. The GGA+U method was used to calculate the properties in all analyzed doped systems to correct the strong Coulomb interactions of 3d electrons. Doping atom Mn had Hubbard U_{eff} value set as 3.5 eV. The total energy convergence threshold of 10^{-8} eV was used. All the calculations used the cut-off energy of 500 eV with the Brillouin-zone integrations of 6 × 6 × 6 for Monkhorst-Pack k-point mesh [9]. Moreover, the spin-polarized mode was considered in all calculations. The electronic configuration for transition metal Mn is [Ar] $3d^5$ $4s^2$. The absolute maximum force and convergence tolerance were set at 0.03 eV/atom and $1x10^{-6}$ eV/atom, respectively.

3. Results and Discussion

3.1. Structural properties

The investigated transition metal doped materials belong to the cubic crystal structure with Pm-3m (no.221) space group. The 3D Visualisation for Electronic and Structural Analysis (VESTA) [10] software was used to visualize the crystal structures. The pure compound unit cell is composed of five atoms, with the Cs atom occupying the body-centered 1b Wyckoff site (0.5, 0.5, 0.5) of the crystal. The B-site cations occupy the corners of the cubes at the 1a Wyckoff site (0, 0, 0), while I atoms occupy the face cantered positions at the 3d Wyckoff site (0.5, 0, 0). Figure 1 represents the undoped and Mn doped $CsPbX_3$ (X = I or Br) supercell crystal structure. The impurity added in the pure supercell of $CsPbX_3$ (X = I or Br) by substituting the Pb atom with Mn transition metal dopant with a concentration doping of 12.5 %. The computed structural parameters and bandgap energy of pure and doped $CsPbX_3$ (X = I or Br) perovskite compared with available experimental and experimental values are shown in Table 1. Table 1 shows that replacing Pb with Mn atom decreases the volume and lattice constants for both $CsPbI_3$ and $CsPbBr_3$. This is due to the ionic radius difference between the dopant and host, where Mn (0.83 Å) has a smaller ionic radius compared to the Pb (2.20 Å) atom. This change in lattice parameters is also explained by the electron configuration of the transition metal (Mn), which replaces the Pb atom in both halides' materials. The introduction of 3d electrons in the valence band (VB) and conduction band (CB) close to the Fermi level alters the crystal symmetry. Their bond lengths and angles of pure and Mn doped $CsPbX_3$ (X = I or Br) perovskite was also calculated and compared in Table 1. The bond length of these materials demonstrated that Pb-I and Pb-Br increase with the average distance of 3.312 Å because of the lattice distortion and gives rise to Mn-I and Mn-Br bond lengths between 2.898 - 3.009 Å, respectively as seen in Table 1. The bond angles do not change even after the incorporation of manganese atom.

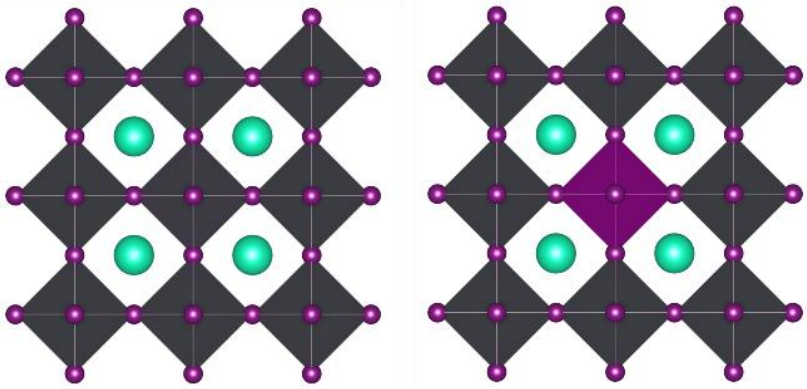


Figure 1. The polyhedron crystal structure supercell (2 x 2 x 2) of (a) undoped and (b) Mn-doped (purple octahedron) $CsPbX_3$, X=I and Br perovskites.

Table 1. The calculated lattice parameters (Å), volume (Å^3), angle (°), bond length (Å), and bandgap of pure and Mn doped $CsPbI_3$ and $CsPbBr_3$.

	Lattice constant (Å)			Volume (Å^3)	Angle (°)	Bond Length (Å)	Band gap (eV)
	Cal		Exp				
$CsPbI_3$	a = 6.398	a = 6.38 [11] a =12.791	a = 6.29 [12]	215.40	α=β=γ=90.00	Pb-I = 3.199	1.488 [11]
Mn-$CsPbI_3$	a = 6.341	a = 6.335 [5]		210.98	α=β=γ=90.00	Pb-I = 3.332 Mn-I = 3.009 Mn-I=3.0 [3]	1.274 1.24 [5]
$CsPbBr_3$	a = 5.949	a = 5.98 [13]	a = 5.870	210.50	α=β=γ=90.00	Pb-Br = 3.126	1.688 1.796 [13]
Mn-$CsPbBr_3$	a = 5.914			206.78	α=β=γ=90.00	Pb-Br = 2.969 Mn-Br = 2.629	1.350

3.2. Electronic properties

The electronic band structure and density of states (DOS) of a material gives a better understanding of the electronic behaviour of the material. The calculated band structure of pure and Mn doped $CsPbI_3$ and $CsPbBr_3$ along the high symmetry including F, γ, B, G, and γ are shown in Figure 2 (a) - (d). The figure shows the band structure of the undoped $CsPbI_3$ and $CsPbBr_3$, the valence band maximum (VBM) and conduction band minimum (CBM) are located at the γ point of Brillouin zone implying that these materials have a direct bandgap of 1.488 and 1.688 eV, respectively. As a result, these materials have the potential to be applied in photovoltaic devices. Dopant introduction into the semiconductor host reduces the bandgap due to the formation of dopant levels at the conduction or valence band edge. The bandgap of Mn doped $CsPbI_3$ and $CsPbBr_3$ materials have an indirect bandgap value of 1.274 and 1.350 eV depicted in Figure 2, respectively. The spin-polarized DOS and PDOS for pure and Mn-doped $CsPbX_3$ (X = I and Br) were shown in Figure 3 (a) - (d). These density of states for pure $CsPbX_3$ (X = I and Br) reveals that the upper portion of the VB is influenced by Pb 6s states and halides (I 5p and Br 4p). The lower energy region, 8 eV are primarily influenced by Cs 5p states, with minor contributions

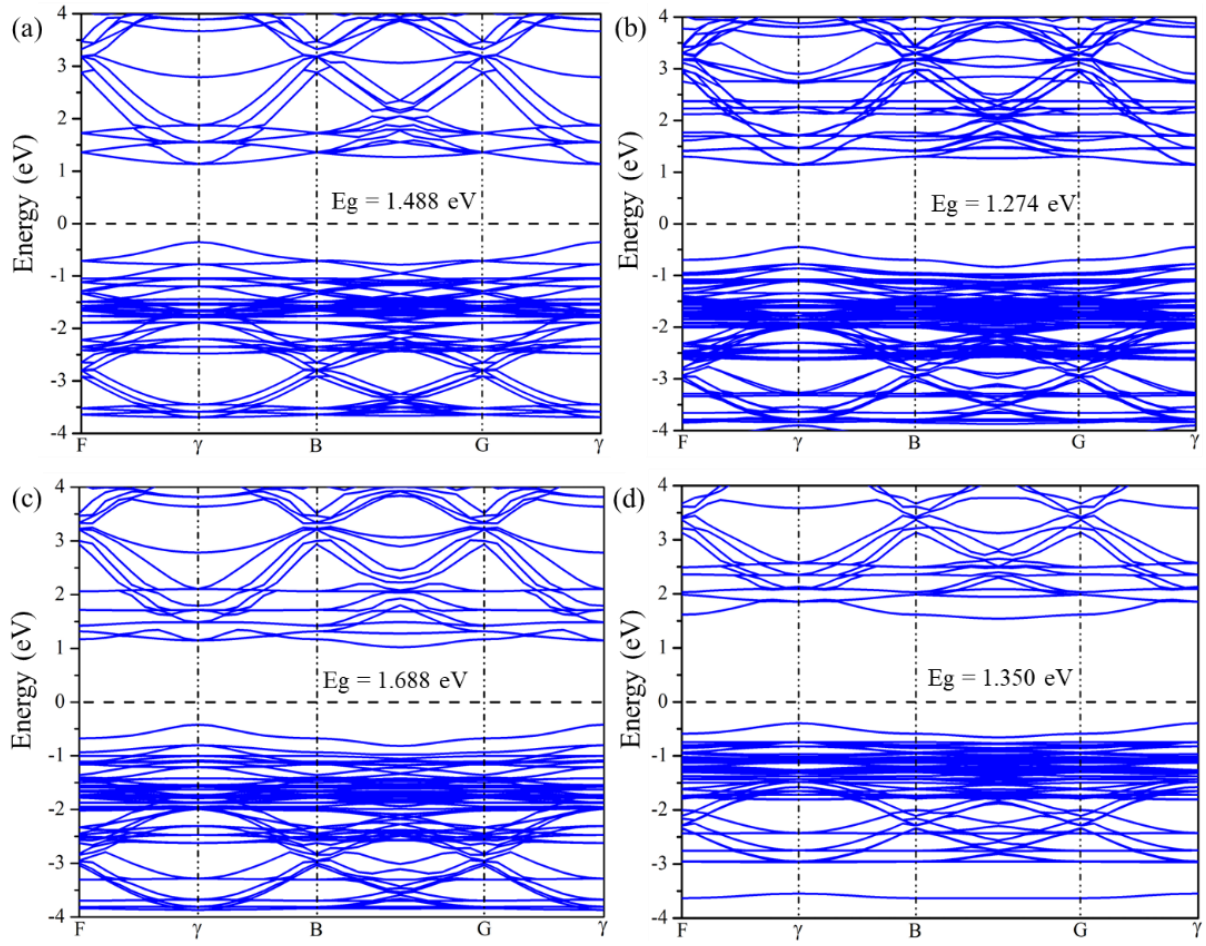


Figure 2. The calculated electronic band structures of undoped and Mn doped $CsPbX_3$ (X = I and Br).

from Pb 6s states. Additionally, notable hybridization occurs between Pb 6s and halides (I 5p and Br 4p) states. The lower edge of the CB is predominantly comprised of Pb 6p states, weakly interacting with Br 4p states. Mn 3d orbitals contributes to both VB and CB for both doped compounds.

3.3. Optical properties

The study of optical parameter including dielectric function, absorption, conductivity, reflectivity, refractive index, and extinction coefficient are important for optoelectronic and solar cell applications which are illustrated in Figure 4. The static frequency limit of dielectric function $\varepsilon_1(0)$ started at 5.97, 5.84, 4.89, and 4.76 for undoped $CsPbI_3$, Mn-doped $CsPbI_3$, undoped $CsPbBr_3$ and Mn-doped $CsPbBr_3$ perovskites, respectively. Beyond zero frequency limit, $\varepsilon_1(\omega)$ of the compounds rises and eventually reaches a maximum value of 9.05 at 2.09 eV, 8.76 at 2.28 eV, 7.37 at 2.18 eV, and 6.67 at 2.76 eV for undoped $CsPbI_3$, Mn-doped $CsPbI_3$, undoped $CsPbBr_3$, and Mn-$CsPbBr_3$ perovskites, respectively. Figure 3 that the value of ε_2 become zero at about 22.8 eV predicting that these materials will become transparent above 22.8 eV. Generally, ε_2 becomes nonzero when the absorption occurs [14].

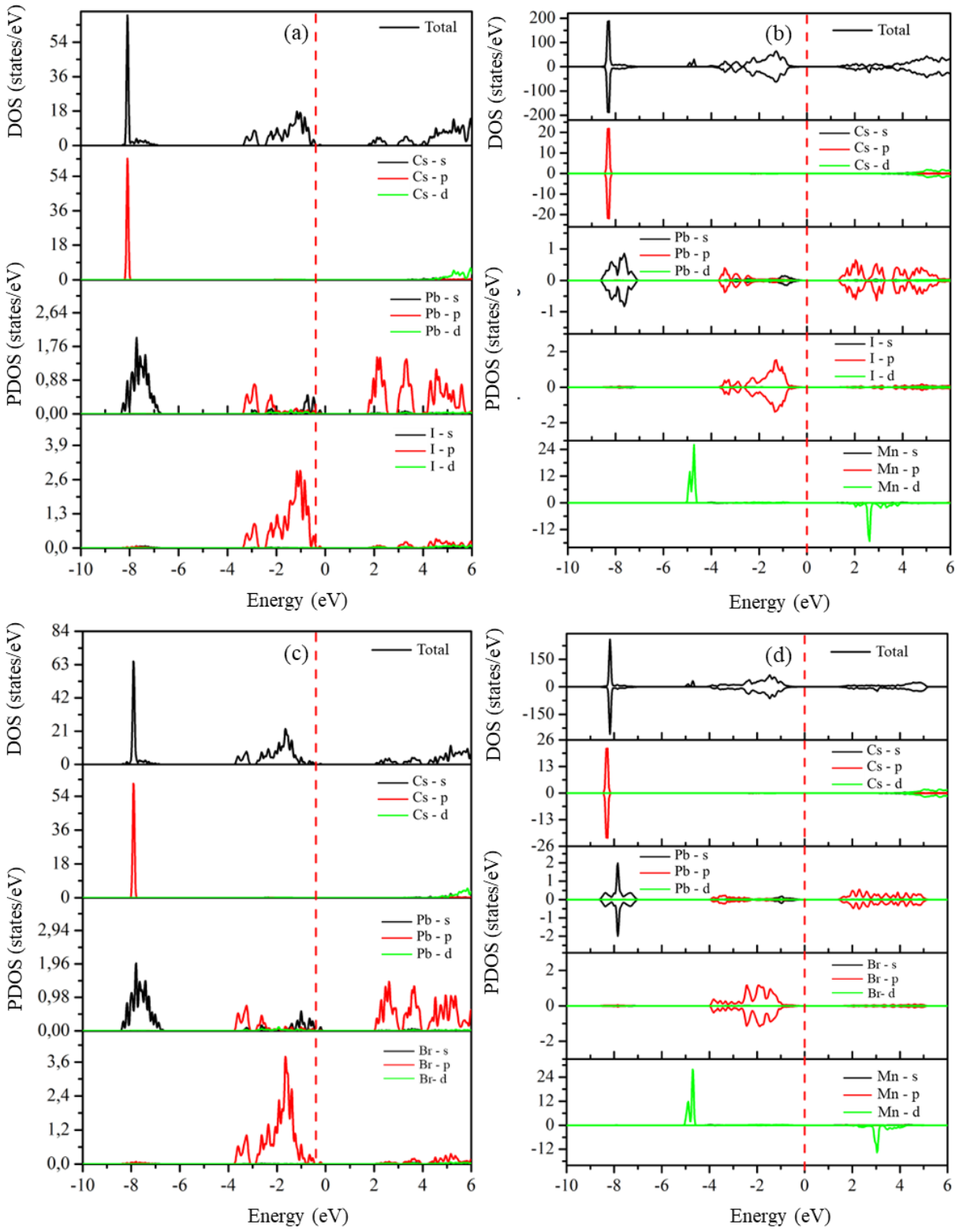


Figure 3. The total and partial density of states for undoped (a) undoped $CsPbI_3$, (b) Mn-doped $CsPbI_3$ (c) undoped $CsPbBr_3$ and (d) Mn-doped $CsPbBr_3$ perovskites.

The calculated refractive index, n(ω) of undoped and Mn-doped $CsPbX_3$ (X = I or Br) is shown in Figure 4 and determines the refracted and reflected light through materials. The static frequency limit n (0) is 3.6, 2.40, 2.37, and 2.46 for undoped $CsPbI_3$, Mn-doped $CsPbI_3$, $CsPbBr_3$, and Mn-$CsPbI_3$ perovskites, respectively. The K(ω) plots are demonstrated in Figure 4, which illustrates the absorbed radiation and behaves in a similar way as $\varepsilon_2(\omega)$. The threshold (critical) energy of K(ω) is 1.40, 1.42, 1.72, and 1.54 eV for undoped $CsPbI_3$, Mn-doped $CsPbI_3$, $CsPbBr_3$, and Mn-$CsPbI_3$ perovskites respectively. The absorption coefficient, α(ω) determines a material's capacity to produce optimal solar energy conversion and to absorb light photons. The optical absorption for undoped and doped Mn-doped perovskites is illustrated in Figure 4. The critical energies for undoped $CsPbI_3$, Mn-doped $CsPbI_3$, $CsPbBr_3$, and Mn-$CsPbI_3$ are 1.404, 1.315, 1.23 and 1.453 eV, respectively. Further, as energy increases, the value of α(ω) increases and reaching its maximum value around 5 eV, respectively. These compounds can absorb photons in the visible region, making them suitable for use in optoelectronic devices.

The reflectivity, R(ω) quantifies a compound's ability to reflect incident light from the surface and demonstrated in Figure 4 for undoped and Mn-$CsPbX_3$ (X=I and Br) perovskites. The calculated static frequency limit of reflectivity R(0) were 0.165 for undoped $CsPbI_3$, 0.171 for $CsPb_{0.875}Mn_{0.125}I_3$, 0.175 for undoped $CsPbBr_3$, and 0.168 for $CsPb_{0.875}Mn_{0.125}Br_3$ perovskite compounds. The optical conductivity, denoted as σ(ω) and depicted in Figure 4, illustrates the generated current density in response to an incident photon with a specific frequency. The critical threshold initiated at approximately at 1.19, 1.91, 1.88, and 1.86 eV for undoped $CsPbI_3$, Mn-doped $CsPbI_3$, $CsPbBr_3$, and Mn-doped $CsPbBr_3$ perovskites, respectively. The intense peaks of σ(ω) for undoped and Mn doped were observed around 6 eV.

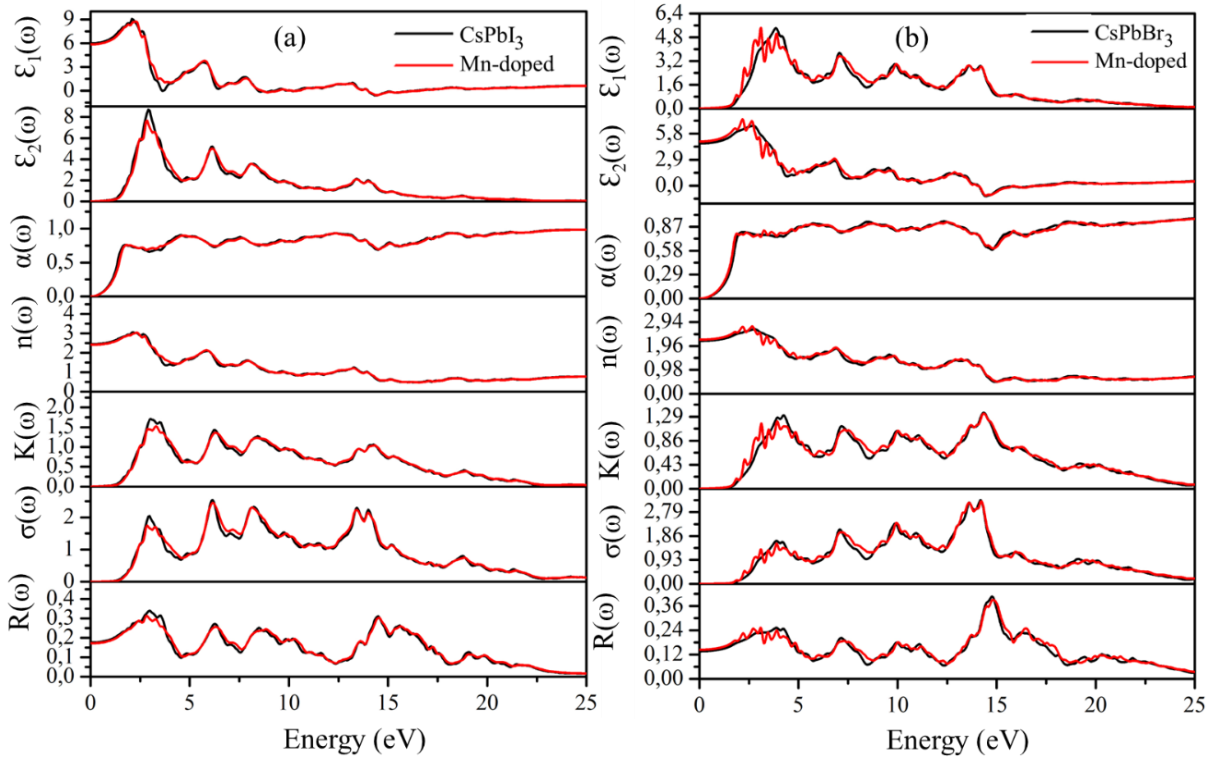


Figure 4. The calculated optical properties of (a) undoped and Mn doped $CsPbI_3$ (b) undoped and Mn doped $CsPbBr_3$ perovskites.

3.3. Elastic properties

The elastic properties of the crystal structures are important to understanding the macroscopic and microscopic behaviour of the materials based on these three elastic constants can be used to determine C_{11}, C_{12}, and C_{44}, for the cubic structure and their criteria are $C_{11} > 0$, $C_{44} > 0$, $C_{11}-C_{12} > 0$, $C_{11}+2C_{12} > 0$. The independent elastic constants calculated for both undoped and Mn doped $CsPbX_3$ (X=I, and Br) are shown in Table 2 which satisfy the Born criteria for cubic structure and implying that these perovskites are mechanically stable. The elastic modulus for these perovskite materials were determined using bulk (B), shear (G) and Young (Y) modulus as illustrated in Table 2 and these moduli were calculated using Voigt-Reuss-Hill approximation. The computed B of the investigated undoped and Mn doped $CsPbX_3$ (X=I and Br), both these compounds exhibit a small value, indicating their flexibility and softness. The introduction of Mn doping led to a small enhancement in the bulk modulus for both compounds. These compounds can be easily fabricated into thin films, offering potential applications in optoelectronics and solar cells. The shear modulus characterizes a material's resistance to deformation, while Young's modulus gauges its rigidity, a higher Y value signifies greater stiffness. As shown in Table 2, both halides experienced a slight increase in stiffness upon Mn doping, rendering them more resilient against deformations induced by the presence of the metal dopant. Pugh and Poisson ratio determines the ductility (<1.75 and <0.26) and brittleness (> 1.75 and >0.26) of a material, respectively. Based on the results presented in Table 2, it can be inferred that undoped and Mn-doped $CsPbX_3$ (X=I and Br) displayed characteristics of ductile behaviour.

Table 2. The computed elastics constants, C_{ij} and modulus for undoped and Mn $CsPbX_3$ (X=I and Br) doped perovskites.

C_{ij} (Gpa)	Compounds			
	$CsPbI_3$	Mn-doped	$CsPbBr_3$	Mn-doped
C_{11}	43.02 34.41[2]	44.76	49.39 43.53	52.76
C_{12}	5.92 4.71	8.34	7.63 7.44	9.33
C_{44}	2.86 3.35	4.24	4.27 4.24	4.24
B	18.27	19.80	21.55 19.47 [15]	23.72
G	6.72	8.75	8.58	8.78
B/G	2.72	2.72	2.51	2.70
Youngs modulus	17.74	20.01	22.56	23.29
Poison's ratio	0.34	0.34	0.32	0.34

4. Conclusion

The structural, opto-electronic and mechanical stability of pure and Mn doped $CsPbX_3$ (X=I and Br) were successfully investigated using DFT. The findings revealed that the addition of Mn reduces the bandgap of both halides and enhances the photon absorption for optoelectronic and solar cell devices. Furthermore, these calculated perovskite materials are ductile and mechanical stable according to elastic properties.

Acknowledgements

The authors acknowledge the Council for Scientific and Industrial Research (CSIR) and Department of Science and Innovation (DSI) Foundational Digital Capability Research for financial support and Center for High-Performance Computing for computing resources.

References

[1] Mao X, Sun T, Wu T, Chu T, Deng W and Han K 2018 *J. Phys. Chem*. C. **122** 7670-7675.
[2] Maleka P M , Dima R S, Ntwaeaborwa O M and Maphanga R R 2023 *Phys. Scr*. **98** 045505.
[3] Murtaza G and Ahmad I 2011 *Phys. B: Condens. Matter*. **406** 3222-3229.
[4] Maleka P M , Dima R S, Ntwaeaborwa O M and Maphanga R R 2022 *Mater. Today: Proc*. **62** S12-S22.
[5] Quinten D M, Akkerman A, Dang Z, De Angelis F and Manna L 2017 *ACS. Energy. Lett*. **9** 2183-2186.
[6] Zhang Y Z, Chen J and Yang L 2020 *Chem. Phys. Lett.* **10** 18542-18576.
[7] Perdew J P, Ruzsinszky A, Tao J, Staroverov V N, Scuseria G E and Csonka G I 2005 *J. Chem Phys.* **123** 062201.
[8] Perdew J P and Zunger A 1981 *Phys. Rev*. B. **23** 5048.
[9] Monkhorst H J and Pack J D 1976 *Phys. Rev*. B. **13** 5188.
[10] Momma K and Izumi F 2011 *J. Appl. Crystallogr.* **44** 1272-1276.
[11] Chang Y, Park C H, Matsuishi K 2004 *J. Korean. Phys. Soc.* **44** 889-893.
[12] Filip M R, Eperon G E, Snaith H J and Giustino F 2014 *Nat. Commun*. **5** 1-9.
[13] Jie Meng Z L, Lin W, Liang M, Zou X, Zhao Q, Geng H, Castelli I E, Canton S E, Pullerits T and Zhen K 2021 *Chem. Sci*. **13** 1734-1745.
[14] Islam M, Hadi M and Podder J 2019 *AIP. Adv*. **9** 13-16.
[15] Naher M and Naqib S 2021 *Condens. Mater. Sci*. **6** 2103.

Derivation of empirical interatomic potentials for doping Co and Ni into $LiMn_2O_4$ spinel

Donald Hlungwani, Phuti Ngoepe, and Raesibe Ledwaba
Materials Modelling Centre, University of Limpopo, Private Bag X 1106, Sovenga, 0727, SA

E-mail: **donald.hlungwani01@gmail.com**

Abstract. Buckingham interatomic potentials for interactions arising from doping $LiMn_2O_4$ spinel with Co and Ni have been determined successfully. The potentials are derived in spinel systems with structure properties calculated with the Density Functional Theory (DFT) code, VASP. The quality of the derived potentials was assessed by comparing the fitted structure properties with the structure properties calculated from the derived potentials. All the structures were found to have percentage differences of < 7.8% between the fitted and calculated structure properties (lattice constants, cell volume, C_{ij} elastic constants, and bulk modulus). The derived potentials were further validated through molecular dynamics simulations of $LiCo_2O_4$ and $LiNi_2O_4$ spinel structures at various temperatures. Moreover, the state of these spinel structures was captured through atomic-level snapshots, radial distribution functions (RDFs), and total energy vs. temperature graphs. The derived potentials were able to describe the $LiCo_2O_4$ and $LiNi_2O_4$ spinel structures at various temperatures.

1. Introduction

Li-Mn-O spinel continues to evoke enormous interest as a future positive electrode material for high-energy-density lithium-ion batteries. $LiMn_2O_4$ spinel crystallizes into a cubic structure with a space group of Fd-3m, where lithium atoms are situated in tetrahedral sites and manganese atoms occupy octahedral sites. Consequently, oxygen atoms at the 32e sites form a close-packed array, resulting in three-dimensional pathways for easy intercalation of Li^+ ions [1]. The three-dimensional channels facilitate high-rate capabilities suitable for electric vehicles (EVs) or hybrid-electric vehicles (HEVs). Moreover, it is cost-effective, thermally stable, and environmentally benign [2, 3]. However, during prolonged charge/discharge cycles and at elevated temperatures (> 55 °C), the material deteriorates [4, 5]. Numerous studies have ascribed the cause of this structural degradation to the irreversible structure and phase change due to the Jahn-Teller effect and the disproportionation reaction of Mn^{3+} ($2Mn^{3+} \rightarrow Mn^{4+} + Mn^{2+}$). Consequently, the formed Mn^{2+} is known to dissolve in the electrolyte, resulting in severe capacity fading [6, 7].

It has been shown that cation doping can stabilize the $LiMn_2O_4$ spinel structure through the replacement of a fraction of Mn^{3+} ions by divalent or trivalent cations such as Co, Ni, Zr, Al, etc. [8, 9]. Wang Y and co-workers carried out an experimental study of Mg-doped $LiMn_2O_4$ spinel and found that 2% Mg dopant returns ~81% discharge capacity after 500 cycles. Moreover, the structure remained intact in its cubic form after the introduction of Mg [10]. Ma Y et al. explored the mechanism of doping Li-Mn-O spinel with Ti [11]. Furthermore, it was demonstrated that the stability of the Li-Mn-O spinel structure can be improved by the introduction of small percentages of Ti. Moreover, a study by Cai Z

and co-workers suggested that a strong Al-O bond can result in improved structural integrity of an Al-doped $LiMn_2O_4$ spinel compared to the undoped structure [12].

The positive results set cation doping as a significant solution for improving the electrochemical performance of $LiMn_2O_4$ spinel. However, there is insufficient knowledge of the effect of cation doping on the microstructural properties of doped $LiMn_2O_4$ spinel. Especially the impact of doping $LiMn_2O_4$ spinel with Co and Ni on its microstructure and on the mobility of Li^+ in the structure at various temperatures. Such studies are either computationally expensive or impossible in the Density Functional Theory (DFT) framework due to its unfavourable scaling with system size, which practically only allows studies of systems that contain a few hundred atoms.

Molecular dynamics techniques such as amorphization and recrystallization (A&R) can introduce structural features such as micro-twining, point defects, and grain boundaries, which solely depend on the underlying forcefield [13]. The features mentioned are significant to electrochemical performance and contribute to the understanding of the relationship between the microstructure and its properties, which are essential for the optimization of material performance. However, the accuracy of an MD simulation is determined by the forcefield function and its parameters, which approximate the quantum potential energy. The insufficiency of accurate interatomic potentials (forcefields) has become a barrier to large-scale MD simulations due to their ambiguous derivation process, often referred to as an art rather than a science. The derivation of the forcefield parameters to capture the interactions in the material in question involves optimization of the forcefield parameters to reduce the difference between experimental or DFT structural properties and the structural properties calculated from the forcefield. However, the final parameters of the forcefield function depend on the initial parameters. As such, to derive accurate and physical forcefield function parameters, a physical starting point is required (initial parameters that relate to the interatomic interaction in question). In this regard, we have employed a curve fitting (CV) technique and the General Lattice Utility Program (GULP) to derive precise interatomic potentials for interactions that result from doping Li-Mn-O spinel with nickel and cobalt [14]. CV is used to determine a reasonable starting point from a potential energy surface calculated within the DFT framework for the interatomic interaction in question. Consequently, these parameters will be refined in GULP [14].

2. Methodology

The study makes use of first-principles methods based on quantum mechanics, lattice dynamics methods based on classical mechanics, and curve fitting (CV) methods. The derived potentials are tested with the molecular dynamics (MD) code, DLPOLY [15].

2.1. Derivation of forcefield

The Buckingham interatomic potential function was chosen to describe the short-range interaction in accordance with the Born Model of ionic solids. The Born model of ionic solids provides a way of describing the energy of a lattice. The Buckingham potential function is given in Equation 1 below:

$$V(r) = Ae^{-Br} - \frac{\lambda}{r^6} \tag{1}$$

where, r is the separation distance between the two interacting atoms, and A and B are the parameters of the repulsive potential stopping the two atoms from overlapping. Furthermore, λ is the potential parameter for the attractive interaction. The potential energy surfaces are fitted to a Buckingham potential function using the SciPy scientific module consisting of well-known mathematical functions [16]. The curve_fit function was chosen for finding optimal Buckingham potential parameters that best reproduce a potential energy surface of a given interaction. The potentials obtained through curve fitting are further refined with GULP [14]. GULP is widely used to fit empirical forcefields using energy surfaces and data from traditional experiments or first-principles calculations [14]. All the potentials were fitted at 300 K and 0 Pa.

2.2. Molecular dynamics

All molecular dynamics calculations in this work are performed using the DL_POLY code [17]. The Li-Li, Li-O, O-O, Mn-Mn, Mn-O, and Mn-Li Buckingham potentials used in this study were derived and tested elsewhere [17, 18]. The Nose-Hoover thermostat was applied, and a time per step of 0.0005 ps for 100 000 steps was chosen [19]. The systems were allowed to equilibrate for 50 000 steps, and a cutoff radius of 10 Å was chosen.

3. Results and Discussions

3.1. Interatomic Potentials

The Buckingham interatomic potentials for the interactions arising from doping $LiMn_2O_4$ spinel with Co and Ni have been derived successfully and are listed in Table 1. The results for the derivation of the Buckingham potentials of the Co-Co, Co-Li, Co-O, and Co-Mn interactions are illustrated in Table 2. GULP was utilized in determining the final Buckingham interatomic potentials illustrated in Table 2 [14]. However, the initial potential parameters were derived from potential energy surfaces calculated with the first-principles code FHI-aims. The Buckingham interatomic potentials for Co-Co, Co-Li, and Co-O were fitted in $LiCo_2O_4$. The fitted structure properties referred to as experimental in the table are reproduced to a percentage difference of less than 5%. The fitted lattice parameter compares well to the lattice parameter of $LiCo_2O_4$ calculated from the derived potentials. Moreover, there is a good comparison between the calculated volume of 537.127 $Å^3$ and the fitted volume of 536.063 $Å^3$, as shown by a percentage difference of 0.2%. The Co-Mn interaction was derived from a Co-doped Li-Mn-O spinel structure. The fitted volume and lattice parameters are reproduced within a percentage difference of less than 2.02%. Furthermore, the bulk modulus and elastic constant are reproduced to a percentage difference of less than 3.5%.

Table 3 shows the results for the derivation of the Buckingham interatomic potentials for Ni-Ni, Ni-Li, Ni-O, and Ni-Mn interactions in Ni-doped Li-Mn-O spinel and the Ni-Co interaction in Ni-doped Li-Co-O spinel. The fitted $LiMn_{1.875}Ni_{0.125}O_4$ spinel structure is reproduced by the derived Buckingham interatomic potentials, as evidenced by a percentage difference of less than 0.5% between the experimental (fitted) and calculated lattice parameters and volume. Furthermore, the calculated and fitted elastic constants and bulk modulus compare well, with a percentage difference of less than 6.7%. The Buckingham interatomic potentials for the Co-Ni interaction in Ni-doped Li-Co-O spinel have also been derived successfully. The fitted lattice parameters and volume are reproduced within a percentage difference of less than 0.2%. Moreover, a good comparison between the fitted and calculated elastic constants and bulk modulus is also observed, with a percentage difference of less than 7.2%.

Table 1. The derived Buckingham interatomic potentials emerging from doping $LiMn_2O_4$ spinel with Co and Ni.

Interactions	$A\ (eV)$	ρ(Å)	$C(eVÅ^{-6})$
Co - Co	48021.956	0.162	63.937
Co - O	68670.662	0.166	21.732
Co - Li	17210.534	0.148	10.158
Co - Mn	120663.933	0.256	1695.250
Ni - O	84002.097	0.169	49.086
Ni - Mn	610056.862	0.234	3378.044
Co - Ni	31071.223	0.165	54.294

Table 2. Comparison of the VASP (Fitted) and GULP (calculated) structure properties of $LiCo_2O_4$ and $LiMn_{1.875}Co_{0.125}O_4$, results for the derivation of the Buckingham interatomic potentials for Ni – Ni, Ni – Li, Ni – Mn and Ni – O interactions at temperature of 300 K and a pressure of 0.00 GPa [14].

	$LiCo_2O_4$		$LiMn_{1.875}Co_{0.125}O_4$	
Properties	Fitted	Calculated	Fitted	Calculated
α (Å)	8.129	8.123	8.153	8.208
Volume ($Å^3$)	537.127	536.063	541.981	552.980
C_{11} Elastic constant	222.585	229.713	209.040	211.364
C_{12} Elastic constant	104.944	110.176	97.840	94.682
C_{44} Elastic constant	82.350	85.279	85.120	84.215
Bulk Modulus (GPa)	145.518	150.021	134.58	132.740

Table 3. Comparison of the Fitted and calculated structure properties of $LiMn_{1.875}Ni_{0.125}O_4$ and $LiCo_{1.875}Ni_{0.125}O_4$, results for the derivation of the Buckingham interatomic potentials for Ni – Ni, Ni – Li, Ni – Mn and Ni – O interactions at temperature of 300 K and a pressure of 0.00 GPa [14].

	$LiMn_{1.875}Ni_{0.125}O_4$		$LiCo_{1.875}Ni_{0.125}O_4$	
Properties	Fitted	Calculated	Fitted	Calculated
α (Å)	8.155	8.167	8.126	8.121
Volume ($Å^3$)	542.380	544.682	536.496	535.576
C_{11} Elastic constant	214.540	223.853	212.477	228.188
C_{12} Elastic constant	96.460	103.118	111.564	111.520
C_{44} Elastic constant	86.430	88.995	81.600	86.570
Bulk Modulus (GPa)	135.820	141.255	144.212	150.409

3.2. Molecular dynamics

The derived Buckingham interatomic potentials for Ni and Co doping into $LiMn_2O_4$ have been tested through molecular dynamics simulations. Figures 1 and 2 show structural snapshots of $LiCo_2O_4$ and $LiNi_2O_4$, respectively, at various temperatures. The increase in temperature results in an increase in the vibrations of atoms about their lattice positions. The crystalline to amorphous phase change is observed at ~1800 K for $LiNi_2O_4$ and at ~1900 K for $LiCo_2O_4$. This is in line with the fact that Co has a higher melting point than Ni [20]. The crystalline-to-amorphous phase change can also be observed on the total energy vs. temperature graphs in figures 3 (a) and (b) for $LiCo_2O_4$ and $LiNi_2O_4$, respectively. The increase in temperature for the $LiCo_2O_4$ structure at a temperature greater than 1900 K facilitates the crystallin-to-amorphous phase change until the change is complete. This is indicated by a sharp energy change and an infinitesimal temperature change from ~1900 K. Phase-change is observed in the atomic snapshots illustrated in figure 1, where the loss of long-range atomic range is rapidly increasing from the temperature of ~1900 K. Similarly, at ~1800 K, the increase in temperature does not increase the temperature of the system for the $LiNi_2O_4$ spinel structure. The total energy vs. temperature plot for $LiNi_2O_4$ is in alignment with the atomic snapshots in figure 2. Figures 4 and 5 capture the distribution of atoms in the Li-M-O (M=Ni, Co) spinel structures at various temperatures. The RDF peaks are broadening with increasing temperature, suggesting a disruption of long-range atomic arrangements in the spinel structures, as generally expected. The RDFs of $LiCo_2O_4$ and $LiNi_2O_4$ in figures 4 and 5, respectively, show the greatest probability of finding an atom with respect to a given reference atom at any position in the separation distance between 3.5 and 10 Å. Moreover, RDF peaks for the two structures ($LiCo_2O_4$ and $LiNi_2O_4$) are broadening with increasing temperatures. At a separation distance of 4 Å, $LiCo_2O_4$ at 1900 K has a higher probability of finding an atom with respect to a reference atom when compared to $LiNi_2O_4$ at 1900 K.

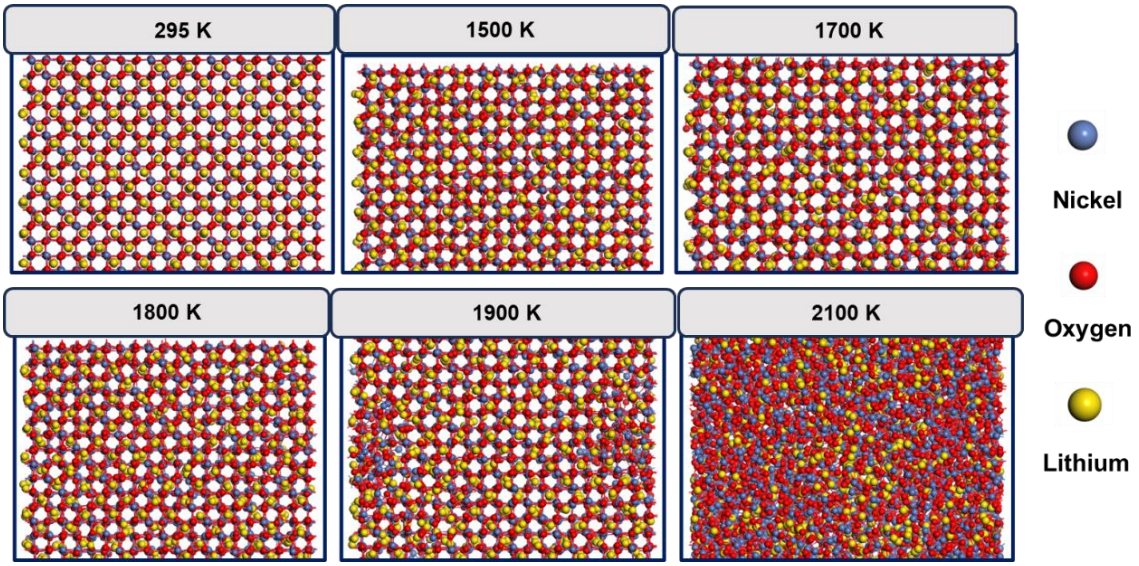


Figure 1. Atomic-level structure images of $LiCo_2O_4$ spinel gradually heated to a temperature of 2100 K.

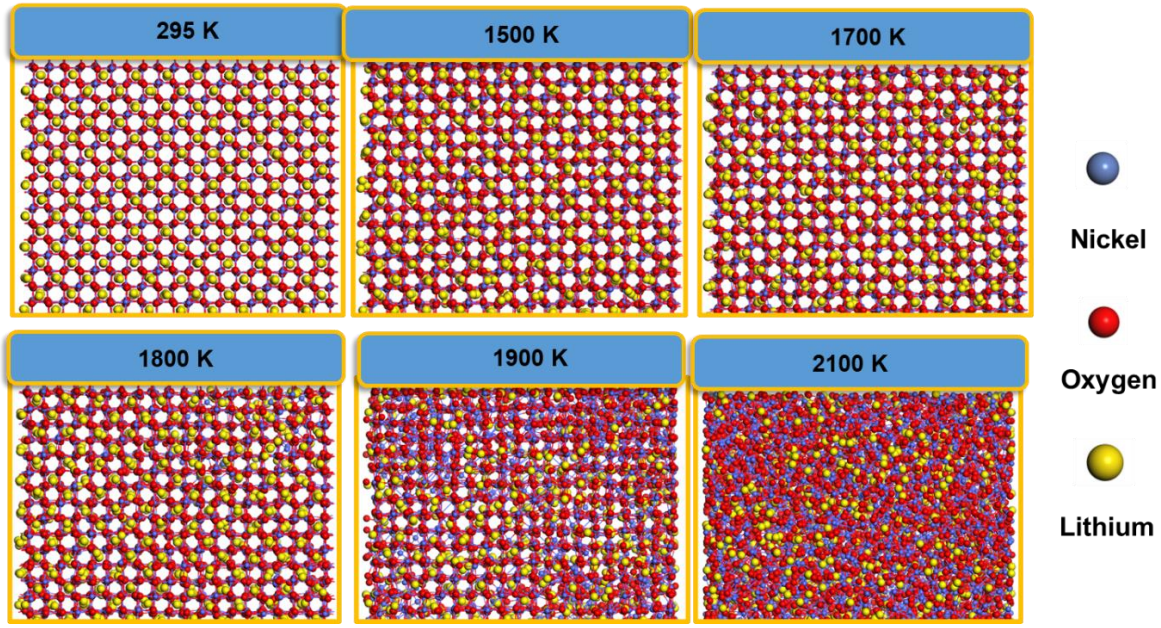


Figure 2. Illustration of gradual heating of a $LiNi_2O_4$ spinel bulk supercell to a temperature of 2100 K.

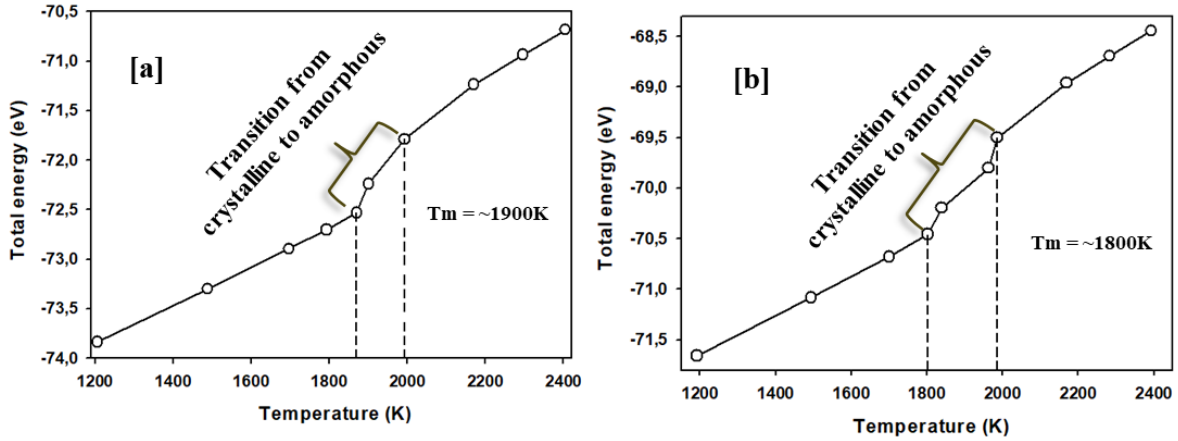


Figure 3. Total energy against temperature of (a) $LiCo_2O_4$ and (b) $LiNi_2O_4$ spinel graphs that capture the transition from crystalline to amorphous.

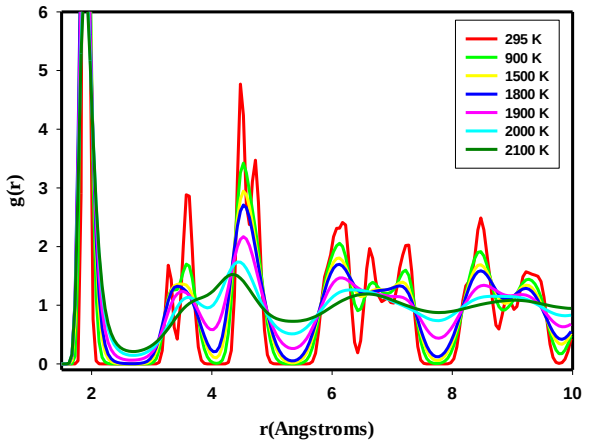


Figure 4. RDF graph of the Co-O interaction in $LiCo_2O_4$.

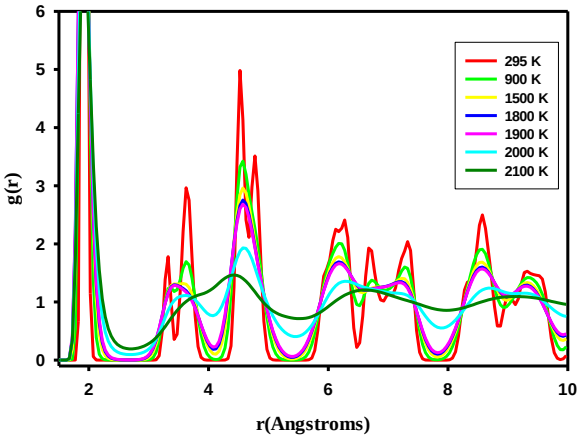


Figure 5. RDF graph of the Ni-O interaction in $LiNi_2O_4$.

4. Conclusion

The Buckingham interatomic potentials for Ni and Co interactions for doping $LiMn_2O_4$ spinel have been developed successfully, as evidenced by a good agreement between the fitted and calculated structural properties. The potentials can describe the Li-M-O (M=Co, Ni) spinel structures at various temperatures, which is significant for studies of electrochemical properties that are temperature-dependent. Furthermore, the potentials will be essential for high-temperature techniques such as simulated amorphization and recrystallization. Accordingly, the potentials can be used to perform large-scale molecular dynamics simulations that will yield insights on the effect of the introduction of a small amount of Ni and Co in the Li-Mn-O spinel structure.

References

[1] Goodenough J B, Thackeray M M, David W I F and Bruce P G 1983 *Mater. Res. Bull.* **18** 461-472

[2] Li W, Song B and Manthiram A 2017 *Chem. Soc. Rev.* **46** 3006-3059

[3] Jiang J, Du K, Cao Y, Peng Z, Hu G and Duan J 2013 *J. Alloys Compd* **577** 138-142

[4] Haruna A B, Barrett D H, Rodella C B, Erasmus R M, Venter A M, Sentsho Z N and Ozoemena K. I 2022 *Electrochim. Acta.* **426** 140786-140786

[5] Hlungwani D, Ledwaba R S and Ngoepe P E 2022 *Mater.* **15** 5678

[6] Appiah W A, Ryou M H and Lee Y M 2018 *J. Electrochem. Soc.* **166** A5109

[7] Leong V G H, Hong S S and Castro R H 2022 *Mater. Lett.* **325** 132824

[8] Xu W, Zheng Y, Cheng Y, Qi R, Peng H, Lin H and Huang R 2021 *ACS Appl. Mater. Interfaces.* **13** 45446-45454

[9] Michalska M, Ziółkowska D A, Jasiński J B, Lee P H, Ławniczak P, Andrzejewski B, Ostrowski A, Bednarski W, Wu S H and Lin J Y 2018 *Electrochim. Acta*. **276** 37-46

[10] Wang Y, Luo S, Li P, Zhang H, Yan S, Li H, Tian Y, Teng F, Wang Q, He A and Zhuge F 2022 *Mater. Today Commun* **33** 104391

[11] Ma Y, Dai Y, Zhou Q, Cheng J, Li H and Hu W 2022 J. Electron. Mater. **5** 77-83

[12] Cai Z, Ma Y, Huang X, Yan X, Yu Z, Zhang S, Song G, Xu Y, Wen C and Yang W 2020 *J. Energy Storage.* **27** 101036

[13] Sayle D C and Johnston R L 2003 Curr. Opin. Solid State Mater. Sci. **7** 3-12

[14] Gale J D 1997 *J. Chem. Soc., Faraday Trans.* **93** 629-637

[15] Smith W and Forester T R 1996 *J. Mol. Graph.* **14** 136-141

[16] Virtanen P, Gommers R, Oliphant T E, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J and Van Der Walt S J 2020 *Nat. Methods* **17** 261-272

[17] Maphanga R R, Sayle D C, Sayle T X and Ngoepe P E 2011 *Phys. Chem. Chem. Phys.* **13** 1307-1313

[18] Sayle T X, Catlow C R A, Maphanga R R, Ngoepe P E and Sayle D C 2005 *J. Am. Chem. Soc.* **127** 12828-12837

[19] Hoover W G 1985 *Phys. Rev. A.* **31** 1695

[20] Sato J, Omori T, Oikawa K, Ohnuma I, Kainuma R and Ishida K 2006 *Sci.* **312** 90-91

Optimization c-Titanium (II) Oxide Electron Transport Layer for Solar Cell Application

S B Sibiya[1], N Nombona[2], M Diale[1]*
[1]Department of Physics, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa
[2]Department of Chemistry, University of Pretoria, Private Bag X20, Hatfield 0028, South Africa

E-mail: **sibiya.sizwe@up.ac.za**

Abstract. The electron transport layer plays a vital role in extracting and transporting photo-generated electrons, modifying the interface, and minimizing the charge recombination in perovskite solar cells. This study reports on the influence of titanium (II) oxide as an electron transport layer in perovskite solar cells. The metal oxide used as electron transport material in planar perovskite solar cells has proved to be effective in improving devices' performance, enhancing electron mobility, and energy alignment. XRD diffractograms confirmed an anatase phase crystal structure with an average crystallite size of ±24.44 nm from the 4 prepared thin films. Structural micro-strains were calculated directly from the modified Williamson-Hall equation to be ± 9.75×10^{-4}. SEM morphology revealed a mixture of compact grains with full surface coverage. UV-Vis absorbance spectrum showed an onset absorption at 338 nm and a well-aligned bandgap of 3.9 eV for the examined titanium dioxide thin films.

1. Introduction

Perovskite solar cells (PSCs) are gaining traction as a prospective technology in thin-film photovoltaics due to their impressive properties which include high light absorption capacity [1], lengthy carrier diffusion length [2], and solution processability [3]. At present, perovskite solar cells (PSCs) have achieved a maximum power conversion efficiency (PCE) of 25.5% [4], surpassing the efficiency record of crystalline silicon solar cells. Over the span of a decade, the advancement in PCE for PSCs has surged from 3.8% to 25.5% [5].

Furthermore, the traditional architecture of the planar perovskite solar cell has an active layer sandwiched between an electron transport layer (ETL) and a hole transport layer (HTL) [6]. *Choi et al.,* conducted a study wherein they devised two fluorinated isomeric counterparts of the well-established hole transport compound spiro-OMeTAD. The goal was to modify the energy level alignment, hydrophobic nature, and hole extraction capability within perovskite solar cells (PSCs). Notably, they discovered that introducing a fluorine group at the meta-position on the benzene ring (referred to as spiro-mF) led to a reduction in the highest occupied molecular orbital (HOMO) position from -4.97 to -5.19 eV. This adjustment facilitated a more favorable alignment of energy levels with the valence band maximum (VBM) of the $FAPbI_3$ perovskite absorber, which rests at -5.40 eV. Consequently, this alignment minimized interfacial energy losses, resulting in an enhancement of the open-circuit voltage (V_{OC}) of the device. Furthermore, the introduction of fluorine (F) atoms induced a more compact arrangement of molecules in the solid state due to non-covalent intra-molecular interactions. This, in turn, enhanced the electronic connection between spiro-mF and the surface of the perovskite thereby

facilitating improved hole transport and extraction capabilities. Consequently, both the short-circuit-current (J_{SC}) and fill factor (FF) experienced a minor augmentation. These combined effects led to an efficiency increase in perovskite solar cells (PSCs), elevating it from 23.44% to 24.84%.

In addition, a deeper understanding of materials and structures for the ETL is crucial in improving device performance and long-term stability [7]. The electron transport layer plays an important role in extracting and conveying photo-generated electrons from the perovskite layer to either the cathode or anode [8]. At the same time, the ETL also modifies the interface between the perovskite layer and the electrode [9], minimizing charge recombination dynamics [10].

In this study, compact titanium (II) oxide c-TiO_2 was used as an ETL. The results of the experiment of spray-deposited c-TiO_2 are discussed. We will summarize the XRD, SEM, and UV-Vis absorption spectral results obtained to date and discuss future opportunities to move the field forward.

2. Experimental

Preparation of c-TiO_2 thin films using the spray pyrolysis technique

Thin film based on c-TiO_2 were prepared from a mixture of 0.5 M titanium tetra-isopropoxide (TTIP) sourced from Sigma Aldrich with a purity of 97.0% and 50 mL solution of ethanol obtained from Sigma-Aldrich Reagent. The solution was stirred at ambient temperature (25 ℃) until achieving a homogeneous, colorless mixture. To prepare for c-TiO_2 deposition, the edge portions of the substrates was shielded with thermal tape to protect the electrodes of the samples. The substrates were positioned onto a preheated hotplate, gradually raised to the desired temperature of 250 ℃ to avoid thermal stress on the glass substrates. The thin films were fabricated using spray pyrolysis technique, depicted in Figure 1, by spraying the c-TiO_2 solution mixture onto the heated substrates. This process employed a spray pressure of 3 kPa, utilized air as a carrier gas, employed a 20 cm nozzle to substrate diameter, featured 0.5 mm nozzle diameter, and involved rapid sweeping motions with intervals of 60 s and delays of 15 s delays between sweeps. The substrates well allowed to naturally cool to ambient temperature. Subsequently, the c-TiO_2 thin films underwent annealing at 450 ℃ for one hour for further characterization.

3. Results and Discussion

3.1 Structural analysis for titanium dioxide properties

Figure 2. shows the XRD diffractograms of c-TiO_2 deposited on glass/FTO substrate using the spray pyrolysis (SP) technique. The measurements were conducted using a Bruker D2-Phaser X-ray diffractometer employing Cu Kα radiation with a wavelength of 0.15405 nm. Observed peaks for the scale 2θ of TiO_2 occurred at 26.53°, 33.76°, 37.76°, 51.53°, 54.57º, 61.58°, 65.55°, 77.33°, and 78.33°. These peaks are closely related to those of the polycrystalline anatase TiO_2 of the standard pattern (JCPDS no.: 88-1175 and 84-1286). Moreover, corresponding peaks relate to the crystallographic planes (101), (103), (004), (200), (105), (213), (116), (220), and (206), respectively. Furthermore, the orientation of the 2θ diffraction peaks at 26.53° and 37.75° corresponds to that of anatase phase material and is in good agreement with the JCPDS 21-1272 file number.

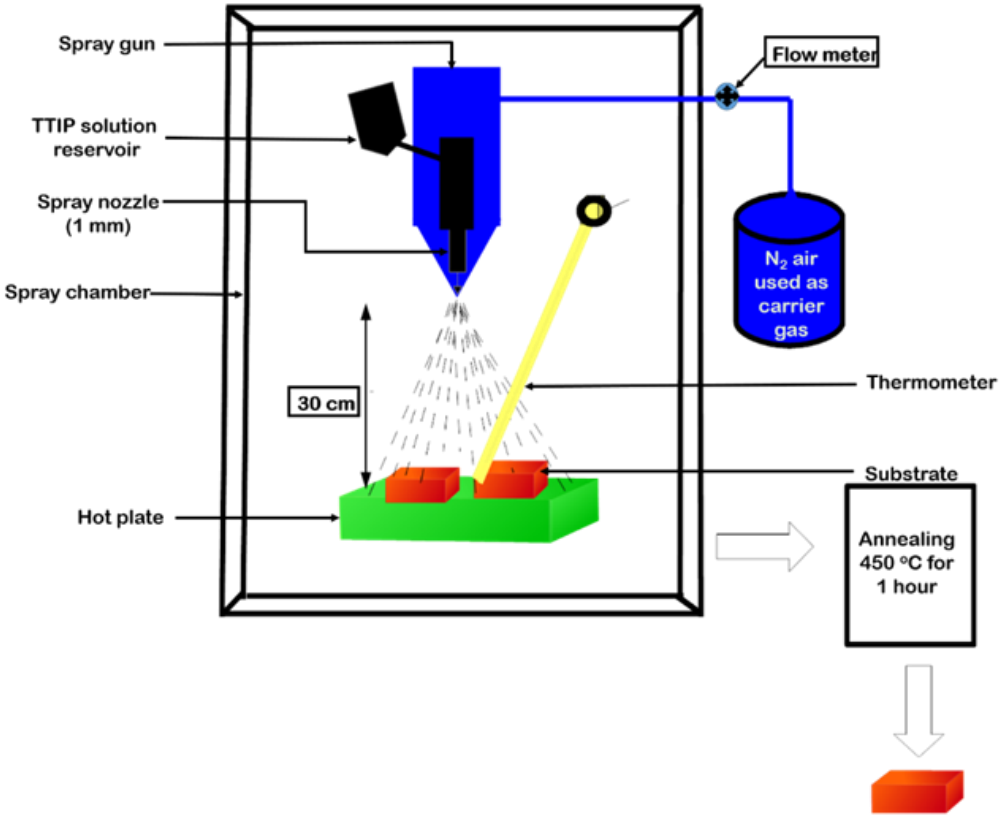


Figure 1. Schematic of spray pyrolysis

The principal peak at $2\theta = 26.53^{\circ}$ corresponds to the (101) crystallographic plane of the anatase phase in c-TiO_2 nanoparticles, denoting their exclusive contribution to the pure anatase phase, excluding rutile or brookite phases. The average crystallite size of c-TiO_2 was determined utilizing the Debye-Scherrer formula as described in Equation [1]:

$$D = \frac{k\lambda}{\beta \cos \Theta}, \tag{1}$$

where D is the crystal Size (nm), β the full width at half maxima of the peak (FWHM) in radians, k is the dimensionless shape factor close to unit 0.9, while θ is the Bragg's angle and λ the wavelength. The average crystal size was calculated and found to be 24.44 nm. Stress data regarding the crystallite strains produced by the crystal defects were calculated using a modified Williamson-Hall Equation [2], and the calculated average micro-strain for the samples was consistent with 9.75×10^{-4}.

$$\beta \cos \Theta = \frac{k\lambda}{D} + 4\varepsilon \sin \Theta, \tag{2}$$

3.2 Titanium dioxide morphological properties

Synthesized thin films' surface morphology was studies using from the using field-emission scanning electron microscopy (FE-SEM). Figure 3(a) demonstrates an agglomerated surface morphology on the top layer of the films. Conversely, Figure 3(c) reveals complete surface coverage at lower magnification. Notably, Figure 3(b) displays nanoparticles as well-defined grains, supporting the tetragonal crystal structure as evidenced by the X-ray diffraction (XRD) peak broadening illustrated in Figure 2

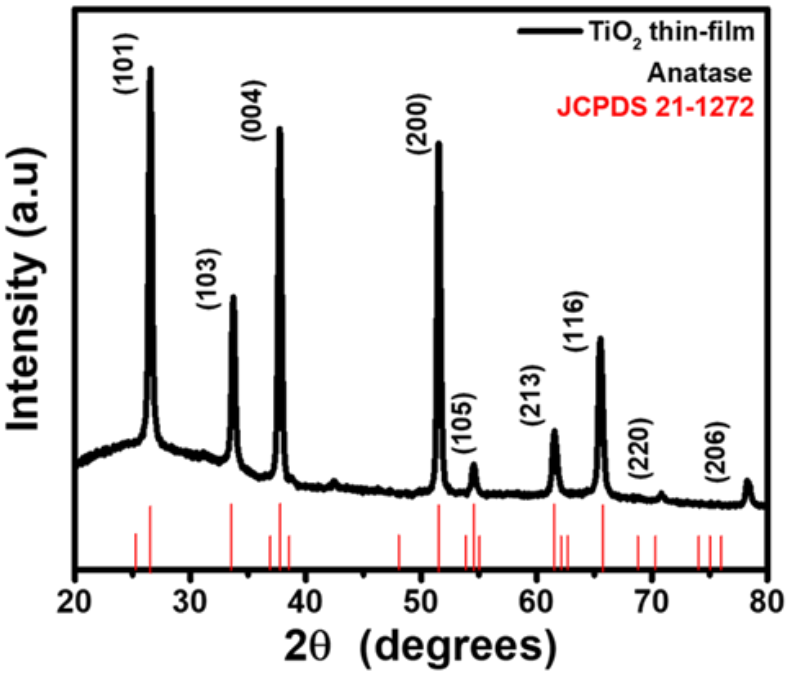


Figure 2. XRD pattern of c-TiO_2 thin film on FTO/glass substrate.

. However, the impact of crystalline materials on the broadening of XRD patterns in nano-sized TiO_2 is found to be negligible. The crystal sizes within the films ranged between 25 and 29 nm, resembling those observed in mesoporous TiO_2 thin films reported by *Kim et al*. These films exhibit a notably large, defect-free surface topography. The XRD diffractograms is in good agreement with the SEM micrographs, affirming a homogeneous structure without traces of precursor impurities. Consequently, the successful synthesis of high-quality c-TiO_2 thin films is demonstrated.

3.3 UV-Vis absorption properties of c-TiO_2 thin films

The absorption spectral properties of the c-TiO_2 thin films were examined from the UV-Vis spectroscopy setup. In Figure 4(a), the absorption spectrum depicts the thin films exhibiting maximum absorption directly at the wavelength of 338 nm. Furthermore, the calculated bandgap, illustrated in Figure 4(b) and determined as 3.9 eV for the engineered electron transport layer (ETL) thin films, aligns well with the wide bandgap previously investigated by Pal, M., et al.,[11] . Additionally, Tauc's plot, derived from equation [3] below, served as a confirmation method for the bandgap value.

$$(\alpha h\nu)^n = K\left(h\upsilon - E_g\right) \tag{3}$$

This equation describes the relationship between the material's absorption coefficient (α), its optical bandgap (E_g) and the frequency (ν) of the incident light. Where h (6.62×10^{-34} m^2 $kg.s^{-1}$) know as plank's constant, n the nature of transmission (2 for direct bandgap and ½ for indirect bandgap materials).

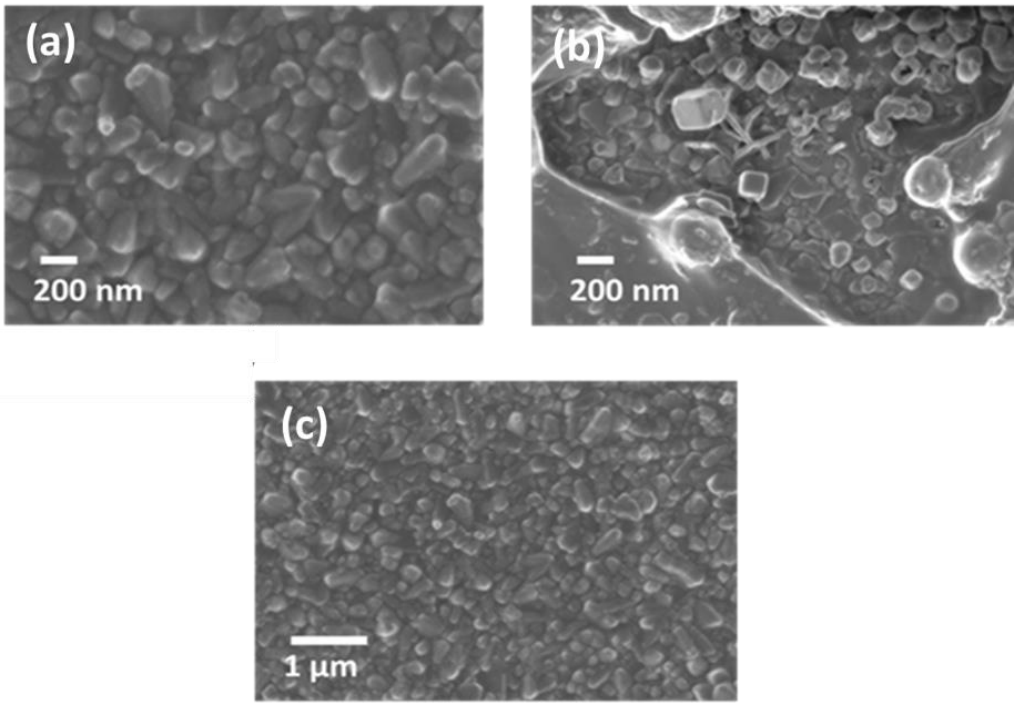


Figure 3. Top-view morphology images of c-TiO_2.

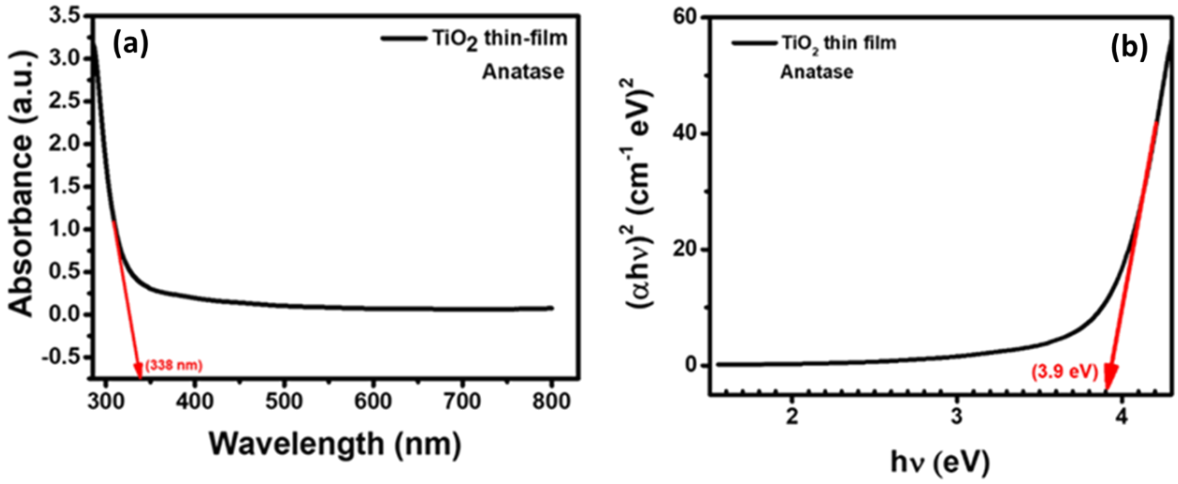


Figure 4. UV-Vis (a) absorbance spectrum and (b) the Tauc's plot for anatase c -TiO_2 thin films.

4. Conclusion

This study assessed c-TiO_2 as an electron transport layer, showed its effectiveness in enhancing device performance, electron mobility, and energy alignment when we used in a solar cell as transport layer. Insight of the electron mobility and energy alignment is articulated in the work of *Chio et al*., and colleagues. The XRD analysis confirmed an anatase phase crystal structure with an average crystallite size of ±24.44 nm. Micro-strain analysis revealed a value of $\pm 9.75\times10^{-4}$, while SEM images demonstrated full surface coverage by compact grains. The UV-Vis absorbance spectrum indicated an onset absorption at 338 nm and a bandgap of 3.9 eV for the c-TiO_2 thin films. Future research could

explore alternative electron transport materials, optimize layer thickness, and investigate interlayer interactions for enhanced perovskite solar cell efficiency and stability. However, in our case we are looking at further studying the resistivity of the thin films in this study and better channel ion migration at the interface.

5. References

[1] Ansari, M.I.H., A. Qurashi, and M.K. Nazeeruddin, Frontiers, opportunities, and challenges in perovskite solar cells: A critical review. Journal of Photochemistry and Photobiology C: Photochemistry Reviews, 2018. 35: p. 1-24.

[2] Shrestha, S., X. Li, H. Tsai, C.-H. Hou, H.-H. Huang, D. Ghosh, Long carrier diffusion length in two-dimensional lead halide perovskite single crystals. Chem, 2022. 8(4): p. 1107-1120.

[3] Zhang, L., X. Zhou, C. Liu, X. Wang and B. Xu, A review on solution-processable dopant-free small molecules as hole-transporting materials for efficient perovskite solar cells. Small Methods, 2020. 4(9): p. 2000254.

[4] Kim, H.-J., H.-S. Kim, and N.-G. Park, Progress of perovskite solar modules. Advanced Energy and Sustainability Research, 2021. 2(6): p. 2000051.

[5] Ye, M., X. Hong, F. Zhang and X. Liu, Recent advancements in perovskite solar cells: flexibility, stability, and large scale. Journal of Materials Chemistry A, 2016. 4(18): p. 6755-6771.

[6] Choi, H., C.-K. Mai, H.-B. Kim, J. Jeong, S. Song, G. C. Bazan , Conjugated polyelectrolyte hole transport layer for inverted-type perovskite solar cells. Nature Communications, 2015. 6(1): p. 7348.

[7] Zheng, S., G. Wang, T. Liu, L. Lou, S. Xiao and S. Yang, Materials and structures for the electron transport layer of efficient and stable perovskite solar cells. Science China Chemistry, 2019. 62: p. 800-809.

[8] Elseman, A.M., C. Xu, Y. Yao, M. Elisabeth, L. Niu, L. Malavasi, Electron Transport Materials: Evolution and Case Study for High-Efficiency Perovskite Solar Cells. Solar Rrl, 2020. 4(7): p. 2000136.

[9] Sinha, N.K., D.S. Ghosh, and A. Khare, Role of built-in potential over ETL/perovskite interface on the performance of HTL-free perovskite solar cells. Optical Materials, 2022. 129: p. 112517.

[10] Wu, W.Q., J. F. Liao, J. X. Zhong, Y. F. Xu, L. Wang and J. Huang, Suppressing Interfacial Charge Recombination in Electron-Transport-Layer-Free Perovskite Solar Cells to Give an Efficiency Exceeding 21%. Angewandte Chemie, 2020. 132(47): p. 21166-21173.

[11] Pal, M., Garcia Serrano, P. Santiago and U. Pal, Size-controlled synthesis of spherical TiO_2 nanoparticles: morphology, crystallization, and phase transition. The Journal of Physical Chemistry C, 2007. **111**(1): p. 96-102.

Using the sf-model to describe spintronic devices

V Nolting
Vaal University of Technology, Vanderbijlpark

E-mail: **volkmarn@vut.ac.za**

Abstract. Crystals containing Rare Earth atoms are interesting candidates for spintronic devices due to the fact that magnetism and the electric current are caused by different electron groups. The exactly solvable zero bandwidth limit of the sf-model is used to calculate the conduction electron spin polarization $P\,(T,n)$, the chemical potential $\mu\,(T,n)$, and the 4f-magnetization $m\,(T,n)$. The Curie temperature T_C strongly depends on band occupation n; an approximate calculation of $T_C\,(n)$ demonstrates that it can be enhanced well into the region of room temperature. As an application to 4f-antiferromagnetism the alloy series $Ce\,Ni_{1-x}\,Co_x\,Ge_2$ is investigated. It is shown that the susceptibility $X\,(T)$ displays a peak at the Neel temperature T_N thereby indicating a phase transition into the paramagnetic phase. The existence of spin dependent electron transport properties confirm the suitability of ferromagnetic semiconductors with respect to spintronic applications.

1. Introduction

In spintronics or spin transport electronics the results of conventional magnetism and semiconductor physics are correlated. Storage and processing of data is brought together on a single chip utilizing both the electron's charge and its spin. The aim is to make appliances smaller and smaller while at the same time optimizing efficiency and reducing costs. To achieve this spin dependent electron transport phenomena are desirable [1,2] and in this respect crystals with atoms from the group of Rare Earth elements become important. These atoms contain partially filled 4f-shells that are localized at the lattice site and are responsible for the magnetic moment of the atom, while the 6s-electrons become the quasi-free conduction electrons that can move through the entire system. Magnetism and the electric current are thus caused by two different electron groups and it is precisely the interaction between these electron groups that is responsible for the many interesting phenomena in these materials, for example RKKY interaction, Kondo and spinglass behavior.

Current research on spintronics is often applied to nanoscale devices to enhance memory, computing applications, and data processing capabilities [3,4]. In semiconducting devices transport properties and spin filtering capabilities are influenced and controlled by magnetic layers utilized as either spin polarizers or analyzers. The key parameters in this regard are doping density and spin waves. While the former influences the Curie temperature of the material the latter -via the electron-magnon interaction- is responsible for many interesting phenomena of magnetic semiconductors that via the sf-exchange interaction play a crucial part of the Hamiltonian to be discussed in Section 2 below. It is the aim of this article to further investigate the role and influence of these key parameters on the properties of conventional ferromagnetic semiconductors. Even though they turn out to be promising candidates for spintronic devices, their low Curie temperatures T_C generally limit their practical applications [5]. The Curie temperature can be enhanced by increasing the band occupation n of the material. This is going to be discussed in more detail in Section 3 of this article.

It is well documented in the literature that the sf-model describes magnetism in magnetic semiconductors reasonably well [6,7]. Such model calculations are less accurate than first principles methods based on DFT calculations and often only allow a qualitative comparison with the experiment. However, model calculations have the advantage that phenomena like magnetism can be explained from specific interaction terms in the Hamiltonian; similar conclusions cannot be drawn with the same conviction from DFT calculations. W Nolting et al [8,9] combine the many body problem of the sf-model with a selfconsistent band structure calculation based on DFT to obtain highly realistic results for the Curie temperature T_C of ferromagnetic 4f-systems.

The many body problem defined by the sf-model is generally not exactly solvable. An interesting exactly solvable limiting case, namely the zero bandwidth limit [10] is presented in Section 3. Note that the zero bandwidth limit represents the strong coupling region of the sf-model where the parameter ratio $g/W \gg 1$ and the quasiparticle subbands are separated by energy gaps. On the other hand, molecular field type of approximations are only reasonable for weakly coupled systems when $g/W < 1$ and the energy bands partially overlap.

Even more promising candidates regarding spintronic devices are antiferromagnetic 4f-systems discussed in Section 4. Concluding remarks are added at the end of the article.

2. The sf-Model

The sf-model describes the local interaction between the two different electron groups in magnetic semiconductors and is defined by the Hamiltonian:

$$H = H_s + H_f + H_{sf},$$
$$H_s = \sum_{ij\sigma} T_{ij}\, c_{i\sigma}^{+}\, c_{j\sigma} + \frac{1}{2}\, U \sum_{i\sigma} n_{i\sigma}\, n_{i-\sigma}. \quad (1)$$

Here $c_{i\sigma}^{+}$ denotes the creation operator for a σ-electron at lattice site R_i, $c_{i\sigma}$ is the corresponding annihilation operator and

$$n_{i\sigma} = c_{i\sigma}^{+}\, c_{i\sigma}$$

is the operator of the occupation number. U is the intraatomic Coulomb matrix element while the T_{ij} represent the hopping integrals. They are related to the Bloch energies via Fourier transformation, namely:

$$T_{ij} = \frac{1}{N} \sum_{k} \varepsilon\,(k)\;\; e^{ik\cdot(R_i - R_j)}$$

H_s describes the system of itinerant conduction electrons that are treated as s-electrons and has the well known form of the Hubbard model. The subsystem of localized magnetic moments is described in a realistic manner by the Heisenberg model:

$$H_f = -\sum_{ij} J_{ij}\, S_i \cdot S_j \quad (2)$$

The spins at R_i and R_j interact via the exchange integrals J_{ij} that are conveniently restricted to nearest neighbors. The two subsystems are coupled by an sf-exchange, i.e. a local interaction between the 4f-spin S_i and the conduction electron spin σ_i:

$$H_{sf} = -g \sum_i \sigma_i \cdot S_i = -\frac{1}{2}\, g \sum_{i\sigma} (z_\sigma\, S_i^z\, n_{i\sigma} + S_i^\sigma\, c_{i-\sigma}^{+}\, c_{i\sigma}) \quad (3)$$

g is the intraatomic sf-exchange constant. Note that the second term of Eq (3) describes spinflip processes where an electron at lattice site R_i and spin σ is annihilated and a corresponding electron of opposite spin is created; the first term is a diagonal contribution containing the z-component of the spin operator.

The Hamiltonian of equation (1) describes a non-trivial many body problem that is generally not exactly solvable. However, there are a couple of exactly solvable limiting cases available, namely the $T = 0\ K$ case of one electron in an otherwise empty conduction band [11] and the above mentioned zero

bandwidth limit. The former case explains the existence of the magnetic polaron [12] while the latter is discussed in the following section.

Exactly solvable limiting cases are important to test the accuracy of standard approximative procedures. In the case of an exactly solvable limit one can numerically determine the discrepancy between the exact and the approximate result and from that draw conclusions regarding the more general many body problem.

3. The zero bandwidth limit

In this interesting special case [10] the dispersion of the energy states $\varepsilon\,(k)$ is neglected. In a thought experiment the lattice constant is chosen to be so large that the conduction band degenerates to a single level T_0 that due to the sf-exchange splits into four quasiparticle energies E_i with:

$$E_1 = T_0 - \frac{1}{2} g\, S$$
$$E_2 = T_0 + \frac{1}{2} g\,(S+1)$$
$$E_3 = T_0 + U - \frac{1}{2} g\,(S+1)$$
$$E_4 = T_0 + U + \tfrac{1}{2} g\, S \qquad (4)$$

with temperature and particle number dependent spectral weights. To what extent these energy levels are occupied is determined by the chemical potential $\mu\,(T,n)$ that only weakly depends on band occupation n. We obtain:

$$\mu = T_0 - \tfrac{1}{2}\, g\, S = E_1 \text{ for } n < 1$$
$$\mu = T_0 + U + \tfrac{1}{2}\, g\, S = E_4 \text{ for } n > 1. \qquad (5)$$

The other two energy levels, i.e. E_2, E_3 are either completely empty or fully occupied. However, the chemical potential never lies in the vicinity of either of these energy levels. The spin dependent band occupation numbers n_σ also depend on both T and n.

$$n_\uparrow = \tfrac{n}{2} + \langle S^z \rangle \; \tfrac{n}{2\,(S+1-n)},$$
$$n_\downarrow = \tfrac{n}{2} - \langle S^z \rangle \; \tfrac{n}{2\,(S+1-n)} \qquad (6)$$

The 4f-magnetization $\langle S^z \rangle$ is itself temperature and band occupation dependent but it cannot be fully self consistently calculated within the sf-model. Instead we use the corresponding result from the Heisenberg model and conclude that its saturation value:

$$\langle S^z \rangle\,(T=0,n) = \tfrac{S\,(S+1-n)}{S+1} \qquad (7)$$

decreases as a function of band occupation n. On the other hand, the Curie temperature T_C is enhanced according to:

$$k_B\, T_C \sim z\, J_0 \sim z\, n^{4/3}, \qquad (8)$$

as the effective coupling between the localized magnetic moments is mediated by the conduction electrons of the material [13]. Here z denotes the number of nearest neighbors in the crystal and J_0 the corresponding exchange integral. The dependence of T_C on band occupation n is plotted in Figure 1 below. Compared to the empty conduction band the Curie temperature at finite band occupations can be increased by at least a factor of 5. In the case of EuO with $T_c\,(n=0) = 66.8\,K$ in this way Curie temperatures above room temperature can be achieved. This confirms the suitability of ferromagnetic semiconductors with regards to spintronic devices. In the experiment the T_C enhancement is realized in electron doped EuO and similar increases of T_c with n have been reported in the literature [14, 15, 16].

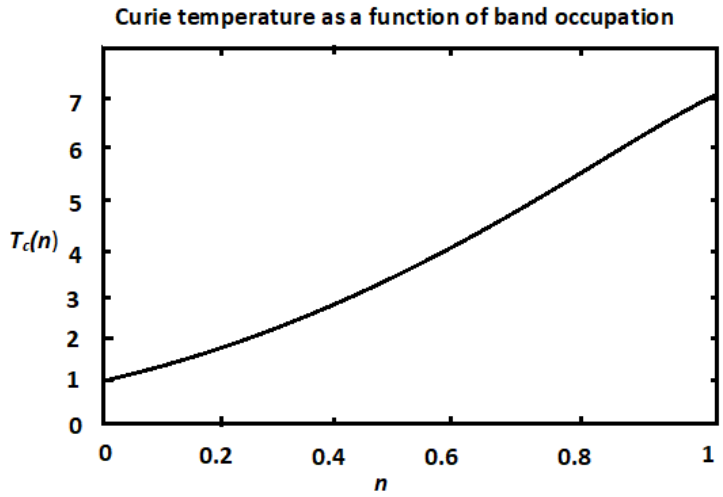


Figure 1. Curie temperature $T_c\,(n)/T_c\,(n=0)$ as a function of band occupation n.

Another interesting quantity in the context of spintronics is the conduction electron spin polarization:

$$P\,(T,n) = \frac{n_\uparrow - n_\downarrow}{n_\uparrow + n_\downarrow} \tag{9}$$

It is plotted in Figure 2 below for the case $T = 0$.

The system of conduction electrons is for practically all band occupations n at least to some degree polarized which again proves the applicability of ferromagnetic semiconductors as candidates for spintronic devices. The only exception is the case $n = 2$ which represents a fully occupied conduction band. For this case we obtain $n_\uparrow = n_\downarrow$.

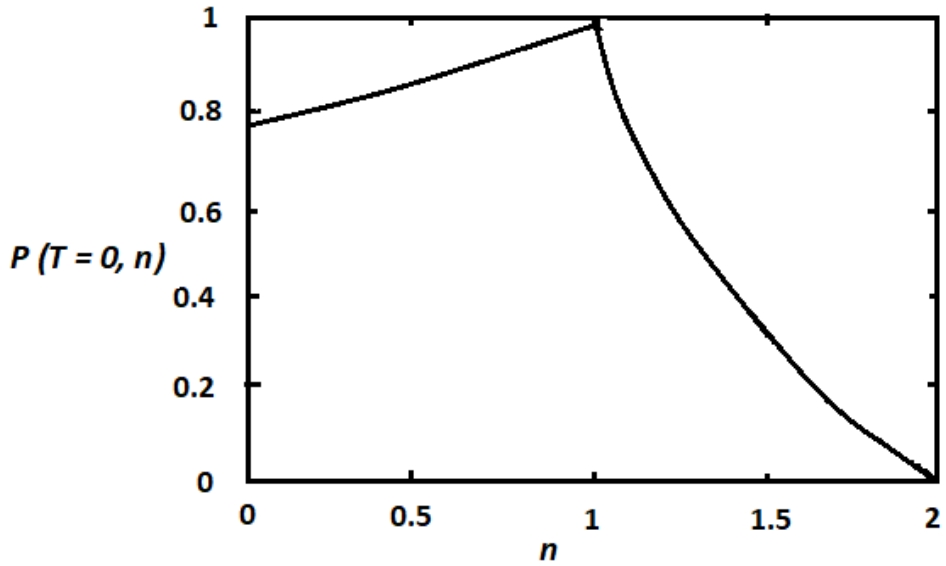


Figure 2. Conduction electron spin polarization $P(T = 0, n)$ as a function of band occupation n.

4. Applications to 4f-antiferromagnetism

Even more promising candidates for spintronic devices are antiferromagnetic 4f-systems as here the magnetic moments of neighbouring atoms point in opposite directions. The indirect coupling between the magnetic moments is of RKKY type and is already quite well described by the Heisenberg model. In the molecular field approximation the magnetization of the two sublattices is calculated from:

$$M_i\,(T,B) = M_{0i}\;B_S\left(\beta\;S\;g_J\;\mu_B\;(B + \mu_0\;\textstyle\sum_j \lambda_{ij}\;M_j\;)\right) \tag{10}$$

Here λ_{ij} denote the exchange parameters of the antiferromagnetic material while $M_{01} = M_{02}$ are the saturation magnetizations of the two sublattices. The function $B_S\,(x)$ defines the Brillouin function and its argument B is the external magnetic field. The magnetization at vanishing magnetic field is known as the spontaneous magnetization. It is selfconsistently calculated from the equation:

$$M_{is}\,(T) = M_{oi}\;B_S\;\left(\beta\;S\;g_J\;\mu_B\;\mu_0\;(\lambda_{11} - \lambda_{12})\;M_{is}\,(T)\right) \tag{11}$$

for $i = 1{,}2$. Note that the total spontaneous magnetization:

$$M_S\,(T) = M_{1s}\,(T) + M_{2s}\,(T) = 0$$

at all temperatures T. On the other hand, if $B \neq 0$ the total magnetization:

$$M\,(T,B) = M_1\,(T,B) + M_2\,(T,B) \neq 0$$

and obeys at high temperatures the Curie-Weiss law with a negative paramagnetic Curie temperature $\Theta < 0$. From the maximum value of the susceptibility the Neel temperature can be estimated. This is done for the alloy series $CeNi_{1-x}\;Co_x\;Ge_2$ which describe heavy fermion systems with Kondo behaviour and hybridization between the localized 4f and itinerant conduction electrons suggesting a possible coexistence between magnetism and an intermediate valence state. This hybridization and the magnetic state of the alloy are controlled via the angular momentum by the parameter x. Using $x = 1$ and the data of reference [17] the susceptibility $X\,(T)$ is plotted as a function of temperature T in Figure 3 below.

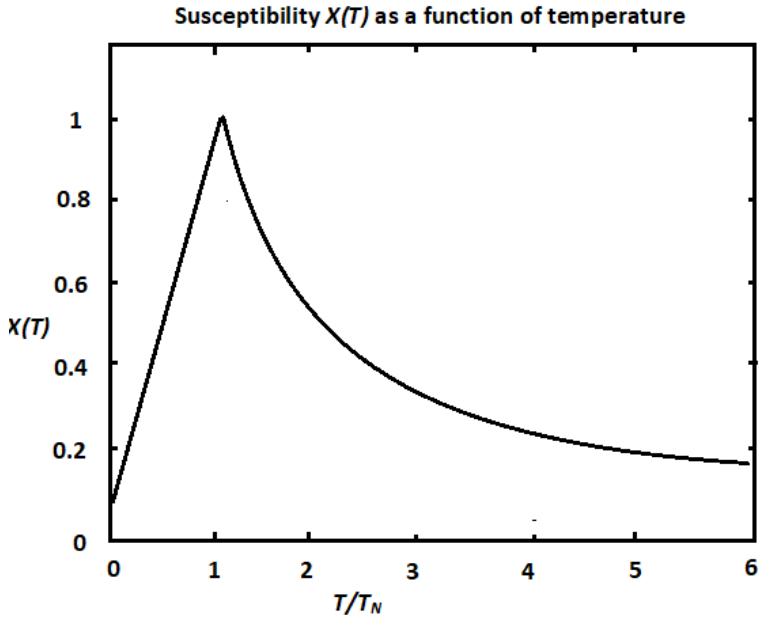


Figure 3. Susceptibility $X\,(T)$ as a function of the reduced temperature T/T_N.

$X(T)$ has the typical form of an antiferromagnetic material with:

$$X\,(T) \underset{T\to 0}{\longrightarrow} 0.$$

Note the peak in the susceptibility at $T = T_N$ thereby indicating a phase transition into the paramagnetic phase. The peak weakens with decreasing x and then finally disappears for $x < 0.6$. This

behaviour is due to increased Kondo interactions in the critical range of x-values and confirms the dependence of the magnetic state on the doping density.

5. Conclusions

In this article the sf-model is used to theoretically describe ferromagnetic semiconductors and confirm their applicability for spintronic devices. Crystals containing atoms from the group of Rare Earth elements are identified as promising candidates in this respect as their Curie temperature T_C as a function of band occupation n is found to be in the region of room temperature. Results from the exactly solvable zero bandwidth limit furthermore show that ferromagnetic semiconductors display spin dependent electron transport properties.

Antiferromagnetic 4f-systems exhibit a peak in the susceptibility at $T = T_N$. The dependence of the magnetic state on the doping density is confirmed.

References

[1] S A Wolf et al, IBM Journal of Research and Development 50, 101 (2006)
[2] S A Wolf, Science 294, 1488 (2001)
[3] A Hirohoto et al, J of Magnetism and Mag Mat 509, 166711 (2021)
[4] E R Hedin and Y S Joe, Spintronics in Nanoscale Devices, CRC Press (2013)
[5] W Liu et al, Materials TodayPhysics 21 (2021)
[6] W Nolting, phys stat sol (b) 96, 11 (1979)
[7] M I Auslender and V I Irkhin, J Phys C18 (1985)
[8] W Nolting et al, Phys Rev B35, 7015 (1987)
[9] W Nolting et al, Phys Rev B35, 7025 (1987)
[10] W Nolting and M Matlak, phys stat sol (b) 123, 155 (1984)
[11] W Nolting, U Dubil, and M Matlak, J Phys. C18, 3687 (1985)
[12] W Nolting, Fundamentals of Many Body Physics, Springer (2009)
[13] W Nolting, Quantum Theory of Magnetism, Teubner (1986)
[14] I Kuryliszyn-Kudeska et al, Semiconductor Science and Technology 21, 8 (2006)
[15] T Stollenwerk and J Kroha, Phys Rev B92, 205119 (2015)
[16] V Nolting, Proceedings of SAIP2016, 60 (2016)
[17] B K Lee et al, Phys Rev B71, 214433 (2005)

Thermodynamic stability and formation energies of hydrogen and carbon vacancy centres in hydrogenated graphene

H Mapingire and R E Mapasha
Department of Physics, University of Pretoria, Pretoria, 0002, Republic of South Africa

E-mail: **edwin.mapasha@up.ac.za**

Abstract. We utilise Quantum espresso simulations to compute formation energies, thermodynamic transition levels and effective U-parameters for the mono-vacancies V_H, V_C and V_{CH} in hydrogenated graphene. Our first-principles calculations show that the formation processes of all the three vacancy point defects are endothermic, hence they need activation energy to form. The point defects give rise to defect states that exhibit spin-polarisation and are possible candidates for novel nano-technology applications. The derived U-parameters are all positive showing the stability of the point defects in the positive and negative charged states.

1. Introduction

The comprehension of the properties of point defects such as lattice vacancies is of fundamental importance because of their various applications in nano-technology. Defects have a bearing on the life-time perfomance and efficiency of technological devices. A defect may make a material useful for various technological innovations or conversely may make a material to be unusable [1]. Considerable research on the properties of point defects in two-dimensional materials is crucial for gaining a comprehensive understanding of these defects, with the goal of utilising them in nano-scale technological devices.

In this contribution, we focus on hydrogen and carbon vacancies in hydrogenated graphene because they generally induce defect states within the bandgap of some two-dimensional materials and may be possible colour centers [2]. Graphene is a two-dimensional material that is composed of carbon atoms that are covalently bonded to each other in such a way as to form a structure that is hexagonal in nature [3]. If hydrogen atoms are chemisorbed to the carbon atoms, we form hydrogenated graphene- a two-dimensional material that has generated intense interest because of the tunable bandgap which is formed at the gamma point [4]. Hydrogenated graphene is a saturated hydro-carbon that is $sp^3-\mathrm{hybridized}$[5]. During the growth or formation processes of hydrogenated graphene [6], atomically confined point defects can form. A lattice structure can have an ion or an atom missing from a particular lattice site, giving rise to a lattice vacancy. Lattice vacancies can be formed as a result of less than perfect packing of atoms during the phenomenon of crystallization. The comprehension of the properties of point defects such as impurity atoms, Frenkel defects, Scholtky as well as lattice vacancies is of fundamental importance because of their various applications in nano-technology. Defects have a bearing on the life-time perfomance and efficiency of technological devices. A defect may make a material useful for various technological innovations or conversely may make a material to be unusable. Moreover, the

presence of vacancies in a given structure can amplify electron transport. We also propose that the presence of vacancies in a material is of paramount importance as far as the controlling of diffusion is concerned.

In our current contribution, we examine the thermodynamic stability (formation energy), charge transition level and electronic properties of the mono-vacancies V_H and V_C and V_{CH} in hydrogenated graphene for different charge states. Furthermore, we derive the U-parameters which have a bearing on the stability of these vacancies. To correctly identify the positions of transition levels and defect states, we used hybrid functionals.

2. Characterization methods

The initial and fundamental stage in the characterization of point defects is the determination of the defect formation energy. Utilizing Quantum espresso [7] first principles computations guided by the Kohn-Sham density functional theory [8], we computed the formation energy of all the vacancy configurations housed in a $6 \times 6 \times 1$ supercelll. The hydrogenated graphene supercell layer comprised a total of 143 carbon and hydrogen atoms for the V_H and V_C configurations while the V_{CH} configuration had a total of 142 atoms. With the aid of convergence testing we used wave function and charge density threshold values of $3.30 \times 10^1\ Ry$ and $3.30 \times 10^2\ Ry$ respectively. In all our entire calculations we included spin component and a force convergence threshold value of $1.0 \times 10^{-6}\ Å/eV$. We allowed our calculations to run for a maximum of 200 electron steps. The initial calculations for all the defect configurations was geometric optimization which is crucial as it allows the relaxation of all the atomic species. For better computational accuracy and improved property characterization we used hybrid functionals [9] to run our entire calculations.

We deduced the formation energy using equation 1 [10], where $E_T[V_i^q]$ denotes the supercell total energy in which a vacancy, V, of atomic type i and that has a charge state q is entwined. $E_T[pure]$ represents the pristine supercell total energy of the crystalline structure at play. The integer n_i stands for the atomic species of i type that have been removed from their lattice sites within the supercell during the vacancy creation phenomenon. The chemical potential of atomic species that have been removed from their lattice positions is represented by μ_i. The terms E_F and E_c represents the Fermi energy and the correction energy respectively.

$$E^f[V_i^q] = E_T[V_i^q] - E_T[pure] - \sum_i n_i\mu_i + qE_F + E_c \quad (1)$$

Fundamental characterization information is embedded in defect level diagrams. These diagrams are graphs of formation energy versus Fermi level positions. Transition levels are Fermi level positions whereby the formation energy of two consecutive charge states are equivalent. Thermodynamic transition levels can be deduced from the defect level diagrams. The difference in energy between successive thermodynamic transition levels yields the U-parameters [11].

3. Results

In the sub-sections that follow, we present our findings which give invaluable insights into the stability and electronic structure of the hydrogen vacancy, carbon vacancy and the carbon-hydrogen vacancy in hydrogenated graphene.

3.1. Hydrogen vacancy (V_H)

This point defect is created in hydrogenated graphene by removing a single hydrogen atom that is attached to a carbon atom giving rise to a dangling bond on the parent carbon atom as depicted by figure 1(a). The removal of a hydrogen atom causes a slight distortion of the system's atomic geometry. We derived formation energy of 2.64 eV for this hydrogen mono-vacancy as shown by table 1. This value

is in close agreement with the formation energy derived by Mapasha *et al* [12]. The positive formation energy shows the defect formation process is endothermic and non-spontaneous [13][14].

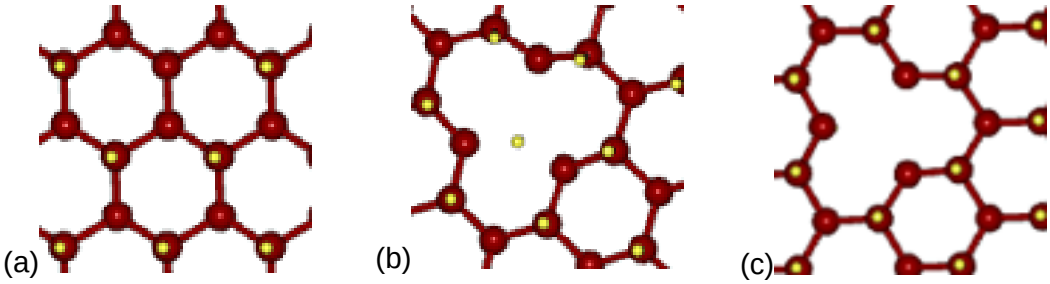


Figure 1. Vacancy point defects in hydrogenated graphene: (a) V_H (b) V_C and (c) V_{CH}. The red spheres represent carbon atoms while the yellow spheres represent hydrogen atoms.

Table1. Formation energies and Fermi-level position for vacancy point defects in graphane.

Defect	Charge state	Formation energy(eV)	Fermi-level position(eV)	Formation energy(eV)[E^F=0]
V_H	0	2.64	-0.38	2.64
	-	4.49	+1.35	3.14
	+	1.59	-2.39	3.98
V_C	0	12.57	- --	12.57
	-	12.36	+1.40	13.76
	+	13.98	-1.42	13.98
V_{CH}	0	13.75	+0.60	13.75
	-	13.41	+1.37	14.78
	+	11.53	-2.33	13.86

Charging the defect negatively increased the formation energy significantly to 4.49 eV as illustrated by table 1. The V_H defect in the neutral charge state has one unpaired electron. The addition of an extra electron may give rise to coulombic repulsive forces between the electrons inducing the noted increase in the formation energy. The formation energy of the singly positive charge state of the hydrogen mono-vacancy is 1.59 eV which is considerably low albeit positive. V_H^0 exhibits $C_S(m)$ symmetry group transformations while both $V_H{}^{-1}$ and $V_H{}^{+1}$ depicts $C_{3v}(3m)$ symmetry transformations. The V_H defect is thus subject to the Jahn-Teller distortion [15, 16] since there is a lowering of symmetry when the defect forms. The $C_{3v}(3m)$ symmetry group is subject to possible intrinsic splittings, hence there is a likelihood of utilising V_H defect for spintronics applications [17]. Analysis of the density of states (DOS) plots as depicted by figure 2 shows that the hydrogen monovacancy induces defect electronic states within the graphane bandgap. The defect states are asymmetrical in contrast to the bulk states that are clearly aligned. We anticipate this defect to show traits of magnetism because of the presence of unpaired electrons. We predict that this defect slightly widens the graphane bandgap. A closer examination of the DOS plots shows that the V_H^0 and $V_H{}^{+1}$ defect states may be coupled with the bulk

states, hence are not clearly distinct as compared to the $V_H{}^{-1}$ states. The DOS plots for $V_H{}^{-1}$ show spin-polarised defect states which we predict to be shallow donor levels close to the conduction band.

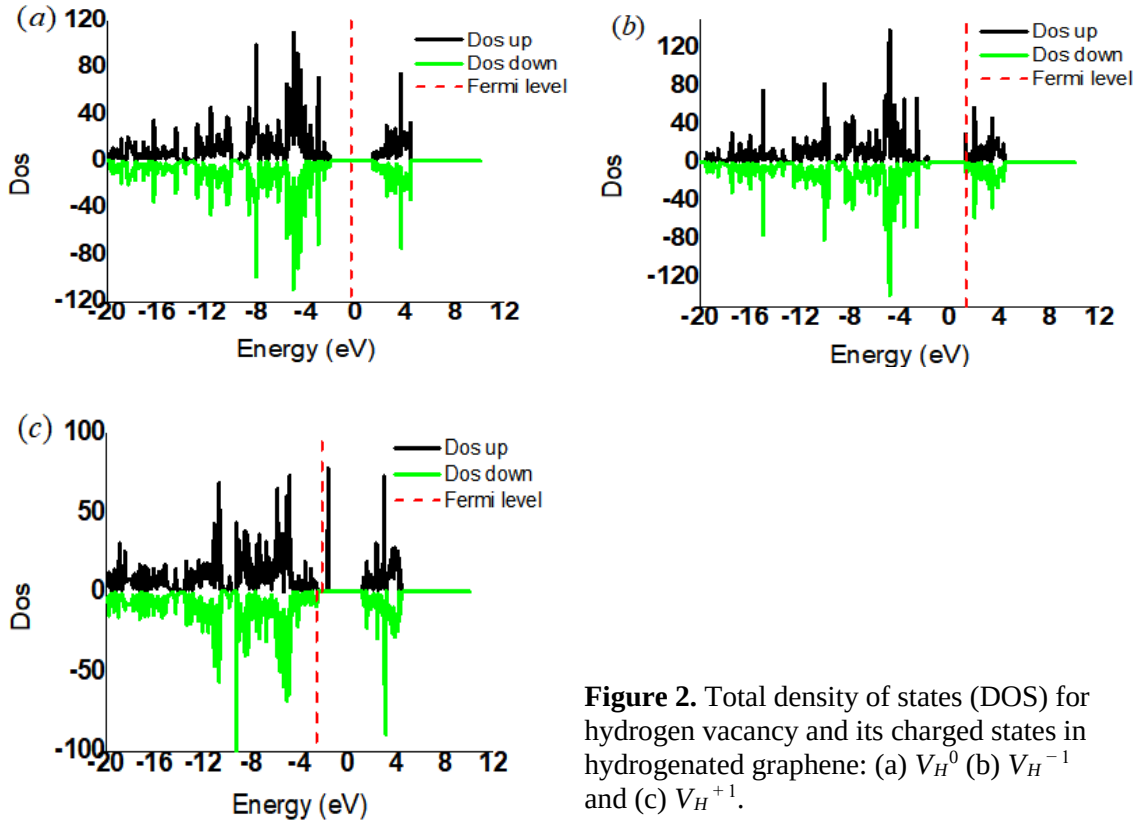


Figure 2. Total density of states (DOS) for hydrogen vacancy and its charged states in hydrogenated graphene: (a) $V_H{}^0$ (b) $V_H{}^{-1}$ and (c) $V_H{}^{+1}$.

The presence of these states close to the conduction band edge raises the possibility of utilising this point defect for nano-electronics and optical applications [18]. The calculated value of the U-parameter for this vacancy point defect is +1.84 eV. The positive U-parameter value for this point defect shows that this defect permits the formation of the singly positive and singly negative charge states without notable traits of instability. The calculated charge transition levels, $(+1/0)$ and $(0/-1)$ for this defect, are respectively -1.34 eV and 0.5 eV. This is in good agreement with the transition levels depicted by the defect level diagram of figure 3.

3.2. Carbon vacancy (V_C)

A carbon monovacancy point defect is formed when a single carbon atom is removed from its site on a hydrogenated graphene monolayer as shown by the image of figure 1 (b). The removal of a carbon atom forms four dangling bonds - three from the nearest three adjacent carbon atoms and one from the hydrogen atom that still exist in the vicinity of the created vacancy. The carbon vacancy in its neutral state, $V_C{}^0$, has four unbonded electrons. We propose that the existence of these four unpaired electrons induce notable electronic repulsive coulombic forces. As a result of this we put it forward that the formation energy of this defect is very high and positive at a value of 12.57 eV as depicted in table 1. Charging the defect negatively to form $V_C{}^{-1}$ lowered the formation energy slightly to 12.36 eV. We noted that the presence of this defect induce heavy lattice distortion of the atoms in the vicinity of the defect. $V_C{}^0$ and $V_C{}^{-1}$ exhibited the $C_s(m)$ symmerty constraints, hence there is lowering of symmetry from the high symmetry point group of the pristine structure -$C_{3m}(m)$.

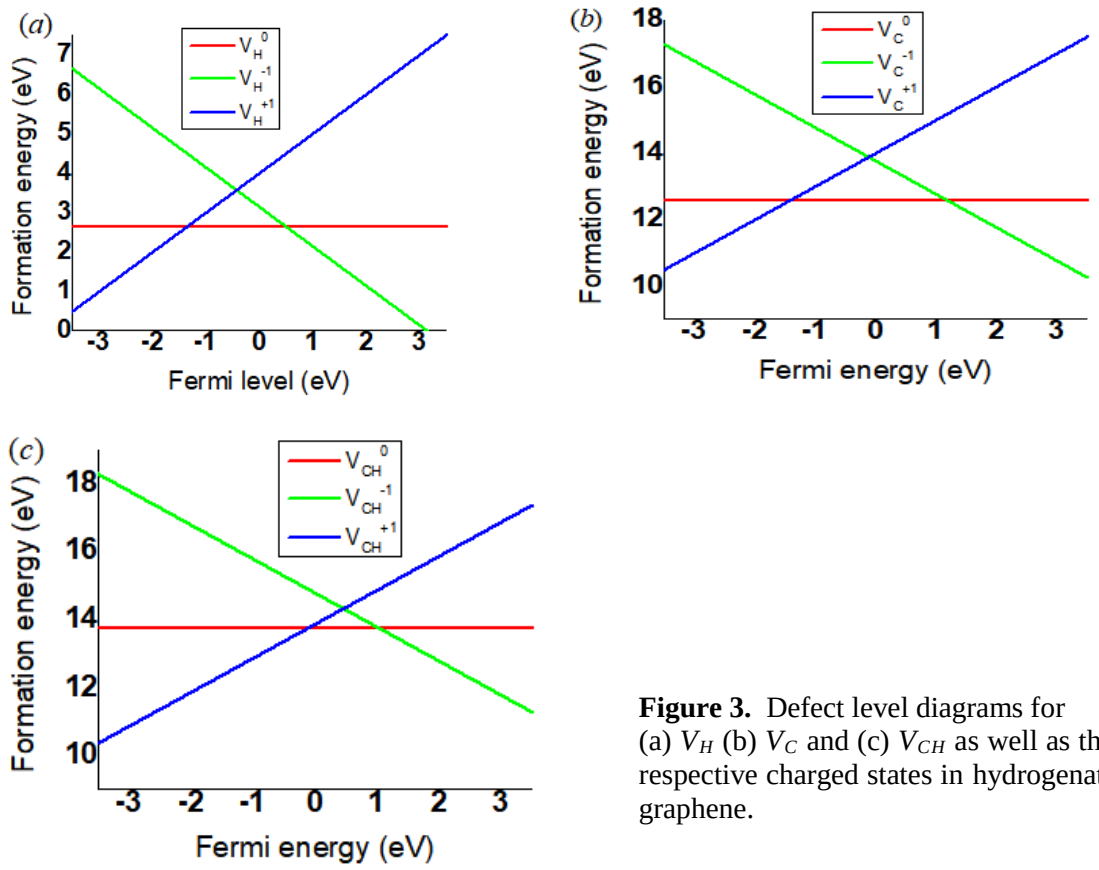


Figure 3. Defect level diagrams for (a) V_H (b) V_C and (c) V_{CH} as well as their respective charged states in hydrogenated graphene.

We thus have noted that the carbon monovacancy formation in the graphane monolayer is subject to the pseudo Jahn-Teller static distortions. The V_C defect induces deep and shallow level defect states within the graphane bandgap. A close examination of DOS plots as shown by figure 4 depicts multiple defect states within the bandgap of graphane. The asymmetric nature of these defect states clearly shows the magnetic traits as shown by figure 4. The bulk states are also asymmetrical showing that this defect can be exploited for various magnetic applications such as magnetic nano-sensors. Charging this defect to form $V_C{}^{-1}$ gave rise to a distinct spin-up state just above the Fermi level. This spin polarised defect state is also deep and considerably isolated from the band edges. Just like the $V_C{}^0$, the DOS plot for $V_C{}^{-1}$ is also non-aligned. The derived thermodynamic transition level values, $(+1/0)$ and $(0/-1)$ are respectively -1.41 eV and 1.19 eV. These values correlate with the values shown by the defect level diagram shown by figure 3(b). V_C yielded a positive value of the U-parameter +2.60 eV showing stability for charged states formation.

3.3 Carbon-hydrogen vacancy (V_{CH})

3.2. A carbon-hydrogen vacancy is formed on hydrogenated graphene when both a carbon atom and the hydrogen atom attached to the very same carbon atom are simultaneously removed from their sites as illustrated by figure 1(c). It is energetically costly to form this defect. The calculated formation energy is 13.75 eV for $V_{CH}{}^0$as shown in table 1. Charging the defect reduces the formation energy to 13.41 eV for $V_{CH}{}^{-1}$ and to 11.53 eV for $V_{CH}{}^{+1}$. The defect symmetry for the neutral and charged states is $C_s(m)$.

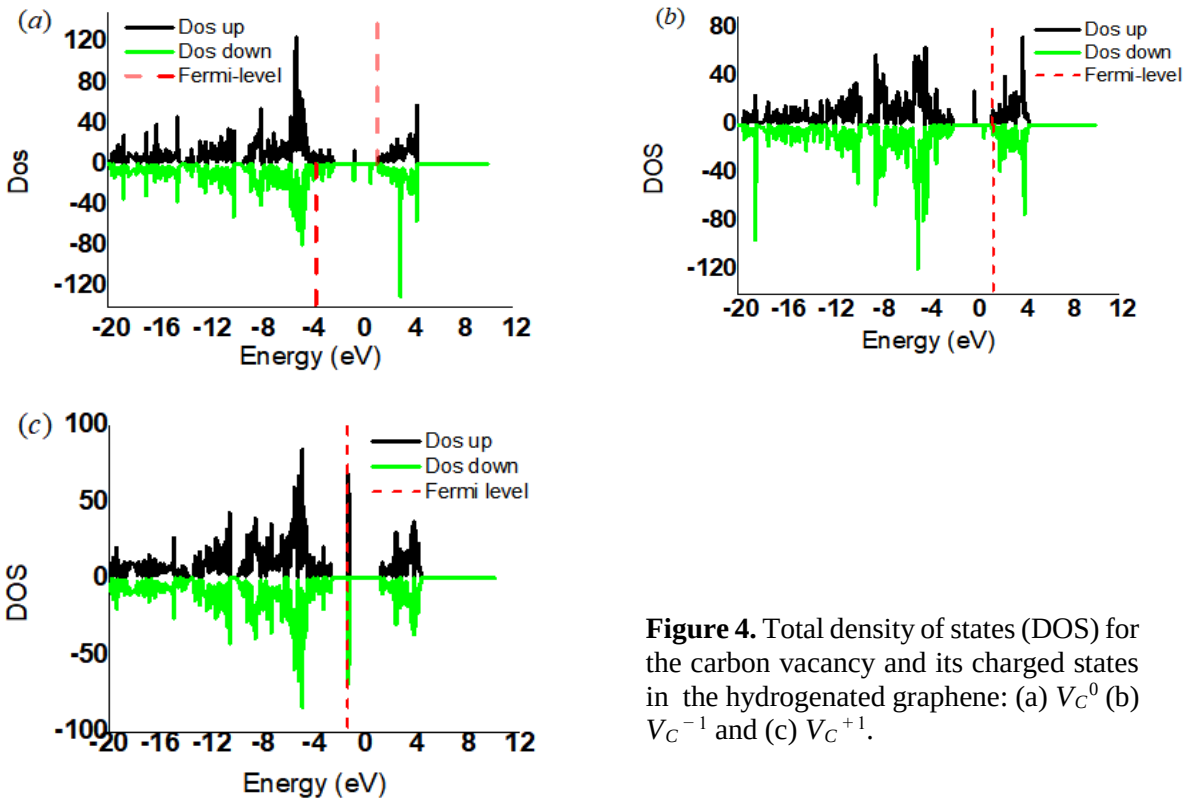


Figure 4. Total density of states (DOS) for the carbon vacancy and its charged states in the hydrogenated graphene: (a) V_C^0 (b) V_C^{-1} and (c) V_C^{+1}.

This symmetry is preserved for all the configurations we considered. We can thus propose that the defect transition from one thermodynamic charge state to another is not subject to Jahn-Teller distortions. V_{CH} point defect in the neutral charge state induces spin-polarised states in the bandgap of graphane. The defect induces magnetism since the both the defect states and the bulk states are asymmetrical as shown by figure 5. Careful analysis of the total DOS plots shows multiple defect states close to the band edges for V_{CH}^0. The defect states clearly induces magnetism, hence there is a real possibility of manipulating the V_{CH} point defect for various magnetic applications. The negatively charged state of this defect , V_{CH}^{-1} , also induces spin-polarised defect states within the hydrogenated graphene bandgap. The calculated charge transition levels $(+/0)$ and $(0/-1)$ are respectively -0.11 eV and 1.03 eV and these values align with those depicted by the defect level diagram of figure 3(c). V_{CH}^{-1} is also magnetic and brings about both donor and acceptor states close to the band edges. The transition from V_{CH}^0 to V_{CH}^{-1} seems to quench the magnetism. We also noted that the transition of the defect from the neutral state to the singly positive charged state seems to completely eliminate the magnetic traits of the defect. However, this transition give rise to spin-up states and spin-down states that are non-spin polarised. The spin-up states in the middle of the bandgap are symmetric to the spin-down states, hence we propose that V_{CH}^{+1} is non-magnetic. V_{CH} yielded a positive value of the U-parameter which is 1.14 eV.

4. Summary and conclusion

With the aid of first-principles computations we studied the effects of the presence of vacancy point defects in hydrogenated graphene. We characterised hydrogen and carbon monovacancies of the type V_H, V_C and V_{CH}. These vacancy point defects yielded relatively high positive formation energies. The implication of this finding is that these point defects may require activation energy for them to form.

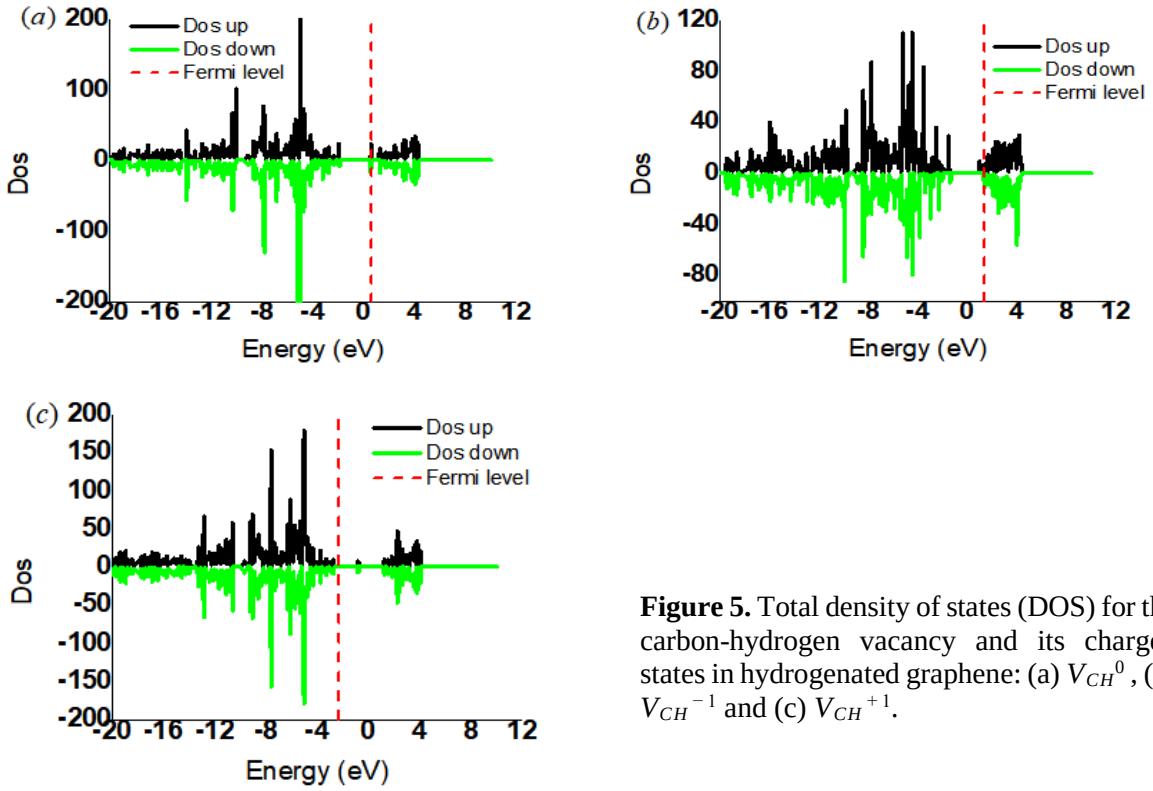


Figure 5. Total density of states (DOS) for the carbon-hydrogen vacancy and its charged states in hydrogenated graphene: (a) $V_{CH}{}^{0}$, (b) $V_{CH}{}^{-1}$ and (c) $V_{CH}{}^{+1}$.

Moreover, the concentration of vacancy point defects in two-dimensional materials may be very low making it difficult to observe and analyse them except with the use of very sophisticated instruments which are not readily available to all and sundry. V_H has lower formation energies as compared to V_C and V_{CH}. We can therefore propose that, of the three types of point defects considered, V_H is likely to form as compared to the other two. We also noted that vacancies in hydrogenated graphene modify the electronic structure of this material because of the creation of spin-polarised defect states between the bandgap. A possible positive implication of this is the utilization of these point defects for bandgap engineering in hydrogenated graphene [19]. We also found out that these vacancy point defects alter the chemical and optical properties of this two-dimensional material, raising the possibility of utilising them as for novel nano-technology applications. These point defects are also generally subject to the pseudo Jahn-Teller distortion [20] which encompass a decrease in the lattice system's energy by way of eliminating the lattice structure's degeneracy. One critical finding of this contribution is that all the considered point defects yielded positive effective U-parameters showing that they permit the formation of singly positive and singly negative charged states without giving rise to any form of instability. The acceptor levels $(0/-)$ of the vacancy point defects have higher energies as compared to the donor levels $(+/0)$ and hence the arrangement of the thermodynamic transition states is not inverted showing that these defects are stable in their respective charged states. Among other challenges, a comprehensive analysis of vacancy point defects is affected by factors such as temperature, characterization techniques as well as their inevitable interaction with some other defects. As an ongoing study, it is imperative to investigate the nitrogen-vacancy complexes in hydrogenated graphene in order to determine if it is possible to reduce the high activation energy.

Acknowledgements

We would like to express our heart-felt gratitude to the University of Pretoria post-graduate funding for providing the sponsorship for this project. We also would like to thank the Centre for High Perfomance Computing (CHPC) for availing first principle simulation resources. REM also thanks NiTheCS for financial support.

References

[1] Ashcroft N W, Mermin N D and Wei D 2016 *Solid State Physics* (L. S. Han, ed.) Cengage Learning
[2] Yu P Y, Cardona M 1996 *Fundamentals of Semiconductors: Physics and Materials Properties*
[3] Novoselov K S, Geim A K, Morozov S V, Jiang D, Zhang Y, Dubonos S V, Grigorieva I V and Firsov A 2004*Science***306**666-69
[4] Lebegue, S.; Klintenberg, M.; Eriksson, O.; Katsnelson, M. I. 2009 *Phys. Rev. B: Condens. Matter Mater. Phys.***79** (24), 245117
[5] Podlivaev A I and Openov L A 2017 *Condensed Matter***106**110-15
[6] Whitener K E 2018 *J. Vac. Sci. Technol.* ***A*** **36**, 05G401, 1-15
[7] Giannozzi P *et al* 2009 *Condens. Matter* **21**395502
[8] Kohn W and Sham L J 1965*Phys. Rev.*, **140**, A1133
[9] Heyd J, Scuseria G E and Ernzerhorf M 2003*Journal of Chemical Physics***118** 8207-8215.
[10] Zhang S B and Northrup J E 1991*Phys. Rev. Lett.***67** 2339
[11] Dreyer C E, Alkauskas A, Lyons J L, Janotti A and Van de Walle C G 2018 *Annu. Rev. Mater. Res.* **48** 2.1-2.26
[12] Mapasha R E, Molepo M P and Chetty N 2016*Physica E***79** 52-58
[13] Kittel C 2005 *Introduction to Solid State Physics* (Stuart Johnson, ed.). John Wiley and Sons, Inc.
[14] Patterson J and Bailey B 2010 *Solid-State Physics* (J. D Patterson and B. C. Bailey, ed.). Springer, 2010.
[15] Gehring G A and Gehring K A 1975*Rep. Prog. Phys.***38**1-89
[16] Bersuker I B 2001*Chem. Rev*.101, **4** 1067-1114
[17] Gali A 2019*Nanophotonics***8** (11)
[18] Freysoldt C, Grabowski B, Hickel T, Neugebauer J, Kresse G, Janotti A and Van de Walle C G, 2014*Rev. Mod. Phys.***86**. 253
[19] Lebegue S, Klintenberg M, Eriksson O, and Katsnelson M I 2009*Phy. Rev*. **B79** 245117
[20] O'Brien M C M and Chancey C C, 1993 *Am. J. Phys.***61**688-697

A study of the thermal response of chromium-tin (Cr/Sn) bimetal films using *in situ* RBS by Artificial Neural Networks

I.K. Segola[1,2], G. Magchiels[3], K. Lebesana[1,2], Z. Khumalo[2], C.B Mtshali[2], L. Kotsedi[2], A. Vantomme[3], M. Msimanga[1,4]

1 Tshwane University of Technology, Department of Physics, P Bag X680, Pretoria 0001, South Africa
2 NRF-iThemba LABS, Tandetron Laboratory, P.O Box 722, Somerset West 7129, Cape Town, South Africa
3 Quantum Solid-State Physics, KU Leuven, Celestijnenlaan 200D, 3001 Leuven, Belgium
4 NRF-iThemba LABS, TAMS Laboratory, Private Bag WITS 2050, Johannesburg, South Africa

E-mail: **kutlwanosegola@gmail.com**

Abstract. Bimetallic thin films of chromium-tin were prepared using an electron beam evaporator, the films were prepared such that a layer of chromium was deposited onto a thin film of tin. The first sample was then annealed in the Rutherford Backscattering Spectrometry (RBS) chamber under high vacuum while probing the sample with He^{2+} ions. The RBS spectra were collected *in situ* while annealing at a ramp rate of 3°C/min from room temperature to 600°C, using a 3 MeV He^{2+} ion beam to probe atomic interdiffusion with depth. The collected spectra were analysed using Artificial Neural Networks (ANNs). The ANN model was trained using a set of simulated RBS spectra generated using the SIMNRA software package. The results of the stochiometric analysis showed that the layers of Cr-Sn start phase change at temperatures of about 275°C and 475°C for the CrSn and Cr_2Sn phases, respectively. The results of this study demonstrate the potential of ANNs in the analysis of *in-situ* RBS spectra and the importance of considering the thermal effects during thermal treatment.

1. Introduction

Metal bilayers are structures composed of two layers of different metals that are typically stacked on top of each other. These bilayers have a wide range of applications in various fields due to their unique properties and the ability to tailor their characteristics based on the chosen metals [1]. Chromium-tin (Cr/Sn) bilayers can offer unique properties, with Cr being resistant to oxidation, along with Sn having a low melting point, they can be combined for improved solder performance in specialized applications [2]. Furthermore, these properties can be enhanced through thermal annealing. Monitoring the reaction kinetics during a thermal anneal using *in situ* real-time Rutherford Backscattering Spectrometry (RBS) allows observation of the solid phase reactions that the bimetal layers undergo; by tracking stoichiometric changes, new phases and/or compounds formed and the rate at which they react [3].

RBS is a powerful well-established non-destructive characterization technique for analysing material composition and atomic depth profiling. RBS is widely used for thin film characterization and enables discrimination of the atomic depth profiling of the near-surface area of solid materials [4] [5]. RBS works on the principle of accelerating charged particles (typically He^{2+} ions) and impinging them upon the target of a material, the particles are backscattered at an angle and detected by a solid-state detector. Notably, *in situ* real-time RBS enables the tracking of diffusion kinetics and stoichiometric changes

during the thermal treatment of thin films. However, the analysis of spectra collected in real-time can be challenging and time-consuming due to the large amount of data involved [6]–[9].

The recent progress in machine learning and Artificial Neural Networks (ANNs) provides a solution to this challenge. These networks are trained to address specific problems by acquiring the ability to establish connections between a wide range of potential inputs and their corresponding output parameters. By leveraging the capabilities of ANNs, *in-situ* real-time RBS spectra can be efficiently and accurately analysed [9]–[15]. Earlier research conducted by N.P. Barradas [10], N.R. Nené [11], J. Demeulemeester [13], and T. Silva [15] employed ANNs to analyse RBS data, yielding significant outcomes. Their work demonstrated the potential of ANNs in effectively handling RBS data, indicating the feasibility of utilizing these neural networks for RBS analysis. Consequently, this approach holds promise for addressing the laborious process of dealing with the massive data that is associated with the *in situ* RBS measurements [14], [15].

This paper reports on the RBS data collected in real-time during a thermal annealing of chromium-tin (Cr/Sn) bimetal layers, and utilising ANNs to analyse the collected data.

2. Experimental Data

Bimetal layers of Cr/Sn were deposited sequentially onto a silicon dioxide (SiO_2) substrate using an electron beam evaporator (e-beam) system. The deposition was carried out under a vacuum to minimise any cross-contamination. A constant deposition rate of approximately 0.9 $Å.s^{-1}$ was maintained, with the working pressure of the system fluctuating between 1×10^{-5} and 1×10^{-6} mbar, and the filament current kept at about 7.9 mA.

The *in-situ* RBS measurements were carried out using a 3 MeV monoenergetic $4He^{++}$ beam from the Tandetron accelerator. The sample was mounted on the annealing stage and tilted at an angle of 10^{o} towards the detector that was positioned at a scattering angle of 165^{o} with a base pressure of the chamber measured to be 10^{-7} mbar. A beam current of roughly 50 nA was used throughout the measurement to avoid ion beam-induced heating of the samples. The thermal annealing stage is connected in the RBS chamber to enable *in situ* measurements, and the temperature was linearly ramped from room temperature to 600 ℃, with a ramp rate of 3 ℃ per minute.

3. Artificial Neural Networks

ANNs are a subset of machine learning and are great for generalising arbitrary functions without the knowledge of the underlying physics or the scientific principles. Neural networks are non-parametric, and this in turn enables the models to be developed without any prior knowledge of the distribution of the given data population, and how the variables (features) of the data are connected [16]. Various kinds of neural networks such as recurrent neural networks (RNNs), and convolutional neural networks (CNNs) are available to address diverse problems.

In this investigation, we opted for a fully interconnected feedforward neural network, selecting it based on its suitability rather than its complexity. A feedforward neural network is constructed with layers of neural nodes that establish sequential connections. These connections have initial random weights that undergo constant updates and adjustments throughout the training process. One of the important parts of a neural network is the architecture which is defined by the number of input nodes, hidden layers, and output layer. Additionally, activation functions in between the layers play a pivotal role in introducing non-linearity to the neural network's operations to help it capture the patterns that appear in the data [16]–[18].

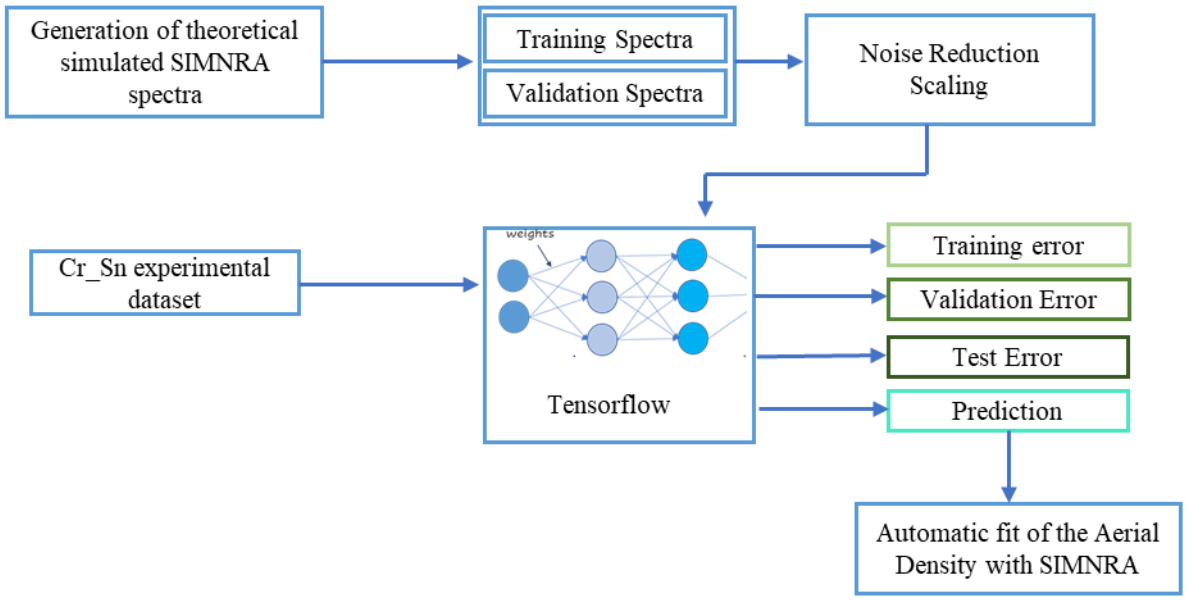


Figure 1. Automated cycle of the Artificial Neural Networks training and data processing. A Python code developed to create massive synthetic spectra by SIMNRA through Microsoft OLE automation [*adapted from Guimarães et al* [14]].

Training the ANN requires a training set, and 15000 synthetic spectra were generated from SIMNRA [19] for this task. The examples were generated by connecting SIMNRA via its Microsoft OLE automation functionality provided in the SIMNRA users' guide. After generating the examples and subsequently rescaling them to match the elemental peaks in the spectra of interest, the ANN was trained with a custom TensorFlow [20] library class (see Figure 1). The neural network's inputs included the RBS signal data, specifically the channels and corresponding counts. Meanwhile, the network's output was focused on determining the thickness of the Cr, and Sn layers within the thin films. Figure 2, shows the different architectures that were trained with the same number of examples, and the mean squared error was used as a benchmark in choosing which architecture to work with. The architecture (x,100,150, y) was chosen as the architecture of choice, with x being the number of input nodes and y the number of output nodes.

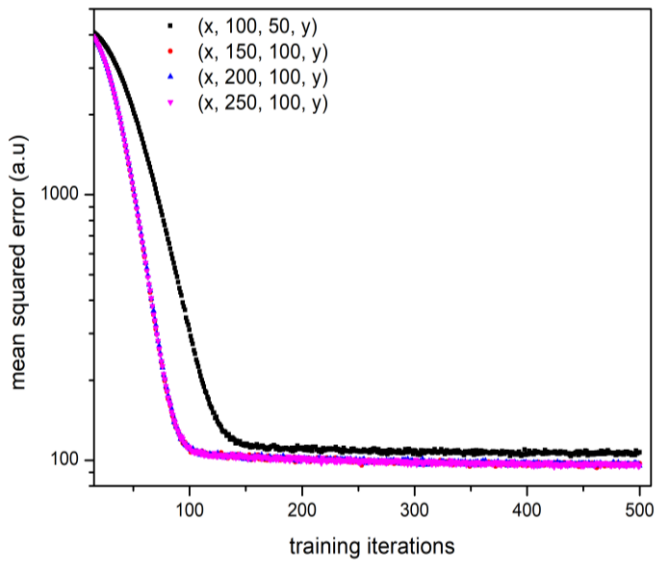


Figure 2. The mean squared error (mse) of neural networks trained in different architectures with the same number of inputs (x) and outputs (y) against the training iterations (epochs).

4. Results and Discussion

In the real-time RBS contour plot below (Figure 3), three distinct regions have been marked, these regions correspond to the three RBS graph measurements shown adjacent to the contour plot. During this analysis of the RBS plot, it was shown that the thin films were deposited with an unintended palladium (Pd) contamination, which inadvertently participated in the thermal diffusion process along with the intended bimetal layers thus resulting in an unforeseen compositional variation in the thin films.

The initial low-temperature region (Fig 3. c) induced no observable reaction. The mid-range temperatures (Fig 3.b) initiated the reaction involving Pd and Sn as seen by the development of a shoulder peak at the higher channel side of the spectrum at temperatures of about 245 °C. The third region (Fig 3.a) represents high temperatures of above 475 °C at which significant changes were observed, signifying the formation of compounds from the subsequent reaction between Sn, Pd, and Cr. Figure 4.a below presents the RBS spectra before and after the thermal annealing of the films using *in situ* RBS setup. The spectra indicate that Sn and Pd reacted resulting in the formation of a singular compound marked by a single distinct peak. Subsequently, this compound engaged in a reaction with Cr. Additionally, only one oxygen peak from the substrate is visible at low temperatures as compared to two peaks that were observed at a high temperature, signifying the release of oxygen.

Figure 4.b illustrates the variation in the thickness (depth) of the Cr and Sn (Pd) layers as a function of the annealing temperature. This thickness, measured in atoms/cm², has been determined through analysis using ANNs. The figures also demonstrate how real-time RBS enables accurate quantitative assessment of the phases formed during reactions. Notably, despite the absence of conventional analytical methods applied to this dataset, the outcomes derived from the ANN analysis have proven to be aligned with the anticipated results from the thermal annealing. Gaining an understanding into the stoichiometric changes can help directly influence the stability and performance of the bilayers in specialized applications.

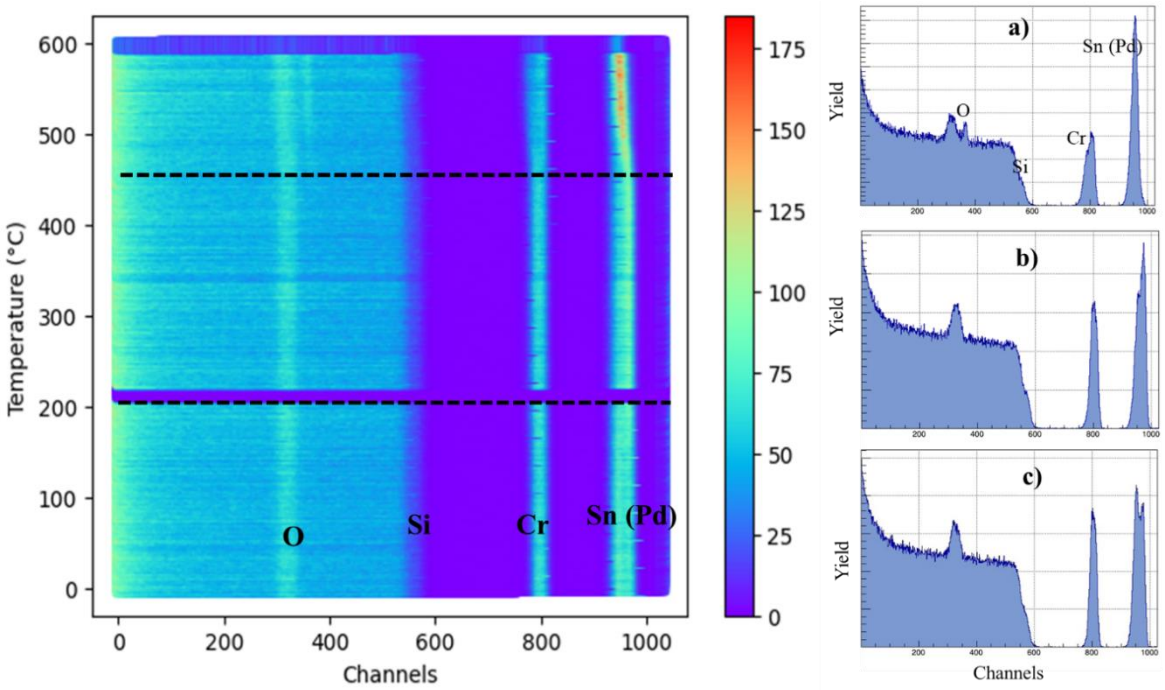


Figure 3. The contour plot of the temperature effects on the depth evolution of the Cr/Sn thin films. The different spectra at each temperature range. With a) at the high temperatures, b) mid-level temperatures, and c) at the low temperature.

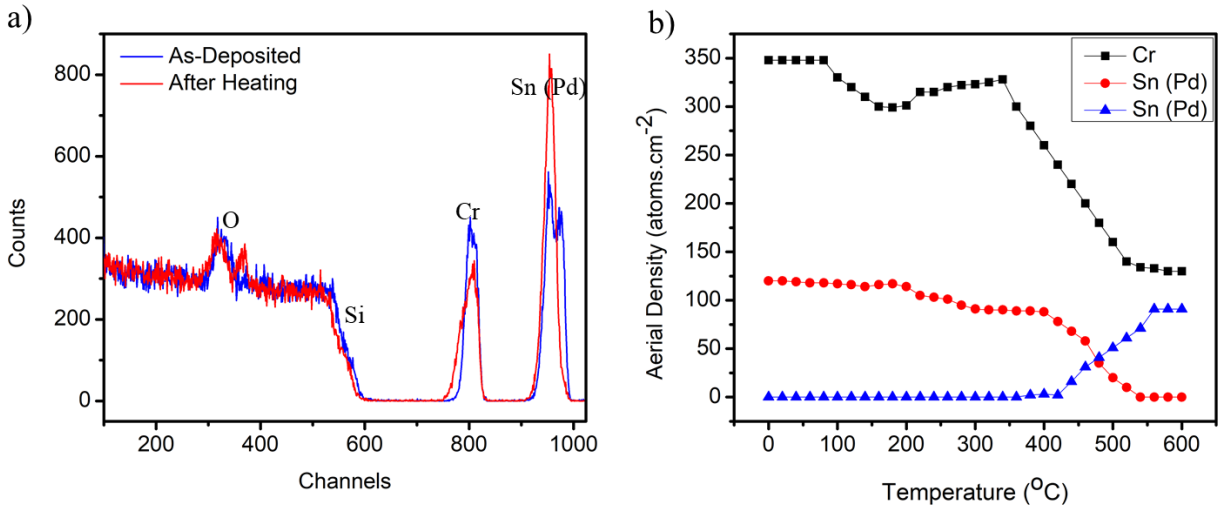


Figure 4. a) The RBS plot of the as-deposited films and after thermally annealing the films from room temperature to 600 ℃, and b) The ANN evaluation of the depth profiling of the Sn (Pd) and Cr during the temperature ramp.

5. Conclusions and Future Work

The incorporation of ANNs as an analytical tool proved to be highly valuable and efficient in analysing RBS data and studying the diffusion kinetics of Cr/Sn. While ANNs and RBS are established tools but by utilising ANNs, we successfully monitored the diffusion process in real-time and tracked the stoichiometric changes of the various elements involved in the Cr/Sn system. This indicates the effectiveness and accuracy of the ANNs in providing valuable insights into the material transformations during the process of thermal annealing. Moving forward, conducting studies without the presence of contaminants will undoubtedly amplify the utility of this methodology. Additionally, future investigations can explore the optimization of the ANN model, data collection techniques, and analysis algorithms to further improve the accuracy and reliability of the results obtained.

References

[1] A. Bandyopadhyay, Y. Zhang, and B. Onuike, 'Additive manufacturing of bimetallic structures', *Virtual and Physical Prototyping*, vol. 17, no. 2, pp. 256–294, Apr. 2022, doi: 10.1080/17452759.2022.2040738.

[2] R. Rashidi and H. Naffakh-Moosavy, 'The influence of chromium addition on the metallurgical, mechanical and fracture aspects of Sn–Cu–Bi/Cu solder joint', *Journal of Materials Research and Technology*, vol. 15, pp. 3321–3336, Nov. 2021, doi: 10.1016/j.jmrt.2021.10.015.

[3] G. L. P. Berning, C. C. Theron, and H. C. Swart, 'A real-time RBS study of the reaction of Tb and Dy with SiO2', *Applied Surface Science*, vol. 157, no. 3, pp. 129–134, Apr. 2000, doi: 10.1016/S0169-4332(99)00514-0.

[4] W.-K. Chu, *Backscattering Spectrometry, Wei-Kan Chu, James W. Mayer, and Marc-A. Nicolet, Academic Press, 1978*. 1978. doi: 10.13140/RG.2.1.1948.0807.

[5] W. K. Chu and J. R. Liu, 'Rutherford backscattering spectrometry: reminiscences and progresses', *Materials Chemistry and Physics*, vol. 46, no. 2–3, pp. 183–188, Nov. 1996, doi: 10.1016/S0254-0584(97)80012-0.

[6] C. C. Theron, J. A. Mars, C. L. Churms, J. Farmer, and R. Pretorius, 'In situ, real-time RBS measurement of solid state reaction in thin films', *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 139, no. 1, pp. 213–218, Apr. 1998, doi: 10.1016/S0168-583X(97)00946-4.

[7] C. C. Theron, J. C. Lombaard, and R. Pretorius, 'Real-time RBS of solid-state reaction in thin films', *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 161–163, pp. 48–55, Mar. 2000, doi: 10.1016/S0168-583X(99)00882-4.
[8] D. Smeets, B. C. Johnson, J. C. McCallum, and C. M. Comrie, 'Real-time in situ study of hydrogen diffusion in amorphous Si formed by ion implantation', *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 269, no. 22, pp. 2657–2661, Nov. 2011, doi: 10.1016/j.nimb.2011.07.103.
[9] K. Kantre *et al.*, 'SIGMA: A Set-up for In-situ Growth, Material modification and Analysis by ion beams', *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 463, pp. 96–100, Jan. 2020, doi: 10.1016/j.nimb.2019.11.007.
[10] N. P. Barradas and A. Vieira, 'Artificial neural network algorithm for analysis of Rutherford backscattering data', *Phys. Rev. E*, vol. 62, no. 4, pp. 5818–5829, Oct. 2000, doi: 10.1103/PhysRevE.62.5818.
[11] N. R. Nené, A. Vieira, and N. P. Barradas, 'Artificial neural network analysis of RBS and ERDA spectra of multilayered multielemental samples', *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 246, no. 2, pp. 471–478, May 2006, doi: 10.1016/j.nimb.2006.01.016.
[12] N. P. Barradas, A. Vieira, and R. Patrício, 'Artificial neural networks for automation of Rutherford backscattering spectroscopy experiments and data analysis', *Phys. Rev. E*, vol. 65, no. 6, p. 066703, Jun. 2002, doi: 10.1103/PhysRevE.65.066703.
[13] J. Demeulemeester *et al.*, 'Artificial neural networks for instantaneous analysis of real-time Rutherford backscattering spectra', *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 268, no. 10, pp. 1676–1681, May 2010, doi: 10.1016/j.nimb.2010.02.127.
[14] R. D. S. Guimarães *et al.*, 'Processing of massive Rutherford Back-scattering Spectrometry data by artificial neural networks', *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 493, pp. 28–34, Apr. 2021, doi: 10.1016/j.nimb.2021.02.010.
[15] T. S. Silva *et al.*, 'Processing of massive Rutherford Back-scattering Spectrometry data by artificial neural networks', 2019, doi: 10.1016/j.nimb.2021.02.010.
[16] C. E. Lawson *et al.*, 'Machine learning for metabolic engineering: A review', *Metabolic Engineering*, vol. 63, pp. 34–60, Jan. 2021, doi: 10.1016/j.ymben.2020.10.005.
[17] L. Giambagli, L. Buffoni, T. Carletti, W. Nocentini, and D. Fanelli, 'Machine learning in spectral domain', *Nat Commun*, vol. 12, no. 1, p. 1330, Feb. 2021, doi: 10.1038/s41467-021-21481-0.
[18] A. El-Shahat, 'Introductory Chapter: Artificial Neural Networks', 2018. doi: 10.5772/intechopen.73530.
[19] M. Mayer, *SIMNRA User's Guide for version 7.03, Max-Planck-Institut für Plasmaphysik*. Garching, Germany, 2020.
[20] M. Abadi *et al.*, 'TensorFlow: A System for Large-Scale Machine Learning', in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, Savannah, GA: USENIX Association, Nov. 2016, pp. 265–283. [Online]. Available: https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi

Magnesium-silicon alloy phase stability predictions: a cluster expansion study

K Chuene, K Phoshoko, P Ngoepe
Materials Modelling Centre, University of Limpopo, Private bag X1106 Sovenga 0727, South Africa

E-mail: **katlego.phoshoko@ul.ac.za**

Abstract. Rechargeable batteries can help meet the demand for portable energy and ease energy challenges in nations like South Africa. Hence, rapid development of innovative high-energy battery electrode materials is crucial. Magnesium-silicon (Mg-Si) anodes emerged as alternatives to graphitic anodes used in current lithium-ion battery technologies. However, their commercialisation is deterred by obstacles including incompatibility with conventional battery electrolytes. This study explored the material space of the Mg-Si alloy system through crystal structure predictions using a machine learning (ML) powered cluster expansion technique (CE). The aim was to predict and analyse novel and stable phases of the magnesium-silicon system whose compatibility with electrolytes shall be interrogated in future work. The genetic algorithm within CE was used to accelerate the material discovery by performing ground-state searches for stable structures in the Mg-Si system. Mg_5Si (Imma), Mg_2Si (P-3m1), MgSi (P4-mmm), Mg_3Si (Pm-3m), $MgSi_3$ (P2-m), $MgSi_3$ (Im-3m), and Mg_3Si (P4/mmm) were predicted to be thermodynamically stable using the BCC and FCC-based parent lattices. These structures' geometric properties are presented. The orthorhombic magnesium-rich phase, Mg_5Si, was found to be structurally stable and a promising candidate for Mg-Si-based rechargeable batteries. This study's findings are crucial to the advancement of high-energy battery electrode materials for next-generation batteries.

1. Introduction

Rechargeable batteries have significant potential for meeting the ongoing need for portable energy and relieving energy problems, especially in developing countries such as South Africa. To address this issue effectively, the research of novel and improved high-energy battery electrode materials must be accelerated. In this context, magnesium-silicon (Mg-Si) anodes have emerged as potential substitutes for traditional graphitic anodes in existing lithium-ion battery technologies [1]. However, the widespread commercialisation of Mg-Si alloy-based batteries is impeded by a variety of impediments, including Mg-Si anode incompatibility with standard battery electrolytes [2,3].

Traditional approaches to studying electrode materials rely heavily on time-consuming experimental methods. However, increases in computing power and a growing number of methodologies have transformed the examination of even the most complicated systems. In contrast to experimental procedures, intelligence-based techniques, such as those built on artificial intelligence frameworks, allow the determination of material properties that would otherwise be challenging to accomplish in such short amounts of time [4 - 7].

In this work, the material space of the Mg-Si alloy system was explored through crystal structure predictions conducted via a machine learning (ML) powered cluster expansion technique (CE) [8]. The genetic algorithm was used to accelerate the material discovery process by performing the ground-state search of stable structures from the Mg-Si alloy system.

2. Methodology

Cluster expansion (CE) is a computer approach capable of expressing the energy of a material system as a function of occupation variables at each lattice position. The lattice's distribution of atomic species (including vacancies) is changed, and the energies of the resulting configurations are calculated. The energy $E(\vec{\sigma})$ of structure $\vec{\sigma}$ is broken down into clusters with their associated effective interaction energies. The clusters are then chosen with material attributes determined using ab initio methods and reference values. This makes it possible to predict novel materials, as such CE is used to perform ground-state phase structure predictions on the Mg-Si system.

2.1. Cluster Expansion

The cluster expansion calculations were performed using the Universal Cluster Expansion Code (UNCLE) [9, 10]. The grid search was based on the FCC and BCC – based parent lattices. The genetic algorithm was used to optimize the cluster expansion and determine the optimal set of effective cluster interactions. The leave-one-out cross-validation score fitness criterion is used to qualify the accuracy of the cluster expansion prediction.

3. Results and discussions

3.1. Phase stability predictions

As seen in the binary ground-state diagrams shown by Figures 1 and 2, the cluster expansion was used to predict new structures. Predicted structures are represented by the green squares with a cross inside. The red line represents the density functional theory (DFT) ground-state line. Structures that are located on the ground-state line are considered thermodynamically stable.

3.1.1. *Binary ground-state diagrams.*

The parent lattice from which the cluster expansion is performed influences what the genetic algorithm produces as progeny (i.e. new structures). Hence, to broaden the range of potential configurations, a BCC and two FCC parent lattices are used. For the BCC-based predictions, Si atoms were added to the 2a Wyckoff position of the Mg (Im-3m), and for the FCC-based predictions, the Si atoms were added to the 4a Wyckoff position of the Si (Fd-3m).

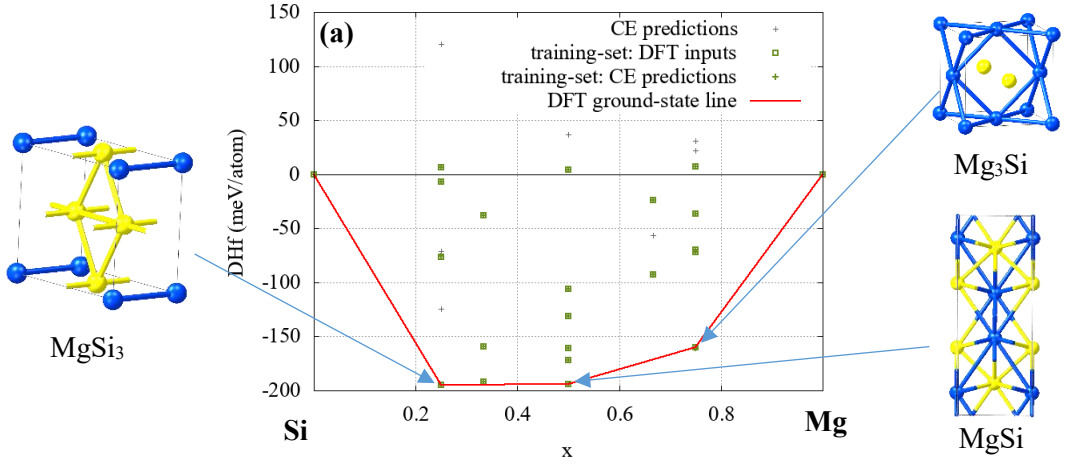


Figure 1. Binary ground-state diagram of Mg-Si emanating from (a) a grid search using the BCC-based system (Im-3m) considering structures of up to 8 basis atoms.

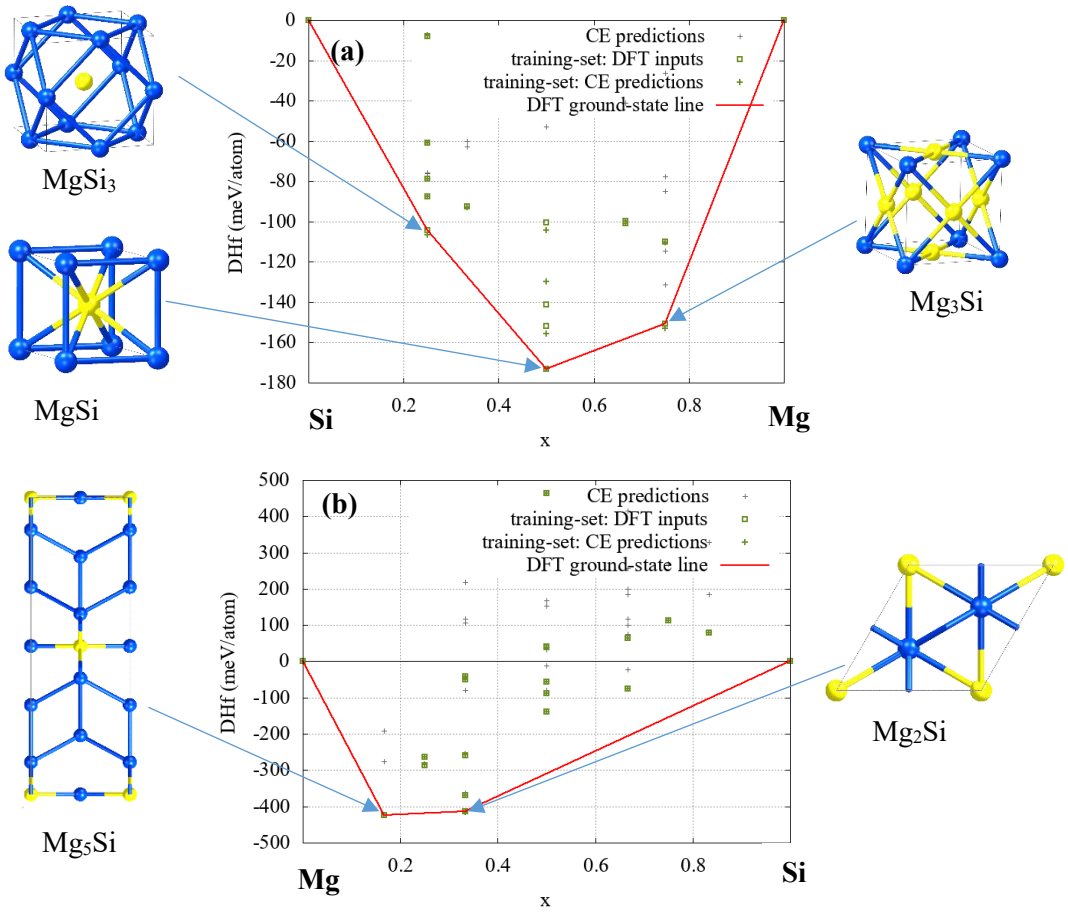


Figure 2. A binary ground-state diagram of Mg-Si derived from (a) a grid search using the FCC-based system (Fm-3m), which evaluated structures of up to 16 basis atoms, and (b) a grid search using the FCC-based system (Fd-3m), that considered structures of up to 24 basis atoms.

3.1.2. Quantitative and qualitative details of cluster expansions. Details of the cluster expansion predictions are shown in Tables 1 through 3.

Table 1. A summary of the BCC cluster expansion iterative optimization progress.

Iteration	No. of struc.	No. of new struc.	CVS [meV/pos.]	% struc. with SD below 5 meV	New structures
0	0	2	-	-	ce1 ce2
0	0	11	-	-	ce17 ce19 ce20 ce15 ce8 ce4 ce26 ce5 ce22 ce1 ce2
1	11	2	0.7	-	ce24 ce9
2	13	2	0.34	-	ce27 ce13
3	15	3	0.0049	-	ce3 ce6 ce12
4	18	1	0.37	-	ce28
5	19	2	0.22	-	ce10 ce18
6	21	1	0.41	-	ce29
7	22	0	0.27	-	

Table 1 summarises the cluster expansion iterative optimization progress. It displays the number of new structures predicted as well as the cross-validation score (CVS) for each iteration. The cluster expansion's convergence threshold was set to 5meV/atomic position. The calculation produced 22 novel structures, satisfying the convergence requirement with a CVS of 0.27 meV/atomic position.

Table 2. Summary of the iterative optimization progress of the Fm-3m based FCC cluster expansion.

Iteration	No. of struc.	No. of new struc.	CVS [meV/pos.]	% struc. with SD below 5 meV	New structures
0	0	2	-	-	ce1 ce2
0	0	12	-	-	ce23 ce3 ce21 ce19 ce20 ce6 ce4 ce15 ce27 ce1 ce2
1	12	4	0.31	-	ce29 ce9 ce7 ce28
2	16	0	3.4	-	

As shown in Table 2, the Fm-3m-based FCC cluster expansion calculation produced 16 new structures that met the convergence criteria of 5meV/atomic position with a CVS of 3.4 meV/atomic position.

Table 3. Summary of the iterative optimization progress of the Fd-3m based FCC cluster expansion.

Iteration	No. of struc.	No. of new struc.	CVS [meV/pos.]	% struc. with SD below 5 meV	New structures
0	0	2	-	-	ce1 ce3
0	0	12	-	-	ce5 ce20 ce22 ce9 ce42 ce33 ce28 ce6 ce1 ce3
1	12	4	2	-	ce10 ce26 ce4 ce11
2	16	4	0.61	-	ce7 ce31 ce15 ce14
3	20	0	0.64	-	

Table 3 shows that 20 novel structures that met the convergence requirement of 5meV/atomic position with an exceptional CVS of 0.64 meV/atomic position were produced by the Fd-3m-based FCC cluster expansion.

3.2. *Geometric properties of predicted thermodynamically stable Mg-Si phases*

Over all, 22 new structures were predicted from the BCC parent lattice, and a total of 36 new structures for the FCC parent lattice (*i.e.* 16 for Fm-3m and 20 for the Fd-3m). However, of these structure, 8 structures where predicted to be most thermodynamically stable and are listed in Table 4 along with their predicted space groups, crystal types, lattice constants, volume and heats of formation. As indicated in Table 4, the BCC (Im-3m) parent lattice produced three phases, namely $MgSi_3$, MgSi, and Mg_3Si. Their related space groups are P2/m, Imma, and P4/mmm, respectively. The Fm-3m-based FCC cluster expansion, like the BCC (Im-3m) parent lattice, created three phases, namely $MgSi_3$, MgSi, and Mg_3Si, but their corresponding space groups are distinct and are given as P2/m, Imma, and P4/mmm, respectively. The Fd-3m produced two thermodynamically stable Mg-Si phases, as opposed to the Im-3m and Fm-3m based cluster expansions which yielded three. But intriguingly, this parent lattice produced an Mg-rich phase, Mg_5Si, with an Imm2 space group and the lowest formation energy (DHf), -0.423479583 eV per active position. The Mg-rich Si phase is more relevant to our research since it would be better suited for Mg ion-based anodes with higher energy densities.

3.3. Electronic density of states and structural stability correlations of the Mg-rich Mg-Si phase

To determine structural stability, we look at the position of the Fermi level, E_f, on the total density of states (DOS). The Fermi level is represented by the line going through 0 eV. Figure 3 shows that the Fermi level falls into a pseudo gap, indicating that Mg_5Si is structurally stable.

Table 4. The geometric details of the thermodynamically stable structures

Parent lattice	System	Space group	Crystal type	Lattice parameters			Volume ($Å^3$)	DHf eV/active position
				a (Å)	b (Å)	c (Å)		
Im3m	$MgSi_3$	P2/m	Primitive monoclinic	2.75217	4.45261	5.16692	63.34	-0.19475550
	MgSi	Imma	Body-centred orthorhombic	3.60095	4.25879	9.55630	146.55	-0.19430725
	Mg_3Si	P4/mmm	Primitive tetragonal	4.23639	4.23639	4.28780	76.95	-0.16004850
Fm-3m	$MgSi_3$	Pm-3m	Primitive cubic	3.99175	3.99175	3.99175	63.60	-0.10419525
	MgSi	P4/mmm	Primitive tetragonal	2.99478	2.99478	3.91037	35.07	-0.17309850
	Mg_3Si	Pm-3m	Primitive cubic	4.25466	4.25466	4.25466	77.01	-0.15069150
Fd-3m	Mg_5Si	Imm2	Body-centred orthorhombic	14.882	4.93418	4.13538	303.678	-0.42347958
	Mg_2Si	P-3m1	Primitive hexagonal	4.5444	4.5444	6.79354	121.503	-0.41430200

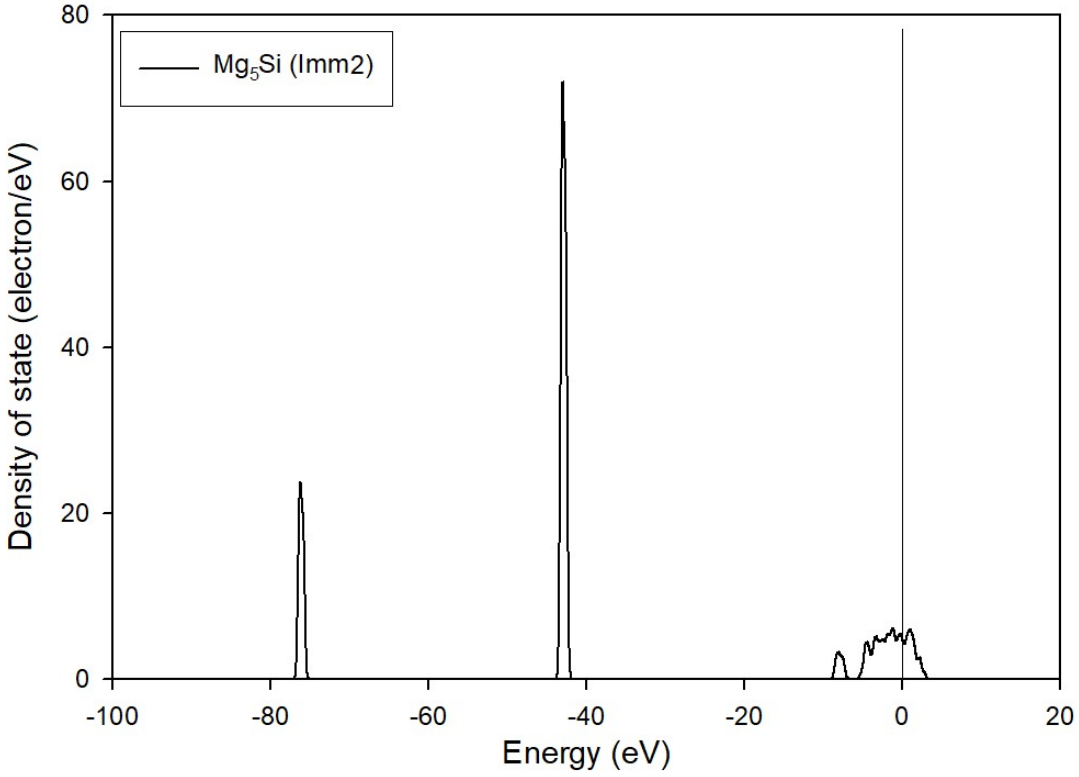


Figure 3. The total electronic density of states for (a) Mg_5Si in the absence of pressure and temperature.

4. Conclusion

Using the cluster expansion approach, the material space for the Mg-Si system was screened. A total of 58 new structures were predicted, of which 8 were found to be thermodynamically stable, namely Mg_5Si (Imm2), Mg_2Si (P-3m1), MgSi (P4-mmm), Mg_3Si (Pm-3m), $MgSi_3$ (P2/m), MgSi (Imma), and Mg_3Si (P4/mmm). The related geometrical properties of the 8 thermodynamically Mg-Si phases have been shown. Mg_5Si, an orthorhombic magnesium-rich phase, was shown to be structurally stable and a good contender for Mg-Si-based rechargeable batteries. The findings from this study are crucial to the advancement of high-energy battery electrode materials for next-generation batteries.

References

[1] Zhang D, Fu J, Wang Z, Wang L, Corsi JS and Detsi E, (2020), Perspective-reversible magnesium storage in silicon: an ongoing challenge, *J. Electrochem. Soc*, **167**, 050514:1-6

[2] Deivanayagam R, Ingram BJ and Shahbazian-Yassar R, (2019), Progress in development of electrolytes for magnesium batteries, *Energy Stor. Mater.*, **21**, 136-153.

[3] Attias R, Salama M, Hirsch B, Goffer Y and Aurbach D, (2019), Anode-electrolyte interfaces in secondary magnesium batteries, *Joule*, **3**, 27-52.

[4] Phoshoko KW, (2020), Density functional tight-binding and cluster expansion studies of lithiated/sodiated silicon anodes for high-energy-density batteries, PhD (Thesis), South Africa, University of Limpopo, Available from: http://hdl.handle.net/10386/3345

[5] Fang J, Xie M, He X, Zhang J, Hu J, Chen Y, Yang Y and Jin Q, (2022), Machine learning accelerates the materials discovery, *Matr. Today. Comm.*, **33**, 104900: 1-36.

[6] Fu Z and Zhang D, (2023), Universal machine learning potential accelerates atomistic modelling of materials, *J. Energy Chem.*, **83**, 1-2

[7] Farhadi B, You J, Zheng D, Liu L, Wu S, Li J, Li Z, Wang K and Liu S, (2023), Machine learning for fast development of advanced energy materials, *Next Materials*, **1**, 100025:1-19

[8] Wu Q, He B, Song T, Gao J and Shi S, (2016), Cluster expansion method and its application in computational materials science, *Comput. Mater. Sci.*, **125**, 243-254.

[9] Lerch D, Wieckhorst O, Hart GL, Forcade RW and Müller S, (2009), UNCLE: a code for constructing cluster expansions for arbitrary lattices with minimal user-input, *Model. Simul. Mat. Sci. Eng.*,**17**, 055003: 1-19

[10] Muller S, (2003), Bulk and surface ordering phenomena in binary metal alloys, *J. Phys.: Condens. Matter*, **15**, R1429-R1500

Optimization of the hole transport layer in polymer solar cell

M M Ntobeng, T E Seimela and M Diale
Department of Physics, University of Pretoria, Pretoria, Private bag, X20 Hatfield 0028. Republic of South Africa.

E-mail: **modisamaatseke@gmail.com.**

Abstract. Recently, organic solar cells (OSC) have been recognized as a potential addition to devices that harness solar energy due to their affordability, lightweight, and adaptability advantages. Despite these advantages, there is still room for improvement in OSCs power-conversion efficiency (PCE), which is crucial for their commercialization. In this study, silver nanoparticles (Ag NPs) were incorporated in different amounts to the Poly (3,4-ethylene dioxythiophene): poly (styrene sulfonate) (PEDOT: PSS) layer of the ITO/ PEDOT: PSS: Ag NPs/P3HT: PCBM/Ag-based device. X-ray diffraction (XRD) patterns of PEDOT: PSS at various volumes of Ag NPs exhibit distinct peaks located at 44° which correspond to the crystallographic plane (200) of Ag's face-centered cubic (FCC) phase. The surface characteristics of the NPs were investigated using transmission electron microscopy (TEM). TEM results revealed that Ag NPs have an average diameter of 22(2) nm. The results from the UV-Vis analysis demonstrated that the integration of Ag NPs influenced the PEDOT: PSS absorption; the newly generated peaks at 435 nm can be interpreted as their plasmonic resonance.

1. Introduction

Organic solar cells (OSCs) have been the center of attention in recent years due to their immense potential as an approach for collecting solar energy in a highly processable and affordable manner [1]. The primary factor behind the increased interest in organic semiconducting materials is the possibility that they will lead to the development of a renewable energy device that is flexible, inexpensive, and simple to manufacture [2]. In this era, the energy industry is seeking a competitive device to collect solar energy that is stable, low-cost, and easy to fabricate. Low efficiency and lifetime instability prevent OSCs from becoming available in scaled-up quantities for commercialization [3]. As a result, they require further study to make them a less expensive alternative to the widely used inorganic solar cells. Conducting polymers (CPs) are widely recognized as innovative materials that merge the specific features of metals, which include electrical and optical properties and charge transfer [4]. PEDOT: PSS is a conductive polymer that offers excellent transparency, low density, and high thermal stability [5]. This is a deep blue micro dispersion hole transport layer (HTL) in solar cell devices [6].

Plasmonic nanostructures with tuneable optical resonance properties improve the absorption of light by amplifying photocurrents in polymer solar cells (PSCs) through light-trapping methods [4]. Surface plasmon resonance (SPR) is a behavior at the frequency of the dipole surface plasmon caused by the surface's confinement of the nanoparticle's (NPs) internal conduction electrons and the creation of a powerful restorative force [7]. Plasmons can result from the interaction of conduction electrons and electromagnetic waves with metallic NPs (localized plasmon) [8]. The metallic NPs are encapsulated within a matrix of dielectric material like polymers since their dipole surface plasmon frequency enables them to interact with light very intensely, as a result of the NPs' collective electron velocity being

excited. In PSCs, various metal NPs including platinum (Pt), gold (Au), and silver (Ag) are being researched since they can improve light absorption [9].

Most of the research on the improvement of PSCs using SPR in devices has centered on metal NPs of various sizes doped in the active layer. Therefore, in this study, we demonstrate a doping method for adding Ag NPs to the HTL of PSCs to improve the absorption of light. We also investigate the impact of Ag NPs on the structural and optical properties of our PSCs.

2. Experimental

2..1. Materials

Silver nitrate ($AgNO_3$), trisodium citrate, ethanol amine, and sodium borohydride ($NaBH_4$) were obtained from Sigma Aldrich. Indium tin oxide (ITO) coated glass substrates and PEDOT: PSS were obtained from Ossila.

2.2.1. Synthesis of Ag NPS.

Ag NPs were synthesized using the chemical reduction method. An aqueous solution of $NaBH_4$ (10 mL, 20 mM) was added to a 100 mL solution of $AgNO_3$ (0.25 mM) and trisodium citrate (0.25 mM). The reaction was stirred for 30 minutes until a yellow colloidal silver solution appeared.

2.2.2. PEDOT: PSS samples fabrication

PEDOT: PSS was first filtered through a hydrophilic filter, it was then doped with 1 mL of ethanol ammine before being spin-coated on glass at 3 000 rpm for 30 s. Different amounts of Ag NPs were prepared by adding 10, 20, and 30 µL of Ag NPs solution into a 100 ml blend of PEDOT: PSS and ethanol amine. The first solution was without Ag NPs while the other solutions contained varying amounts of Ag NPs. The thin films underwent a 15 minute annealing process at a temperature of 120°C, then followed a cooling process to get them to room temperature.

2.2.3. Characterisation

Field emission gun transmission electron microscope (FEG-TEM, Jeol 2100) and field emission scanning electron microscope (FE-SEM Zeiss SEM Microscope Crossbeam-540) were used to characterize the morphology of the samples. The optical absorbance was measured using an Agilent Cary 60 UV–Vis spectrophotometer; the scan range for UV–Vis's absorption was between 200 – 800 nm. The crystallographic structure of the films was determined by XRD using a PAN alytical X'Pert Pro powder diffractometer with an X'Celerator detector. The system used Fe-filtered Co-Kα radiation (λ = 1.789Å), adjustable divergence, and fixed receiving slits. WITec Raman microscope model alpha300 RAS was used for examining the films' structural organization and Raman spectra. The LWX100 microscope objective was used to focus on an excitation wavelength of 532 nm, using a power of 0.012 W.

3. Results and discussion

3.1. Morphological properties.

Figure 1 shows transmission electron microscopy (TEM) image of water-soluble citrate stabilized Ag NPs and their size distribution. The TEM image (Figure 1a) shows that the Ag NPs have elongated spherical morphology and are widely dispersed. The particle size distribution shown in Figure 1b indicates that the produced NPs have an average diameter of 22(2) nm, with a size range of 10 – 70 nm.

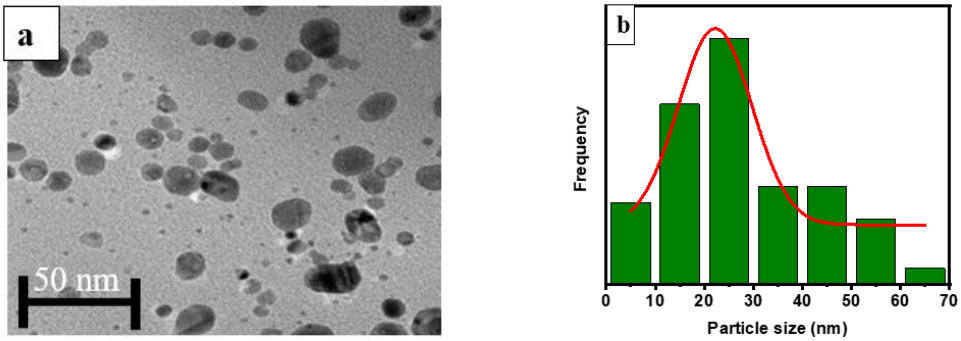


Figure 1. (a) TEM image of Ag NPs and (b) graph of Particle size distribution curve of Ag NPs.

SEM analysis was conducted to examine the surface morphology of PEDOT: PSS at different volumes of Ag NPs. Figure 2 shows SEM images of the PEDOT: PSS layer onto the ITO substrate. The images show full coverage of the polymer on top of the ITO; ethanol amine had an impact on this complete coverage since it enabled PEDOT: PSS to attach to the ITO substrate. Various volumes of incorporated Ag NPs in PEDOT: PSS are illustrated as white dots in Figure 2b through d. The visual images in our study are comparable to those of Liu *et al.* [10].

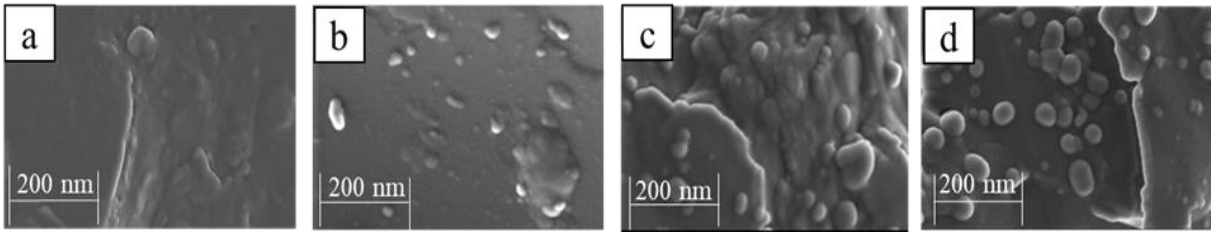


Figure 2. SEM images of (a) pristine PEDOT: PSS and the integration of Ag NPs into PEDOT: PSS using (b) 10 µL, (c) 20 µL (d) 30 µL volumes.

Figure 2a is a pristine PEDOT: PSS, the image shows a coarse surface. As shown in Figure 2b and c, the Ag NPs were dispersed uniformly across the PEDOT: PSS layer with no clear aggregation, while Ag NPs with the highest volume aggregate into clumps of NPs (Figure 2d). Moreover, the NPs exhibit a spherical shape.

3.2. Structural properties

Figure 3 shows XRD patterns of pristine PEDOT: PSS and PEDOT: PSS incorporated with various Ag NPs volumes. In general, PEDOT: PSS exhibits an amorphous nature. According to several research reports [11, 12], a peak associated with PEDOT: PSS that was supposed to exist at about 25° was absent. Our observations however show a broad diffraction peak at $2\theta \approx 28°$. This implies that there is a very low crystallinity among the polymer chains because of the relatively wider peak [13].

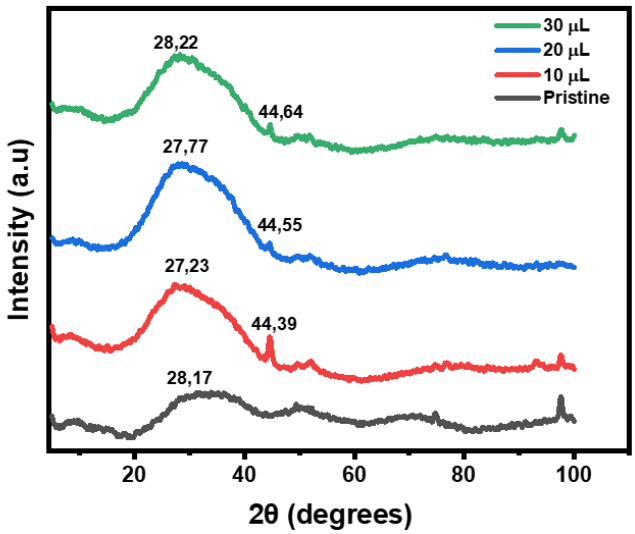


Figure 3. XRD patterns of PEDOT: PSS for pristine and at different volumes of Ag NPs.

After incorporation of the NPs, a sharp diffraction peak emerged at $2\theta \approx 44°$ which corresponds to the (200) crystallographic plane of the face-centered cubic (FCC) phase of Ag [14]. Peak intensity decreases as Ag NPs volume increases due to reduced crystallinity in the film [15].

Figure 4 shows Raman spectra of PEDOT: PSS at various volumes of Ag NPs. Pristine PEDOT: PSS thin film presents three vital bands at 446, 930, and 1567 cm^{-1} related to (C-S-C), (O-S), and (C_∞ -C_β) asymmetric stretching, respectively. After the incorporation of Ag NPs, for a volume of 10 µL the spectrum locations change to 436, 925, and 1579 cm^{-1}; however, for 20 µL and 30 µL of Ag NPs, the band positions move to 446, 930, and 1575 cm^{-1} and 441, 925, and 1571 cm^{-1}, respectively, indicating a blue shift. The findings from Sarkar *et al*. [16] and ours are in good agreement.

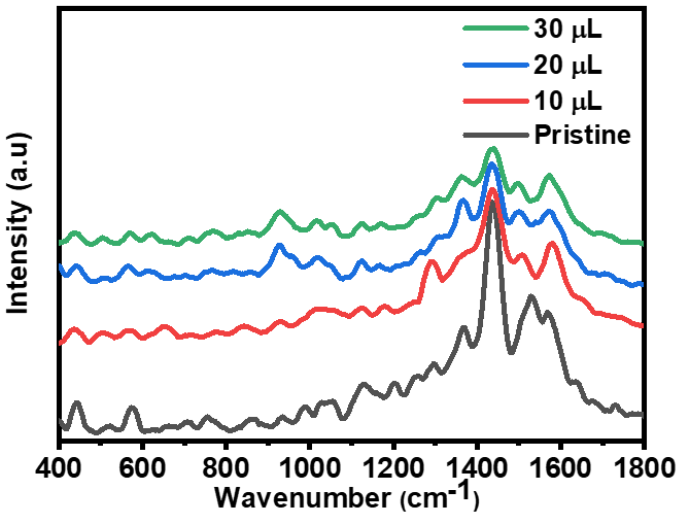


Figure 4. Raman spectra of PEDOT: PSS at different volumes of Ag NPs.

According to studies, the band between 1400 and 1500 cm^{-1} is crucial for determining the structure of the PEDOT: PSS chain. Raman spectra's strongest peak at 1435 cm^{-1} is generated by the ($C_\infty = C_\beta$) symmetric stretching vibration of PEDOT: PSS, indicating the emergence of a shape resembling a

quinoid. As a result, it is identified through the analysis of Raman spectra that PEDOT: PSS structure transitions from a benzoid to a quinoid form [17].

3.3. Optical properties

The UV/Vis spectrum of the Ag NPs in Figure 5 shows an intense peak at 389 nm indicating the formation of Ag NPs. $NaBH_4$ successfully converted the Ag^+ ions into Ag NPs since the SPR of the metal's conducting electrons causes colloidal silver to develop a characteristic absorption spectrum between 350 and 450 nm. The 389 nm intense peak does not expand, indicating that there is little, or no agglomeration as seen in Figure1a, a narrow size distribution is shown by a decrease in the width of the SPR band [10].

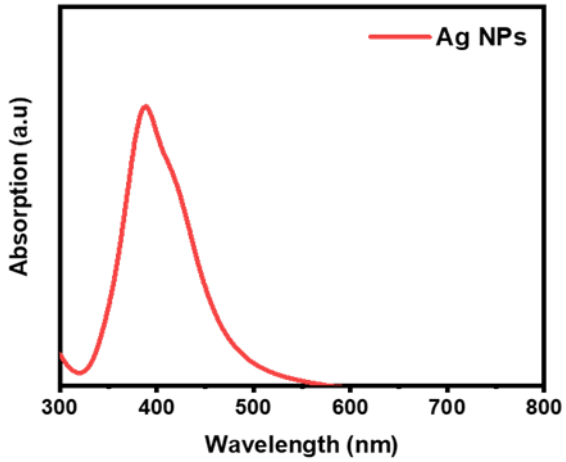


Figure 5. UV/ Vis spectrum of Ag NPs.

Figure 6 shows UV-Vis results of PEDOT: PSS at various volumes of Ag NPs. The transparent polymer pristine PEDOT: PSS is shown by the absence of an absorption peak in Figure 6a. There is one primary plasmon absorption band at 435 nm in PEDOT: PSS with various volumes of Ag NPs illustrated in Figure 6b, which demonstrates that their integration had an impact on the PEDOT: PSS absorption.

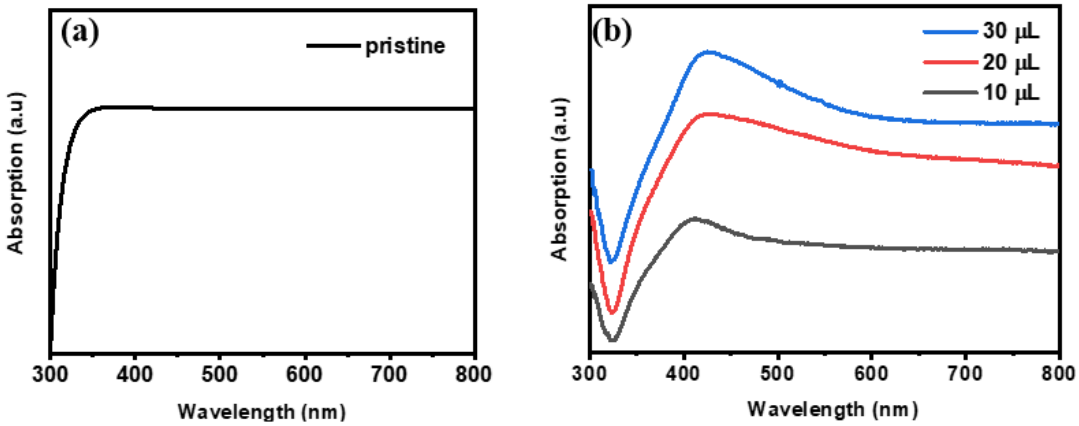


Figure 6. UV/ Vis spectra of PEDOT: PSS (a) pristine and (b) at different volumes of Ag NPs.

The peak intensity of 30 µL is higher than the 20 and 10 µL, implying that increasing the volume of Ag NPS results in a proportional increase in the intensity of the maximum plasmon peak. The volume of Ag NPs in the colloidal solution directly correlates with the intensity of an absorption peak [18].

4. Conclusion

Ag NPs were incorporated in the PEDOT: PSS layer to improve light absorption in PSC. The TEM analysis indicates that the Ag NPs vary in size from 10 to 70 nm. Based on the TEM images, the NPs are widely scattered and have a primarily spherical morphology. The UV-Vis absorption spectra show that the Ag NPs exhibit a surface plasmon absorption band at a maximum of 389 nm, providing evidence that there are spherical Ag NPs present. The absorption peak in Figure 6b demonstrates lower light absorption at 10 µL volume of Ag NPs. In contrast, the absorption peak at 30 µL volume of Ag NPs exhibits increased light absorption, implying that the incorporation of Ag NPs enhanced PEDOT: PSS absorption.

5. Acknowledgments

The authors would like to acknowledge, the University of Pretoria Physics Department, SARCHI Chair (IUD: 115463), and National Research Foundation (NRF) for funding.

6. References

[1] Otieno F, Shumbula NP, Airo M, Mbuso M, Moloto N, Erasmus RM, Quandt A and Wamwangi D 2017 Improved efficiency of organic solar cells using Au NPs incorporated into PEDOT: PSS buffer layer. *AIP Adv*. **7** 8

[2] Helgesen M, Sondergaard R and Krebs FC 2010 Advanced materials and processes for polymer solar cell devices. *J. Mater. Chem*. **20** 36

[3] Brabec CJ 2004 Organic photovoltaics: technology and market. *Sol. Energy Mater Sol. Cells*. **83** 273

[4] Kaur G, Adhikari R, Cass P, Bown M and Gunatillake P 2015 Electrically conductive polymers and composites for biomedical applications. *RSC Adv*. **5** 37553

[5] Zhu Z, Liu C, Jiang F, Xu J and Liu E 2017 Effective treatment methods on PEDOT: PSS to enhance its thermoelectric performance. *Synth. Met.* **225** 31

[6] Ahmad Z, Azman AW, Buys YF and Sarifuddin N 2021 Mechanisms for doped PEDOT: PSS electrical conductivity improvement. *Mater. Adv.* **2** 7118

[7] Duche D, Torchio P, Escoubas L, Monestier F, Simon JJ, Flory F and Mathian G 2009 Improving light absorption in organic solar cells by plasmonic contribution. *Sol. Energy Mater Sol. Cells*. **93** 1377

[8] Barnes WL, Dereux A and Ebbesen TW 2003 Surface plasmon subwavelength optics. *Nature*. **424** 824

[9] Yao M, Jia X, Liu Y, Guo W, Shen L and Ruan S 2015 Surface plasmon resonance enhanced polymer solar cells by thermally evaporating Au into buffer layers. *ACS Appl. Mater. Interfaces.* **7** 18866

[10] Liu XH, Hou LX, Wang JF, Liu B, Yu SZ, Ma LQ, Yang SP and Fu GS 2014 Plasmonic-enhanced polymer solar cells with high efficiency by addition of silver nanoparticles of different sizes in different layers. Sol Energy. 1 627

[11] White MS, Kaltenbrunner M, Głowacki ED, Gutnichenko K, Kettlgruber G, Graz I, Aazou S, Ulbricht C, Egbe DA, Miron MC and Major Z 2013 Ultrathin, highly flexible and stretchable PLEDs. *Nature Photonics*. **7** 811

[12] Peera SG, Tintula KK, Sahu AK, Shanmugam S, Sridhar P and Pitchumani S 2013 Catalytic activity of Pt anchored onto graphite nanofiber-poly (3, 4-ethylenedioxythiophene) composite toward oxygen reduction reaction in polymer electrolyte fuel cells. *Electrochim Acta*. **108** 95

[13] Tsujimoto T, Takayama T and Uyama H 2015 Biodegradable shape memory polymeric material from epoxidized soybean oil and polycaprolactone. *Polymers*. **7** 2165

[14] David L and Moldovan B 2020 Green \synthesis of biogenic silver nanoparticles for efficient catalytic removal of harmful organic dyes. *Nanomater.* **10** 202

[15] Lee H, Kim Y, Cho H, Lee JG and Kim JH 2019 Improvement of PEDOT: PSS linearity via controlled addition process. *RSC Adv.* **9** 17318

[16] Sarkar S, Bhowal AC, Kandimalla R and Kundu S 2021 Structural and electrical behaviors of PEDOT: PSS thin films in the presence of negatively charged gold and silver nanoparticles: A green synthesis approach. *Synth. Met.* **279** 116848

[17] Sze PW, Lee KW, Huang PC, Chou DW, Kao BS and Huang CJ 2017 The investigation of high quality PEDOT: PSS film by multilayer-processing and acid treatment. *Energies*. **10** 716.

[18] Medina-Ramirez I, Bashir S, Luo Z and Liu JL 2009 Green synthesis and characterization of polymer-stabilized silver nanoparticles. *Colloids Surf. B.* **73** 185

Microstructural characterization of low elastic modulus β-Ti alloy fabricated by arc melting process

M Madigoe[1*] and R Modiba[1,2]
[1]Advanced Materials Engineering, Manufacturing Cluster, Council for Scientific and Industrial Research (CSIR), Meiring Naudé, Pretoria, South Africa
[2]Department of Physics, Sefako Makgatho University, P.O. Box 94, Medunsa, Ga-Rankua, 0204, South Africa

E-mail: MSeerane@csir.co.za

Abstract. Development of low elastic modulus β-Ti alloys for implant application has grown significantly in recent years. In this work, alloy Ti-28.3Nb-13.8Ta-6.9Zr calculated using first-principles computations, was produced by button arc melting (BAM) process. Pure metal powders of Ti, Nb, Ta and Zr were pre-mixed and prepared as feedstock powder in a Copper-Hearth BAM furnace. The alloy was manufactured, cut and heat treated (solution treated + aged). The samples were analyzed for microstructure using the following techniques – OM, ImageJ, scanning electron microscopy, and x-ray diffraction analysis. The samples mechanical properties such as hardness and elastic Young's modulus were also evaluated. The alloy showed dominant β-phase microstructure with a dendritic morphology in both the as-cast and heat treated condition. The alloy showed satisfactory hardness and low elastic Young's modulus. The strength of the alloy improved after heat treatment indicative of precipitation hardening.

1. Introduction

In series of titanium alloys, β titanium (β-Ti) alloys are very important as bioimplants owing to their unique properties such as low elastic Young's modulus, high strength, excellent biocompatibility and shape memory effect [1]. The widely used titanium-based bioimplants such as commercially pure titanium (CP-Ti) and Ti6Al4V extra-low interstitial (ELI) alloy have elastic modulus ranging between 100 – 112 GPa. This made them first-choice metallic bioimplants compared to stainless steel and Cr-Co based bioimplants. Even though the alloys are currently used as bioimplants, their elastic Young's modulus is still higher than that of human cortical bone (10 – 40 GPa). Recently, there has been attempts to further reduce the modulus of Ti alloys closest to that of the bone by stabilizing the beta (β) phase at the expense of the alpha (α) phase. Alloying elements like Nb, Ta, Mo and Zr have been identified as toxic-free strong β stabilizers and potential candidates to lower the modulus [2]. Excessive addition of these elements result in formation of unwanted phases such as martensite and omega (ω), and straining of lattice structure, thereby compromising the mechanical properties [3].

According to literature data, the microstructural morphology of β-Ti alloys may be changed suitably by controlling the heat treatment parameters, that is, temperature and time [4]. Effect of heat treatment on Ti-Nb-Ta-Zr (TNTZ) alloys has been explored. Li et al. [5] found that under isothermal aging condition, the ω and α phases precipitate from the β phase for Ti-29Nb-13Ta-4.6Zr alloy. Moreover, the α phase could not precipitate during short time aging for Ti-39Nb-13Ta-4.6Zr alloy due to increased amount of Nb. The aim of this work was to assess the microstructure of Ti-28.3Nb-13.8Ta-6.9Zr alloy at specific heat treatment conditions and to measure the elastic Young's modulus.

2. Methodology

First-principles calculations (BIOVIA Material Studio) were performed using density functional theory (DFT) implemented in the CASTEP code to predict elastic properties of the alloy Ti-28.3Nb-13.8Ta-6.9Zr wt.% (Ti-Nb_{20}-Ta_5-Zr_5, at.%). Cut-off energy convergence = 669.4 eV and Monkhorst-Pack k-point grid = 18×18×18 (bcc β-Ti unit cell) were used. The elastic Young's modulus was estimated based on Voigt-Reuss-Hill (VRH) approximation.

The alloy was produced by arc melting using a water-cooled copper hearth from stoichiometrically weighted pure Ti, Nb, Ta and Zr metal powders. The as-cast sample was cut and solution treated at 950°C for 1 h and water quenched. The solution treated sample was aged for 12 h at 480°C followed by furnace cooling. The heat treatment was performed under argon atmosphere to avoid oxidation.

The samples were cut, ground, polished and finally etched in a 10 vol.% HF, 10 vol.% HNO3 + glycerol solution to reveal the microstructure. The as-cast and heat treated samples were analysed for phases and elemental distribution using optical microscopy (OM Leica DMI5000 M), scanning electron microscopy (SEM Jeol – 6510), ImageJ software package and, x-ray diffraction (XRD, EMPYREAN diffractometer system). X-ray diffraction measurements were carried out at 45 kV and 40 mA using monochromatic Cu Kα radiation (λ = 0.154 nm). The hardness and elastic Young's modulus was measured by a Vickers micro-hardness tester (Future Tech. Corp., FM-700) and a nano-indenter (Anton Paar, TTX-NHT3), respectively. Hardness test was done at 500 gf at a dwell time of 15 seconds. Nano-indentation was done at 400 mN load at a dwell time of 20 seconds.

3. Results

3.1. *First-principles calculations*

Figure 1 shows a bcc unit cell of β-Ti with space group Im-3m. The structure was proportionally built using Vesta 3D visualization program in a virtual crystal approximation (VCA). Table 1 summarizes the elastic constants ($c_{ij's}$) of the alloy along with the elastic Young's modulus (E) predicted using first-principles calculations. The lattice constant of bcc β-Ti crystal structure is 0.319 nm.

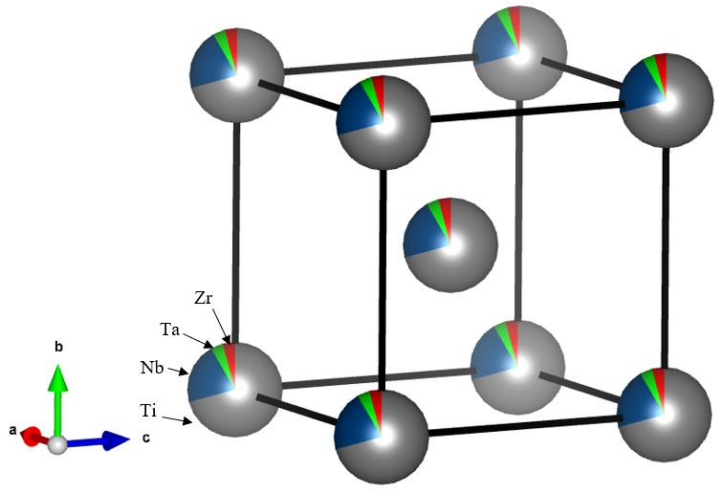


Figure 1. A VCA bcc β-Ti unit cell of alloy Ti-28.3Nb-13.8Ta-6.9Zr proportionally built by Vesta**.**

Table 1. The elastic properties of the alloy predicted from first-principles calculations.

Alloy	a (nm)	c_{11}	c_{44}	c_{12}	c'	c''	A	E (GPa)
Ti-28.3Nb-13.8Ta-6.9Zr	0.331	126.6	43.7	122.6	2.0	78.9	21.8	44.7

3.2. *Experimental results*

Figures 2 and 3 show the SEM microstructures of the alloy captured in backscatter imaging. Table 2 gives measured chemical compositions of the alloy before and after heat treatment.

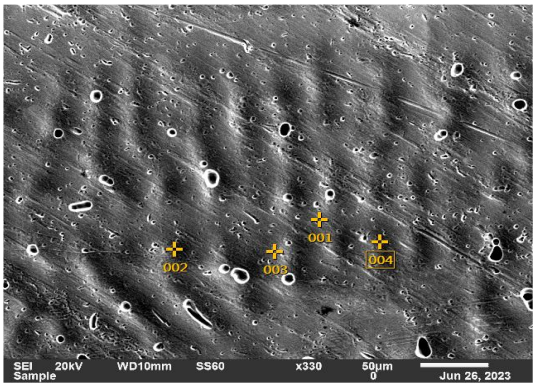


Figure 2. SEM analysis of alloy Ti-28.3Nb-13.8Ta-6.9Zr in the as-cast condition.

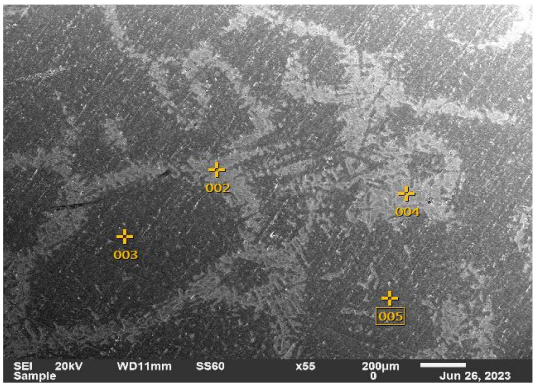


Figure 3. SEM analysis of alloy Ti-28.3Nb-13.8Ta-6.9Zr after heat treatment.

Table 2. Chemical composition of the alloy Ti-28.3Nb-13.8Ta-6.9Zr.

Element	As-cast (mass%)	Heat treated (mass%)
Nb	27.6	27.1
Ta	14.4	14.7
Zr	4.77	5.12
Ti	Bal.	Bal.
α content	26.2%	31.2%
β content	73.8%	68.8%

Figures 4 and 5 depict the microstructural morphologies of the alloy in the as-cast condition and heat treated state along with ImageJ micrographs. Table 2 shows α-β fractions representative of ImageJ micrographs.

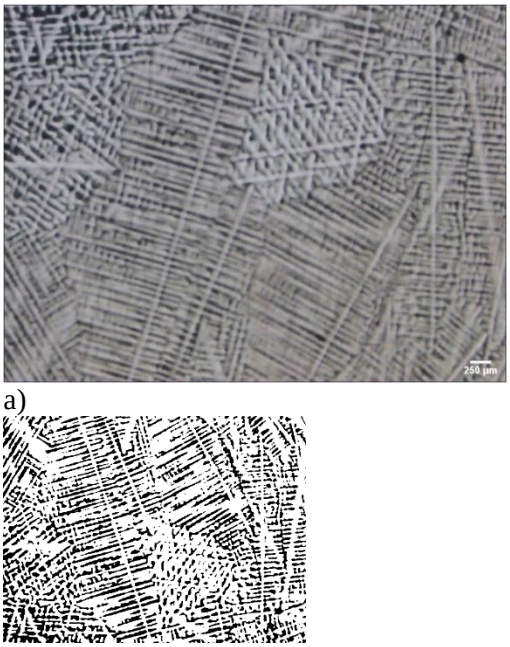


Figure 4. a) Optical and b) ImageJ micrographs of as-cast Ti-28.3Nb-13.8Ta-6.9Zr alloy.

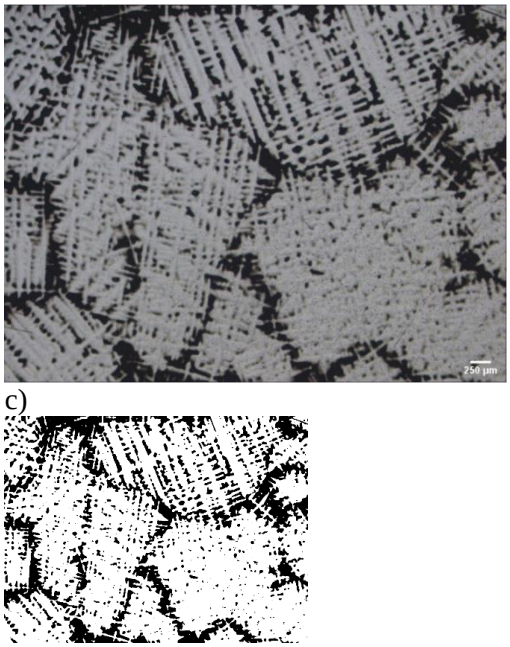


Figure 5. c) Optical and d) ImageJ micrographs of heat treated Ti-28.3Nb-13.8Ta-6.9Zr alloy.

To confirm the phases present in the microstructures, XRD patterns were measured and indexed using Rietveld refinement analysis. Figure 6 shows indexed XRD patterns of as-cast and heat treated samples. The plastic recovery curves in figure 7 gave an indication of the samples elastic Young's modulus and ductility. The samples elastic moduli were 69.8 GPa and 70.8 GPa in the as-cast condition and heat treated state, respectively. The indent depth decreased after heat treatment indicative of precipitation hardening. Hardness values of the samples were measured as 238 HV as-cast and 272 HV heat treated.

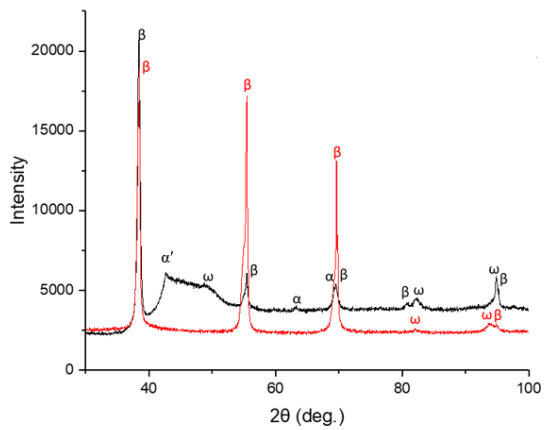


Figure 6. XRD curves of the as-cast (black) and heat treated (red) samples.

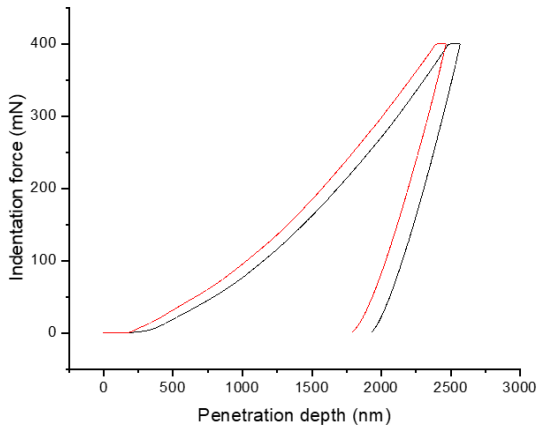


Figure 7. Nano-indentation plastic recovery curves of the as-cast (black) and heat treated (red) samples.

4. Discussion of results

4.1. *First-principles calculations approach*

A virtual 3D lattice crystal structure was built in Vesta code representative of bcc (β) Ti-Nb_{20}-Ta_5-Zr_5 alloy. The lattice parameter (*a)* increased slightly after elemental doping suggesting that the lattice structure of the alloy was strained to a certain degree. For mechanical stability of the structure, a well-known Born stability criteria [6] was followed, which for cubic crystals $c_{11} > 0$, $c_{44} > 0$, $c_{11} > |c_{12}|$ and $c_{11} + 2c_{12} > 0$. The alloy met the criteria. The shear modulus c' and Cauchy pressure c'' were assessed to ascertain mechanical stability and ionic bonding of the alloy, respectively, after doping. The VRH approximation gave a low elastic Young's modulus of 44.7 GPa closer to that of human cortical bone (40 GPa) suggesting that the elemental composition has potential to lower the elastic Young's modulus.

4.2. *Experimental approach*

A dendritic microstructure was observed in both as-cast and heat treated samples representative of TNTZ microstructures (figure 3). SEM-EDX analysis showed a slight deviation of the chemical composition after melting. The deviation was attributed to elemental diffusion and inhomogeneity. The as-cast alloy showed porosity indicative of kirkendal effect. Heat treatment resulted in closing pores indicative of a homogeneous microstructure.

The XRD analysis predominantly detected β phase. Small peaks of martensite (α″) and ω phase were detected but diminished after heat treatment. The amount of β phase decreased to 68.8% after heat treatment indicative of precipitation hardening (PH). The hardness of the alloy increased from 238 HV to 272 HV confirming samples' loss of ductility. The measured elastic Young's modulus at 69.8 GPa in the as-cast alloy showed no significant change after heat treatment (70.8 GPa).

4.3. *First-principles and experimental approaches*

A deviation between first-principles calculations and experimental measurements was observed. This can be attributed to the calculations approach. It was assumed in the VCA approach that atoms of the alloying elements occupied the same lattice positions as titanium atoms during doping to retain a stable bcc structure. This claim will tested using *ab-initio* calculation methodology.

5. Conclusion

First-principles calculations predicted the mechanical stability of the alloy Ti-28.3Nb-13.8Ta-6.9Zr in the VCA approach and the elastic Young's modulus was predicted as 44.7 GPa. The microstructure of the alloy was predominantly β phase which is characteristic in TNTZ alloys. ImageJ analysis and XRD patterns confirmed the dominance of β phases. Heat treatment resulted in increased hardness indicative of precipitation hardening and improved strength. Elastic Young's modulus increased slightly up to 70 GPa. A synergy between first-principles calculations and experimental results was not clear. A VCA approach appeared to have overestimated the calculations and elastic properties of the alloys. Ab-initio approach will be used in the future to test the claim.

References

[1] L. Wang, Z. Lin, X. Wang, Q. Shi, W. Yin, D. Zhang, Z. Liu and W. Lu, Effect of aging treatment on microstructure and mechanical properties of Ti27Nb2Ta3Zr β titanium alloy for implant application, *Materials Transactions*, vol. 55, No. 1 (2014), pp. 141-146.

[2] M. Geetha, A.K. Singh, R. Asokamani and A.K. Gogia, Ti based biomaterials, the ultimate choice for orthopaedic implants – A review, *Progress in Materials Science*, vol. 54, Issue 3 (2009), pp. 397-425.

[3] Y. Li, C. Yang, H. Zhao, S. Qu, X. Li and Y. Li, New developments of Ti-based alloys for biomedical applications – Review, *Materials*, vol. 7 (2014), pp. 1709-1800.

[4] A. Baran-Sadleja, J. Krawczyk, M. Motyka, L. Frocisz, Microstructure and hardness of aged Ti-10V-2Fe-Al near-beta titanium alloy, *Materials Engineering*, vol. 4, Issue 218 (2017), pp. 158-162.
[5] S. Li, Y. Hao, R. Yang, Y. Cui and M. Niinomi, Effect of Nb on microstructural characteristics of Ti-Nb-Ta-Zr for biomedical applications, *Materials Transactions*, vol. 43, No. 12 (2002), pp. 2964-2969.
[6] M. Born, On the stability of crystal lattices I, *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 36, No. 2, 1940, pp. 160-172.

DIVISION

B

NUCLEAR, PARTICLE, AND RADIATION PHYSICS

Health Risk Assessment of Toxic Heavy Metals in Irrigation Water, Rustenburg

Peter Oluwadamilare OLAGBAJU[1] and Olanrewaju Bola WOJUOLA[1]

[1]Department of Physics, North-West University, Mafikeng Campus, Mmabatho, South Africa

E-mail: **olagbajupetero@gmail.com**

Abstract. Several water sources and their state in Africa are a major concern due to pollution from mining, industrial, and other anthropogenic sources. Water is essential to life and healthy living, and also serves as home to various species of edible fishes. Unfortunately, its contamination has led to significant deaths resulting from diseases associated with the consumption of contaminated water. Globally, toxic heavy metals are among the numerous contaminants associated with water pollution. In this study, concentrations of toxic heavy metal in irrigation water used in Rustenburg was measured using inductively coupled plasma-mass spectrometry, and its associated carcinogenic and non-carcinogenic health risks were assessed. The measured concentration of toxic heavy metals in sampled water, ranges from ND (not detected) to 1.20×10^{-04}, 1.00×10^{-05} to 4.00×10^{-04}, ND to 1.24×10^{-02}, 1.20×10^{-04} to 1.84×10^{-02}, ND to 5.47×10^{-02}, 1.30×10^{-04} to 2.75×10^{-02}, and ND to 1.12×10^{-01} for Cd, Pb, Zn, Cu, Ni, Co, and Fe respectively, with corresponding mean values of 6.24×10^{-05}, 2.33×10^{-04}, 4.15×10^{-03}, 8.68×10^{-03}, 3.37×10^{-03}, 3.33×10^{-03}, and 5.99×10^{-02}. Non-carcinogenic risks resulting from dermal contact and ingestion of toxic heavy metals in water are low for adults and children of the exposed populations. Estimated carcinogenic risk due to dermal contact and ingestion of toxic heavy metals were also less than 10^{-6}, indicating negligible carcinogenic risk, thus making the sampled water safe for agricultural and domestic use of dwellers.

1. Introduction

Mining and industrial activities, though essential in the economic development of South Africa among other nations of the world, are globally associated with environmental pollution, posing significant risks to human health [1]. This is due to the production of non-biodegradable, toxic heavy metals in the form of waste and tailings [2]. They are continuously released into the environment and transferred to neighbouring communities, including agricultural sites and water bodies.

According to Sharma, Singh [3], heavy metals are metals and metalloids with relatively high densities ranging from 3.5 to 7 g/cm^3. They can pose significant threat to biotic systems even at low concentrations, and they include Cadmium (Cd), Chromium (Cr), Zinc (Zn), Mercury (Hg), Arsenic (As), Nickel (Ni), Copper (Cu), and lead (Pb). They constitute a major class of environmental contaminants in the developing world.

Environmental Protection Agency has reported 0.1 mg/l, 0.1 mg/l, 1.0 mg/l, 5 mg/l, 0.005 mg/l and 0.002 mg/l, as the maximum allowable limit for exposure to Pb, Ni, Cr, Cu, Zn, Cd and Hg in drinking water [3]. Exposure to mercury above the allowable limit can result in infertility, dementia and kidney damage; cadmium in renal dysfunction, liver and blood damage, lung and prostrate cancer,

lymphocytosis; arsenic in brain damage, skin cancer, cardiovascular and respiratory disease; chromium in chronic bronchitis, diarrhea, emphysema, liver damage, lung cancer, kidney failure; zinc in neurological disorder, kidney and liver failure, weakness, prostrate cancer; nickel in renal disorder, lung and nasal cancer; copper in nose, eyes and mouth liver and kidney damage; and lead in infertility, brain and kidney [4-6].

Water is one of the major routes through which humans accumulate toxic heavy metals, directly via contact and ingestion or through consumption of plant and animal irrigated and fed with contaminated water [7, 8]. Polluted water sources are harmful to human health, as a result of exposure to toxic heavy metals from mining, agriculture and industrial waste, and the consumption of toxic heavy metals in aquatic organisms, or the use of contaminated surface water for domestic and recreational purposes such as swimming and washing [8, 9].

According to the World Health Organization (WHO), approximately 844 million people lack access to a fundamental source of drinking water source and about 230 million people spend more than 30 minutes per day collecting water from an improved water source, which may include boreholes, protected wells, piped water and rainwater, springs and packaged/delivered water [2, 10]. Insufficient water and wastewater treatment, coupled with increase industrial activity have resulted in increased toxic heavy metal pollution in rivers, lakes and other water sources in African countries [2, 11]. The impact of increased water pollution is a challenge in developing countries due to the lack of resources and capacity to treat contaminated water effectively for domestic and recreational use [2].

According to World Health Organization, about 1.6 million people die every year from preventable water-related illnesses, such as diarrhea, and 90% of these fatalities are children under 5 years of age. Therefore, it is essential to continuously monitor and control state of water bodies to avoid the huge cost and process of contaminants removal (remediation). This study assesses the health risk associated with toxic heavy metals in irrigation waters in Rustenburg, South Africa.

2. Methodology

2.1. Study area

The study area, Rustenburg, is a mining and industrial city in the North -West Province of South Africa. It is recognized for its mining and industrial activities, which have left farmlands and communities littered with radioactive waste [12]. It also has one of the most diverse and extensive farming systems in Africa [12, 13]. South Africa has about 12 percent of its land suitable for crop farming, with 3 percent considered fertile, and about 1.3 million hectares of land under irrigation, making it a water-scarce country. This accounts for approximately 60 percent of the total water usage across various sectors. It plays a substantial role in the food chain, serving as a significant source of exposure through the consumption of food crops according to Baleta and Pegram [14]. Rustenburg is known for its platinum operation and serves as home to the Bushveld Igneous Complex (BIC), reportedly containing a significant amount of base metals such as copper (Cu), iron (Fe), cobalt (Co), chromium (Cr), nickel (Ni) and zinc (Zn) in its host ores [15].

2.2. Sampling and sample preparation

Water samples were collected from irrigation dams, wells and boreholes from farms around mines in Rustenburg, South Africa. Collected samples were well-labelled, and the coordinates of each sampling location were noted and documented using a Global Positioning System (GPS) for later referencing and to ensure traceability.

In seventeen irrigation water sources around mines in Rustenburg, about 1 liter of water sample was collected in polyethene bottles from each source. To prevent radiotoxic absorption on the container walls, the used polyethene bottles were thoroughly washed, and 10 ml of HCl was introduced to the sampled water after collection [17]. All the samples were well-labelled for ease of identification, and the Global Positioning System was used to record the coordinates of each sampling location, to ensure traceability for later referencing.

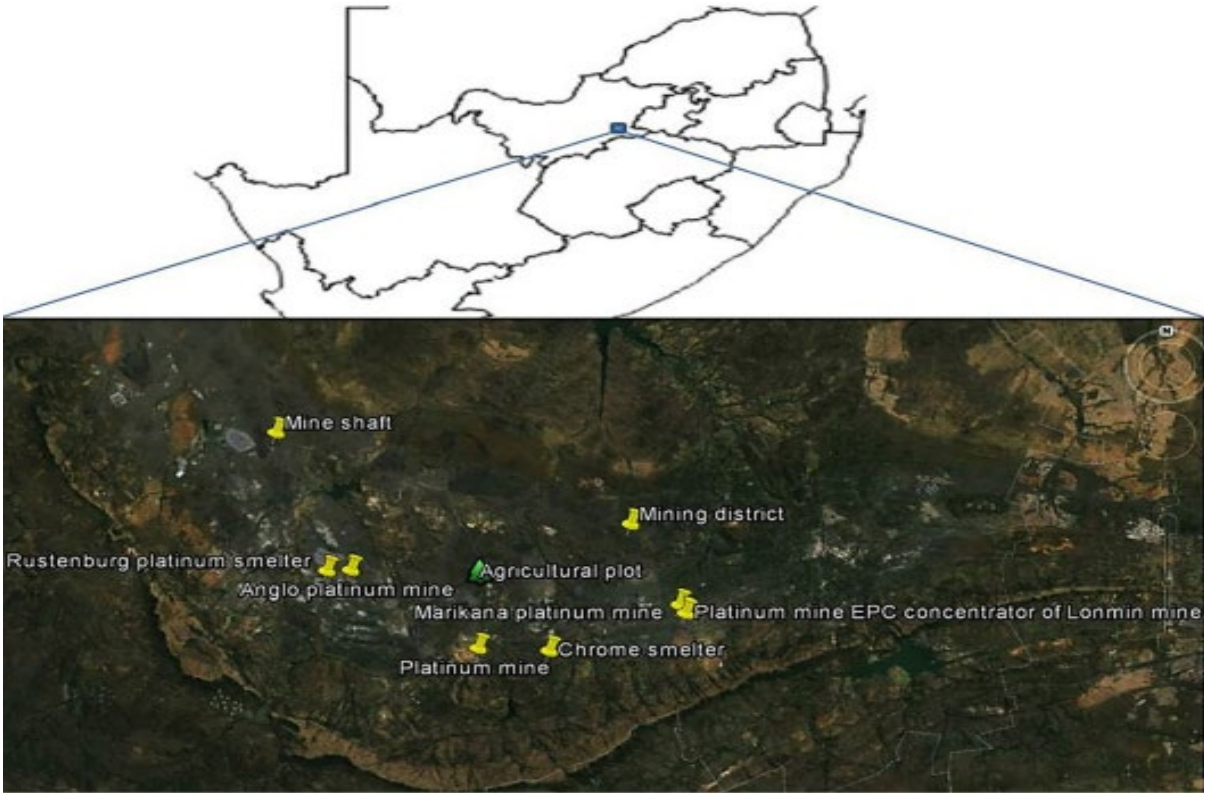


Figure 1. Map of South Africa showing the agricultural plots in Rustenburg [16].

2.3 Measurement method

The water samples were analysed in triplicate using Inductively Coupled Plasma - Mass Spectrometry (ICP-MS), comprising of a sample introduction system and data processing unit, ions source, the mass analyser and the detector. Due to its wide linear dynamic detection range, high sensitivity and simultaneous multi-element analysis, total quant method was employed to analyse the digested samples [8, 9]. Inductively coupled plasma mass spectrometry calibration was done by measuring the instrumental response to reference solution (a 10 mg/L multi-element calibration standard of Al, Ag, Ba, As, Cd, Bi, Be, Cs, Ca, Cr, Co, Cu, Li, Fe, Ga, In, K, Mg, Mn, Ni, Pb, Ne, Sr, Rb, Se, U, Tl, V and Zn).

To ensure there is no drift, cross-contamination of samples and for quality control, a certified standard solution and blank solution were run after every eight sample measurements [8, 18]. The elemental concentrations of heavy metal in mg/l was estimated using the expression [19]:

$$Concentration\left(\frac{\mu g}{g}\right) = C \times \frac{V_f}{w \times s} \times \frac{DF}{1000} \qquad (1)$$

where, C = instrument value in µg/l (the average of all replicate integrations), W = initial aliquot amount (g), S = % solids/100, and V_f = final digestion volume (ml).

3. Health risk assessment

Health risk assessment is essential in environmental monitoring and protection purposes. It is use to estimate the health effects resulting from exposure to carcinogenic and non-carcinogenic contaminants [20]. The health risk of toxic heavy metals in sampled irrigation water was estimated using the United State Environmental Protection Agency guidelines [21].

Non-carcinogenic risk is defined as the likelihood of an individual suffering adverse effect from exposure to toxic substance [21]. Average daily intake due to dermal contact and ingestion of toxic heavy metals as well as hazard quotient was used in this work to estimate the non-carcinogenic effects of toxic heavy metal in sampled water. Similarly, carcinogenic risk is defined as the probability of an individual developing cancer as a result of toxic heavy metals (carcinogens) exposure. It was estimated using cancer risk via dermal contact and ingestion pathway. Cancer slope factor (CSF) and reference Doss (RfD) used for carcinogenic and non-carcinogenic assessment of toxic heavy metals were derived from United State Environmental Protection Agency and South Africa, Department of Environmental Affairs [22].

4. Results and discussion

The concentrations of Fe, Mn, Co, Ni, Cu, Zn, Pb, and Cd, measured in samples of irrigation water, range from ND to 4.52×10^{-02} mg/l, ND to 3.51×10^{-02} mg/l, ND to 1.12×10^{-01} mg/l, 1.30×10^{-04} to 2.75×10^{-02} mg/l, ND to 5.47×10^{-02} mg/l, 1.20×10^{-04} to 1.84×10^{-02} mg/l, ND to 1.24×10^{-02} mg/l, 1.00×10^{-05} to 4.00×10^{-04} mg/l, and ND to 1.20×10^{-04} mg/l respectively. Cr, As and Hg were not detected in any of the water samples, which is attributed to spectral interference as a result of the total quant method adopted for measurement [23, 24].

The average concentrations of toxic heavy metals were found to be less than South Africa, WHO and USEPA permissible limits for toxic heavy metals in drinking water presented in Table 1. Toxic heavy metal concentration in sampled irrigation water increases in the order of Cd (5.47×10^{-06} mg/l) < Pb (1.85×10^{-05} mg/l) < Zn (5.39×10^{-04} mg/l) < Cu (8.50×10^{-04} mg/l) < Co (1.08×10^{-03} mg/l) < Al (1.50×10^{-03} mg/l) < Ni (2.27×10^{-03} mg/l) < Mn (2.92×10^{-03} mg/l) < Fe (4.54×10^{-03} mg/l).

Pearson correlation matrix was used to investigate the relationship between measured toxic heavy metals in sampled water. From the results presented in Table 3, high positive correlation coefficients indicate a linear relationship between selected toxic heavy metals, while a high negative correlation implies a non-linear relationship between them. Thus, Mn and Pb, Mn and cd, Fe and Cu, Fe and Pb, Fe and Cd, Co and Ni, Cu and Zn, Cu and Pb, Pb and Cd are positively correlated, suggesting similar contaminants sources for each pair.

Non-carcinogenic risk resulting from dermal contact and ingestion of toxic heavy metals in water for adults and children was estimated using reference doses (RfD) and presented in Figure 2. The results of the average daily intake (ADI), hazard quotient (HQ) and hazard index (HI) for ingestion and dermal pathways shows there is no apparent risk to exposed populations because HQ and HI values are found to be less than one [26, 27]. Estimated carcinogenic risk due to dermal contact and ingestion of Ni, Cd, Pb and Co, is presented in Figure 3. They are found be less than 10^{-6}, indicating negligible carcinogenic risk.

Table 1. Permissible limits of toxic heavy metal in water.

Country	Permissible limits for toxic heavy metals in water (mg/l)												
	Mg	Cr	Mn	Fe	Co	Ni	Cu	Zn	As	Cd	Hg	Pb	Ref
South Africa	NA	NA	NA	NA	NA	NA	1.0	NA	0.01	0.003	0.001	0.01	[25]
WHO	NA	0.05	NA	NA	NA	0.07	2.0	NA	0.01	0.003	0.006	0.01	[10]
USEPA	NA	0.1	NA	NA	0.1	NA	1.3	0.5	0.01	0.005	0.002	0.015	[22]

NA = Not Available

Table 2. Toxic heavy metal concentration in irrigation water.

	mg/l							
Sample Id	Cd	Pb	Zn	Cu	Ni	Co	Fe	Mn
W01	4.00×10^{-05}	2.40×10^{-04}	1.24×10^{-02}	7.78×10^{-03}	1.60×10^{-04}	1.30×10^{-04}	5.60×10^{-02}	2.68×10^{-02}
W02	5.00×10^{-05}	1.30×10^{-04}	6.77×10^{-03}	7.83×10^{-03}	3.00×10^{-05}	5.80×10^{-04}	5.15×10^{-02}	2.72×10^{-02}
W03	7.00×10^{-05}	2.30×10^{-04}	ND	1.20×10^{-04}	5.47×10^{-02}	2.75×10^{-02}	ND	ND
W04	5.00×10^{-05}	2.40×10^{-04}	5.19×10^{-03}	1.57×10^{-02}	ND	2.19×10^{-03}	5.64×10^{-02}	2.82×10^{-02}
W05	5.00×10^{-05}	3.10×10^{-04}	8.44×10^{-03}	1.84×10^{-02}	3.00×10^{-04}	3.10×10^{-03}	8.42×10^{-02}	4.75×10^{-03}
W06	5.00×10^{-05}	2.00×10^{-04}	1.91×10^{-03}	7.02×10^{-03}	3.00×10^{-05}	1.04×10^{-03}	7.11×10^{-02}	3.47×10^{-03}
W07	7.00×10^{-05}	3.30×10^{-04}	6.45×10^{-03}	1.39×10^{-02}	2.40×10^{-04}	2.86×10^{-03}	6.69×10^{-02}	2.85×10^{-02}
W08	4.00×10^{-05}	1.10×10^{-04}	1.50×10^{-03}	3.18×10^{-03}	8.00×10^{-05}	6.10×10^{-04}	6.80×10^{-02}	3.59×10^{-03}
W09	1.00×10^{-04}	2.30×10^{-04}	2.99×10^{-03}	6.97×10^{-03}	ND	1.16×10^{-03}	4.67×10^{-02}	2.55×10^{-02}
W10	6.00×10^{-05}	2.80×10^{-04}	3.41×10^{-03}	1.01×10^{-02}	1.60×10^{-04}	2.17×10^{-03}	7.62×10^{-02}	5.57×10^{-03}
W11	1.00×10^{-04}	2.90×10^{-04}	4.89×10^{-03}	1.08×10^{-02}	4.90×10^{-04}	2.96×10^{-03}	1.12×10^{-01}	1.50×10^{-02}
W12	1.20×10^{-04}	4.00×10^{-04}	4.37×10^{-03}	1.19×10^{-02}	5.10×10^{-04}	2.22×10^{-03}	7.85×10^{-02}	4.35×10^{-02}
W13	7.00×10^{-05}	3.30×10^{-04}	3.50×10^{-03}	9.91×10^{-03}	1.90×10^{-04}	1.85×10^{-03}	6.46×10^{-02}	4.33×10^{-02}
W14	1.00×10^{-04}	3.20×10^{-04}	4.28×10^{-03}	1.11×10^{-02}	2.40×10^{-04}	2.69×10^{-03}	7.37×10^{-02}	4.52×10^{-02}
W15	8.00×10^{-05}	2.80×10^{-04}	3.01×10^{-03}	8.70×10^{-03}	2.50×10^{-04}	2.86×10^{-03}	6.76×10^{-02}	4.39×10^{-02}
W16	ND	1.00×10^{-05}	ND	1.19×10^{-03}	ND	8.80×10^{-04}	2.05×10^{-02}	9.90×10^{-04}
W17	1.00×10^{-05}	3.00×10^{-05}	1.44×10^{-03}	3.06×10^{-03}	ND	1.78×10^{-03}	2.33×10^{-02}	2.38×10^{-03}
Maximum	1.20×10^{-04}	4.00×10^{-04}	1.24×10^{-02}	1.84×10^{-02}	5.47×10^{-02}	2.75×10^{-02}	1.12×10^{-01}	4.52×10^{-02}
Minimum	ND	1.00×10^{-05}	ND	1.20×10^{-04}	ND	1.30×10^{-04}	ND	ND
Average	6.24×10^{-05}	2.33×10^{-04}	4.15×10^{-03}	8.68×10^{-03}	3.37×10^{-03}	3.33×10^{-03}	5.99×10^{-02}	2.05×10^{-02}

ND = Not Detectable

Table 3. Correlation coefficient of toxic heavy metals in water sample.

	Mn (mg/l)	Fe (mg/l)	Co (mg/l)	Ni (mg/l)	Cu (mg/l)	Zn (mg/l)	Pb (mg/l)	Cd (mg/l)
Mn (mg/l)	1.000							
Fe (mg/l)	0.327	1.000						
Co (mg/l)	-0.268	-0.511	1.000					
Ni (mg/l)	-0.305	-0.576	0.990	1.000				
Cu (mg/l)	0.441	0.698	-0.348	-0.439	1.000			
Zn (mg/l)	0.344	0.424	-0.336	-0.337	0.619	1.000		
Pb (mg/l)	0.612	0.599	0.084	0.002	0.703	0.402	1.000	
Cd (mg/l)	0.620	0.502	0.132	0.070	0.395	0.105	0.816	1.000

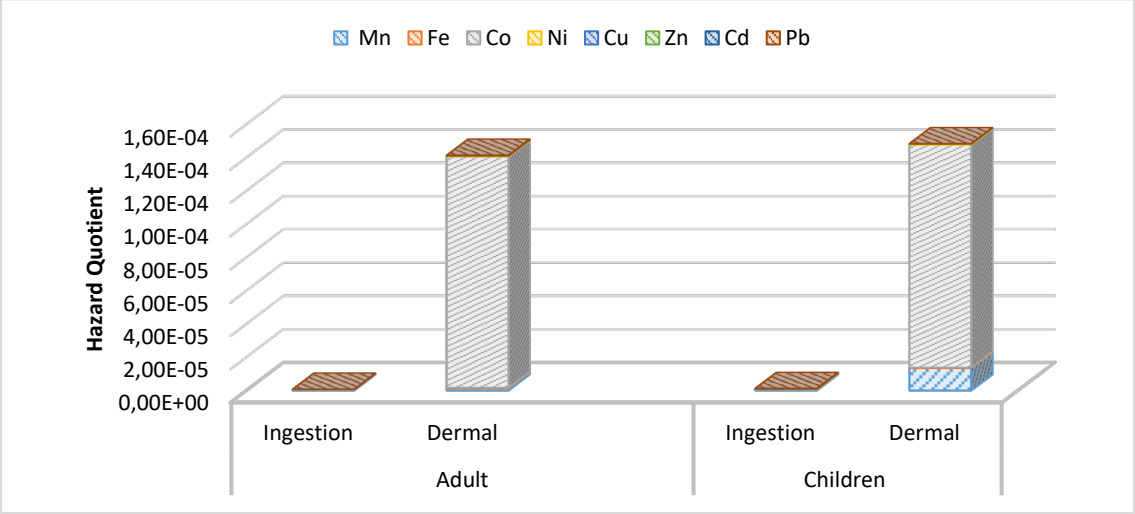


Figure 2. Estimated hazard quotient for sampled irrigation water.

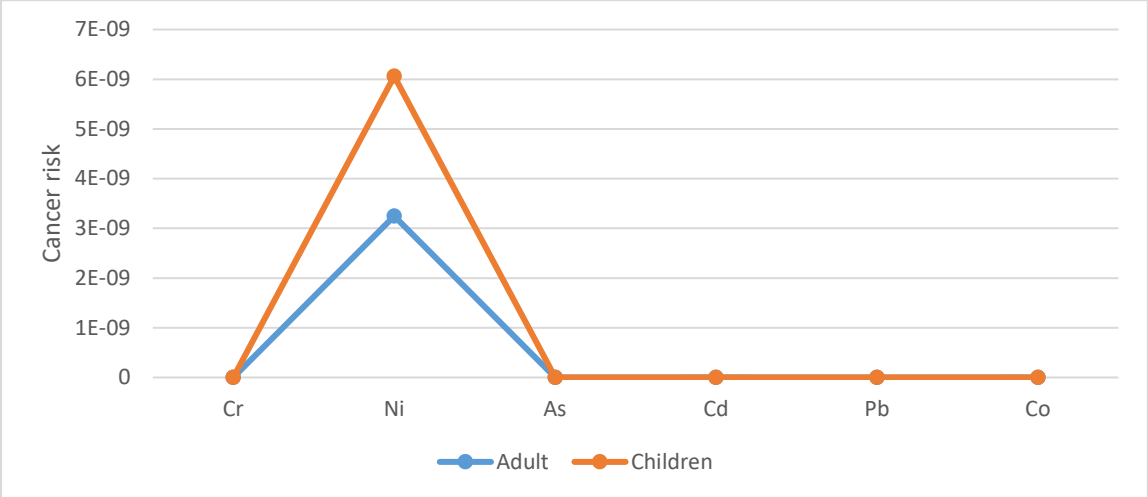


Figure 3. Estimated cancer risk of toxic heavy metals for adults and children in sampled irrigation water.

5. Conclusion

In this study, the concentration of toxic heavy metal in irrigation water has been measured, and the carcinogenic and non-carcinogenic health risk assessed. The average concentration of measured toxic heavy metals was found to be less than South Africa, WHO, and USEPA permissible limit for toxic heavy metals in drinking water. The average daily intake (ADI), hazard quotient (HQ) and hazard index (HI) for ingestion and dermal pathways show there is no apparent risk to exposed populations. Also, estimated carcinogenic risk due to dermal contact and ingestion of toxic heavy metals indicates negligible carcinogenic risk. Thus, investigated irrigation water show insignificant health risk, making irrigation water safe in the study area for domestic and agricultural purposes.

6. Acknowledgement

This work is based on the research supported in part by the National Research Foundation of South Africa (Grant No: 138066). The authors hereby express gratitude to the National Research Foundation and the North-West University, South Africa for supporting this work.

References

[1] Olagbaju, P.O., O.B. Wojuola, and V. Tshivhase. *Radionuclides Contamination in Soil: Effects, Sources and Spatial Distribution*. in *EPJ Web of Conferences*. 2021. EDP Sciences.

[2] Joseph, L., et al., *Removal of heavy metals from water sources in the developing world using low-cost materials: A review.* Chemosphere, 2019. **229**: p. 142-159.

[3] Sharma, S., B. Singh, and V. Manchanda, *Phytoremediation: Role of Terrestrial Plants and Aquatic Macrophytes in the Remediation of Radionuclides and Heavy Metal Contaminated Soil and Water.* Environmental Science and Pollution Research, 2015. **22**(2): p. 946-962.

[4] Ayangbenro, A.S. and O.O. Babalola, *A New Strategy for Heavy Metal Polluted Environments: A Review of Microbial Biosorbents.* International Journal of Environmental Research and Public Health, 2017. **14**(1): p. 94.

[5] Jensen, A.A. and F. Tuchsen, *Cobalt Exposure and Cancer Risk.* Critical Reviews in Toxicology, 1990. **20**(6): p. 427-439.

[6] Li, C., et al., *A Review on Heavy Metals Contamination in Soil: Effects, Sources, and Remediation Techniques.* Soil and Sediment Contamination: An International Journal, 2019. **28**(4): p. 380-394.

[7] Edokpayi, J.N., et al., *Evaluation of Water Quality and Human Risk Assessment due to Heavy Metals in Groundwater Around Muledane Area of Vhembe District, Limpopo Province, South Africa.* Chemistry Central Journal, 2018. **12**(1): p. 1-16.

[8] Olagbaju, O.P., O.B. Wojuola, and M.V. Tshivhase, *Radiological assessment of irrigation water used in Rustenburg, South Africa.* Radiation Protection Dosimetry, 2023. **199**(8-9): p. 781-784.

[9] Nde, S.C., et al., *Modelling the Dynamics of the Cancer Risk due to Potentially Toxic Elements in Agricultural Soils, in the Upper Crocodile River Catchment, North-West Province, South Africa.* Ecotoxicology and Environmental Safety, 2021. **211**: p. 111961.

[10] WHO, *Guidelines for Drinking-water Quality: Second Addendum. Vol. 1, Recommendations.* 2008.

[11] Edokpayi, J.N., et al., *Evaluation of Temporary Seasonal Variation of Heavy Metals and their Potential Ecological Risk in Nzhelele River, South Africa.* Open Chemistry, 2017. **15**(1): p. 272-282.

[12] Dudu, V.P., M. Mathuthu, and M. Manjoro, *Assessment of Heavy Metals and Radionuclides in Dust Fallout in the West Rand Mining Area of South Africa.* Clean Air Journal, 2018. **28**(2): p. 42-52.

[13] Olagbaju, P., O. Wojuola, and M. Tshivhase, *Activity Concentration Measurement of Naturally-Occurring Radionuclides in Various Vegetation plots in Rustenburg, South Africa.* 2021.

[14] Baleta, H. and G. Pegram, *Water as an Input in the Food Value Chain: Understanding the Food Energy Water Nexus.* WWF-SA, South Africa, Sustainable Development. Proceedings of the National Academy of Sciences, 2014. **100**(14): p. 8086-8091.
[15] Ochieng, L., *Mobility Studies of Heavy Metals in Mining Polluted Soil, Rustenburg Section of Anglo Platinum*. 2002, University of the Witwatersrand.
[16] Mosai, A.K., E.N. Bakatula, and H. Tutu, *Adsorption of Cadmium, Copper and Chromium by an Agricultural Soil Impacted by Mining Activities.* Water, Air, & Soil Pollution, 2017. **228**(8): p. 1-12.
[17] Guidebook, A., *Measurement of Radionuclides in Food and the Environment.* Vienna: International Atomic Energy Agency. Retrieved from https://www.iaea.org/publications/1398/measurement-of-radionuclides-in-food-and-the-environment, 1989.
[18] Olawoyin, R., S.A. Oyewole, and R.L. Grayson, *Potential Risk Effect from Elevated Levels of Soil Heavy Metals on Human Health in the Niger delta.* Ecotoxicology and Environmental Safety, 2012. **85**: p. 120-130.
[19] Kaur, M., et al., *Human Health Risk Assessment from Exposure of Heavy Metals in Soil Samples of Jammu District of Jammu and Kashmir, India.* Arabian Journal of Geosciences, 2018. **11**(15): p. 1-15.
[20] EPA, U. *A Review of the Reference Dose and Reference Concentration Processes*. in *Washington, DC: Risk Assessment Forum, U*. 2002.
[21] Mathuthu, M., C. Kamunda, and M. Madhuku, *Modelling of Radiological Health Risks from Gold Mine Tailings in Wonderfonteinspruit Catchment Area, South Africa.* International Journal of Environmental Research and Public Health, 2016. **13**(6): p. 570.
[22] USEPA, S., *Guidelines for Water Reuse.* Special Restricted Crop Area in Mendoza, Argentina, 2012.
[23] Amaral, C.D., et al., *A novel strategy to determine As, Cr, Hg and V in drinking water by ICP-MS/MS.* Analytical methods, 2015. **7**(3): p. 1215-1220.
[24] Thomas, R., *Chapter 14: Review of Interferences.* Practical Guide to ICP-MS: A Tutorial for Beginners, 2013. **3**: p. 133-143.
[25] Hodgson, K. and L. Manus, *A Drinking Water Quality Framework for South Africa.* Water SA, 2006. **32**(5).
[26] EPA, *Risk Assessment Guidance for Superfund. Volume I: Human Health Evaluation Manual (Part a)*. 1989, EPA/540/1-89/002.
[27] DEA, *The Framework for the Management of Contaminated Land, South Africa.* . Department of Environmental Affairs, 2010.

Decline and fall of nuclear β- and γ-vibrations

J F Sharpey-Schafer

Department of Physics, University of Zululand, Private Bag X1001, KwaDlangezwa, 3886, South Africa

E-mail: **johnsharpeyschafer@gmail.com**

Abstract. The conjecture of Bohr and Mottelson that the lowest excited $K^\pi = 0_2^+$ and $K^\pi = 2_\gamma^+$ rotational bands in deformed nuclei, with even numbers of both protons and neutrons (even-even nuclei), could be identified as "vibrations" of the quadrupole shape, turns out to be fallacious. The vast amount of current experimental data shows that the $K^\pi = 0_2^+$ rotational bands are 2p-2h states and that the $K^\pi = 2_\gamma^+$ rotational bands are due to the energy being lowered by the nucleus favouring triaxiality and rotating about its intermediate axis. The experimental data mitigates against current, but still fashionable, nuclear models such as the Interacting Boson Model (IBM) and interpretations involving phonon excitations. It is significant that the deformed mean-field and single-particle behaviour is valid all the way up to fission.

1. Introduction

The Collective Model of deformed nuclei was formulated by Aage Bohr and Ben Mottelson back in 1953 [1] and is expounded in detail in their iconic book [2]. Any object that is not perfectly spherical can rotate, so deformed nuclei rotate to give rise to discrete energy levels forming a 'band'. The separation of the energy levels depends on the effective moment-of-inertia of the rotating object. Deformed nuclei rotate collectively with all nucleons in the configurations/wavefunctions of their initial state. Besides rotations Bohr and Mottelson also postulated that the quadrupole shape could 'vibrate'. Their model assumed deformed nuclei had a quadrupole shape using Lund Convention parameters β and γ to describe the deformation (Fig. 1): axial deformation is described by β and the triaxiality by γ. At $\gamma = 0°$ the nucleus is axially symmetric with a prolate shape. At $\gamma = 60°$ the nucleus is axially symmetric but with an oblate shape. In the region $0° \leq \gamma \leq 60°$ the nucleus is triaxial. The model postulates that there is a potential $V(\beta,\gamma)$ which is simplified by taking $V(\beta,\gamma) = V(\beta)V(\gamma)$ so that β and γ are treated as orthogonal. For simplicity $V(\beta)$ and $V(\gamma)$ are taken to be simple harmonic oscillators with oscillation frequencies ω_β and ω_γ. The problem is that the mass parameters are not known so that they have to be taken from experiment assuming that the first excitation in the $V(\beta)$ potential is the first excited 0_2^+ state at $\hbar\omega_\beta$ and the first excited $K = 2_\gamma^+$ band starting at $\hbar\omega_\gamma$; the β and γ bands. With the limited data available in 1953 this was a very reasonable assumption. Unfortunately it is not correct. However it is the interpretation of these structures that is given in all textbooks, the Nuclear Data Tables and the Firestone Data Sheets.

In 1879 John William Strut (Lord Rayleigh) worked out the frequency of vibration of a superfluid incompressible liquid sphere obtaining $\omega^2 = (\lambda-1)\lambda(\lambda+2)\gamma/\rho R^3$ where λ is the multipolarity of the vibration and is 2 for quadrupole, 3 for octupole etc., γ is the surface tension, ρ the density of the liquid and R is the radius of the drop [3]. In convenient units for a nucleus this becomes:

$$\omega^2 = (\lambda-1)\lambda(\lambda+2)C_S/3R_A^2mA \qquad (1)$$

where C_S is the surface term in the Weizsäcker Binding Energy formula ≈ 18 MeV, $R_A \approx 1.3$ fm and m is the nucleon mass. Quantizing the classical vibration using $E_x = \hbar\omega$ gives the results which are shown in Fig. 2, indicating that classical quadrupole and octupole vibrations are well above the Weizsäcker pairing gap, unlike the experimental 0_2^+ states and the first excited $K = 2_\gamma^+$ bands.

2. "β-vibrations"

Shiro Yoshida [4] pointed out in 1962 that the standard monopole pairing would decant all the strength of two-nucleon transfer reactions to excited $0_{n>1}^+$ states into the transition to the ground 0_1^+ state. Maher

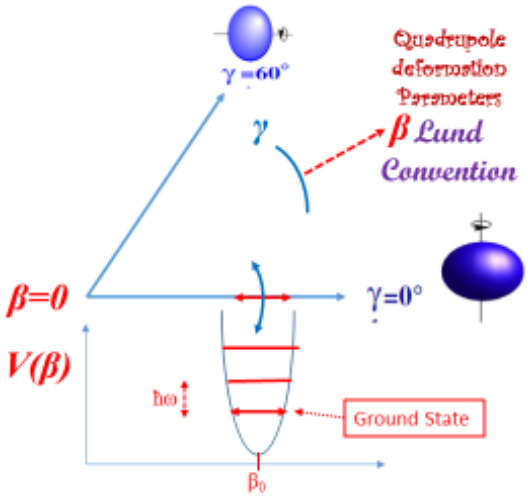


Figure 1. Diagram illustrating the β and γ quadrupole variables of the Lund convention.

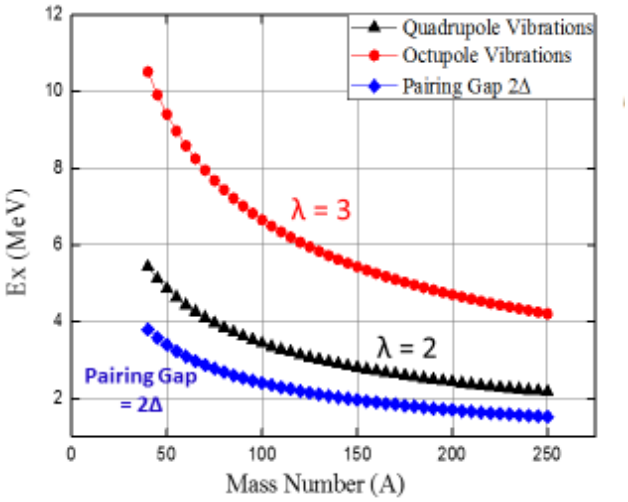


Figure 2. Classical vibrational energies of a liquid drop using eq. (1). The pairing energy $\Delta = 12/A^{1/2}$ MeV [2].

et al. [5] made (p,t) two neutron pick-up measurements on U and Th isotopes. They found that the 0_2^+ states were unexpectedly strongly populated. In the region of the nuclear chart near $A \approx 150$ both (t,p) two-neutron stripping and (p,t) pick-up reactions significantly populate the 0_2^+ states and higher lying 0_n^+ states [6,7], indicating that simple monopole pairing is too crude an interaction to describe these experiments. A better way of treating the pairing interaction was suggested by Griffin, Jackson and Volkov [8]. It was to split the pairing interaction into two; one for Nilsson orbitals in low-K prolate driving orbitals (low Ω), and a separate pairing interaction for high-K oblate driving orbitals (high Ω) as illustrated in Fig. 3. For two identical nucleons in the same j state, the monopole pairing interaction has the effect of lowering the coupled 0^+ state leaving all the other configurations at the same energy. Splitting the pairing gives two 0_2^+ states! But, in even-even nuclei with only two nucleons outside a closed shell, the reality is as illustrated on the right hand side of Fig. 3. The split pairing suggestion was adopted by others [7,9-12] and the 0_2^+ states formed were baptised "Pairing Isomers" by Ricardo Broglia [13]. The Bohr and Mottelson 'vibrational' states should also be strongly populated in inelastic scattering reactions. Experimentally it is found that while the $K = 2_\gamma^+$ bands are indeed strongly populated, the 0_2^+ states are not. As the 0_2^+ states are populated by $L = 0$ exchange of two nucleons they have to be 2p-2h states and not 'vibrations'.

If the experimentally observed states were 'vibrations', collective motion of all the nucleons, then they should emit enhanced quadrupole photons collectively between the 'vibrational' bands and the ground state band. Relatively recently Paul Garrett reviewed data on the *B(E2)* enhanced radiation from 0_2^+ states [14] and concluded, *"While it is clear that there are very few good examples of a β vibration, the question arises as to the nature of the 0_2^+ states, indeed of all low-lying 0^+ states."* Both Garrett [14] and Bohr and Mottelson [2, pp. 168ff] say that the best example of a β 'vibration' is ^{174}Hf referencing Ejiri and Hageman [15] who measured lifetimes of the excited states in the ground band and 0_2^+ band.

The conclusion that the low-lying 0_2^+ states have a 2p-2h structure means that the Nilsson orbital, that the pairs occupy, will be blocked in the neighbouring odd nuclei from coupling to the 0_2^+ state in the core nucleus. An example is illustrated in Fig. 4 where the neutron high-K 'oblate' Nilsson orbital $\nu[505]11/2^-$ in ^{155}Gd does not couple to the 0_2^+ excitation in the core ^{154}Gd [16]. Schmidt et al. [17] had previously shown that the low-K 'prolate' neutron orbitals $\nu[521]3/2^-$ and $\nu[651]3/2^+$ do indeed couple to the core 0_2^+ configuration. All these neutron configurations couple to the $K = 2_\gamma^+$ "γ vibration". It becomes clear that the 0_2^+ states near neutron number $N = 90$ meet the criteria for being Pairing Isomers.

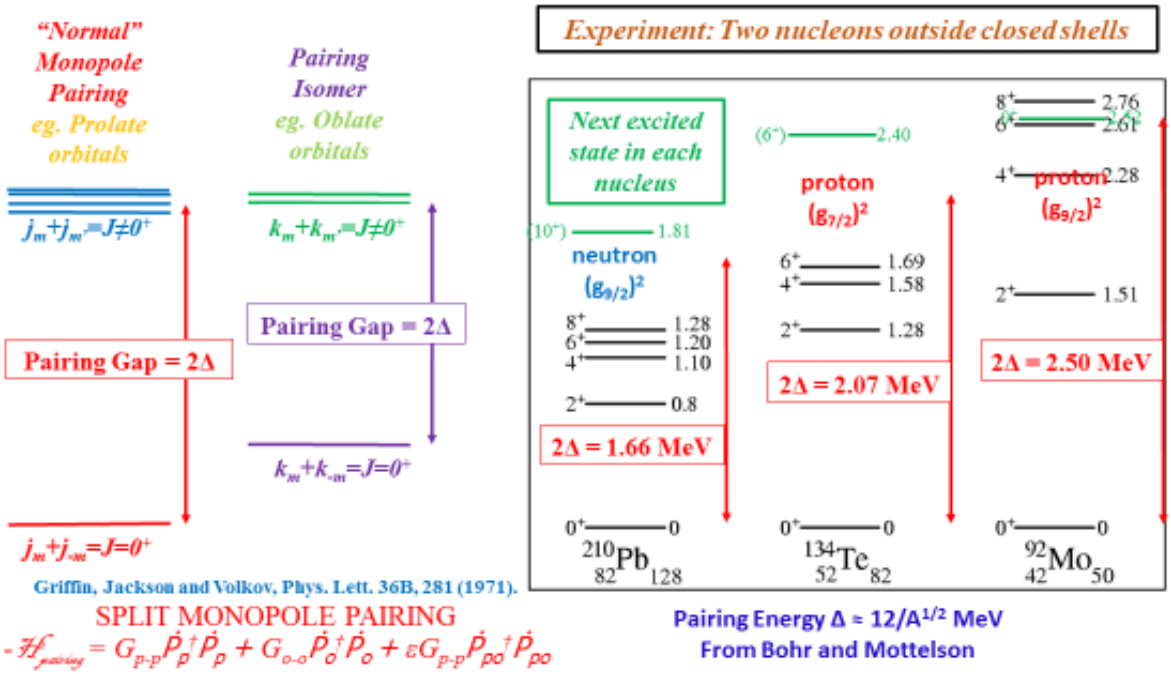


Figure 3. Left; illustration of splitting the monopole pairing. Right; Experimental reality.

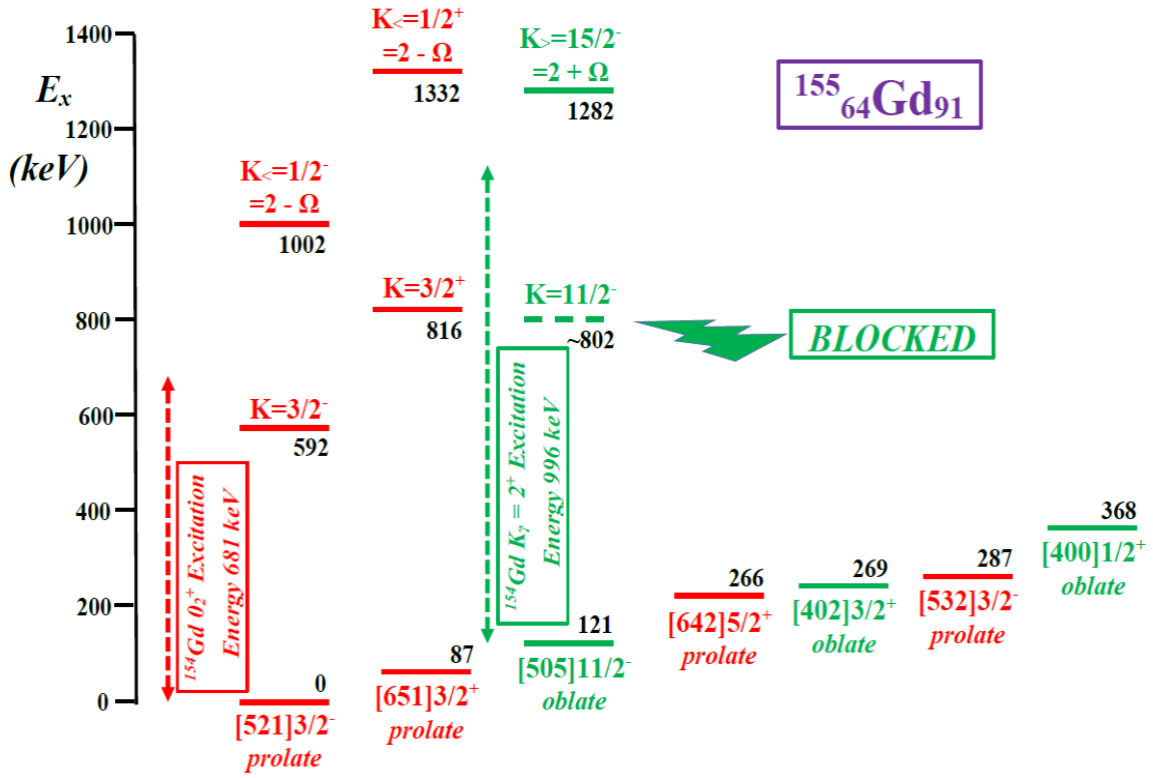


Figure 4. Schematic showing that the ν[505]11/2$^-$ orbital in ^{155}Gd is blocked from coupling to the $K = 2_\gamma^+$ core excitation in ^{154}Gd [16,17].

The question arises, are Pairing Isomers only found where high-K states are extruded to the Fermi surface by the adding of neutrons after the $N = 82$ shell closure and the increasing deformation? In Fig.5 we show the 0_1^+ ground-state and 0_2^+ state energy levels, minus the energies of a rigid rotor, plotted against spin for the nuclei $^{154}_{64}Gd_{90}$, $^{156}_{66}Dy_{90}$ and $^{174}_{72}Hf_{102}$ [18]. Fig. 5 strongly suggests that the underlying physics determining the structure of these three isotopes must be very similar. The population in inelastic (d,d') scattering of the 0_2^+ states in both $^{154}_{64}Gd_{90}$ [19] and $^{174}_{72}Hf_{102}$ [20] is very weak in comparison to that of the $K = 2_\gamma^+$ and 3_{oct}^- bands, confirming the different structures. The 0_2^+ state in $^{174}_{72}Hf_{102}$ cannot be a neutron pairing isomer, as in $^{154}_{64}Gd_{90}$ and $^{156}_{66}Dy_{90}$, because there are no high-K neutron Nilsson orbitals at the Fermi surface. But there is the high-K proton Nilsson orbital $\pi[404]7/2^+$ available which is involved in the structure of neighbouring isotopes and isomers in ^{174}Hf itself. This 0_2^+ state in ^{174}Hf is the first proton Pairing Isomer to be identified [18]. It shows that Pairing Isomers can exist for both neutron and proton 2p-2h configurations. It also shows that Pairing Isomers can occur mid-shell when there are suitable Nilsson levels close to the Fermi surface.

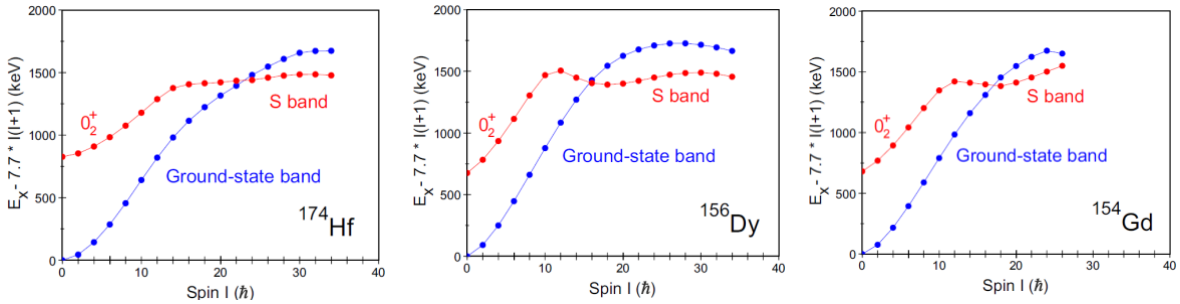


Figure 5. Excitation energies of levels, spin I, in the ground-state and 0_2^+ bands in ^{174}Hf, ^{156}Dy and ^{154}Gd minus a rigid rotor reference $E_r = 7.7I(I+1)$ keV [18].

3. "γ-vibrations"

Bohr and Mottelson offer two explanations for the first excited $K = 2_\gamma^+$ bands found in the pairing gap of even-even deformed nuclei: *"Such a collective mode could have the character of a vibration around an axially symmetric equilibrium or might be associated with an equilibrium shape deviating from axial symmetry."* [2, page 164]. And later [2, page 166*ff*] *"The available evidence, though inconclusive, thus appears to favour an interpretation in terms of a γ vibration…"*. They favour the vibrational model as it dovetails with their interpretation of β vibrations. The data, these days, indicate that β vibrations are an endangered species suggesting that the triaxiality explanation should be preferred.

An excellent account of approaches to the understanding of $K = 2_\gamma^+$ bands has been given by Shiekh et al. [21]. Wood, Allmond and colleagues [22,23] have extensively examined the effects of quadrupole triaxiality using the triaxial rotor model. A deformed axially asymmetric shape will rotate about its intermediate axis, as that axis has the largest moment-of-inertia (Fig. 6). The rotation can then tilt with respect to the intermediate axis giving rise to $K = 2_\gamma^+$, $4_{\gamma\gamma}^+$ ... bands which require excitation energy with respect to the ground state $K = 0_1^+$ band aligned with the intermediate axis [24,25]. Some of the most recent calculations of potential energy surfaces (PES) and of probability density distributions (PDD) are shown in Figs. 7, 8 and 9 [26-28]. These calculations demonstrate that while the ground state can be axially symmetric the $K = 2_\gamma^+$ band is triaxial.

Data supporting this interpretation are: that the lowest $K = 2_\gamma^+$ bands invariably track the ground-state band with increasing spin [27,29]; they are definitely collective as they are populated strongly in inelastic collisions [20]; they couple to all single-particle states (Fig. 4) showing they do not have a simple p-h structure; they decay by strong $B(E2)$ transitions to the ground-state band so that their internal structures must be similar (as in allowed β-decay); their intrinsic spin must be tilted with respect to the principal axis in order to have a projection $K = 2_\gamma^+$ onto that axis [25]. Finally, if there are no β vibrational bands, it becomes impossible to explain the $K = 2_\gamma^+$ bands as 'vibrational'.

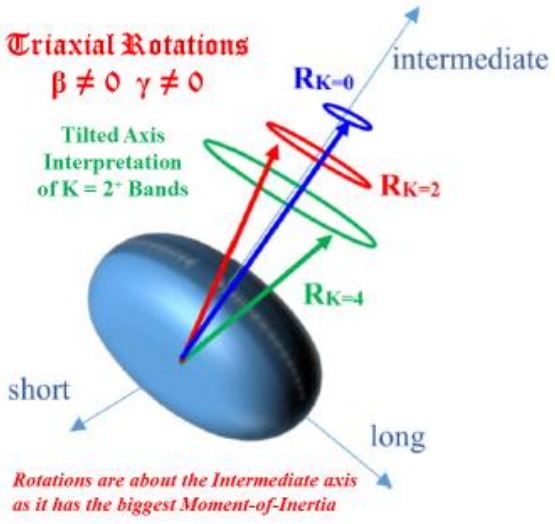


Figure 6. The rotational angular momentum around the long and short nuclear axes generates a tilt of the total rotational angular momentum [26].

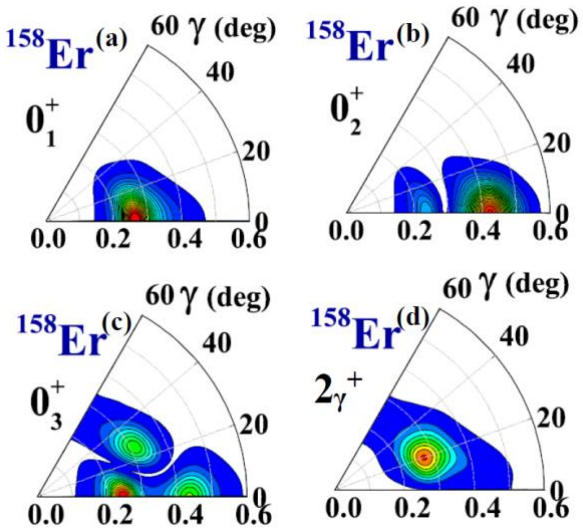


Figure 7. 5DCH-CDFT calculations of the Probability Density Distributions for the first band heads in ^{158}Er [27].

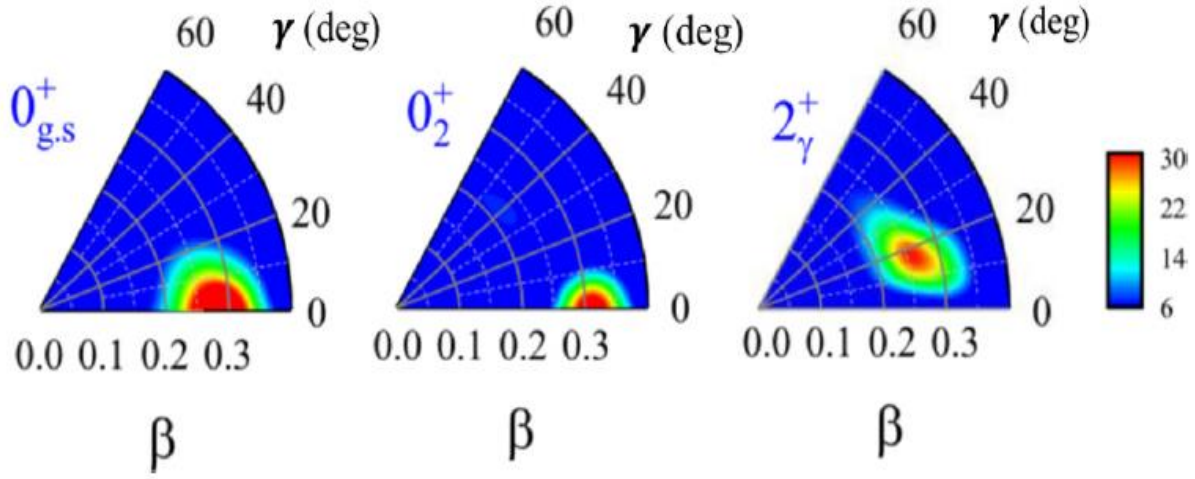


Figure 8. Probability Density Distributions for the first three band heads in ^{176}Os [26].

4. Conclusions

1. The $K = 0_2^+$ and $K = 2_\gamma^+$ bands are totally different phenomena and are not "vibrations".
2. The 0_2^+, 0_3^+.... are 2p-2h states (Pairing Isomers).
3. The $K = 2_\gamma^+$, $4_{\gamma\gamma}^+$ …. have triaxial quadrupole shapes with tilted rotation.
4. Boson (IBM) and other phonon interpretations do not describe the data.
5. Much of the data presented here can be found in greater detail in the review [29].

Acknowledgements

I would like to thank all my South African colleagues, who have worked on this topic with me, plus John Simpson who improved some of the figures.

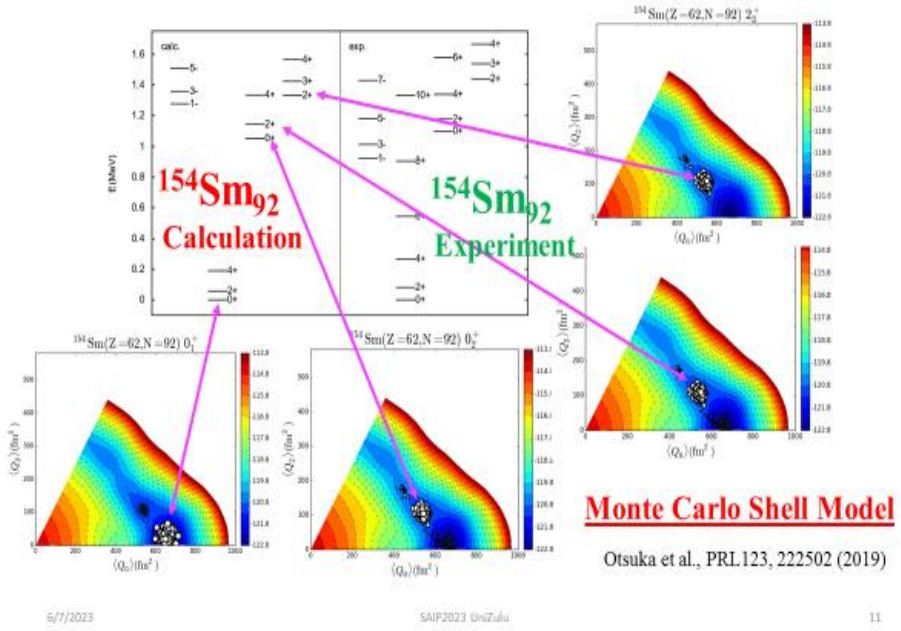


Figure 9. Probability Density Distributions for excited states in ^{154}Sm compared with experiment [28].

References

[1] Bohr A and Mottleson B R 1953 *Dan. Mat. Fys. Medd.* **27**, 16
[2] Bohr Aage and Mottelson B R 1975 *Nuclear structure Vol.II* (Benjamin)
[3] Lord Rayleigh (John William Strutt) 1879 *Proc. Roy. Soc.* **29** 71 Appendix II Eq.40
[4] Yoshida Shiro 1962 *Nucl. Phys.* **33** 685 (1962).
[5] Maher J V et al. 1970 *Phys. Rev. Lett.* **25** 302
[6] Shahabuddin M A M et al. 1980 *Nucl, Phys.* A **340** 109
[7] Kolata J J and Oothoudt M 1977 *Phys. Rev.* C **15** 1947
[8] Griffin R E, Jackson A D and Volkov A B 1971 *Phys. Lett.* B **36** 281
[9] Abdulvagabova S K, Ivanova S P and Pyatov N I 1972 *Phys. Lett.* B **38** 251
[10] B`es D R, Broglia R A and Nilsson B 1972 *Phys. Lett.* B **40** 338
[11] van Rij W I and Kahana S I 1972 *Phys. Rev. Lett.* **28** 50
[12] Ragnarsson I and Broglia R A 1976 *Nucl. Phys.* A **263** 315
[13] Ragnarsson I 2018 private communication.
[14] Garrett P E 2001 *J. Phys. Nucl. Part. Phys.* G **27** R1
[15] Ejiri H and Hagemann G B 1971 *Nucl. Phys.* A **161** 449
[16] Sharpey-Schafer J F et al. 2011 *Eur. Phys. J.* A **47** 6
[17] Schmidt H H et al. 1986 *J. Phys. G* **12** 411
[18] Sharpey-Schafer J F 2023 *Eur. Phys. J.* A **59** 183
[19] Bloch R, Elbek B and Tjøm P O 1967 *Nucl. Phys.* A **91** 576
[20] Walker P M 1983 *Phys. Scr.* **T5** 29
[21] Sheikh J A et al. 2008 *Phys. Rev.* C **77** 034313
[22] Wood J L et al. 2004 *Phys. Rev.* C **70** 024308
[23] Allmond J M et al. 2008 *Phys. Rev.* C **78** 014302
[24] Frauendorf S and Meng Jie 1997 *Nucl. Phys.* A **617** 131
[25] Lawrie E A, Shirinda O and Petrache C M 2020 *Phys. Rev.* C **101** 034306
[26] Mdletshe L et al. 2022 *Phys. Rev.* C **106** 044325
[27] Majola S N T et al. 2019 *Phys. Rev.* C **100** 044324
[28] Otsuka Takaharu 2018 *Nucl. Spin Workshop.* Valencia June
[29] Sharpey-Schafer J F, Bark R A, Bumbvi S P, Dinoko T R S and Majola S N T 2019 *Eur. Phys. J.* A **55** 15

Performance of the C10 cells of the Tile Calorimeter of the ATLAS detector during Run 2 data taking period

N P Rapheeha[1], **B Mellado**[1,2]

[1] School of Physics and Institute for Collider Particle Physics, University of the Witwatersrand, Johannesburg, Wits 2050, South Africa

[2] iThemba LABS, National Research Foundation, PO Box 722, Somerset West 7129, South Africa

E-mail: ntsoko.phuti.rapheeha@cern.ch

Abstract. The performance of the C10 cells of the Tile calorimeter is studied by their response to muons originating from $W \to \mu\nu_\mu$ events collected by the ATLAS detector during the Run 2 data taking period. The response is quantified by measuring the amount of energy deposited per unit length of each cell, in both the data and Monte Carlo simulations. The ratio of the response in data and Monte Carlo is used to quantify the calorimeter response, with deviations from unity hinting that there might be poor electromagnetic energy scale calibration in the experimental data.

1. Introduction

This study aims to determine the performance of C10 cells of the Tile Calorimeter (TileCal) [1] of the ATLAS detector [2]. The performance is determined by studying the response of the cells to isolation muons originating from the decay of the W boson to a muon and a muon neutrino $W \to \mu\nu_\mu$. When a charged particle passes through a cell, the response of the cell (R) is quantified as the amount of energy (ΔE) deposited per unit length (Δx) in that particular cell. The energy loss of muons as they transverse through matter is a well understood process [3]. Ionisation is the dominant energy loss in muons with energies below 100 GeV and the probability distribution of the response follows the Landau function for thin scintillators. The $W \to \mu\nu_\mu$ events used in this study are produced during proton-proton (pp) collisions at the Large Hadron Collider [4] observed by the ATLAS detector during the Run 2 data taking period (2015-2018) at centre-of-mass energy of $\sqrt{s} = 13$ TeV with 139 fb^{-1} integrated luminosity. .

Muons stemming from the $W \to \mu\nu_\mu$ process are selected due to their distinct and clear experimental characteristics within the detector. Background contributions from muons originating from other processes can be easily suppressed. In the $W \to \mu\nu_\mu$ channel, the electroweak background is predominantly composed of $W \to \tau\nu_\tau$ and $Z/\gamma^* \to \mu^+\mu^-$ processes. Background events from the $Z/\gamma^* \to \mu^+\mu^-$ process are suppressed by requiring the selected events have exactly one muon candidate. The decay of the τ-lepton into a muon and a τ neutrino is a complex process that involves multiple particles, leading to a displaced τ decay vertex and, on average, a muon with low transverse momentum denoted as p_{T} [5, 6]. The p_{T} of the selected muon candidates in addition to the position of the primary vertex is used to reject

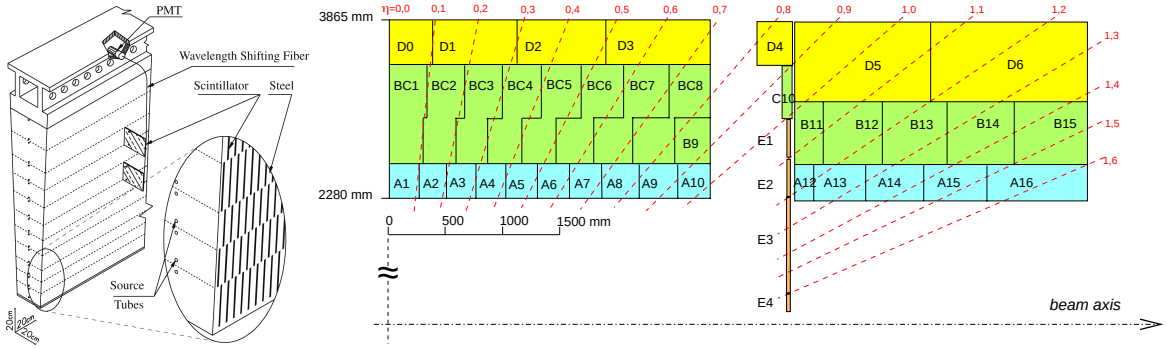


Figure 1. Left: Schematic of one TileCal wedge [7], showing the mechanical assembly and optical readouts. Right: Schematic of the TileCal cell layout in a plane parallel to the beam axis [8] and scintillator structure on the A-side of the detector.

these background events. In events where heavy quarks undergo semi-leptonic decays, there's a possibility of hadrons being misidentified as leptons. These multi-jet background processes substantially contribute to the total background. The presence of missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) in the events is to suppress these multi-jet backgrounds.

2. The ATLAS Tile Calorimeter

The ATLAS experiment at the Large Hadron Collider is one of the two general-purpose detectors designed to reconstruct events from pp or heavy-ion collisions. The hadronic calorimeter system of the ATLAS detector is formed by the TileCal, which plays a major role in the reconstruction of hadrons, jets, τ-leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ arising from non-interacting particles.

It covers the central region of the ATLAS detector up to $|\eta| < 1.7$. The TileCal is divided into three segments along the z-axis (the beam axis). It has one fixed long barrel (LB) 5.8 metres in length covering the central area of the collision point $|\eta| < 1.0$ and two 2.6 meters in length extended barrels (EB) in the forward and backward region $0.8 < |\eta| < 1.7$. The region $0.8 < |\eta| < 1.0$ of the Intermediate Tile Calorimeter is called the plug. The region between $0.9 < |\eta| < 1.0$ has 96 mm stacks (C10) in the z-direction. The TileCal is symmetric with respect to the interaction point, the positive η side is called the A-side while the negative η side is called the C-side. Figure 1 shows the schematic layout and the scintillator structure of the TileCal cell assembly on the A-side of the detector.

The full azimuthal (ϕ) coverage around the beam axis is achieved using 64 wedge-shaped modules, a schematic of such module of the TileCal is shown in Figure 1, each covering $\Delta\eta = 0.1$ radians. The modules are oriented radially and normal to the beam line. The source tubes are used to circulate a capsule of ^{137}Cs radioactive source for calibration purposes. The TileCal has total of 256 of such modules, 128 on the A side and the other 128 on the C side. When a charged particle traverses through the scintillating tiles, light is emitted and collected at the scintillator edges. The light generated by the scintillating tiles is then transported to two photomultiplier tubes (PMTs) by wavelength shifting (WLS) fibres [9] in standard TileCal cells. In special TileCal cells formed by a single piece of scintillator, the scintillation light is transported by just a single PMT. The TileCal cells are defined by a volume defined by the η, ϕ and radial segmentation of the TileCal.

The Technical Design Report [10] provides a detailed description of the ATLAS TileCal and its construction, optical instrumentation and installation into the ATLAS detector are detailed in Refs. [9, 11].

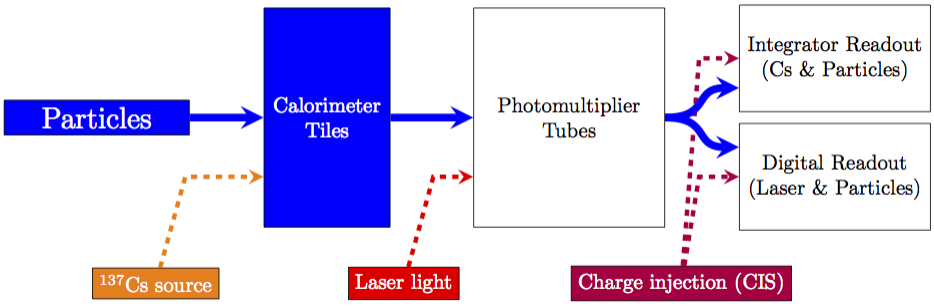


Figure 2. A schematic of the TileCal signal reconstruction path [15]. The path of the physics signal is shown by the thick solid lines and calibration paths of the different systems are indicated with dashed lines.

3. Calibration Procedure

Three calibration systems are used to monitor the signal reconstruction in TileCal cells. The calibration systems are used to maintain a time-independent electromagnetic (EM) energy scale, and account for changes in the hardware and electronics due to radiation damage, ageing and faults. The Caesium (Cs) calibration system uses global electromagnetic scale to monitor the optical and electrical response of each PMT in the ATLAS environment [12]. A hydraulic system moves a capsule containing the ^{137}Cs radioactive source through stainless steel tubes that pass through every TileCal module, the holes for the tubes are illustrated in Figure. 1. The laser calibration system is used to monitor and correct for PMT response variations in between Cs scans and to monitor time stability during data-taking periods [13, 14]

A set of calibration constants f are then used to convert the reconstructed channel amplitude in units of ADC counts to units energy in GeV. The conversion is given by:

$$E[\text{GeV}] = \frac{A[\text{ADC}]}{f_{\text{pC}\to\text{GeV}} \cdot f_{\text{Cs}} \cdot f_{\text{Las}} \cdot f_{\text{ADC}\to\text{GeV}}}, \tag{1}$$

where $f_{\text{pC}\to\text{GeV}}$ is a calibration constant determined in dedicated electron and muon test beam analyses. The constants f_{Cs}, f_{Las} and $f_{\text{ADC}\to\text{GeV}}$ are obtained from independent Caesium, laser and CIS calibration systems, respectively.

Figure 2 shows a flow diagram summarising the different calibration systems along with the paths followed by the signals different calibration sources. The study of the response of the TileCal cells allows us to confirm whether the calibration of TileCal modules was applied correctly. The three calibration systems aid in identifying potential sources of the problematic channels. If a module has problems originating from readout electronics, then the laser and CIS calibration systems will be able to detect that.

4. Event Selection

Table 1 summarises the selected events used in this analysis. The selections were chosen to optimise the number of events with a single muon originating from the $W \to \mu\nu_\mu$ process while minimising the number of events from background processes. The selected events are required to contain a single reconstructed muon candidate that passes either the "medium" or "tight" identification requirement [16] to eliminate electroweak background events. A cut on the $E_{\text{T}}^{\text{miss}}$ is required to exclude the multi-jet background processes. Cuts on the track ($\sum p_{\text{T}}|_{\Delta R=0.4}$) and calorimeter ($E_{\text{LAr}}|_{\Delta R=0.4}$) isolation variables are applied to ensure that the selected muon candidate is well isolated from the hadronic activity around it.

Table 1. Event selection based on the W decays

	Variable	Requirement
1	Number of Muons	$\mathrm{N_{muons}} = 1$
2	Transverse invariant mass	$40 < M_{\mathrm{T}} < 140$ GeV
3	Missing transverse energy	$30 < E_{\mathrm{T}}^{\mathrm{miss}} < 120$ GeV
4	Track isolation	$\sum p_{\mathrm{T}}\vert_{\Delta R=0.4} < 1$ GeV
5	Calorimeter isolation	$E_{\mathrm{LAr}}\vert_{\Delta R=0.4} < 1.5$ GeV
6	Momentum of the muon	$p^{\mu} <= 80$ GeV
7	Transverse momentum of the muon	$p_{\mathrm{T}}^{\mu} > 28$ GeV
8.	Average interactions per bunch crossing	$10 < \mu > 40$

In addition to the event selections listed in Table 1, the energy deposited by the passing muon candidate in the gap and crack cells is required to be above 60 MeV in order to suppress signals that are caused by electronic noise. The muon is required to traverse a minimum path length of 100 mm in the C10 cells. The absolute azimuthal angle between the muon's track and cell's centre $|\Delta\phi(\mu, \mathrm{cell})| < 0.046$.

The experimental pp data collected during the LHC Run 2 data taking at $\sqrt{s} = 13$ TeV is used. Integrated luminosities for the respective data taking periods are as follows: $\int \mathcal{L} = 36.2$ fb^{-1} for 2015+2016, $\int \mathcal{L} = 44.3$ fb^{-1} for 2017, and $\int \mathcal{L} = 58.5$ fb^{-1} for 2018. Monte Carlo (MC) simulated samples of the $W \to \mu\nu_{\mu}$ events are used as a reference when studying the response from experimental data. Three sets of MC event samples are generated to match pile-up conditions of the three data taking periods: 2015-2016, 2017, and 2018. The event generator Sherpa [17] was used to generate the W boson production events which are then interfaced with PYTHIA8 [18] for parton showering.

The event selection from Table 1 is applied to both the data and MC samples. The events from the simulated samples are normalised to the $W \to \mu\nu_{\mu}$ production cross-section multiplied by the integrated luminosity of the corresponding data taking period. The simulated events are also reweighted to account for the different pile-up conditions.

5. Results

The R uniformity of each cell is evaluated over the entire azimuthal direction on the A and C sides of the detector. A Gaussian likelihood function,

$$\mathcal{L}_c = \prod_{m=1}^{64} \frac{1}{\sqrt{2\pi}\sqrt{\sigma_{c,m}^2 + s_c^2}} \exp\left[-\frac{1}{2} \frac{(R_{c,m} - \mu_c)^2}{\sigma_{c,m}^2 + s_c^2} \right], \tag{2}$$

where $R_{c,m}$ and $\sigma_{c,m}$ are the observed R and its statistical uncertainty for a given cell (c) in module m. The uncertainty $\sigma_{c,m}$ is obtained by propagating the uncertainties of the dE/dx distributions from data and Monte Carlo,

$$\sigma_{c,m} = \left|\frac{r^{\mathrm{data}}}{r^{\mathrm{MC}}}\right| = \sqrt{\left(\frac{\sigma_r^{\mathrm{data}_r}}{r^{\mathrm{data}}}\right)^2 + \left(\frac{\sigma_r^{\mathrm{MC}_r}}{r^{\mathrm{MC}}}\right)^2}, \tag{3}$$

where r is just short hand for truncated mean of the dE/dx distribution. The uncertainty of the truncated mean σ_r is defined as the standard mean error of the mean in each module.

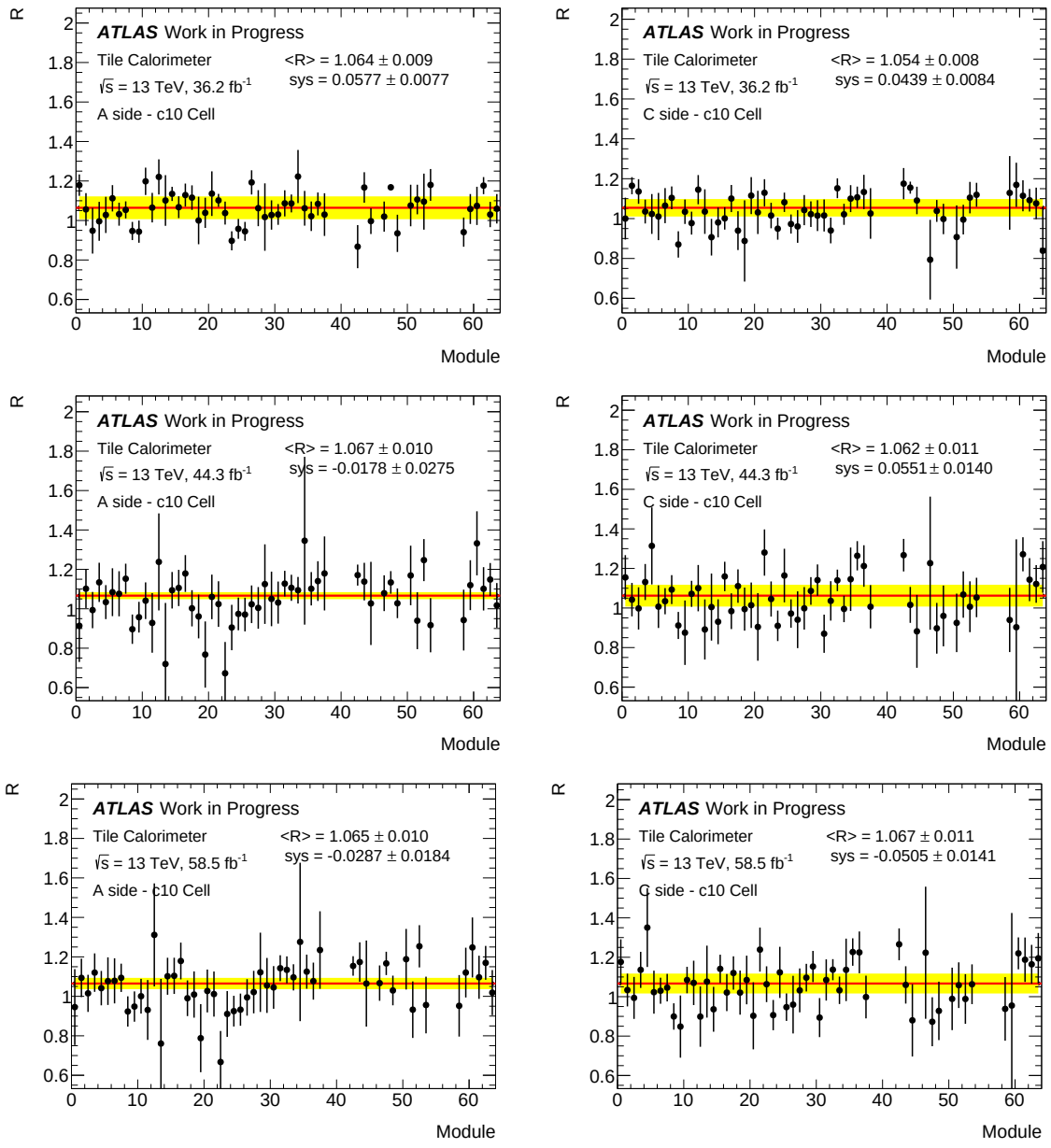


Figure 3. Distribution of module number vs dE/dx in C10cells. The red line represents the fitted average response across all modules and the yellow band is the systematic uncertainty associated the fitted average value of R. The plots show results for 2015-2016 (top), 2017 (middle) and 2018 (bottom) data taking periods.

The unbinned log likelihood $-2\log\mathcal{L}$ is minimised using minuit [19] package to find $\hat{\mu}_C$ and $\hat{s}_C$ which are the average response $< R >$ and the systematic uncertainty attributed to the non-uniformity across the modules.

The response of the cells could change over time due to factors like PMT drift response, ageing effects of the scintillators and wavelength shifting fibres [20]. The average response C10 cell on the A-side shows the biggest drop in 2018 data taking period, this can be attributed the drop in response of modules 14, 20 and 23. A significant deviation in the response on the C-side

of the detector is seen in 2017 data, this can be attributed a drop in the response of modules 48 and 49.

6. Conclusions

In this study, we investigated the performance of the C10 cells within the TileCal of the ATLAS detector during the Run 2 data-taking period. By analysing the response of these cells to muons originating from $W \to \mu\nu_\mu$ events, we aimed to assess the calorimeter's performance and its calibration accuracy.

An unbinned negative log likelihood function was minimised to obtain the average response of each cell during different data taking years. The average response of C10 cells vary slightly across modules, raising questions about the ϕ uniformity of the response. In the preliminary results, the drop of the average response suggested potential issues such as PMT drift response, scintillator ageing, or wavelength shifting fibre effects.

References

[1] Henriques Correia A M (for the ATLAS collaboration) 2015 The ATLAS Tile Calorimeter Tech. rep. CERN Geneva URL https://cds.cern.ch/record/2004868

[2] Aad G *et al.* (ATLAS) 2008 *Journal of Instrumentation* **3** S08003–S08003 also published by CERN Geneva in 2010 URL https://doi.org/10.1088/1748-0221/3/08/s08003

[3] Zyla P *et al.* (Particle Data Group) 2020 *PTEP* **2020** 083C01 and 2021 update

[4] Evans L and Bryant P 2008 *Journal of Instrumentation* **3** S08001–S08001 URL https://doi.org/10.1088/1748-0221/3/08/s08001

[5] Belous K *et al.* (Belle) 2014 *Phys. Rev. Lett.* **112** 031801 (*Preprint* 1310.8503)

[6] Feldman G J *et al.* 1982 *Phys. Rev. Lett.* **48**(2) 66–69 URL https://link.aps.org/doi/10.1103/PhysRevLett.48.66

[7] Aad G *et al.* (ATLAS) 2010 *Eur. Phys. J. C* **70** 1193–1236 (*Preprint* 1007.5423)

[8] Manzoni S 2019 *Electron and Photon Energy Calibration* (Cham: Springer International Publishing) pp 115–148 ISBN 978-3-030-24370-8

[9] Abdallah J *et al.* (ATLAS) 2013 *Journal of Instrumentation* **8** T11001–T11001 URL https://doi.org/10.1088/1748-0221/8/11/t11001

[10] 1996 *ATLAS tile calorimeter: Technical Design Report* Technical design report. ATLAS (Geneva: CERN) URL http://cds.cern.ch/record/331062

[11] Abdallah J *et al.* (ATLAS TileCal) 2013 *JINST* **8** P01005

[12] Starchenko E, Blanchot G, Bosman M, Cavalli-Sforza M, Karyukhin A, Kopikov S, Miagkov A, Nessi M, Shalimov A, Shalanda N, Soldatov M, Solodkov A, Soloviev A, Tsoupko-Sitnikov V and Zaitsev A 2002 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **494** 381–384 ISSN 0168-9002 proceedings of the 8th International Conference on Instrumentatio n for Colliding Beam Physics URL https://www.sciencedirect.com/science/article/pii/S0168900202015073

[13] Blanchot G, Bosman M, Budagov J, Cavalli-Sforza M, Efthymiopoulos I, Isaev A, Ivanyushenkov Y, Karyukhin A, Kopikov S, Nessi M, Senko V, Shalanda N, Soldatov M, Solodkov A, Solovyanov O, Starchenko E, Tsoupko-Sitnikov V, Vichou I and Zaitsev A 2020 **15** P03017 URL https://dx.doi.org/10.1088/1748-0221/15/03/P03017

[14] Agaras M N *et al.* 2023 *JINST* **18** P06023 (*Preprint* 2303.00121)

[15] Vaslin L (ATLAS) 2021 The ATLAS Tile Calorimeter Performance and its upgrade toward High-Luminosity LHC Tech. rep. CERN Geneva URL https://cds.cern.ch/record/2790583

[16] Aad G *et al.* (ATLAS) 2016 *Eur. Phys. J. C* **76** 292 (*Preprint* 1603.05598)

[17] Bothmann E *et al.* (Sherpa) 2019 *SciPost Phys.* **7** 034 (*Preprint* 1905.09127)

[18] Sjostrand T, Mrenna S and Skands P Z 2008 *Computer Physics Communications* **178** 852–867 ISSN 0010-4655 (*Preprint* 0710.3820)

[19] James F and Roos M 1975 *Comput. Phys. Commun.* **10** 343–367

[20] Aaboud M *et al.* (ATLAS) 2018 *Eur. Phys. J. C* **78** 987 (*Preprint* 1806.02129)

A search for tWZ production at $\sqrt{s}$=13TeV with the ATLAS detector

Alex Veltman, James Keaveney

Department of Physics, University of Cape Town, South Africa, 7701

E-mail: vltale001@myuct.ac.za

Abstract. The production of a single top quark in association with a W boson and a Z boson (tWZ) is a rare Standard Model process that has never before been measured. The process is a valuable input into global Standard Model Effective Field Theory (SMEFT) fits due to its sensitivity to top-electroweak SMEFT coefficients in regions of high W boson and Z boson transverse momenta. It is also a relevant background in other top quark related measurements such as the $t\bar{t}Z$ cross-section measurement. A search has been performed for tWZ production using $140\,\text{fb}^{-1}$ of proton-proton collision data at a centre of mass energy of 13 TeV. The search explores the trilepton and tetralepton tWZ final states. Regions are defined using physics object multiplicities and graph neural networks are employed to perform signal-background discrimination in the regions. The signal strength of tWZ production μ_{tWZ} is extracted using a profile likelihood fit with a full systematic model describing experimental and modelling uncertainties. Preliminary blinded measurements of μ_{tWZ} are shown using an Asimov data set for each channel. A combined signal extraction across both channels will also be presented. The measurements of μ_{tWZ} shown will include their associated expected significances and expected upper limits.

1. Introduction

The production of a single top in association with a W boson and a Z boson (tWZ) is a rare process predicted by the Standard Model (SM). The tWZ process has never been observed due to a small expected cross-section of 161 fb and difficult to discriminate sources of backgrounds. An example of a leading order diagram of tWZ production is shown in Figure 1, which includes two electroweak vertices. Due to the presence of top-electroweak vertices, the process is seen as valuable for global Standard Model Effective Field Theory (SMEFT) fits [1]. The rate of tWZ production at high energies is sensitive to certain top-electroweak SMEFT operators, which will allow tWZ production to provide a unique constraint in global SMEFT fits. The process is also a background in other top quark related analyses such as the cross-section measurement of $t\bar{t}Z$ production.

The analysis is divided into the trilepton and the tetralepton decay channels, where both channels require the final state Z boson to decay leptonically. The trilepton decay channel requires either the prompt W boson or the W boson originating from the decaying top quark to decay leptonically. In the trilepton channel, $t\bar{t}Z$ production and the production of a WZ pair with additional jets(WZ+jets) contribute background events that are difficult to distinguish from signal events. The tetralepton channel requires both the prompt W boson and the W boson originating from the decaying top quark to decay leptonically. In the tetralepton channel,

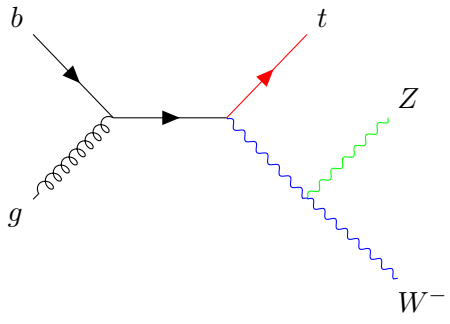


Figure 1. Example of a leading order Feynman diagram for tWZ production. The final state products of a t quark (red), a Z boson (green) and a W boson (blue) are shown.

the major backgrounds are $t\bar{t}Z$ production and the production of a ZZ pair with additional jets (ZZ+jets). The background events in the tetralepton channel are easier to distinguish from signal events in the trilepton channel. The tetralepton channel has fewer total events, which increases the statistical uncertainty associated with the channel compared to the trilepton channel.

2. Samples and event selection

Data and simulated samples were prepared for the trilepton and tetralepton channels. The search was performed using $140\,\mathrm{fb}^{-1}$ of 13 TeV proton-proton collisions recorded using the ATLAS detector [2]. This dataset was recorded at the Large Hadron Collider between 2015 and 2018. A comparable dataset of simulated Standard Model processes was generated to compare the event counts between ATLAS data and the Standard Model prediction. The simulated processes include tWZ production as well as the background processes of $t\bar{t}Z$ production and diboson (WZ+jets/ZZ+jets) production with additional jets. Simulated samples were also created for minor multibosonic and top-related background processes.

Leptons, jets and missing transverse momentum were reconstructed from experimental signatures in the various subdetectors of the ATLAS detector. Only electrons and muons were considered. Jets that contain B-hadrons were identified using a machine learning discriminator DL1r [3], where b-tagged jets are defined as jets that pass the 77% DL1r b-jet tagging efficiency operating point.

Events in the trilepton channel were required to have three leptons, where the leading lepton, subleading and subsubleading lepton must have $p_T > 30$GeV, $p_T > 20$GeV and $p_T > 14$GeV, respectively. Muons must be within the region of $|\eta| < 2.5$ and electrons must be within $|\eta| < 2.47$ and not within $1.37 < |\eta| < 1.52$. All jets must have $p_T > 25$ GeV, be in the forward region of $|\eta| < 2.5$ and have a jet-vertex-tagger (JVT) [4] of greater than 0.5. Any combination of oppositely-signed same-flavoured (OSSF) leptons must have a combined mass that is greater than 10 GeV. A combined OSSF lepton pair is a Z candidate if the pair's mass is within 10 GeV of the Z boson mass $m_Z = 91.19$. Events in the trilepton channel were required to have exactly one Z candidate.

Each event in the tetralepton channel must have four leptons, where the highest energy lepton has $p_T > 28$ GeV, the second highest energy lepton has $p_T > 18$ GeV and the third and fourth highest energy leptons have $p_T > 10$ GeV. The jet selection requirements and the lepton η requirements are the same as for the trilepton channel. The sum of the lepton charges must also be neutral and all OSSF lepton pairs in the event must have a mass greater than 10GeV.

All events were partitioned into regions based on the multiplicity. The trilepton channel has three regions: one signal region (SR) and two control regions (CR). The events in the trilepton tWZ SR were required to have three or more jets, where exactly one of the jets is a b-tagged jet. The events in the trilepton $t\bar{t}Z$ CR must have four or more jets, where two or more of the jets are b-tagged jets. The events in the trilepton WZ CR must have one or two jets, where

exactly one jet is a b-tagged jet. Five regions were defined for the tetralepton channel. The tetralepton channel has two signal regions that require events to have one Z candidate, greater than one jet and exactly one b-tagged jet. The tWZ opposite-flavoured (OF) SR and the tWZ same-flavoured (SF) SR differ based on the flavours of the leptons that did not originate from the Z candidate. The tWZ OF SR contains events where the two non-Z leptons have different flavours and the tWZ SF SR contains the events where the two non-Z leptons have the same flavour.

3. Background discrimination with graph neural networks

The signal regions are dominated by background contributions from $t\bar{t}Z$ and diboson production. A pair of graph neural networks [5] were developed to discriminate between signal and background events in the trilepton and tetralepton channels. Each event can be modelled as a graph where the muons, electrons, jets and missing transverse momentum are nodes in the graph. Each node carries information regarding the 4-momentum of the object and what type of object the node represents. The edges between the nodes carry an edge weight that is equal to the angular difference ΔR between the objects associated with the nodes. Additional graph-level information was also included regarding the number of jets in the event. The graph models followed the message-passing neural network architecture [6].

The model used in the trilepton channel was trained and tested using simulated tWZ, $t\bar{t}Z$ and WZ+jets events from the trilepton tWZ SR. The model used in the tetralepton channel was trained and tested using simulated tWZ, $t\bar{t}Z$ and ZZ+jets events from the tetralepton signal regions and the $t\bar{t}Z$ CR. The models are binary classifiers that produce a score between 0 and 1, where the signal tWZ events were assigned a label of 1 and the background $t\bar{t}Z$ and diboson events were assigned a label of 0.

4. Systematic uncertainties

The search considers the experimental uncertainties related to data-taking and the modelling uncertainties associated with the various simulated processes. The recommendations regarding systematic uncertainties provided by the internal ATLAS combined performance groups were followed. The experimental uncertainties associated with the luminosity and pileup conditions of the datasets were considered. A set of uncertainties related to the reconstruction of objects were included in the analysis, such as uncertainties associated with the energy scale and resolution of objects, jet flavour information and lepton efficiency scale factors.

The modelling uncertainties in the search included the estimation of the cross section normalisation factor applied to each of the simulated background samples. The choice of diagram removal scheme used for the tWZ sample [7] was accounted for using a systematic uncertainty. For the tWZ sample, systematic uncertainties were introduced to describe normalisation and factorisation scale variations in the matrix element calculations and to describe different parameterisations of the parton distribution function (PDF).

5. Extraction method

The parameter of interest of the extraction method is the signal strength of tWZ production, $\mu_{tWZ} = \sigma_{\text{obs}}/\sigma_{SM}$, where σ_{obs} is the observed cross-section of tWZ production and σ_{SM} is the expected cross-section of tWZ production according to the SM. A likelihood model [8] was constructed with μ_{tWZ} as the parameter of interest as well as nuisance parameters describing each systematic uncertainty. The parameter values were estimated using a profile likelihood fit [9], which allows for the statistical significance Z_0 of the measurement to be obtained.

In this proceeding, only an *Asimov dataset* [9] was used during the signal extraction. An Asimov dataset is designed to produce best-fit values of the parameters that are exactly equal to the simulated expectation. The Asimov dataset allows for the experimental sensitivity of the

analysis strategy to be evaluated. A *stat-only fit* was performed to evaluate the uncertainty due to statistical fluctuations. This involves first performing a fit to data with the parameter of interest fixed in order to estimate the other parameters. The parameter of interest is then estimated while the other parameters are kept fixed at these new values.

6. Fit results with Asimov dataset

The extraction of the tWZ signal strength μ_{tWZ} was performed individually on the trilepton and tetralepton Asimov datasets and a combined extraction was performed using both Asimov datasets. The measured signal strength μ_{tWZ} was obtained through profile likelihood fits using the trilepton dataset, tetralepton dataset and combined dataset and the results are shown in Figure 2. All μ_{tWZ} were found to be equal to one, which is expected when performing fits to an Asimov dataset.

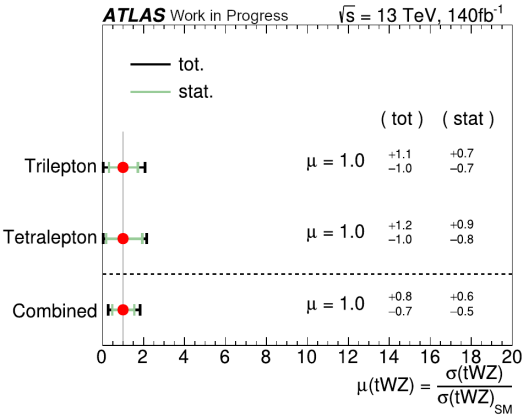


Figure 2. The estimated value and uncertainty of μ_{tWZ} for various fits to different Asimov datasets. This includes the fit results from the trilepton-only, tetralepton-only and combined fits. The total uncertainty of the fit is shown. The uncertainty when a stat-only fit is performed is also shown.

The total uncertainty associated with the measurements in the trilepton channel and the tetralepton channel were similar. The precision of the trilepton measurement was affected by both systematic and statistical uncertainty. The measurement in the tetralepton channel was dominated by statistical uncertainty. The relatively high statistical uncertainty was attributed to the difficulty in separating tWZ events and background events as well as the low number of events with four leptons. The combined measurement has a lower total uncertainty compared to the measurements in the trilepton or tetralepton channels alone.

The post-fit values of the nuisance parameters describing the systematic uncertainties are ranked in terms of impact in Figure 3. The estimation of the cross-section normalisation of $t\bar{t}Z$ production had the largest impact on the measurement, which may be the result of $t\bar{t}Z$ production being a background in both the trilepton and tetralepton channel. The systematic uncertainty associated with the choice of event generator used in simulating the $t\bar{t}Z$ sample is also highly ranked in impact. The uncertainties associated with jet energy resolution and the variation in the PDF calculations for the tWZ samples were the second and third most impactful systematic variables.

The expected significance associated with the μ_{tWZ} measurement in the combined fit was $Z_0^{\text{exp}} = 1.34\sigma$. The expected significance of the measurement suggests that a discovery of tWZ production will not be possible if the observed ATLAS data agrees with the simulated samples. A reduction in the systematic and statistical uncertainty would be required for an observation to be expected.

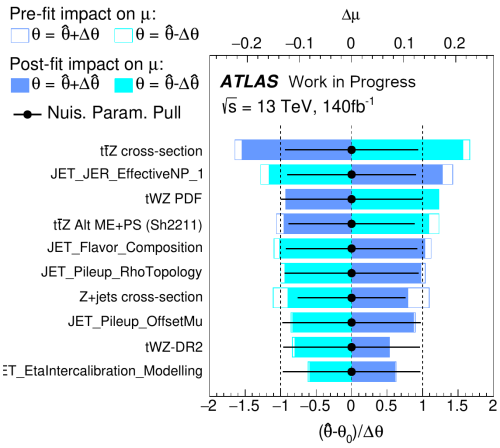


Figure 3. Ranking of pre-fit and post-fit impact values of each nuisance parameter. Only the nuisance parameters with the ten highest impact values are shown. The upper axis is in terms of impact $\Delta\mu$. The pre-fit impacts are shown as unfilled blue rectangles and the post-fit impacts are shown as filled blue rectangles. The lower axis describes pull on the nuisance parameter during the fit, where $\hat{\theta}$, θ_0 and $\Delta\theta$ are the best-fit value, pre-fit value and pre-fit uncertainty of the nuisance parameter. The pull value and post-fit uncertainty of each nuisance parameter are included as black points.

7. Conclusion

The production of a single top quark in association with a W boson and a Z boson (tWZ) is a rare Standard Model process that has never before been measured. A search for tWZ production was performed in the trilepton and tetralepton final states using 140 fb^{-1} of proton-proton collision data taken at the ATLAS detector. The experimental sensitivity of the analysis strategy was demonstrated using Asimov datasets. A signal strength value of $\mu_{tWZ} = 1.0^{+0.8}_{-0.7}$ was found when using a combined trilepton and tetralepton Asimov dataset. The associated expected significance of the measurement was $Z_0^{\text{exp}} = 1.34\sigma$. The measurement was found to be both statistically and systematically limited.

References

[1] Maltoni F, Mantani L and Mimasu K 2019 *JHEP* **10** 004 (*Preprint* 1904.05637)
[2] Collaboration T A 2008 *Journal of Instrumentation* **3** S08003 URL https://dx.doi.org/10.1088/1748-0221/3/08/S08003
[3] ATLAS Collaboration 2019 *The European Physical Journal C* **79** URL https://doi.org/10.1140/epjc/s10052-019-7450-8
[4] 2014 Tagging and suppression of pileup jets Tech. rep. CERN Geneva URL https://cds.cern.ch/record/1643929
[5] Battaglia P W, Hamrick J B, Bapst V, Sanchez-Gonzalez A, Zambaldi V, Malinowski M, Tacchetti A, Raposo D, Santoro A, Faulkner R, Gulcehre C, Song F, Ballard A, Gilmer J, Dahl G, Vaswani A, Allen K, Nash C, Langston V, Dyer C, Heess N, Wierstra D, Kohli P, Botvinick M, Vinyals O, Li Y and Pascanu R 2018 Relational inductive biases, deep learning, and graph networks (*Preprint* 1806.01261)
[6] Gilmer J, Schoenholz S S, Riley P F, Vinyals O and Dahl G E 2017 Neural message passing for quantum chemistry (*Preprint* 1704.01212)
[7] Bessidskaia Bylund O (ATLAS) 2016 Modelling Wt and tWZ production at NLO for ATLAS analyses *9th International Workshop on Top Quark Physics* (*Preprint* 1612.00440)
[8] Cranmer K, Lewis G, Moneta L, Shibata A and Verkerke W (ROOT) 2012 HistFactory: A tool for creating statistical models for use with RooFit and RooStats Tech. rep. New York U. New York URL https://cds.cern.ch/record/1456844
[9] Cowan G, Cranmer K, Gross E and Vitells O 2011 *The European Physical Journal C* **71** URL https://doi.org/10.1140/epjc/s10052-011-1554-0

Extracting the top quark Yukawa coupling from $t\bar{t}$ kinematic distributions in the dilepton final state

Cameron Garvey, James Keaveney

Department of Physics, University of Cape Town, Rondebosch 7701, South Africa

E-mail: cameron.michael.garvey@cern.ch

Abstract. The coupling of the top quark to the Higgs boson is one of the free parameters of the SM and provides a unique window into the electroweak sector of the SM. This analysis aims to extract the top quark Yukawa coupling (Y_t) from the differential cross sections of top quark pair production in the dilepton final state. The data used in this measurement consist of proton-proton collisions at a centre-of-mass energy of 13 TeV recorded using the ATLAS experiment corresponding to 140 fb^{-1}. An event selection was put in place to preferentially select $t\bar{t}$ events with a dilepton final state specifically the decay channels ee,eμ and $\mu\mu$. To simulate variations in Y_t, electroweak corrections were computed which include diagrams involving the exchange of a Higgs boson between the top quarks. Kinematic distributions sensitive to variations of Y_t were constructed with the $m_{\ell\ell bb}$ distribution providing the greatest sensitivity. The extraction was implemented using a binned profile likelihood fit was implemented using template morphing. A reduced set of samples and uncertainties are used with the extracted $Y_t^2 = 1.00^{+0.21}_{0.20}$. A signal injection test was then implemented confirming that the extraction can reliably measure Y_t with minimal bias.

1. Introduction

In the standard model (SM) there are 26 free parameters, of which a large proportion of these parameters relate to the mass of the fundamental fermions (m_f). The mass of these fermions is obtained through the spontaneous symmetry breaking of the Higgs potential [3] and is related to the strength of their coupling to the Higgs field. The coupling is known as the Yukawa coupling (y_f) and is given by equation 1,

$$y_f = \sqrt{2}\frac{m_f}{\nu} \tag{1}$$

with ν corresponding to the vacuum expectation value of the Higgs field. The most massive particle in the SM is the top quark which provides access to the largest Yukawa coupling. The Yukawa coupling of the top quark is expected to be close to unity. We define the parameter $Y_t = y_t/y_t^{SM}$ being the ratio of the measured y_t to the SM expectation. There are several methods which can be used to measure the Y_t. The first method measures the cross section of a top quark pair produced in association with a Higgs boson (ttH) as the rate of production is sensitive to variations in the Yukawa coupling. The second approach investigates the effects of a virtual Higgs boson exchange between a top quark and a top anti-quark and will be the focus of this paper. The exchange of a virtual Higgs affects the kinematics of the top quark pair allowing us to measure Y_t. Two measurements of Y_t have been conducted by CMS in the

lepton+jets [4] and dilepton [5] decay channels of $t\bar{t}$. There are two ongoing analyses within the ATLAS collaboration which aim to measure the Y_t in the lepton+jets and dilepton channels. In this paper I will be presenting the status of the extraction of Y_t from the $t\bar{t}$ differential cross sections in the dilepton final state.

2. Event selection

The data used consists of proton-proton collisions at a centre-of-mass energy of 13 TeV. The data was recorded by the ATLAS experiment during Run 2 of the LHC and is referred to as the full Run 2 data. The full Run 2 data set corresponds to 140 fb^{-1} [8]. At the LHC, top quark pairs are produced via the gg or $q\bar{q}$ initial states at leading order (LO) where the gg production is the dominant initial state. Top quarks will decay to a W boson and a b quark approximately 100% of the time. The dilepton decay channel is shown in equation 2.

$$t\bar{t} \rightarrow bW^+\bar{b}W^- \rightarrow b\ell^+\nu_\ell\bar{b}\ell^-\nu_\ell \tag{2}$$

A reduced set of simulated samples is used which correspond to the signal sample $t\bar{t}$ and the background sample where a top quark is produced in association with a W boson. The Z+jets sample is also expected to contribute and will be included in the next phase of this analysis. A set of selection criteria has been implemented to preferentially select for dilepton $t\bar{t}$ events based on the kinematics described in equation 2. Two leptons were selected which could be either electrons or muons, each with a transverse momentum (p_T) of greater the 25 GeV. The regions of interest are then split into ee, eμ, and $\mu\mu$ channels. A jet multiplicity of two or more is required where each jet is required to have a p_T greater than 20 GeV. Furthermore, two of the jets are required to be tagged as a b-jet using the DL1r b-tagger [6]. The Z+jets process is also expected to contribute in the opposite sign same flavour regions, specifically ee and $\mu\mu$. As such there are requirements in place to veto contributions from Z processes. It is required that the mass of the two leptons ($m_{\ell\ell}$) is not within 10 GeV of the mass of the Z boson as well as the missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$) be greater than 30 GeV. Finally, a second requirement on the dilepton mass is put in place to minimise low energy resonances and as such we require $m_{\ell\ell} \geq 50$ GeV.

3. Electroweak corrections

The $t\bar{t}$ simulation described in section 2 is implemented at next-to leading order (NLO) in QCD. The electroweak (EW) corrections are generally not included in the simulation of $t\bar{t}$ as they have a minimal effect on the $t\bar{t}$ cross section[9]. The effects of these corrections can have a substantial impact on the kinematic distributions of $t\bar{t}$ and need to be accounted for within the simulation. The calculation of the EW corrections to $t\bar{t}$ production are implemented using `Hathor` [2] and follow from the calculations described in Ref. [1]. To obtain the corrections first a double differential cross section is computed which include the LO QCD diagrams as well as certain EW diagrams of order $\alpha_S^2\alpha_{weak}$. The calculation is done in terms of observables that are sensitive to variations in the Y_t specifically the mass of the top quark pair (m_{tt}) and the azimuthal angle of the top quark with respect to the beam direction in the $t\bar{t}$ rest frame ($\cos(\theta^*)$). The EW diagrams which are computed involve the exchange of Z boson, photon or Higgs bosons between the top quark and top anti-quark such as shown in Figure 3.

Taking the ratio of the calculation with the EW diagrams to the LO QCD provides the multiplicative weights $\omega(m_{tt},\cos(\theta^*))$ for each bin of the differential distribution. To apply ω one needs access to m_{tt} and $\cos(\theta^*)$ which are calculated using the kinematics of the top quarks. For this purpose it is important to note that there are several steps in the simulation process but for this case we will use information that corresponds to particles prior to simulating their

(a) **(b)**

Figure 1: Example Feynman diagrams for (a) gg and (b) $q\bar{q}$ initial states in $t\bar{t}$ production with virtual Higgs-boson exchange.

interaction with the detector. This information is called truth level information and provides information regarding the top quark kinematics. Calculating the m_{tt} and $\cos(\theta^*)$ using truth level information on a per-event basis allows us to apply the EW correction to the NLO calculations of $t\bar{t}$. The EW correction can then be calculated for several different values of Y_t specifically. The upper plot of Figure 2 shows the m_{tt} distribution for $t\bar{t}$ at truth level for several different values of Y_t. The y-axis shows the number of entries at truth level. A ratio of each of the histograms to the histogram corresponding to the SM prediction of Y_t =1 is shown in the lower plot. This is shown on a bin-by-bin basis. The greater the ratio varies from a value of 1, indicates a greater sensitivity to variations in Y_t.

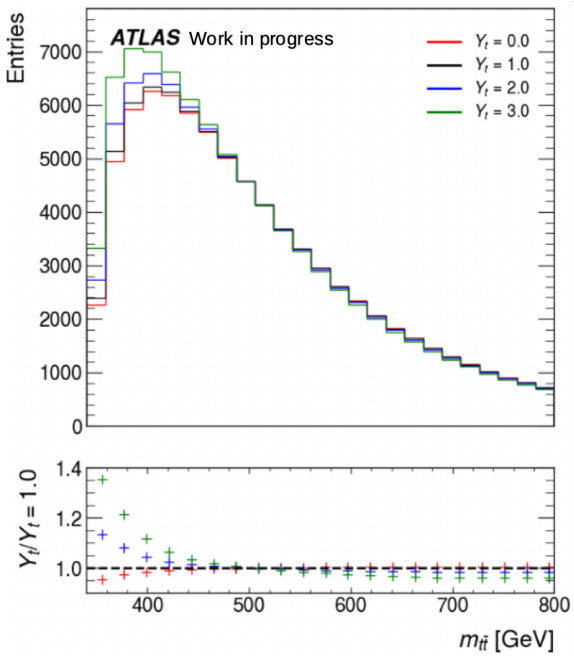
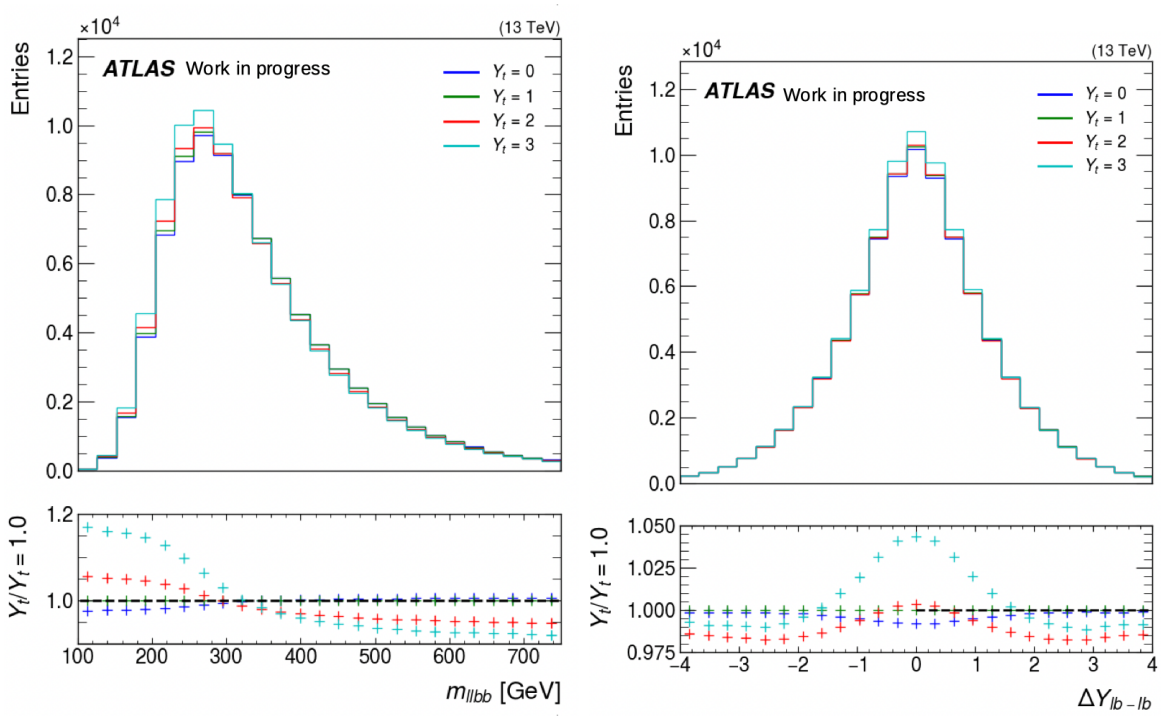


Figure 2: A histogram showing the m_{tt} for the NLO simulation of $t\bar{t}$ at truth level is shown with the multiplicative weights applied for several different values of of Y_t. A ratio of each of the histograms to the histogram corresponding to the SM prediction of Y_t =1 is shown in the lower plot.

From the ratio plot in Figure 2 it can be seen that the largest sensitivity to variations in Y_t can be found at masses around twice the mass of the top quark.

4. Construction of kinematic distributions

To measure the Y_t effectively one needs to construct variables that are sensitive to the EW corrections as described in the previous section. Ideally one would use the observables m_{tt} and $\cos(\theta^*)$. To do so, one would need to reconstruct the kinematics of the top quark as the neutrinos are not measured by the detector. With this in mind observables will be constructed using the measured decay products of the top quarks such as the leptons and b-jets. Proxy variables are created based on the m_{tt} and $\cos(\theta^*)$ which are the $m_{\ell\ell bb}$ and the $\Delta Y_{\ell b\ell b}$, respectively. The $m_{\ell\ell bb}$ is constructed using the combined mass of the two leptons and the two b-jets. The $\Delta Y_{\ell b\ell b}$ is constructed by taking the difference in rapidity between the two lepton b-jet systems where each system is identified by matching the leptons to their closest b-jet in terms of angular distance. Figure 3 shows the $m_{\ell\ell bb}$ and $\Delta Y_{\ell b\ell b}$ distributions for $t\bar{t}$ at detector level for several values of Y_t. The y-axis shows the number of expected events for $t\bar{t}$ at detector level and the ratio plot is as described for Figure 2.



(a) **(b)**

Figure 3: A set of histograms showing the effects of the EW corrections on the (a) mass of the leptons and b-jets and (b) the difference in rapidity between the lepton b-jet systems. This done for $t\bar{t}$ at detector level for several values of Y_t. A ratio of each of the histograms to the histogram corresponding to the SM prediction of Y_t =1 is shown in the lower plot.

Referring to the ratio plots shown in Figures 3a and 3b it can be seen that the greatest variation can be found in the low mass region of the $m_{\ell\ell bb}$ distribution. The larger variations indicates that the $m_{\ell\ell bb}$ distribution should provide enhanced sensitivity to variations in the Y_t over other less sensitive distributions such as the $\Delta Y_{\ell\ell bb}$. It should be noted that the list of variables tested has not been exhaustive and there could be room for improvement. A potential next step would be to reconstruct of the top-quark kinematics to access variables such as m_{tt} and $\cos(\theta^*)$ gain a sensitivity closer to what is found in Figure 2.

5. Results

Due to the early stage of this research the results use a reduced setup and as such do not include a comprehensive set of backgrounds or systematic uncertainties. The extraction is utilises the full run 2 data set, corresponding to 140 fb^{-1}, as measured by the ATLAS experiment but only includes one background, namely tW. The set of uncertainties considered are the statistical uncertainties as well as the uncertainties on the cross sections of $t\bar{t}$ and tW processes. The results shown in this paper are fully blinded and the fit is performed using Asimov data [7]

The extraction of the Y_t is implemented using a binned profile likelihood fit. The specific method employed to extract the Y_t is template morphing which extracts the parameter of interest using the shape of the associated distribution. We will be using the $m_{\ell\ell bb}$ distribution as it is the distribution found with the greatest sensitivity to variations in the Y_t. We then vary the distribution to several values of Y_t by applying the EW corrections for Y_t of 0,1,2 and 3 to create templates as shown in Figure 3. A linear interpolation is then applied between each of the bins in the four templates. Due to the linear interpolation the results extracted are in terms of Y_t^2. The results of the extraction are shown in equation 3.

$$Y_t^2 = 1.00^{+0.21}_{-0.20} \tag{3}$$

The extracted value of$Y_t = 1$ is expected since we are using Asimov data. A signal injection test was implemented to ensure that the extraction method contains minimal biases and is behaving in an expected manner. A custom Asimov data set is created for this purpose and is described as follows. The EW corrections are applied to the NLO $t\bar{t}$ simulation corresponding to a specific value of Y_t and then combined with the background sample. This is then referred to as the custom Asimov data set which will be used to validate the extraction method. The simulation is then fit to the custom Asimov data set and Y_t is extracted. If the extraction is unbiased the extracted Y_t by the fit should correspond exactly to the Y_t injected into the custom Asimov data set. This setup is run for several different values of Y_t where the results of this test are shown in Figure 4. The x and y axes show the injected Y_t^2 and the measured Y_t^2, respectively. The ideal situation is when the injected and measured values of Y_t^2 match exactly and is given by the red dotted line. In Figure 4, there is good agreement between the extracted values of Y_t^2 with the red line and provides that the extraction can reliably measure Y_t with minimal bias.

6. Conclusion

The coupling of the top quark to the Higgs boson is one of the free parameters of the SM and provides a unique window into the electroweak sector of the SM. This analysis aims to extract the top quark Yukawa coupling from the differential cross sections of top quark pair production in the dilepton final state. An event selection was put in place to preferentially select for dilepton $t\bar{t}$ with the analysis containing three regions of interest namely ee,eμ and $\mu\mu$. The EW corrections were then computed by including diagrams that involve the exchange of a Higgs boson between the top quarks. The EW corrections resulted in a multiplicative event weight ω applied to the NLO $t\bar{t}$ simulation. To measure the Y_t effectively one needs to construct variables that are sensitive to the EW corrections. Several distributions were tested for this purpose with the $m_{\ell\ell bb}$ distribution providing the greatest sensitivity. A binned profile likelihood fit was implemented using template morphing to extract the Y_t. The results use a reduced set of samples and uncertainties with the extracted $Y_t^2 = 1.00^{+0.21}_{-0.20}$. Looking forward we plan to include a more comprehensive set of systematic uncertainties. A signal injection test was implemented verifying that the fit procedure is behaving as expected and has negligible bias.

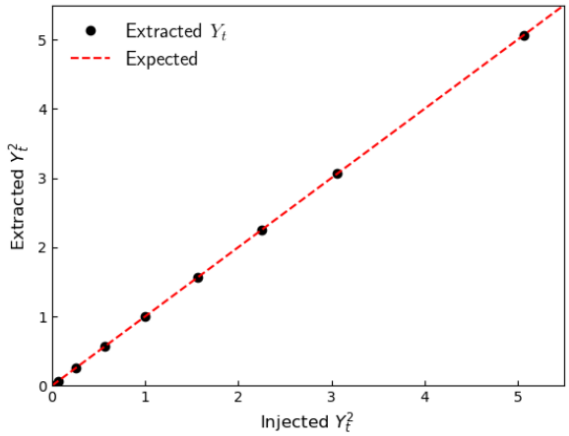


Figure 4: A figure showing the signal injection test of the Y_t extraction. The x and y axes show the injected Y_t^2 and the measured Y_t^2, respectively. The red dotted line indicates the idealised situation where the injected and measured values are equal.

References

[1] Kühn J H, Scharf A and Uwer P 2015 Weak interactions in top-quark pair production at hadron colliders: An update *Phys.Rev.* D**91** 014020
[2] Aliev M, Lacker H, Langenfeld U, Moch S, Uwer P and Wiedermann M 2011 HATHOR - HAdronic Top and Heavy quarks crOss section calculatoR *Comput.Phys.Commun.* **182** 1034-46
[3] Weinberg S 1967 A model of leptons *Phys. Rev. Lett.* **19** 1264
[4] CMS collaboration 2019 Measurement of the top quark Yukawa coupling from $t\bar{t}$ kinematic distributions in the leptonjets final state in proton-proton collisions at $\sqrt{s} = 13TeV$ *Phys. Rev.* D **100**, 072007
[5] CMS collaboration 2020 Measurement of the top quark Yukawa coupling from $t\bar{t}$ kinematic distributions in the dilepton final state in proton-proton collisions at $\sqrt{s}$=13 TeV *Phys. Rev.* D **102** 092013
[6] ATLAS collaboration 2023 ATLAS flavour-tagging algorithms for the LHC Run 2 pp collision dataset *Eur. Phys. J.* C **83** (2023)681
[7] Cowan, G et al. 2011 Asymptotic formulae for likelihood-based tests of new physics *Eur. Phys. J.* C **71** 1554
[8] ATLAS collaboration 2022 Luminosity determination in *pp* collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector at the LHC *Preprint* 2212.09379
[9] Beenakker W et al. 1994 Electroweak one-loop contributions to top pair production in hadron colliders *Nucl.Phys.* B **411** 343-80

Measurements of neutron response functions for a BC-501A liquid organic scintillator between 6.0 MeV and 63.5 MeV at the iThemba LABS fast neutron beam facility

Kutullo Maibane[1], Andy Buffler[1], Tanya Hutton[1], Peane Maleka[2], Zina Ndabeni[1,2]

[1] Metrological and Applied Sciences University Research Unit (MeASURe),
Department of Physics, University of Cape Town, Cape Town, South Africa, 7700.
[2] iThemba LABS, P.O. Box 722, 7129, Somerset West, South Africa.

E-mail: **kutullo.maibane@uct.ac.za**

Abstract. We report on measurements of neutron response functions used for spectrum unfolding in the energy range 6.0 MeV to 63.5 MeV for a BC-501A detector, derived from time-of-flight analyses for the neutron fields produced at the iThemba LABS fast neutron beam facility. The measured detector response functions were validated through the unfolding of the measured pulse height spectrum with MAXED for the neutron field produced by a 66 MeV proton beam irradiating an 8.0 mm Li target, and directly comparing the unfolded neutron energy spectra to those derived from time-of-flight analyses. The demonstrated method will be used to develop a portable scintillator-based neutron detector system for measurements of spectral fluence up to 200 MeV outside of the laboratory environment where time-of-flight is unavailable, whilst maintaining the required standards and traceability.

1. Introduction

The fast neutron beam facility at iThemba LABS (D-line) offers ns-pulsed neutron fields ranging between 5 MeV and 200 MeV, typically produced via the ^{7}Li(p,Xn) or ^{9}Be(p,Xn) reactions [1-3], using proton beams from the separated sector cyclotron (SSC). The D-line is currently undergoing a major upgrade and redevelopment towards ISO/IEC 17025 accreditation, including physical modifications of the vault, new instrumentation for neutron metrology and improved beam monitoring and control systems. At present, the standard procedure for characterising the neutron beams and measuring spectral neutron fluence involves a 2” x 4” BC-501A liquid scintillator and a parallel-plate ^{238}U fission chamber (FC) [2,4]. Neutron energy spectra may either be determined directly from time-of-flight (ToF) measurements using the BC-501A detector, or by unfolding methods if the detector response functions are known sufficiently well over the energy range of interest. The quality of the unfolded neutron energy spectrum is directly related to the knowledge of the detector response to mono-energetic neutrons. For neutron energies below 20 MeV, a detector response matrix can reliably be produced using Monte Carlo radiation transport codes. However, above 20 MeV there are insufficient measurements of cross sections associated with the many possible neutron interactions with carbon to reliably simulate the response functions, and the detector response matrix can only be measured directly. We report on the measurement and validation of neutron response functions used for spectrum unfolding in the energy range 6.0 MeV to 63.5 MeV for a BC-501A detector at the iThemba LABS fast neutron beam facility.

2. Experimental setup

2.1. The iThemba LABS fast neutron facility

The SSC of the iThemba LABS accelerates protons in the energy range from 25 MeV to 200 MeV, and has a dedicated beam-line for producing fast neutron fields. A beam pulse selector can suppress a chosen fraction of proton bunches to enlarge the time interval between pulses, which allows time-of-flight measurements to be carried out. The maximum beam current for 66 MeV is about 1.2 µA in unselected mode.

Figure 1 shows the D-line vault. In the present work, a pulsed proton beam accelerated to 66 MeV at iThemba LABS SSC was directed toward an 8.0 mm natural lithium target situated in the target holder of the D-line vault, as illustrated in Figure 1, producing a neutron field having maximum energy 63.5 MeV via the ^{7}Li(p,n)^{7}Be reaction. Neutrons are collimated to the measurement area through 5 cm x 5 cm square apertures at 0^{o} and 16^{o} relative to the incoming proton beam, with the detectors typically placed 8.00 m away from the target.

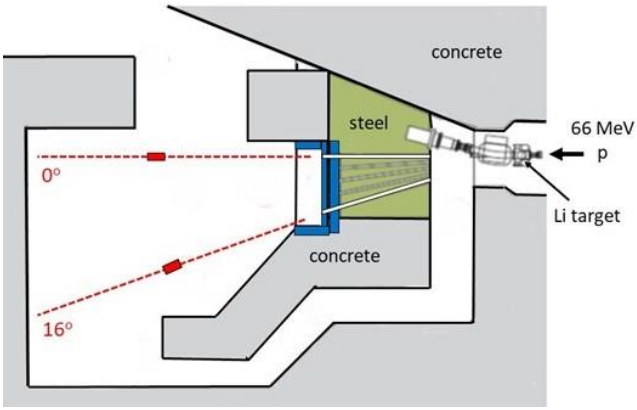


Figure 1. The D-line neutron beam facility at iThemba LABS.

2.2. The BC-501A liquid organic scintillation detector

ToF measurements of the neutron field were carried out using a BC-501A liquid organic scintillator detector (Figure 2), of dimensions 2” diameter x 4” length, coupled to gain-stabilised photomultiplier. The detector offers, excellent neutron-gamma discrimination and 2 ns time-resolution. Events were recorded using the standard NIM-based data acquisition for the facility (Figure 3).

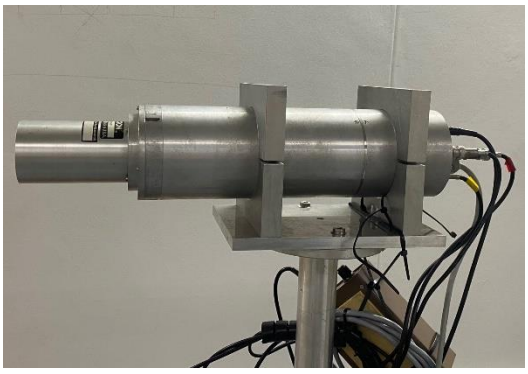


Figure 2. The BC-501A detector.

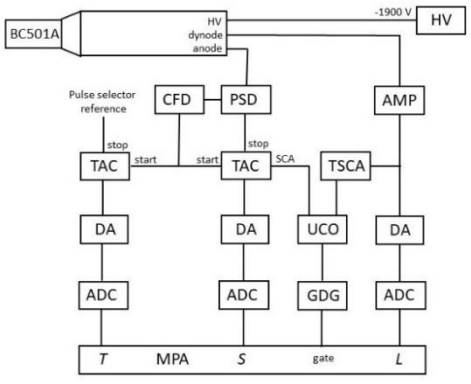


Figure 3. The NIM-based pulse processing system for the BC-501A detector.

2.3. Pulse shape discrimination

The ^{7}Li(p,n)^{7}Be reaction produces neutrons as well as gamma rays that interact in the scintillator detector. In order to separate the measured neutron induced events from the gamma ray induced events, pulse shape discrimination is implemented via the zero-crossover method. In general, anode pulses associated with events induced by gamma rays have a shorter cross-over time than the neutron induced events, and thus allowing for the discrimination.

The fast anode output of the BC-501A detector was directed through a PSD module, where the pulse was integrated and differentiated in a preamplifier to produce a bipolar shape pulse. The Constant Fraction Discriminator (CFD) output from the timing pulse was used to start a Time-to-Amplitude Converter (TAC) in the pulse shaping circuit, the TAC was stopped by a signal from the PSD. The TAC logic output produced the pulse shape parameter *S*. The dynode signal was directed through a preamplifier and a delay line amplifier (DLA) which produced the pulse height parameter *L* [7]. Data are recorded in coincidence in list mode using the standard NIM-based acquisition system (Figure 3).

Figure 4 shows event density as a function of light output parameter *L* and pulse shape *S* for a typical measurement at 0° for 66 MeV protons incident on a natural Li target, with the *L* parameter scaled to MeV electron-equivalent (MeV$_{ee}$) using gamma ray sources. The loci associated with recoiling electrons (e) induced by gamma ray interactions, and the recoiling protons (p), deuterons (d), tritons (t) and alphas (α) particles induced by neutrons interactions are shown. Neutron events are separated from the gamma ray events using software cut, and a lower *L* threshold applied as indicated by the dashed lines.

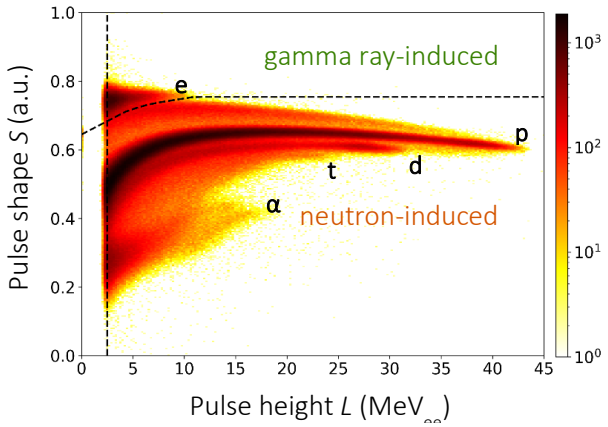


Figure 4. Event density as a function of pulse height parameter *L* and pulse shape parameter *S* measured with a BC-501A scintillator detector at iThemba LABS for 66 MeV protons incident on a natural Li target. Gamma ray (e), and neutron induced events (p, d, t, α) were separated as indicated by the dashed line.

3. Results and Analysis

3.1. Neutron Spectroscopy: Time-of-Flight

The SSC produces a ns-pulsed proton beam, which allows the neutron energy to be determined using time-of-flight method. When protons are incident on the Li target neutrons and gamma rays are produced, and the time the neutrons take to reach the detector is directly related to their energy. The neutron energy E_n can be calculated using:

$$E_n = mc^2 \left[\frac{1}{\sqrt{1-(d/cT)^2}} - 1 \right] \quad (1)$$

where the rest mass of the neutron is represented by *m*, the target-detector distance by *d*, the speed of light by *c*, *T* is the arrival time of the neutron (see Figure 3).

Figure 5 shows the ToF spectra before and after the pulse shape discrimination cut (Figure 4) was applied to veto counts due to gamma ray events in the scintillator. The ToF spectrum for neutrons was used to produce the neutron energy spectrum (Figure 6) using standard relativistic method of neutron time-of-flight (Equation 1). A cut is placed at L = 2.5 MeV_{ee} on the pulse height (see the vertical line in Figure 4), as the BC-501A detector is calibrated for efficiency with a threshold of 2.5 MeV_{ee} on the pulse height parameter L.

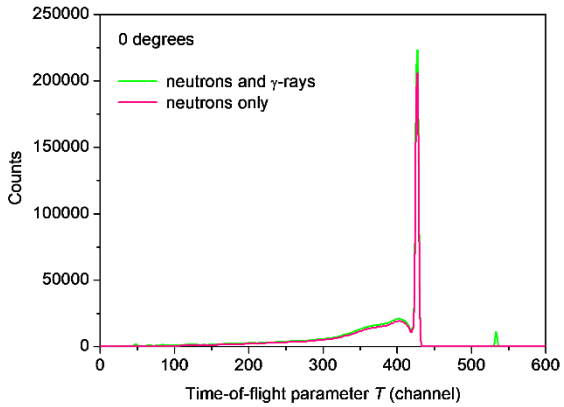


Figure 5. Time-of-flight spectra of neutrons and gamma rays produced at 0° for a 66 MeV proton beam irradiating an 8.0 mm thick lithium target measured with a BC-501A detector.

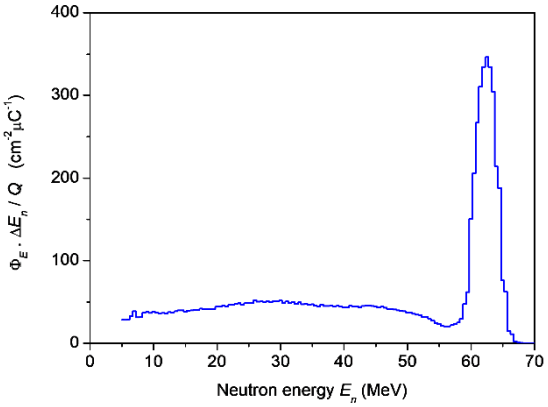


Figure 6. Neutron energy spectrum measured via time-of-flight with a BC-501A detector at 0° for a 66 MeV proton beam irradiating an 8.0 mm thick lithium target, corrected for neutron detection efficiency.

3.2. Neutron Spectroscopy: Unfolding

The neutron energy spectrum was also determined via an unfolding method using the GRAVEL and MAXED codes which are part of the UMG package [5]. The measured light output spectrum z_i is related to the neutron energy spectrum ϕ_j as

$$z_i = \sum_{j=1}^{m} R_{ij}\phi_j. \qquad (2)$$

where R_{ij} is the mono-energetic detector response function for energy j in channel i. The detector response matrix was measured for the BC-501A detector by selecting narrow cuts on the time-of-flight spectrum associated with energies between 6.0 MeV and 63.5 MeV in steps of 2.0 MeV and determining the pulse height spectrum associated with that range. The nearly mono-energetic response functions were each normalised to an area of unity. A subset of detector response functions are shown in Figure 7. This measured response matrix was then used to unfold the measured pulse height spectra with MAXED for 66 MeV protons incident on an 8.0 mm thick Li target at 0° and 16° for the BC-501A detector [6].

Figure 8 shows the measured pulse height spectrum acquired from the detector next to the refolded fit spectrum acquired using the MAXED. Figure 10 shows the solution neutron energy spectra, scaled to beam current and efficiency. Figure 9 shows the measured (ToF and unfolded) neutron energy spectra, the agreement is excellent, with the chi square of 2.1. Uncertainties associated with systematic effects are generally not considered in these unfolding analysis since they are negligible to those associated with counting statistics.

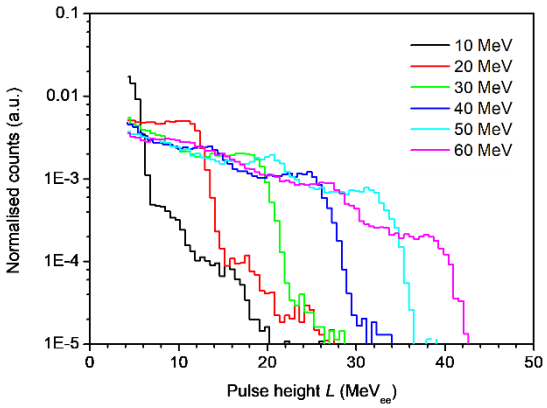


Figure 7. A selection of the response functions over a full energy range (6.0 MeV to 63.5 MeV) for every 2 MeV measured at 0°.

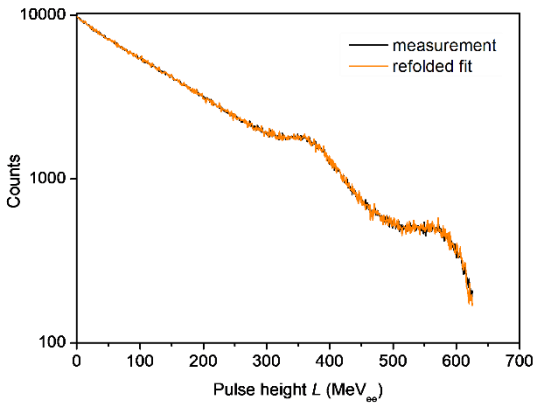


Figure 8. The measured and refolded fit pulse height spectra of neutrons produced at 0° for a 66 MeV proton beam irradiating an 8.0 mm thick lithium target.

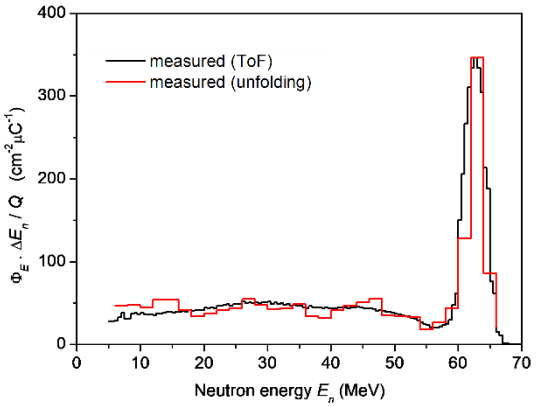


Figure 9. Neutron energy spectra measured (ToF and unfolded) at 0° and adjusted for both the detector efficiency and beam current.

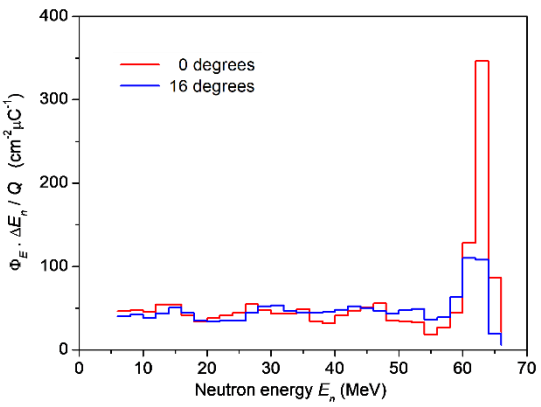


Figure 10. Neutron energy spectra measured via unfolding analyses at (0° and 16°) and adjusted for both the detector efficiency and beam current.

By applying the same scaling factor to the measured and refolded fit pulse height spectra, the resultant spectra can be seen in Figure 8, and there is excellent agreement between the two pulse height spectra. Figure 9 compares the observed neutron energy spectrum with the ToF neutron energy spectrum after the measured pulse height spectrum for neutrons was unfolded using the measured response matrix. The unfolding results are promising, and the unfolded neutron energy spectra are of good quality, with most important features being replicated. However, there are some inconsistencies in the low energy region of the continuum of the two neutron energy spectra, due to the finer energy binning in the ToF neutron energy spectrum (0.5 MeV binning) than the unfolded neutron energy spectrum (2 MeV binning).

4. Discussion and Conclusion

Detector response functions were measured at the iThemba LABS high energy neutron facility using time-of-flight between 6.0 MeV and 63.5 MeV for a BC-501A detector. Good agreement was observed between the neutron energy spectra derived from time-of-flight and unfolding analyses for neutron fields produced by 66 MeV protons incident on a Li target. This supports the methodology employed for

spectrum unfolding and validates the measured response matrix. Unfolding methods and measured response functions enable the use of the BC-501A detector for neutron spectrometry in unknown fields outside of the laboratory, like those present in space and aircraft.

Further detector characterisations are planned for neutron fields up to 200 MeV, with improved statistics to extend the possible range of applications. The minimum energy bin width in the detector response functions is determined by the resolution of the acquired time-of-flight spectrum, as such, further measurements are planned with an increased number of acquisition channels to produce neutron response functions every 0.5 MeV. The process demonstrated here will be used to develop a portable, scintillator-based neutron detector system for measurements of spectral fluence up to 200 MeV outside of the laboratory environment where time-of-flight is unavailable, whilst maintaining the required standards and traceability.

Acknowledgements
This work is based on research that was partially funded by the South African National Research Foundation (Grant Number 138069).

References
[1] Pomp S et al., *EURADOS*, Report 2013-02, May 2013.
[2] Mosconi M et al., *Radiation Measurements.*, **vol 45**, pp. 1342-1345, 2010.
[3] McMurray W R et al., *Nucl. Instrum. Methods Phys. Res. Sect A*, **vol 329**, pp. 217-222, 1993.
[4] Ndlovu N B et al., *JACoW*, Cyclotrons 2019, TUP012, pp. 182 – 185, 2019.
[5] Reginatto M et al., *Nucl. Instrum. Methods Phys. Res. Sect A.*, **vol 476**, pp. 242-246, 2002
[6] Hutton T et al., Measurements of Response Functions. *CYC2022 Proceedings*, 2022.
[7] Musonza E, Measurement of Fluence., MSc Thesis, *University of Cape Town*, 2011

Searches for additional scalars at future e^+e^- colliders

Karabo Mosala[1], Thuso Mathaha[1], Mukesh Kumar[1], Bruce Mellado[1,2]

[1] School of Physics and Institute for Collider Particle Physics, University of the Witwatersrand, 1 Jan Smuts Ave, Braamfontein, Johannesburg, 2000
[2] iThemba LABS, National Research Foundation, PO Box 722, Somerset West 7129, South Africa

E-mail: Karabo.mosala@cern.ch

Abstract. We present growing excesses consistent with a 95 GeV scalar. We provide a comprehensive analysis of the Two Higgs Doublet Model and an additional singlet (2HDM+S) at future e^+e^- collider. In particular, we provide a precise mass reconstruction measurement for the scalar, m_S, using the recoil mass method through $e^+e^- \to ZS$ where $Z \to \mu^+\mu^-$ and $S \to b\bar{b}$ at $\sqrt{s} = 250$ GeV and $\sqrt{s} = 200$ GeV. Furthermore, we employ Deep Neural Network to analyze the properties and behaviour of the scalar particle with a mass most importantly to provide enhanced resolution for the separation between beyond the Standard Model (SM) signal and SM background in the region $95 - 96$ GeV in the $S \to b\bar{b}$ for $\mu^+\mu^-$ channel. A 95 GeV scalar can be observed with 5σ significance at 15(10) fb^{-1} integrated luminosity for $\sqrt{s} = 250(200)$ GeV. This strengthens the discovery of the potential of the future e^+e^- collider.

1. Introduction

The proposal for the development of e^+e^- collider provides a window of opportunity for the field of particle physics. This colliders, equipped with advanced detectors, will provide valuable data for these studies. In previous works [1, 2], the introduction of scalars H and S in an effective model was proposed to explain various features observed in the collider data. These features includes distortions in the Higgs boson (h) elevated associated jet activity, transverse momentum spectrum, and elevated lapton rates in association with b-tagged jets. Additionally, the search for double Higgs boson and weak boson production has also been explored.

Extensions of the standard model (SM) include the 2HDM, which introduce an additional Higgs doublet. This leads to the presence of two CP-even (h, H) scalar bosons, one CP-odd (A) scalar boson, and charged scalar bosons $(H^\pm)$. While 2HDM have been extensively studied in the literature, it has been pointed out that a 2HDM alone cannot explain all the observed features in the data. Therefore, a scalar singlet S has been introduced in conjunction with the 2HDM, resulting in the 2HDM+S model. This model also provides a framework to explore scenarios involving dark matter [3].

In this study, we build upon the phenomenology described in previous works and investigate the 2HDM+S parameter space that can accommodate the observed anormalies in the data [1]. Specifically, we focus on the production of the singlet scalar (S) through the e^+e^- production mode, with $S \to b\bar{b}$ channel.

The CMS experiment at the Large Hadron Collider (LHC) reported a local excess of 2.9σ at 95.4 GeV in the di-photon invariant mass spectrum using Run 2 data [4, 5]. The ATLAS

experiment has recently reported a much smaller excess of 1.7σ at the same mass with the same data set [6]. A local excess of 2.8σ at 95.3 GeV was reported by the CMS experiment in the $\tau^+\tau^-$ final states [7].

2. Model and Framework

A Two Higgs doublet model (`2HDM`) and the addition of a real singlet Φ_S forms the baseline for the `2HDM+S` formalism. Within `2HDM+S` framework, the potential is given by:

$$\begin{aligned} V(\Phi_1,\Phi_2,\Phi_S) = {} & m_{11}^2|\Phi_1|^2 + m_{22}^2|\Phi_2|^2 - m_{12}^2(\Phi_1^\dagger\Phi_2 + h.c.) + \frac{\lambda_1}{2}(\Phi_1^\dagger\Phi_1) + \frac{\lambda_2}{2}(\Phi_2^\dagger\Phi_2) \\ & + \lambda_3(\Phi_1^\dagger\Phi_1)(\Phi_2^\dagger\Phi_2) + \lambda_4(\Phi_1^\dagger\Phi_2)(\Phi_2^\dagger\Phi_1) + \frac{\lambda_5}{2}[(\Phi_1^\dagger\Phi_2)^2 + h.c.] \\ & + \frac{1}{2}m_S^2\Phi_S^2 + \frac{\lambda_6}{8}\Phi_S^4 + \frac{\lambda_7}{2}(\Phi_1^\dagger\Phi_1)\Phi_S^2 + \frac{\lambda_8}{2}(\Phi_2^\dagger\Phi_2)\Phi_S^2. \end{aligned} \tag{1}$$

where Φ_1 and Φ_2 represent $SU(2)_L$ Higgs doublets. The first two lines pertain to the 2HDM potential, while the final line introduces input from the singlet field Φ_S. After electroweak symmetry breaking, the two doublet fields acquire the real VEVs v_1 and v_2 and the singlet field a real VEV v_S and they can be parametrised as

$$\Phi_1 = \begin{pmatrix} \phi_1^+ \\ \frac{1}{\sqrt{2}}(v_1 + \rho_1 + i\eta_1) \end{pmatrix}, \Phi_2 = \begin{pmatrix} \phi_2^+ \\ \frac{1}{\sqrt{2}}(v_2 + \rho_2 + i\eta_2) \end{pmatrix}, \Phi_S = v_S + \rho_S, \tag{2}$$

given in terms of a real neutral CP-add and CP-even fields $\rho_I (I = 1, 2, S)$ and η_i and complex charged fields $\phi_i^+, (i = 1, 2)$, respectively. Within the `2HDM+S` model, we consider the following set of input parameters

$$m_{H_{1,2,3}},\ m_{1,2}^2,\ m_{H^\pm},\ m_A,\ \alpha_1,\ \alpha_2,\ \alpha_3,\ t_\beta,\ v_S,\ v. \tag{3}$$

We refer the reader to Ref.[8, 9] for further details.

3. Analysis, Observable and results

3.1. Discovery potential of CEPC

To assess the capability of the future Circular Electron-Positron Collider(CEPC) in terms of precision measurements related to the Higgs boson, an extensive set of benchmark studies in Higgs physics is underway. These studies encompass detector simulations that incorporate various factors, such as pile-up resulting from background processes like $\gamma\gamma \rightarrow$ hadron interactions, as well as the pertinent background processes from e^+e^- collisions. The simulations are designed to consider the unique CEPC beam spectrum and account for effects like initial state radiation that are relevant to the analysis.

The analysis of events originating from the Higgsstrahlung process $e^+e^- \rightarrow ZS$ serves as a significant method for precision measurements within the context of Higgs physics. This process, dominant at center-of-mass energies below $\sqrt{s} = 250$ GeV, provides a platform for two crucial measurements. Firstly, it enables the precise determination of the coupling between the Higgs boson and the Z boson. Secondly, it facilitates the measurement of the interactions between the Higgs boson and its final decay products.

Through this approach, an unbiased assessment of the coupling g_{HZZ} is carried out, independent of specific models. This involves reconstructing the recoil mass of the Z boson, contributing to the precision of the measurement [10]

$$\mathrm{M_{recoil}} = \sqrt{s + M_{\mu^+\mu^-}^2 - 2(E_{\mu^+} + E_{\mu^-})\sqrt{s}}, \tag{4}$$

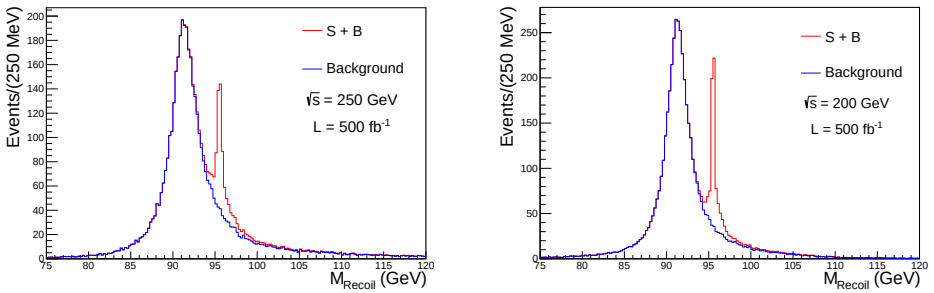


Figure 1. Recoil mass distribution for $e^+e^- \to SZ \to S\mu^+\mu^-$ events with $m_S = 95$ GeV at integrated luminosity $\mathcal{L} = 500$ fb^{-1} for $\sqrt{S} = 250$ GeV (left) and $\sqrt{S} = 200$ GeV (right). Red is S+B where S takes into consideration 10% SM Higgs signal, B is the SM background and Blue is the SM background.

with no reconstruction of the Higgs necessary. $\sqrt{s}$ is the centre of mass energy, $M^2_{\mu^+\mu^-}$, $E_{\mu^+} + E_{\mu^-}$ is the invariant mass and energy of muons, respectively.

To generate SM backgrounds and BSM signals, we specify $\sqrt{s}$, beam polarization (e^+ and e^- beams are 80% left polarized and 80% right polarized, respectively), and $\mathcal{L}$, as well as considering 95 GeV scalar-production modes and their corresponding cross sections for $E_{e^+} = 100/125$ GeV and $E_{e^-} = 100/125$ GeV. To optimise the signal events over the leading backgrounds minimal cuts on leading and sub-leading jets, $j\ j$ and leading and sub-leading muons, $\mu^+\mu^-$ and leading and sub-leading b-tagged jets, $b\bar{b}$ are applied for all processes in this study with appropriate additional cuts; $E_{j,b,\mu} > 5$ GeV, $\mathrm{M}^{\mathrm{Recoil}}_{\mu+\mu-} < 120$ GeV, $\mathrm{M}_{b\bar{b}} < 100$ GeV.

3.2. Signal and Background Discrimination

The event selection criteria described in Section 3.1 eliminate most of the background events coming from different SM physics processes. However, the number of remaining background events after this selection may still have a substantial dilution effect. Therefore, a second selection step was developed by training a Machine Learning (ML) model to further discriminate the signal events from the background events. A SM sample for $e^+e^- \to ZS,\ S \to b\bar{b},\ Z \to \mu^+\mu^-$ together with BSM sample $e^+e^- \to Zb\bar{b},\ Z \to \mu^+\mu^-$ has been generated to be used in the training. The training was performed by using several event and physics object variables as the inputs of the model Figure 2.

In Figure 3, the normalized deep neural network (DNN) model output distributions for both signal and background events separately on the training and validation samples are shown. A significant difference has been observed between the distributions obtained from the training and the validation samples for both signal and background events with the inclusion (left) and exclusion (right) of the recoil mass respectively.

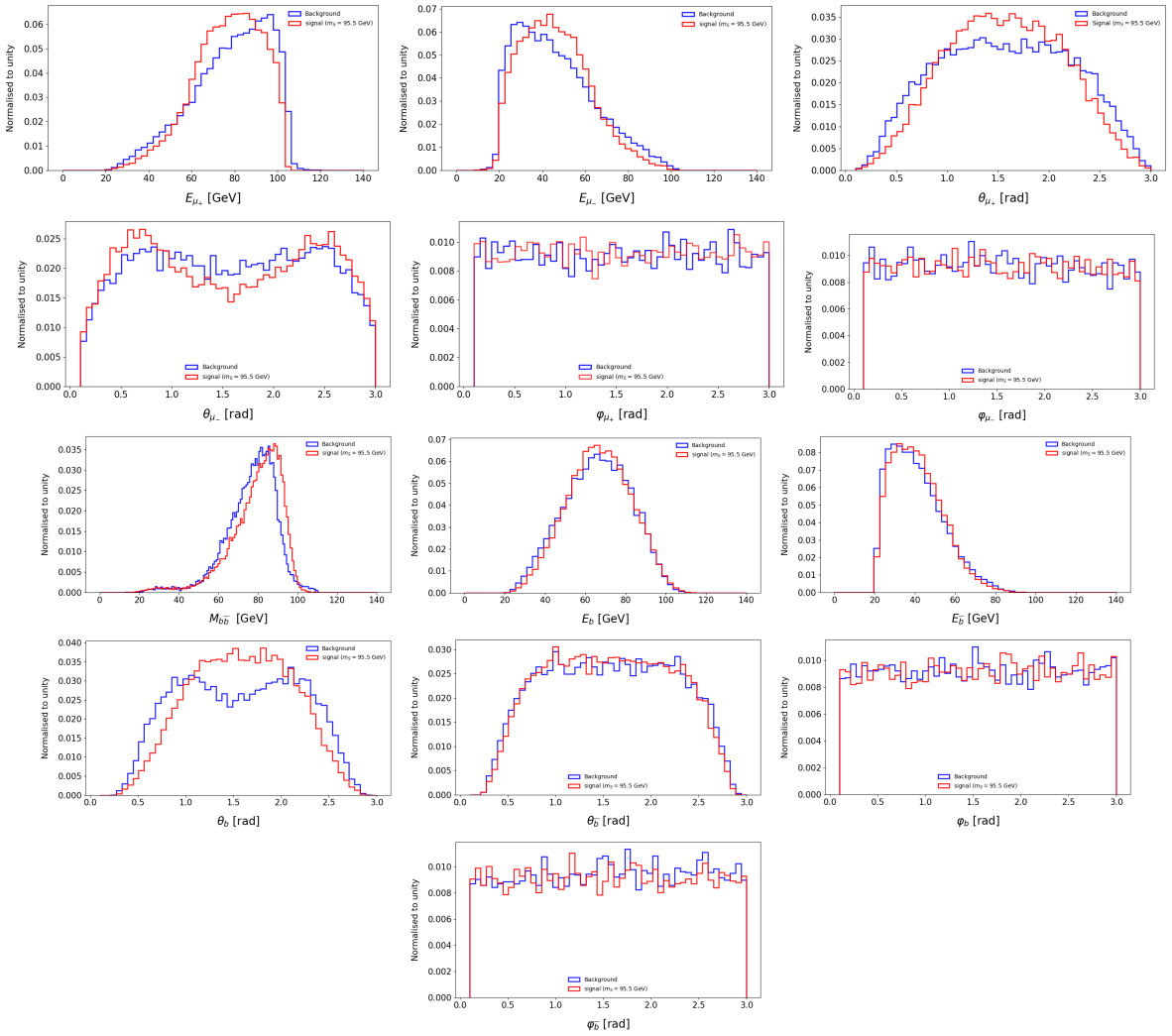


Figure 2. Chosen sensitive kinematic distributions for the muon pair ($\mu_+ = \mu_1$, $\mu_- = \mu_2$) and the b-tagged jets used as input variables for DNN normalized to unity at $\sqrt{s} = 250$ GeV.

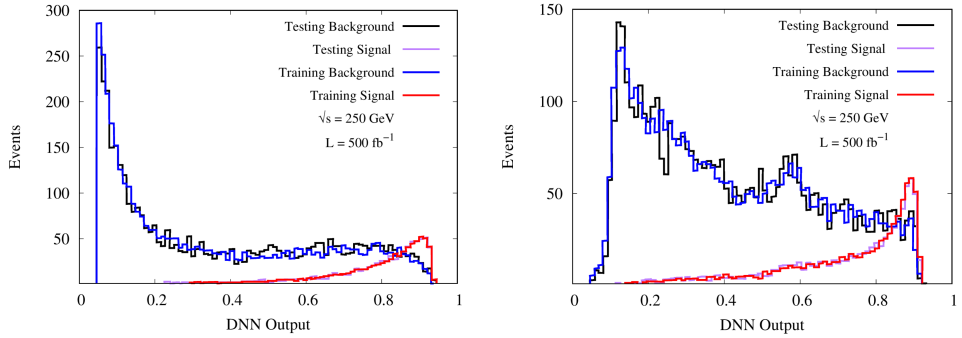


Figure 3. The DNN output with training and testing when the recoil mass is not included (right) and when the recoil mass is included (left) and the respective ROC curves below for $\sqrt{s} = 250$ GeV.

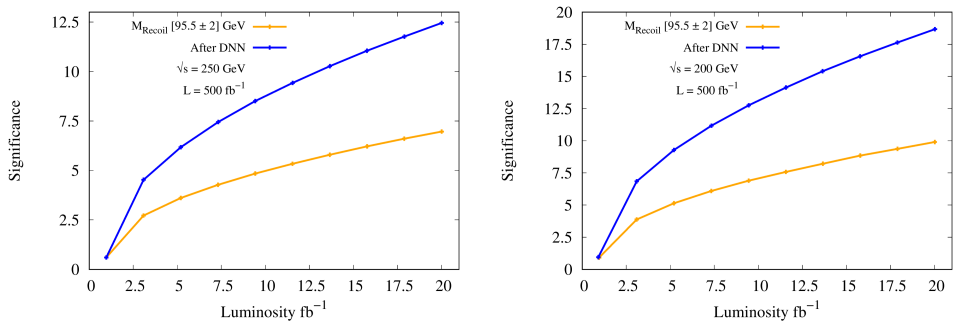


Figure 4. The signal significance as a function of Luminosity ($\mathcal{L}$) when recoil mass window of ± 2 GeV is considered and after DNN for (left) $\sqrt{s} = 250$ and (right) $\sqrt{s} = 200$ at $\mathcal{L} = 500$ fb^{-1} integrated luminosity at CEPC.

4. Conclusions

This proceedings investigated the potential for producing the ≈ 95 GeV Scalar (S) at $\sqrt{s} = 250$ GeV and $\mathcal{L} = 500$ fb^{-1} integrated luminosity CEPC environment. Figure 1 presents recoil mass reconstruction spectrum. DNN was crucial for a clear distinction in the region $m_S \in [90 - 96]$ GeV at $\sqrt{s} = 250$ GeV and $\sqrt{s} = 200$ GeV. Figure 4 displays optimized signal significance relevant to CEPC initial operation at considering 5% systematics.

References

[1] Buddenbrock S V, Chakrabarty N, Cornell A S, Kar D, Kumar M, Mandal T, Mellado B, Mukhopadhyaya B and Reed R G (preprint 1506.00612)
[2] Kumar M, Buddenbrock S V, Chakrabarty N, Cornell A S, Kar D, Mandal T, Mellado B, Mukhopadhyaya B and Reed R G 2017 J. Phys. Conf. Ser. **802** no.1 012007
[3] Bhattacharya S, Dey A, Lahiri J and Mukhopadhyaya B (preprint 2308.12473)
[4] Georges A *et al.* (ATLAS) 2013 Phys. Lett. B 726 88-119
[5] Bahl H, Biekötter T, Heinemeyer S, Li C, Paasch S, Weiglein G and Wittbrodt J 2023 Comput. Phys. Commun. **291** 108803
[6] Butterworth J M, Arbey A, Basso L, Belov S, Bharucha A, Braam F, Buckley A, Campanelli M, Chierici R and Djouadi A *et al.* 2009 (preprint 1003.1643)
[7] Isabell E, Margarete M, and Jonas W 2019 Comp. Phy. Com. 234 256-262
[8] Buddenbrock S V, Cornell A S, Iarilala E D R, Kumar M, Mellado B, Ruan X and Shrif E M 2019 J. Phys. G **46** no.11 115001
[9] Muhlleitner M, Sampaio M O P, Santos R and Wittbrodt J 2017 JHEP **03** 094
[10] Chen Z, Yang Y, Ruan M, Wang D, Li G, Jin S and Ban Y 2017 Chin. Phys. C **41** no.2 023003

Top reconstruction in the dilepton channel for the top Yukawa extraction

James Mitchell[1], **Cameron Garvey, James Keaveney**

Department of Physics, University of Cape Town, Rondebosch 7701, South Africa

E-mail: [1]mtcjam009@myuct.ac.za

Abstract. An overview of using top reconstruction applied to extracting the Yukawa coupling of the top quark (Y_t) in the dilepton final state of top quark pair production, using proton collisions recorded by the ATLAS experiment in LHC run 2 is presented. Electroweak corrections for the exchange of the Higgs boson between the top quark and the top anti-quark for different values of Y_t result in changes to the kinematic distributions of the top quark and top anti-quark system close to the $t\bar{t}$ production energy threshold. This extraction is currently performed using the proxy observable distributions $m_{\ell\ell bb}$, the invariant mass of the final state products, and $\Delta Y_{\ell b \ell b}$, the difference in rapidity in the final state products. The observables that are most sensitive to these changes are the invariant mass of the $t\bar{t}$ system,m_{tt} and the difference in rapidity, ΔY_{tt}, however, to use these distributions, we must reconstruct the kinematics of the $t\bar{t}$ system by estimating the kinematics of the missing neutrinos. These proceedings detail an initial investigation into analytically solving the constrained equations describing the $t\bar{t}$ system to estimate the neutrino kinematics.

1. Introduction

The Standard Model (SM) is recognised as one of the most successful theoretical frameworks developed; it focuses on the seventeen fundamental particles, describing the interactions between them. The discovery of the Higgs Boson in 2012 affirmed the Higgs mechanism, a pivotal component of the Standard Model (SM) that allows elementary particles to gain mass by interacting with the Higgs field [1]. The intensity of the interaction between a fermion and the Higgs field is represented by the Yukawa coupling of the fermion given by:

$$y_f = \sqrt{2}\frac{m_f}{\nu} \tag{1}$$

where ν represents the Higgs field's expected vacuum value, while m_f signifies the fermion's mass. Given that the top quark is the most massive fundamental particle known to us, it exhibits the strongest interaction with the Higgs field according to the SM. By accurately determining the ratio between the observed Yukawa coupling and the predicted SM value, we have the opportunity to evaluate the accuracy of the SM and potentially identify subtle indicators of undiscovered physics. The exchange of a virtual Higgs boson between the top quarks influences the kinematics of these quarks, enabling us to constrain the value of Y_t. The CMS collaboration has already carried out measurements of Y_t in both the lepton + jets [2] and the dilepton channels [3]. Furthermore, ATLAS is not only conducting an investigation in the dilepton channel but is also performing an analysis in the lepton + jets channel.

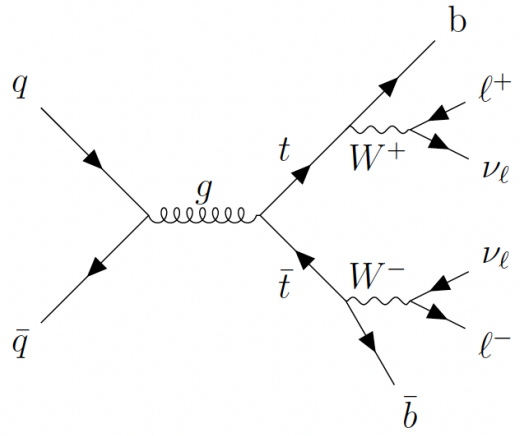


Figure 1: Feynman diagram of the dilepton final state of the $t\bar{t}$ system.

Currently, the analysis is performed using the invariant mass of the measured final state products, $m_{\ell\ell bb}$, however, this kinematic distribution does not contain the full kinematic information of the $t\bar{t}$ system. We can instead use the invariant mass of the $t\bar{t}$ system, m_{tt}, which appears to be more sensitive to variations in Y_t according to simulation. To construct the m_{tt} distribution, we must estimate the kinematics of the final state neutrinos, as they are not detected by the ATLAS detector. This paper will give an overview of the analysis, motivate the need reconstructing the m_{tt} distribution, as well as presenting an initial analytical method for solving this problem.

2. Analysis overview

2.1. $t\bar{t}$ production and decays

Pairs of $t\bar{t}$ originate from a gg or $q\bar{q}$ initial state with gg initiation being the primary method of production, responsible for about 90% of $t\bar{t}$ production at the LHC [4]. The top quark decays into a W-boson and a b-quark, and the b-quark then forms a jet within the detector. The W-boson can decay either hadronically or leptonically. In our analysis, we focus exclusively on the dilepton channel, wherein both W-bosons have undergone leptonic decay, as depicted in Figure 1.

2.2. Data sets and event selection

The utilized data is derived from proton-proton collisions with a center-of-mass energy of 13 TeV. This dataset, gathered by the ATLAS experiment between 2015 and 2018 during LHC's Run 2, is referred to as the full Run 2 dataset, equating to 140 fb^{-1} [5].Corresponding Monte-Carlo simulation samples were used to represent the $t\bar{t}$ signal sample and relevant backgrounds. Selections applied to all of the data consisted of: presence of exactly two leptons (either electron or muon) in an event, each boasting a transverse momentum exceeding 25 GeV; a minimum of two jets, each with a transverse momentum over 20 GeV; specifically two jets identified using the DL1r b-tagger algorithm, the most effective b-tagging algorithm used by the ATLAS collaboration [6]; an invariant mass of the paired leptons that deviates by more than 10 GeV from the Z-boson mass to exclude Z-processes; and a missing transverse momentum surpassing 30 GeV, accounting for the presence of the two undetected neutrinos.

2.3. Electroweak corrections
Electroweak corrections are computed for different values of $m_{\ell\ell bb}$ to account for the effects of the exchange of a virtual Higgs between the two top quarks. At different values of Y_t the exchange of the virtual Higgs boson affects the kinematics of the top quarks and results in observable effects. The electroweak corrections were computed using the HATHOR package [7].

2.4. Kinematic distributions and motivating top reconstruction
Kinematic distributions must now be constructed to observe how the change in Y_t corresponds to a changes in the distributions for our variables of interest. Currently, we are using the proxy variable $m_{\ell\ell bb}$, representing the invariant mass of the measured decay products, the two leptons and the two b-jets. The invariant mass of the $t\bar{t}$ system, m_{tt}, would be the ideal kinematic variable, however, it is impossible to reconstruct this variable without accounting for the missing neutrinos. If the neutrino momentum can be estimated well, this kinematic distribution may be used, and the result of our measurement could be improved.

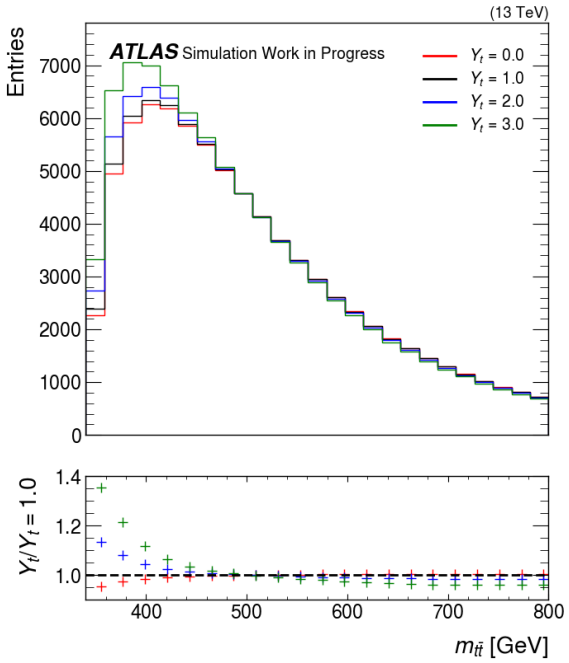


Figure 2: Kinematic distribution of m_{tt} for different values of Y_t created using truth level Monte-Carlo simulated data

Figure 2 shows the simulated truth level data for m_{tt} for different values of the calculated Y_t. m_{tt} is seen to be most sensitive around the threshold region, where $m_{tt} < 2\ m_t$. This increased sensitivity is due to the exchange of the virtual Higgs boson; the further above the threshold region the virtual Higgs boson is produced, the more relative momentum it needs in order to be exchanged by the two top quarks, suppressing the effects of changing Y_t above the threshold region.

2.5. Extracting Y_t
Currently, the value of Y_t is extracted using a binned profile likelihood fit using the $m_{\ell\ell bb}$ distribution. Template morphing is used to interpolate between the different values of Y_t allowing us to extract a value. We would want to perform the same extraction method on

the m_{tt} distribution, to investigate whether including all kinematic information of the system will improve our result.

3. Top reconstruction

The $t\bar{t}$ system can be modelled by the following set of equations [8]:

$$\not{E}_x = p_{\nu_x} + p_{\bar{\nu}x} \tag{2}$$
$$\not{E}_y = p_{\nu_y} + p_{\bar{\nu}y} \tag{3}$$
$$E_\nu^2 = m_\nu^2 + p_{\nu_x}^2 + p_{\nu_y}^2 + p_{\nu_z}^2 \tag{4}$$
$$E_{\bar{\nu}}^2 = m_{\bar{\nu}}^2 + p_{\bar{\nu}x}^2 + p_{\bar{\nu}y}^2 + p_{\bar{\nu}z}^2 \tag{5}$$
$$m_{W^+}^2 = (E_{l^+} + E_\nu)^2 - (p_{l^+x} + p_{\nu_x})^2 - (p_{l^+y} + p_{\nu_y})^2 - (p_{l^+z} + p_{\nu_z})^2 \tag{6}$$
$$m_{W^-}^2 = (E_{l^-} + E_{\bar{\nu}})^2 - (p_{l^-x} + p_{\bar{\nu_x}})^2 - (p_{l^-y} + p_{\bar{\nu_y}})^2 - (p_{l^-z} + p_{\bar{\nu_z}})^2 \tag{7}$$
$$m_t^2 = (E_b + E_{l^+} + E_\nu)^2 - (p_{b_x} + p_{l^+x} + p_{\nu_x})^2 - (p_{b_y} + p_{l^+y} + p_{\nu_y})^2 - (p_{b_z} + p_{l^+z} + p_{\nu_z})^2 \tag{8}$$
$$m_{\bar{t}}^2 = (E_{\bar{b}} + E_{l^-} + E_{\bar{\nu}})^2 - (p_{\bar{b}x} + p_{l^-x} + p_{\bar{\nu_x}})^2 - (p_{\bar{b}y} + p_{l^-y} + p_{\bar{\nu_y}})^2 - (p_{\bar{b}z} + p_{l^-z} + p_{\bar{\nu_z}})^2 \tag{9}$$

These equations can only be solved analytically if additional constraints are applied: W-boson mass fixed to 80 GeV; mass of the top quarks fixed to 172.5 GeV; assuming that all missing transverse momentum is attributed to the neutrinos. Applying these constraints will inevitably lead us to inaccuracy, and there is a particular concern with fixing the top mass; this is due to the area of our m_{tt} distribution that is most sensitive to Y_t being below the threshold region, where at least one of the top quarks must be off-shell, however, fixing the top mass assumes that both of the top quarks are on-shell. To solve these equations, our algorithm varies the values of the unknown neutrino eta iteratively in the range [-5, 5] attempting to solve the kinematic equations for all possible combinations of lepton and b-jets. If solutions are found, each solution is weighted, where the weighting function corresponds to how well the solutions agree with the missing transverse momentum in the event. The solutions with the highest weight are then selected, and are used as the four vectors for the $t\bar{t}$ reconstruction.

4. Results

The solution efficiency of this method is 60.15%, meaning that a solution is only found for 60.15% of the events. This low solution efficiency is to be expected, as finite detector resolutions can lead to mis-measurement of final state particles, leading to the equations becoming unsolvable. Methods, such as applying random smearings to the energies have been shown to increase this solution efficiency, making the method more suitable for purpose [9].

Figure 3 shows a scatter diagram of the predicted values of m_{tt} against the true values. For a performant algorithm, we would expect to see a fuzzy line at a roughly 45 degree incline, indicating that the predictions match the true values. The spread of this scatter plot indicates that there is some erroneous behaviour occurring, and that this method as it stands must be improved to become a part of the entire analysis. This method was only intended to be used as a baseline to compare machine-learning methods against, however, its low solution efficiency must be improved significantly to be used.

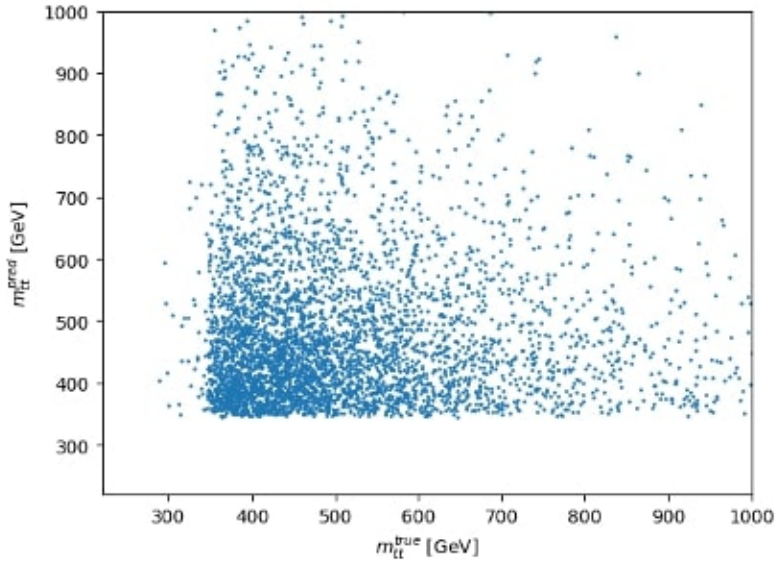


Figure 3: Scatter plot of the predicted values of m_{tt} against the true values.

5. Conclusion

The goal of the analysis is to measure the Yukawa coupling between the top quark and the Higgs boson by using an indirect search method in the $t\bar{t}$ dilepton channel. The m_{tt} distribution contains all of the kinematic information of the system, and appears more sensitive to Y_t than $m_{\ell\ell bb}$ distribution, however, this requires that we estimate the kinematics of the missing neutrinos. An analytical attempt to solve the kinematic equations describing the $t\bar{t}$ system is presented, however, it proved to be ineffective. Further work would involve improving this algorithm using smearing, and attempting to solve this current problem using machine learning methods.

References

[1] Higgs, P. 1964 Broken Symmetries and the Masses of Gauge Bosons *Phys. Rev. Lett.* **13** 508–509.
[2] Sirunyan, A. M. et al. (CMS) 2019 Measurement of the top quark Yukawa coupling from $t\bar{t}$ kinematic distributions in the lepton+jets final state in proton-proton collisions at $\sqrt{s} = 13$ TeV *Phys. Rev.* D **100** 072007.
[3] Sirunyan, A. M. et al. (CMS) 2020 Measurement of the top quark Yukawa coupling from $t\bar{t}$ kinematic distributions in the dilepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV *Phys. Rev.* D **102** 092013.
[4] Ferreira da Silva, P. 2016 Top quark production at the LHC. *arXiv:*1605.05343.
[5] ATLAS Collaboration 2022 Luminosity determination in *pp* collisions at $\sqrt{s} = 13$TeV using the ATLAS detector at the LHC *Preprint* 2212.09379.
[6] Aad, G. et al. (ATLAS) 2023 ATLAS flavour-tagging algorithms for the LHC Run 2 *pp* collision dataset *Eur. Phys. J.* C **83** 681.
[7] Aliev, M., Lacker, H., Langenfeld, U., Moch, S., Uwer, P., & Wiedermann, M. 2011 HATHOR - HAdronic Top and Heavy quarks crOss section calculatoR *Comput. Phys. Commun.* **182** 1034-46.
[8] Sonnenschein, L. 2006 Analytical solution of $t\bar{t}$ dilepton equations *Phys. Rev.* D **73** 054015.
[9] Sirunyan, A. M. et al. (CMS) 2020 Measurement of $t\bar{t}$ differential cross-sections in proton-proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons *J. High Energy Phys.* 2019(2).

Developing the Temperature Mapping Plugin of the Tile Calorimeter of the ATLAS Detector within Tile-in-One

LS Phakathi[1], SS Gumede[1], B Kibirige[1], J Smieško[2] and F Martins[3]

[1]Department of Physics and Engeneering, University of Zululand, kwaDlangezwa campus, kwaZulu Natal, South Africa.

[2]Faculty of Mathematics and Physics, Institute of Particle and Nuclear Physics, Charles University

[3]Laboratório de Instrumentação e Física Experimental de Partículas

E-mail: lungisani.sipho.phakathi@cern.ch

Abstract. The full utilization of the Tile-in-One (TiO) is necessary to continually analyze and visualise the historical temperature data from the Tile hadron calorimeter of the ATLAS off-line (not real time). The Tile Calorimeter (TileCal) is a sampling hadronic calorimeter covering the central region of the ATLAS experiment, with steel as the absorber and plastic scintillators as the active medium. The TiO is a collection of small, independent web tools called plugins which assess the quality of data and conditions for ATLAS TileCal. The TiO web platform needs to be highly flexible and easy to maintain so that plugin developers can also benefit. Detector Control System (DCS) Data Viewer (DDV) in this case helps for historical data. This paper shows the successfulness of developing the temperature plugin and the mapping of it. The work aims to study the variation of temperature of the seven probes inside the drawers. These probes are located on the High Voltage (HV) system side and the readout electronic side of the drawer. The reason for this, is to know the stability of the electronics, and gain stability of the PMTs. Finding a simple way to display not only the state of one specific module but also the status of the entire detector is the main goal of the stable temperature plugin.

1. Introduction

One of the four primary detectors of the Large Hadron Collider (LHC) is the (A Toroidal LHC ApparatuS) ATLAS [1] experiment run by the European Organization for Nuclear Research (CERN). The LHC is the biggest and most powerful particle accelerator that has ever been created by mankind. It is a circular hadron collider with a 26.7 km radius that is situated beneath the French-Swiss border close to Geneva at an average depth of 100 m. At the interaction locations of the ATLAS, CMS, ALICE, and LHCb, the two beams interact and produce proton-proton collisions. The beams are first injected to the LHC from the Super Proton Synchrotron (SPS) and then they circulate in the LHC in opposite directions and collide in interaction points [2]. ATLAS [3]is the general purpose detector at the LHC. The ATLAS detector's primary goal is to look into a variety of physics, including the investigation of the Higgs boson and other Standard Model events, searches for extra dimensions, and particles that might make up dark matter [4]. Sub-detectors and magnets are arranged in cylindrical layers around the beam pipe in the detector [5], which has a barrel shape and is 44 m long by 25 m

wide. However this paper shows in Figure 1 the Tile Calorimeter (TileCal), a sub-detector that is a component of the ATLAS Calorimeter System.

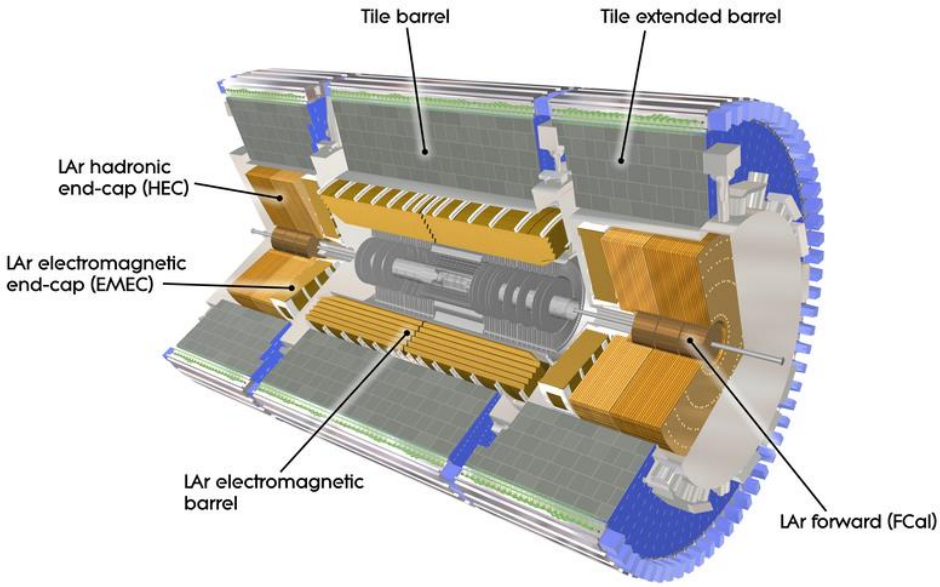


Figure 1. ATLAS TileCal in combination with the sub-detectors surrounding it [6].

TileCal detects hadrons, jets, and taus and assists in muon identification. It also contributes to the reconstruction of jet energy and missing transverse energy [6]. The calorimeter's two extended barrels cover the pseudorapidity range $0.8 < |\eta| < 1.7$, while the calorimeter's center (barrel) component covers the pseudorapidity region $|\eta| < 1.0$. The TileCal is a sampling calorimeter [7] that uses steel plates as the absorber and plastic scintillating tiles as the active medium. When a charged particle traverses through the scintillating tiles, light is emitted and collected at the scintillating edges. The light generated by the scintillating tiles is then transported to two photomultiplier tubes (PMT) by Wave Length Shifting Fibers (WLS) [8]. The PMT send electrical signals to the front-end electronics for it to be shaped, amplified and then digitised at a sampling frequency. Each half central barrel has 22 corresponding detecting cells, and each extended barrel has 16 detecting cells. There are roughly 10,000 readout channels overall since each detection cell is read out independently from both sides [9].

2. Tile-in-One documentation

Tile-in-One (TiO) [10] is a collection of small, independent web applications of plugins, which are used to evaluate the quality of the data and the conditions of ATLAS TileCal. The web platform is designed around a single primary server (TiO/VMO), which is in charge of routing user requests and responses, authenticating users via oauth2proxy, and managing the secure connection to the web platform. This server is located at `ttps:/tio.cern.ch`. Figure 3 shows the architecture of the TiO web platform. The TiO platform aims to unite numerous TileCal web products into a single, shared platform that will share the same processing infrastructure and give access to shared services and data as existing interfaces become outmoded and harder to maintain [11].

It offers standard computing infrastructure and tools, including access to the CERN authentication service or ready-to-use web toolkit libraries like Bootstrap, while yet being

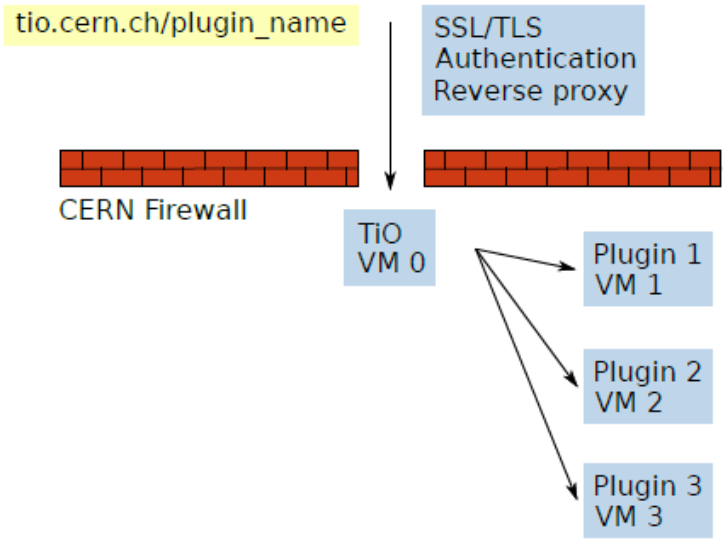


Figure 2. The architecture of the TiO web platform.

adaptable enough to satisfy the unique needs of various applications [12]. TiO is made to be adaptable and straightforward to maintain. This design's objective was to produce a user-friendly web platform with the flexibility to gather information from different data sources [11].

TiO has two template plugins to help new plugins developers. It composes the full stack, meaning that they are in charge of creating the web server and the web page for the typical web application. The plugin generates responses to user requests by loading data from a csv or JSON file on the server side. Several types of responses are possible, but the most important are these two:

- The static file response, in which the client receives entire files from the server..
- Dynamic response via template; in this case, the web page is generated dynamically by the server..

Twenty-two plugins are now active on the TiO platform; two of them have been deprecated and are used by different CERN network users.

3. Temperature Mapping plugin

The Detector Control System (DCS) [13] is responsible of monitoring the low voltage (LV), high voltage (HV), front-end electronics cooling system, and back-end crates in various applications for the detector infrastructure. The temperature of the front-end electronics is monitored by the TileCal DCS at a number of locations, including power supply and monitoring boards, using temperature sensors [14]. In accordance with equation 1, it is essential to monitor the temperature in the TileCal since it is one of the factors that, together with the high voltage delivered, alter the gain of the PMTs.

$$G = HV^{\alpha} \tag{1}$$

G is the gain of the PMT, α is a constant ~ 7. The current study's goal is to create a flexible and reliable web interface environment that will continuously track changes in the temperature

in the module's drawers over a brief period of time, like a few weeks. Utilizing the reliable TiO technology, the project seeks to extract, monitor, and eventually visualize temperature data for stability in the drawers.

Figure 3 shows the the 7 probes that needs to be monitored.

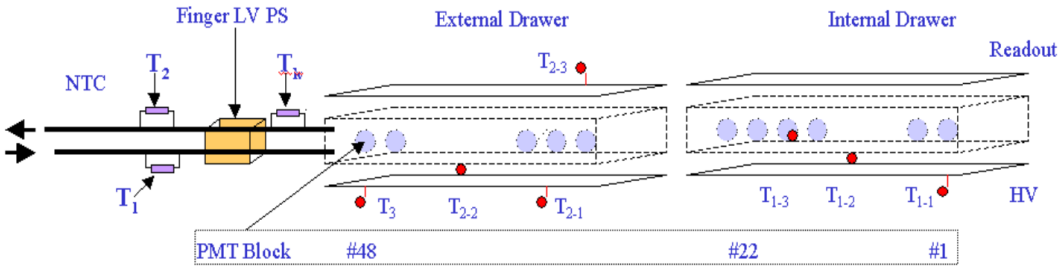


Figure 3. The positions of the temperature probes in the drawers and cooling pipes. In the drawer, the readout electronics are at the top, and the HV system cards are at the bottom. T_1&T_2 are the temperature sensors for entering and exiting cooling water respectively.

- T_{1-1} and T_{2-1}: Temperatures of the internal and exterior HV opto board, respectively.
- T_{1-2} and T_{2-2}: Temperatures inside and outside the drawer.
- The PMT block 22's temperature.
- Card's interface temperature.
- HV micro card temperature.

4. Methodology

Figure 4 shows how the user send the HTTP request to DCS Data Viewer (DDV) the server and how the server respond to the request using python.

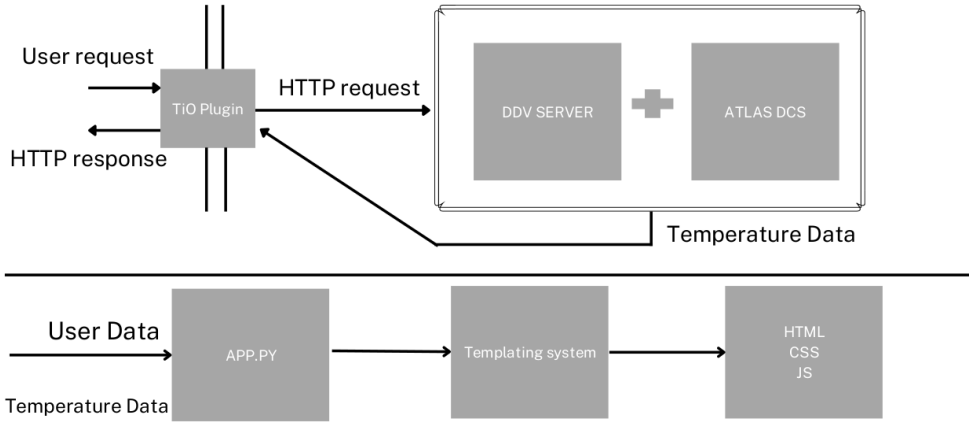


Figure 4. The methodology followed for requesting the data from the DDV server and the response received by the user

5. Preliminary results & Summary

Plugin is successful in terms of being able to extract, analyse and visualise the temperature data. Figure 5 shows the temperature data results of all the four partitions of the TileCal. The data is for the last five days. Colors are mapped based on the minimum and maximum temperatures of the specific probe.

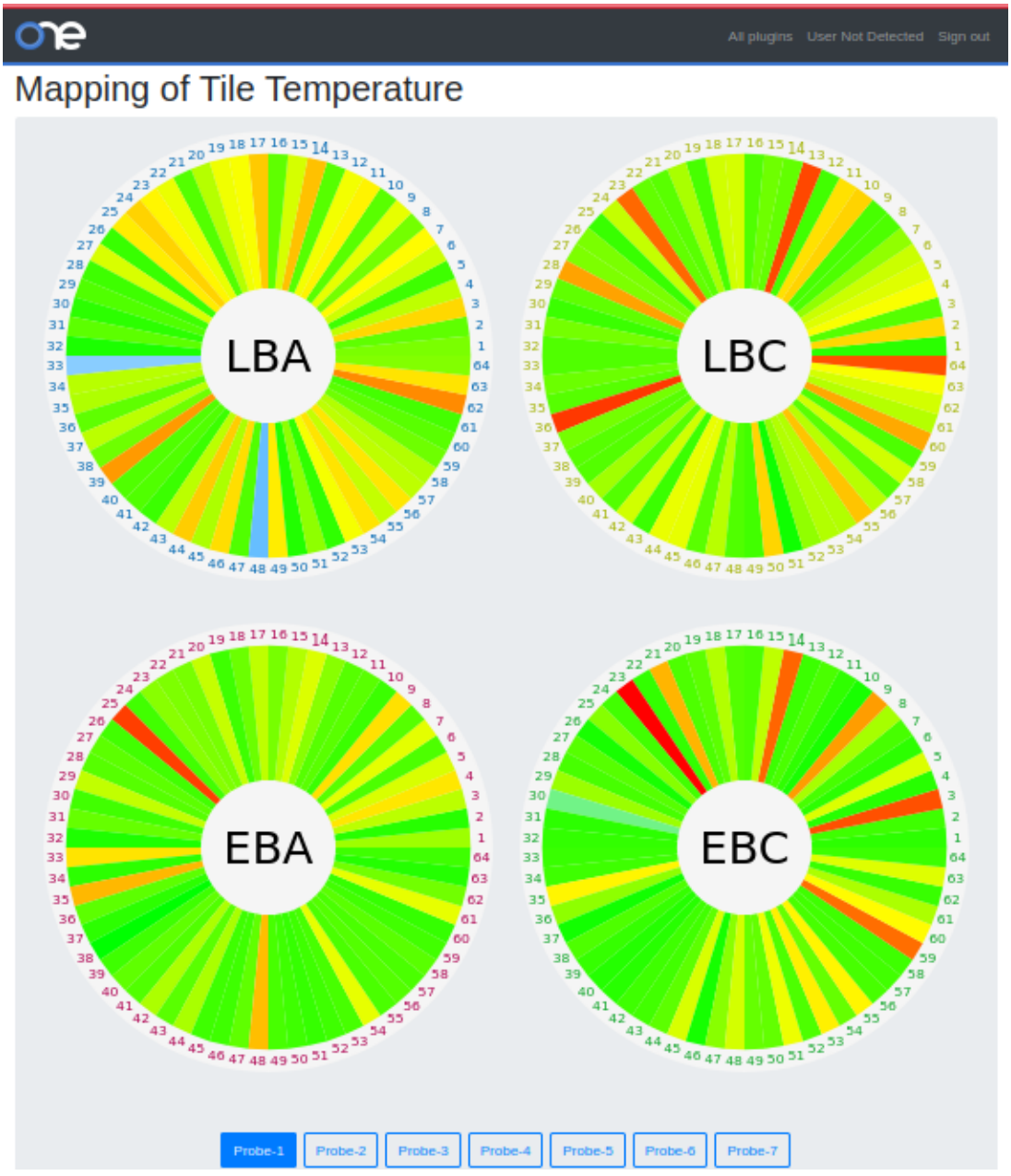


Figure 5. Module for LBA63, showing the fluctuation of 0.2.

This shows the temperature variation over time. The user can view temperature data from TileCal using an interactive plotting library that includes features like downloading.png files for use in reports, zooming in and out of the plots, and recording data at the location where the pointer is positioned.

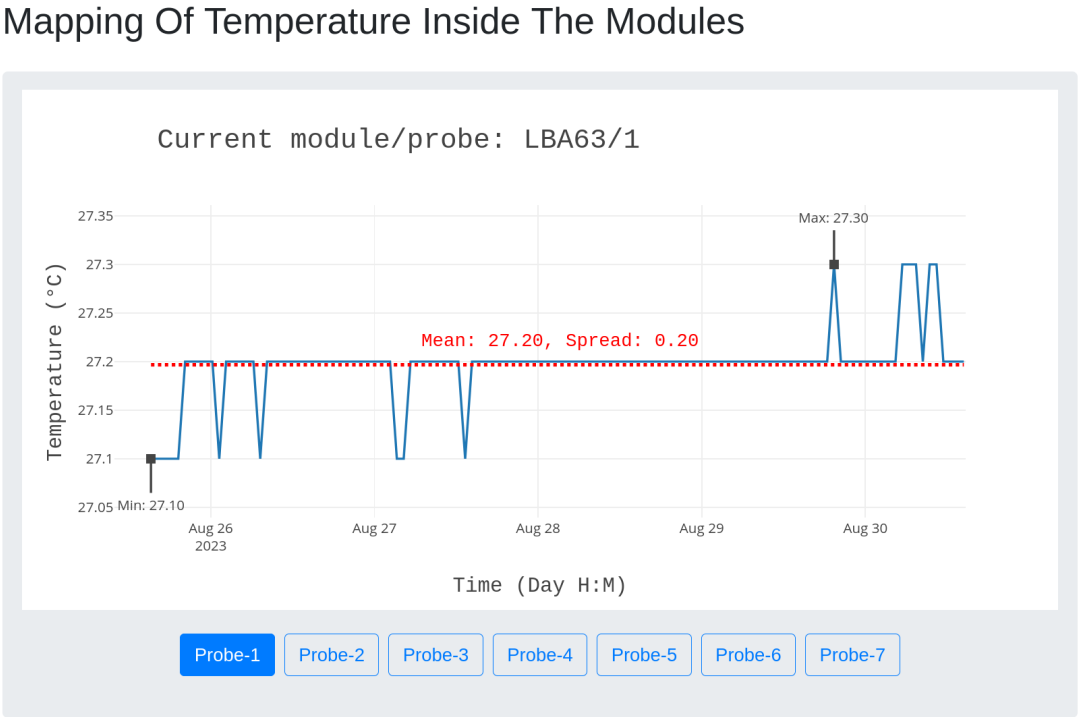


Figure 6. All the partitions of the detector

The gist of the work is to find the probe in the HV system and the readout electronic side inside the drawer where the temperature rises or falls too much over time. By examining the temperature variation over time, the plugin will be able to identify problems with the electronics' cooling system and ensure that they are stable.

References

[1] ATLAS-Collaboration 2008 The atlas experiment at the cern large hadron collider
[2] Ellis J 2004 *The European Physical Journal C-Particles and Fields* **34** 51–56
[3] Aad G, Anduaga X S, Antonelli S, Bendel M, Breiler B, Castrovillari F, Civera J, Del Prete T, Dova M T, Duffin S *et al.* 2008
[4] Giagu S 2019 *Frontiers in Physics* **7** 75
[5] Di Gregorio G 2023 The large hadron collider and the atlas experiment *A Beauty-ful Boson: Search for the Higgs Boson Produced in Association with a Vector Boson and Decaying into a Pair of b-quarks Using Large-R Jets with the ATLAS Detector* (Springer) pp 31–45
[6] Francavilla P, Collaboration A *et al.* 2012 The atlas tile hadronic calorimeter performance at the lhc *Journal of Physics: Conference Series* vol 404 (IOP Publishing) p 012007
[7] ATLAS-Collaboration 2010 Readiness of the atlas tile calorimeter for lhc collisions
[8] ATLAS-Collaboration 1996 Tile calorimeter technical design report, cern/lhcc 96-42
[9] Aaboud M, Aad G, Abbott B, Abdallah J, Abdinov O, Abeloos B, Abhayasinghe D K, Abidi S H, AbouZeid O, Abraham N L *et al.* 2018 *The European Physical Journal C* **78** 1–48
[10] Smirnov Y and Smieško J 2020 Tile-in-one: an integrated system for data quality and condition assessment for the atlas tile calorimeter *EPJ Web of Conferences* vol 245 (EDP Sciences) p 01010
[11] Smieško J 2019 An integrated system for data quality and conditions assessment for the atlas tile calorimeter-tile-in-one *EPJ Web of Conferences* vol 214 (EDP Sciences) p 01030
[12] Scheirich D 2021 The atlas tile calorimeter tools for data quality assessment *EPJ Web of Conferences* vol 251 (EDP Sciences) p 03018
[13] Schlenker S, Arfaoui S, Franz S, Gutzwiller O, Tsarouchas C, Mindur B, Hartert J, Zimmermann S, Talyshev A, Damazio D O *et al.* 2011 The atlas detector control system *Proceedings of ICALEPCS* pp 5–8
[14] Santos S, Galhardo B, Veloso F, Collaboration A *et al.* 2018 *European Physical Journal C* **78** 987

Measurement and simulation of secondary neutron production from a 66 MeV proton beam

J de Klerk and S Peterson

Department of Physics, University of Cape Town, Rondebosch 7700, South Africa

E-mail: dkljos001@myuct.ac.za

Abstract. Prompt gamma rays (PGs) produced by nuclear collisions during proton therapy (PT) can be used for in vivo proton beam range verification through a technique known as prompt gamma imaging (PGI). During PT, the production of PGs within the tissue provides a reliable location map for the proton dose deposition. The detection of these PGs using a Compton Camera (CC) can be used to construct an image of the source of the PGs. Since the reconstructed source image is dependent on the quality of the detection, the detection efficiency of a CC device can be negatively affected by additional secondary radiation (particularly neutrons) produced alongside the prompt gamma rays. The goal of this work is to understand the characteristics of these secondary neutrons during a proton beam irradiation and their potential impact on a CC. The experimental setup consisted of a 66 MeV pencil beam of protons on a variety of targets (HDPE, water, graphite) with the neutron measurements made using an organic liquid scintillator detector (BC-501A) at several angles within the K600 vault (S-line) at iThemba LABS. To obtain neutron energy spectra with the BC-501A, an unfolding procedure is typically used. Here, energy spectra are deconvolved from measured neutron light output spectra that have been calibrated. As the quality of the unfolding is dependent on the prior knowledge of the energy spectrum, a Geant4 simulation was constructed to better understand the produced neutron field and aid the unfolding process. A comparison of the experimental and simulated results will be discussed.

1. Introduction

1.1. Proton Therapy

The concept of precisely targeting localized regions within the body using high-energy protons was first introduced by Robert R. Wilson in the 1940s [1]. Proton therapy (PT) has emerged as a preferred choice compared to more traditional treatments such as photon and electron therapies due to its capacity for generating highly conformal treatments. PT, furthermore, minimizes the impact on healthy tissues and reducing both acute and late side effects [2, 3]. The benefits of PT result directly from the fundamental interactions between charged particles and matter. As protons transverse through tissue, energy is transferred through coulomb interactions between the atomic nuclei and the atomic electrons [4]. The majority of the proton's kinetic energy is transferred at the end of its trajectory, where maximum ionisation takes place, and is known as the Bragg peak. Beyond the Bragg peak, the proton's energy rapidly decreases to zero where its path comes to an end [5].

1.2. PGI and neutrons

Prompt Gamma Imaging (PGI) is a technique used for *in vivo* proton beam range verification during proton therapy [6]. As protons interact within the tissue, nuclei that are in an excited energy state are produced. Within less than 10^{-14} seconds, a characteristic prompt gamma-ray is produced at the location of the proton-nuclei interaction [7]. In this process referred to as prompt emission, these gamma-rays are unique to the atomic composition of the tissue and therefore identifiable by their energy [8]. One device used to detect gamma-rays is a Compton camera (CC) and has been used for *in vivo* proton beam dose delivery and beam range verification [9]. The fundamental workings of a CC relies on having a scatterer and an absorber [10]. At each stage, the energy and position of the gamma-ray is recorded, allowing the reconstruction of the source particle to be confined to the surface of a cone (figure 1). The overlapping of several cones results in an image of the source [11]. In turn, it can be said that the detection efficiency of PGs is linked to a reliable location map for proton dose deposition. There are, however, other secondary particles produced alongside PGs during PT (such as neutrons) which may negatively affect the image quality of the reconstructed source image [12]. The production of neutrons occurs through several proton-tissue reaction channels [13] and from proton interactions with the treatment unit [14]. Coincidental neutrons recorded with gamma-rays produce false scatters which degrade the reconstructed image quality [11, 15].

1.3. Secondary Neutron Detection

The measurement of secondary neutrons within a mixed particle field is of interest as they may have a negative impact in image reconstruction. Various studies have used particle discrimination techniques to produce neutron spectra from a range of incident proton energies. In particular, neutron energies up to 30 MeV have been measured using a plastic scintillator (EJ299-34) from 70 MeV protons incident on a brass target [16]. Additionally, a 60 MeV proton beam incident on a paraffin target has produced up to 35 MeVee neutrons on a solid plastic scintillator (EJ299-33)[17]. In contrast, Higher proton energies of 200 MeV and 150 MeV have produced a maximum of 3.2 MeVee Neutrons with an organic liquid scintillator (EJ-309) from irradiation of a water phantom [18].

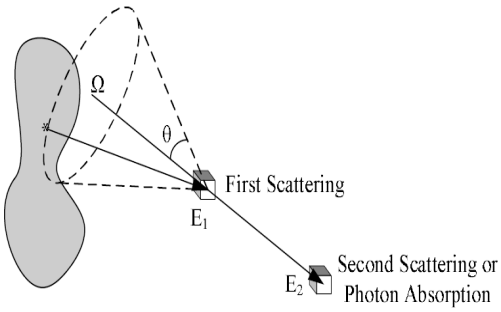


Figure 1: Traditional Compton camera [19].

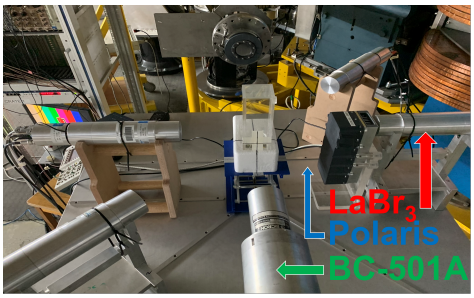


Figure 2: Detector configuration of 66 MeV protons incident on a water phantom.

2. Measurements

To understand the characteristics of secondary neutrons produced during irradiation, measurements were taken within the K600 vault (S-line) at iThemba LABS using a 66 MeV pencil beam on a variety of targets (HDPE, water, graphite). These targets were chosen since they possess similar densities and elemental compositions to that of tissue, namely C, H and O.

Additionally, the various reaction channels between incident protons and the targets result in the production of identifiable prompt gamma rays and neutrons of various energies. Figure 2 displays a schematic of the experimental setup with the UCT Polaris detector placed at 90° and the BC-501A placed at 0° relative to a water target. The beam current was set to a nominal value of 1 nA with a 1 in 5 pulse selection ($\sim$ 300 ns per each pulse).

2.1. Polaris

The UCT Compton camera 'Polaris' build by H3D Inc. [Ann Arbor, MI USA] is a room temperature solid-state detector. The Polaris detectors are comprised of two independent platforms with each one consisting of two cadmium zinc telluride crystals that are arranged side by side. The crystals are pixelated in the x- and y- directions with depth in the z-direction, providing three-dimensional position-sensitive gamma-ray detection. They furthermore have good energy resolution with a 0.8 % FWHM at 662 keV [20]. These attributes make the UCT Polaris detectors a suitable candidate for PGI in PT. Although thermal neutrons are detectable by the Polaris detectors [21], they struggle to detect fast neutrons due to the low energy thresholds that are required [22]. Moreover, the detectors are not equipped to distinguish between neutron and gamma-ray events on an event-by-event basis.

2.2. BC-501A

The BC-501A detector is a liquid organic scintillator detector with a diameter of 5.08 cm and 10.16 cm in length. The detector composition is primarily xylene with small amounts of naphthalene. Its pulse shape (PS) discrimination capability provides the capacity to distinguish between gamma-rays and neutrons [23]. The PS discrimination is produced using a zero cross over method which is dependent on an incident particles' pulse rise time and shape [24]. As a result, the BC-501A detector is suitable to produce neutron spectra in a mixed particle field.

2.3. Results

Figure 3 displays the light output spectrum from the BC-501A detector placed at 90° to a HDPE target. Straight line cuts are placed on the spectrum to separate neutrons from gamma-ray events. The light output spectrum is then projected onto the x-axis to provide a pulse

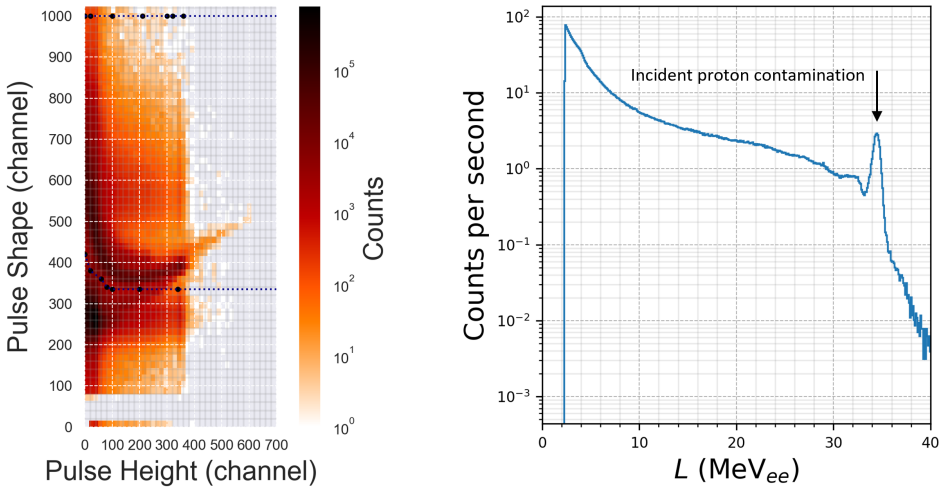


Figure 3: BC-501A light output spectra with corresponding neutron cuts from protons incident on a HDPE target (left). Calibrated neutron pulse height spectrum (right).

height spectrum. The pulse height spectra was calibrated using the 4.44 MeV Compton peak from an Americium-Beryllium (Am-Be) source. This provides an electron energy scale referred to as MeVee (MeV electron equivalent) which is the energy deposited by 1 MeV electron in the detector [25]. To obtain a neutron energy spectrum in MeV, a process referred to as unfolding is required. The unfolding process relies heavily on having a set of well-known response functions (response matrix) [26] which is the unique way a detector behaves to a particle with a specific energy. Due to proton contamination, the unfolding process becomes challenging since there is no response matrix available for protons. To help aid the process of unfolding and develop an understanding of the particle production from the experimental measurements taken at iThemba LABS, a Monte Carlo simulation has been constructed.

3. Geant4 Model

Geant4 is a toolkit designed for Monte Carlo simulation, enabling modelling and analysis of particle transport through various materials. A model of the experimental set up of a 66.4 MeV proton pencil beam has been developed (figure 4). The simulation contains a beam window made of a cobalt alloy (Havar) with a diameter of 4 cm and a thickness of 120 μm. The Havar film was used experimentally to separate the vacuum chamber of the particle accelerator system with the experimental area. The primary protons were placed directly behind the beam window. To gain insight into the experimental proton beam distribution, two radiochromatic

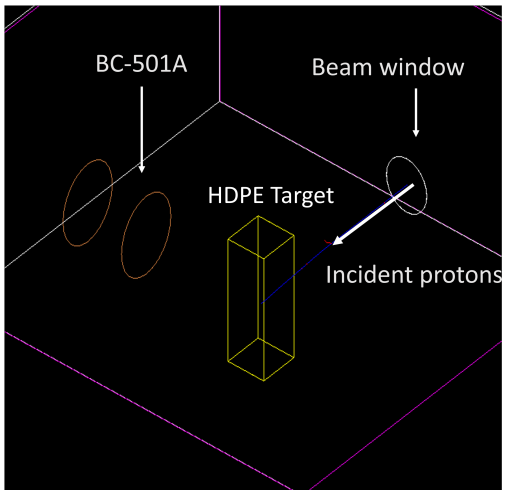


Figure 4: Geometry of Geant4 simulation.

films were used (figure 5). The first film was taped to the beam window and the second film was taped to the target (12.5 cm from the beam window). The simulation sought to mimic the experimental proton beam distribution by adjusting the starting position distribution and momentum distribution of the simulated primary protons. The x,y positions of the simulated protons were recorded within the simulation (figures 6 & 7). The simulated results show a broadening of the proton beam, however, it does not fully capture the shape of the measured proton distribution at the beam window. This is largely due to the none uniform shape observed experimentally. To compare with the BC-501A experimentally measured data, a HDPE target with the BC-501A detector placed at 90° relative to the incident proton beam was simulated. Figure 8 displays the incident energies of gamma-rays, neutrons, primary protons and secondary protons on the detector. A contribution from primary protons that have scattered off the beam window and have hit the detector can be observed. Figure 9 demonstrates the scattering angle

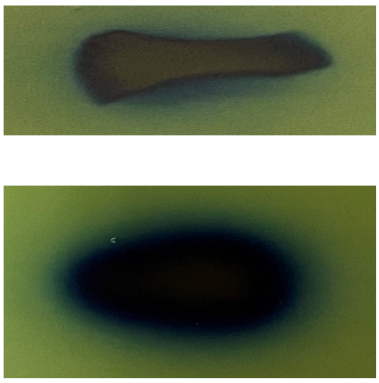


Figure 5: Exiting beam window radiochromatic film (top) & target entry radiochromatic film (bottom).

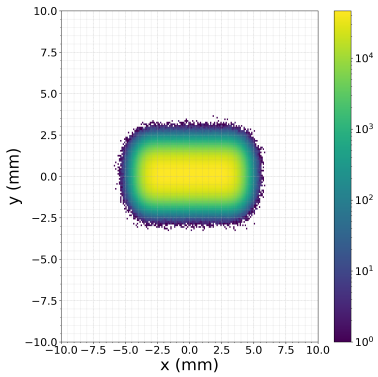


Figure 6: Simulated x,y proton position at beam exit window.

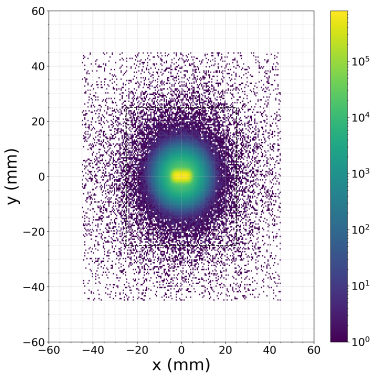


Figure 7: Simulated x,y proton position in front of the target.

distribution of these protons with their respective energies. Secondary protons produced within the target also add to the overall incident energy spectrum. Relative to the measured data, the primary protons are not as prominent in the overall simulated spectrum. Moreover, the simulated secondary proton energies fall within neutron energy range making them difficult to distinguish. Evaluating the overall contribution of protons (primary and secondary) in the light output spectra makes the unfolding process challenging. The identification of proton contribution is significant due to no proton response matrix being available. Thus, proton interactions may be mistaken for neutron interactions, resulting in false neutron interactions. In the attempt to remove proton contamination, neutron interactions may also be removed. Both dangers may contribute negatively to the image quality of the unfolded neutron spectrum.

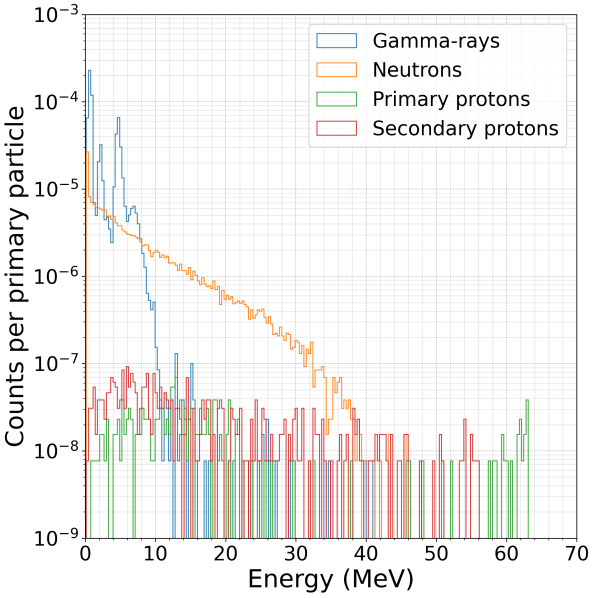


Figure 8: Simulated incident energies on BC-501A from 66.4 MeV protons hitting the HDPE target.

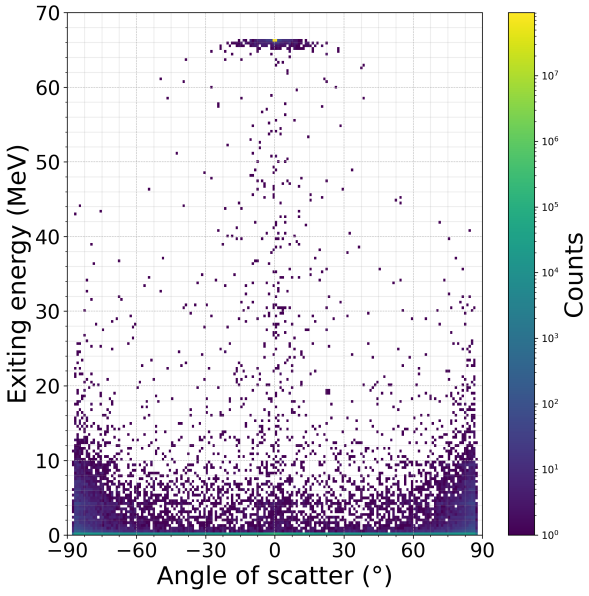


Figure 9: Exiting proton energy with corresponding beam window scattering angle.

4. Conclusion

Calibrated neutron light output spectra have successfully been produced from BC-501A. The measurements include light output spectra from several different targets at various angles. The Monte Carlo simulation provides information regarding primary protons incident on BC-501A. Furthermore, secondary particles produced from proton-target interactions form a continuum of energies that are also detected. Both aspects add valuable contributions for the unfolding process. Obtaining a measured neutron energy spectra requires the completion of an unfolding process through further simulation of particle interaction within the BC-501A detector. The ongoing work involves using the current neutron response matrix together with the measured and simulated BC-501A data to create a new set of response functions. Insights can then be gained into neutrons produced in a PT environment, leading to potential advancements in PGI.

5. References

[1] Wilson R R 1946 *Radiology* **47** 487–491
[2] Olsen D R, Bruland Ø S, Frykholm G and Norderhaug I N 2007 *Radiotherapy and oncology* **83** 123–132
[3] Schneider U, Agosteo S, Pedroni E and Besserer J 2002 *International Journal of Radiation Oncology* Biology* Physics* **53** 244–251
[4] Newhauser W D and Zhang R 2015 *Physics in Medicine & Biology* **60** R155
[5] Martin B R and Shaw G 2019 *Nuclear and particle physics: an introduction* (John Wiley & Sons)
[6] Maggi P, Peterson S, Panthi R, Mackin D, Yang H, He Z, Beddar S and Polf J 2020 *Physics in Medicine & Biology* **65** 125004
[7] Gilmore G 2008 *Practical gamma-ray spectroscopy* (John Wiley & Sons)
[8] Knopf A C and Lomax A 2013 *Physics in Medicine & Biology* **58** R131
[9] Draeger E, Mackin D, Peterson S, Chen H, Avery S, Beddar S and Polf J 2018 *Physics in Medicine & Biology* **63** 035019
[10] Kim S M, Seo H, Park J H, Kim C H, Lee C S, Lee S J, Lee D S and Lee J S 2013 *Physics in Medicine & Biology* **58** 2823
[11] Draeger E, Peterson S, Mackin D, Chen H, Beddar S and Polf J C 2017 *IEEE transactions on radiation and plasma medical sciences* **1** 358–367
[12] Panthi R, Maggi P, Peterson S, Mackin D, Polf J and Beddar S 2020 *IEEE transactions on radiation and plasma medical sciences* **5** 383–391
[13] Tuckwell W and Bezak E 2007 *Physics in Medicine & Biology* **52** 2483
[14] Lin Y C, Lee C C, Chao T C and Tsai H Y 2017 *Radiation Physics and Chemistry* **140** 290–294
[15] S, Peterson P, Maggi R, Panthi S, Beddar D, Mackin J and Polf Investigating the impact of secondaries neutrons on compton camera for medical imaging poster presented at The 2020 Joint AAPM COMP Virtual Meeting
[16] 2019 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **936** 31–33 ISSN 0168-9002 frontier Detectors for Frontier Physics: 14th Pisa Meeting on Advanced Detectors
[17] Chung S, Kacperek A, Speller R and Gutierrez A 2017 *Journal of Instrumentation* **12** C11033
[18] Clarke S, Wieger B, Pryser E, Arghal R, Pozzi S, Hälg R, Bashkirov V and Schulte R 2014 Characterization of secondary neutron production during proton therapy *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)* (IEEE) pp 1–3
[19] Xu D 2006 *Gamma-ray imaging and polarization measurement using three-dimensional position-sensitive cadmium zinc telluride detectors.* Ph.D. thesis
[20] Zhang F, He Z and Seifert C E 2007 *IEEE Transactions on Nuclear Science* **54** 843–848
[21] Martín-Martín A, Iñiguez M, Luke P, Barquero R, Lorente A, Morchón J, Gallego E, Quincoces G and Martí-Climent J 2009 *Radiation protection dosimetry* **133** 193–199
[22] Goodman D, Streicher M, Zhu Y, Brown S and He Z 2017 *IEEE Transactions on Nuclear Science* **64** 2531–2535
[23] Knoll G F 2010 *Radiation detection and measurement* (John Wiley & Sons)
[24] Nakhostin M and Walker P 2010 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **621** 498–501
[25] Mouatassim S, Costa G, Guillaume G, Heusch B, Huck A and Moszyński M 1995 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **359** 530–536
[26] Matzke M 1994 Unfolding of pulse height spectra: the hepro program system Tech. rep.

Increasing the location rate of Positron Emission Particle Tracking (PEPT) measurements

Rayhaan Perin[1], Jonathan P. Shock[2], Stephen W. Peterson[1], Katie Cole[1]

[1]Department of Physics, University of Cape Town, Rondebosch, 7700, South Africa.
[2]Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch, 7700, South Africa.

E-mail: PRNMOG001@myuct.ac.za

Abstract. Positron emission particle tracking (PEPT) is used to track the motion of tracer particles in the mineral separation process of froth flotation. Previous measurements with PEPT were performed at a target location rate of 1.0 kHz due to the 1.0 ms precision of the timestamp recorded by the Siemens ECAT "EXACT3D" HR++ positron emission tomography (PET) scanner at the PEPT Cape Town facility in South Africa. Tracking at the 1.0 kHz location rate reduces the detail in the high frequency component of the Lagrangian path. This work introduces a new time interpolation algorithm that increases the location rate of PEPT measurements with the HR++ to 5.0 - 10.0 kHz. The algorithm also decreases the uncertainty in the time of the line-of-response (LOR) and the uncertainty in the reconstructed path, when comparing the simulation input path to the tracked output path.

1. Introduction

Positron emission particle tracking (PEPT) [1] is used to track the motion of tracer particles in applications such froth flotation from the mineral processing industry [2, 3]. PEPT measurements provide detailed Lagrangian paths of the motion of the tracer particle [2] as well as the time-averaged Eulerian flow properties under different experimental conditions [4]. Simulation tools have emerged, such as GEomtry ANd Tracking 4 (Geant4) Application for Tomographic Emission (GATE) [5], to perform optimisation [6] and modelling [7, 8] of PEPT measurements with a known input path, with the aim of producing improved data analysis techniques for PEPT research.

A PEPT measurement is performed by radiolabelling a tracer particle with a positron emitting nuclide like ^{68}Ga (e.g. [9, 10]). Positrons from the tracer annihilate with local electrons, which emit a pair of almost colinear 511.0 keV gamma rays. If two gamma rays are detected with temporal coincidence by the detectors of a PET scanner, then these detected positions are recorded sequentially in the listmode data file which is the output of the PET scanner. A virtual line-of-response (LOR) is formed by connecting the two gamma ray interaction positions, and with multiple LORs, the location of the tracer particle can be reconstructed with an algorithm such PEPT-ML [11].

PEPT Cape Town is a specialist centre for PEPT at the University of Cape Town and houses the Siemens ECAT "EXACT3D" HR++ positron emission tomography (PET) scanner (referred to here as the HR++). The HR++ records time values for a PEPT measurement by

interspersing a timestamp every 1.0 ms in the continuous feed of detection positions. Previous measurements performed by Cole et al. [3] set a target location rate of 1.0 kHz to correlate to the 1.0 ms timestamp output of the HR++. However in practice, this limits the temporal resolution of the output Lagrangian path. This is a result of attempting to track at a higher location rate to follow the higher frequency fluctuations in the tracer motion (where there are sufficient LORs per location to do so). This will result in multiple tracked locations having the same time value which has an overall impact of increasing the uncertainty in the path. This work will use a validated GATE simulation of the HR++ [8] to generate PEPT data with a ground truth to test a new time interpolation algorithm to increase the maximum tracking location past 1.0 kHz.

2. Methods

2.1. Data generation

A validated GATE simulation of the HR++ was developed by Perin et al. [8] and used to simulate PEPT measurements. The tracer particle used in all tests was an NRW-100 ion exchange resin bead [9] of diameter 300 µm with 1.0 mCi (37 MBq) of radiolabelled ^{68}Ga. Figure 1 shows an image of the GATE simulation geometry of the HR++.

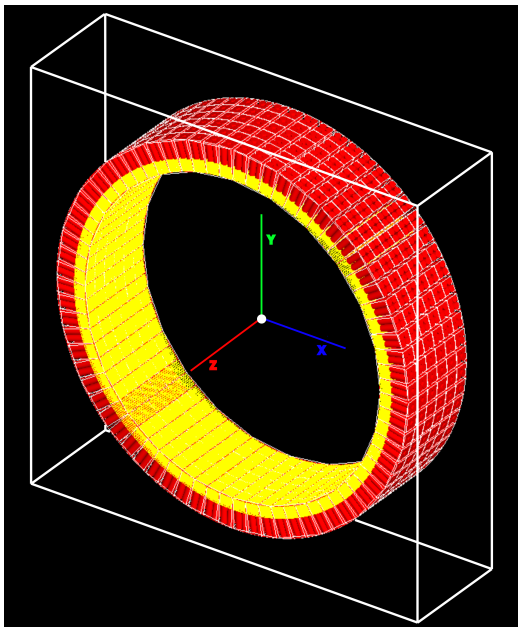


Figure 1. Geometry of the GATE simulation HR++ PET scanner [8]. The white wireframe is the world volume, the red cylinders are the photomultiplier tubes and the yellow regions are the bismuth germanate (BGO) crystals. The Z dimension represents the axial field-of-view (FOV) while the XY plane represents the transaxial FOV.

A pseudo random walk model was created to simulate the path of a tracer particle in the GATE model that is characteristic of motion in a turbulent fluid. The path was formed from a series of steps in each of the three dimensions (X, Y, Z) per unit time as follows,

$$\begin{aligned} S_i \sim \mathcal{N}(0.0, \gamma), \quad i \in (X, Y, Z), \\ X(t+\tau) = X(t) + S_X, \\ Y(t+\tau) = Y(t) + S_Y, \\ Z(t+\tau) = Z(t) + S_Z, \end{aligned} \tag{1}$$

where S_i is the spatial step size in each of the X, Y and Z dimensions which are sampled from a Normal distribution with a mean of 0.0 mm and a standard deviation of γ mm. The previous location in the path is $[X(t), Y(t), Z(t)]$ at time t and $[X(t+\tau), Y(t+\tau), Z(t+\tau)]$ is the next step in the path, at a time $(t+\tau)$ where τ is the time step and is a constant defined by the user. The path is bounded within a cylinder of radius 300.0 mm and height 180.0 mm to make sure that the tracer does not move outside the field-of-view (FOV). Figure 2 shows an example of a path produced by the random walk generator where $\gamma = 1.0$ mm and $\tau = 1.0$ ms.

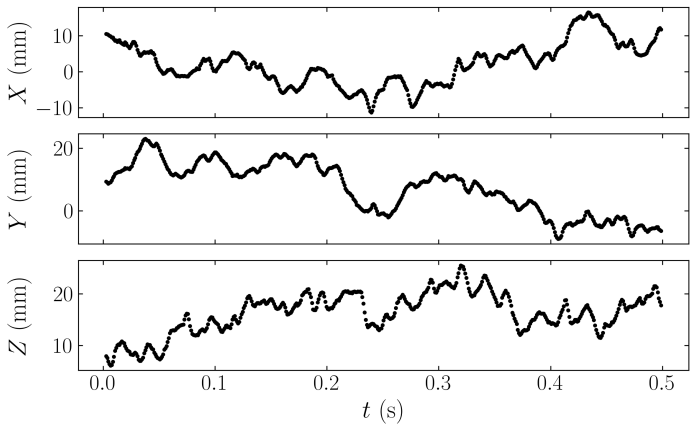


Figure 2. Three dimensions X, Y and Z with time t of a path generated using the pseudo random walk generator with a spatial spread of $\gamma = 1.0$ mm and a time step of $\tau = 1.0$ ms.

2.2. Interpolation algorithm

The impact of the timestamp every 1.0 ms in the listmode data of the HR++ PET scanner can be seen in Table 1 in the "Listmode" column. The new time interpolation algorithm linearly interpolates over all LOR times with the same timestamp which evenly distributes the time of each LOR between each timestamp in the listmode times, as seen in Table 1 in the "Interpolated" column. Interpolation of the listmode times has been previously proposed by Hoffmann et al. [12]. The HR++ GATE simulation can be used to validate the interpolated LOR times by comparing the interpolated LOR times to the simulation LOR times.

3. Results

Figure 3 shows the difference between the simulation LOR times and both the listmode and interpolated LOR times for a GATE simulation of the tracer moving on a random walk.

The time difference between the interpolated and simulation times $\Delta(t)_{i,s}$ shows a significantly narrower distribution for the error in the LOR times as compared with $\Delta(t)_{lm,s}$, the time difference between the listmode and simulation times, which has a uniform distribution with limits between -1 and 0. The distribution of $\Delta(t)_{lm,s}$ results in a standard uncertainty in the LOR time of (maximum time error - minimum time error)$/2\sqrt{3} = 0.29$ ms. The large deviations in the LOR times arise from how the LOR time is assigned to the immediately prior timestamp from the HR++. The distribution of the interpolated time difference $\Delta(t)_{i,s}$ is described by a double exponential function whose two fit parameters were found to be the centre, μ = (-7.3 $\pm$ 1.0)$\times 10^{-4}$ ms, and diversity, $\beta = 0.02014 \pm 0.00014$ ms. The standard uncertainty of the double exponential distribution is $\sqrt{2}\beta = 0.028$ ms. This shows that the interpolation algorithm significantly reduced the uncertainty in the time of each LOR recorded by the HR++.

Table 1. An example of the output of the LOR time interpolation algorithm. The "Simulation" column shows the true interaction time obtained from the simulation, the "Listmode" column shows the recorded time in the listmode data, and the "Interpolated" column shows the new times obtained from the interpolation of the "Listmode" column.

LOR number	Simulation (ms)	Listmode (ms)	Interpolated (ms)
1	0.00	0.00	0.00
2	0.20	0.00	0.25
3	0.49	0.00	0.50
4	0.76	0.00	0.75
5	1.02	1.00	1.00
6	1.29	1.00	1.33
7	1.60	1.00	1.66
8	2.10	2.00	2.00
9	3.20	3.00	3.00

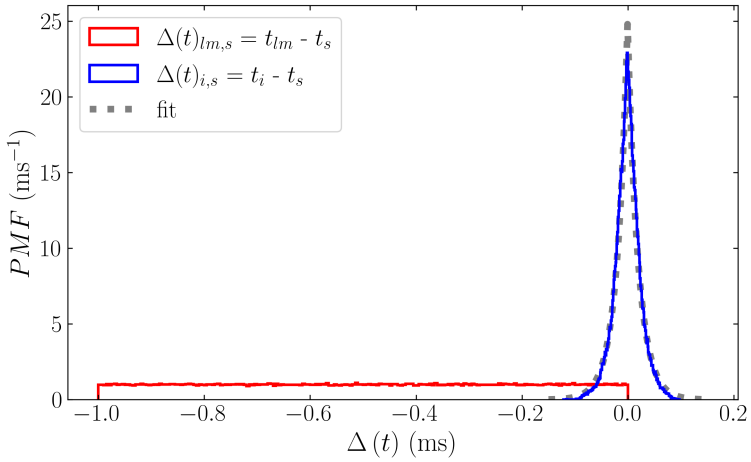


Figure 3. Probability mass function (PMF) of the time difference $\Delta(t)$ between both the listmode (t_{lm}) and interpolated (t_i) LOR times with the simulated LOR times (t_s) for the path shown in Figure 2. The LOR rate was 377 ± 19 kHz and for the fit the reduced χ^2 was 0.28.

Multiple, different, pseudo random walk paths were created as input to the GATE simulation of the PEPT measurement with the time step τ varying from 0.08 ms to 1.0 ms. To optimise the application of the PEPT-ML algorithm to find the tracer locations, the parameters were chosen so that the location rate was $L \approx \tau^{-1}$. The path uncertainty in each dimension was,

$$u(R) = \frac{\sum_{j=0}^{n} |R_i(j) - R_t(j)|}{n}, \quad R \in (X, Y, Z), \tag{2}$$

where n was the number of tracked locations in the path, R_i was the input path dimension and R_t was the tracked path dimension [11]. The path uncertainties in the X, Y and Z dimensions were combined in quadrature, $\Delta(R) = \sqrt{u(X)^2 + u(Y)^2 + u(Z)^2}$, to find the total path uncertainty $\Delta(R)$.

Figure 4 shows the path uncertainty for each of the simulated PEPT measurements with varying τ. In Figure 4(a), the path uncertaintes in both the listmode and the interpolated time

values for a series of locations show a trend for increasing $\Delta(R)$ with decreasing time step. The interpolated time values lead to a path with smaller path uncertainty in general; closer to the size of the tracer particle (300 μm). Figure 4(b) shows the random walk with a time step $\tau = 0.1$ ms and a spatial size of $\gamma = 1.0$ mm to illustrate that the path with interpolated times follows high frequency fluctuations in the input path with higher fidelity than the listmode times.

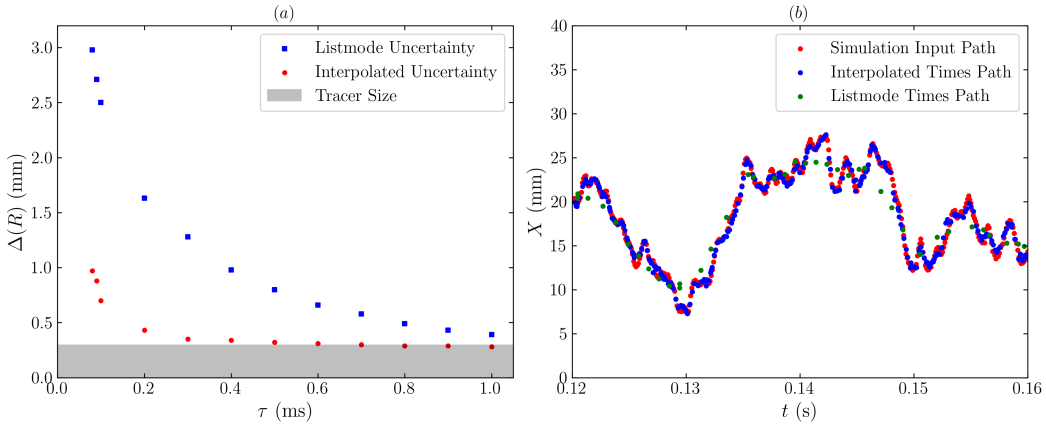


Figure 4. (a) The path uncertainty $\Delta(R)$ for both the listmode and interpolated times as a function of the time step τ of the random walks. The size of the tracer is shown with the shaded region. (b) The simulation input path, the tracked path using the listmode times and the tracked path using the interpolated times in the X dimension with time t. The PEPT-ML tracking parameters can be found in Section 6.

For the types of paths illustrated, this suggests the new maximum location rate is 5.0 - 10.0 kHz which is a factor of 5 - 10 times the previous location rate obtainable using the listmode time values from the HR++.

4. Conclusion

Previous experiments with the HR++ [3] restricted the location rate when tracking with PEPT to 1.0 kHz. A algorithm was proposed to interpolate the time values of the LORs in the listmode data showed that measurements at higher location rates are feasible. Location rates between 5.0 - 10.0 kHz were achieved for the motions presented, which is a factor of 5 to 10 times larger than achieved previously. This reduces the uncertainty in the LOR times to 0.028 ms, as compared to 0.29 ms for the listmode times. This allows for more detailed Lagrangian paths to be reconstructed from the raw listmode data. For the PEPT research community, the implementation of this new time interpolation algorithm will improve PEPT measurements of tracer particles in challenging applications, such as particles undergoing higher accelerations as found in centrifuge pumps and in stirred tanks.

Future work with the time interpolation algorithm entails investigating the effect on the time-averaged velocity distributions of the tracer particle, the effect of tracking in media of varying attenuation and in FOV regions with changing detection efficiency.

5. Acknowledgements

This research was funded by the University Research Council at the University of Cape Town and the National Research Foundation (NRF) through their DSI-NRF First Time Doctoral Scholarships.

6. Appendix

The tracking parameters for PEPT-ML were: sample size 1 ($SS1$) = 320, overlap 1 ($OL1$) = 0, maximum distance (MD) = 0.8 mm, true fraction 1 ($TF1$) = 0.25, sample size 2 ($SS2$) = 5, overlap 2 ($OL2$) = 4 and true fraction 2 ($TF2$) = 0.6 when tracking using the listmode times where the location rate was a constant 1.0 kHz. The tracking parameters for the random walk using the interpolated times can be seen in Table 2. All tracking parameters were optimised using Gaussian processes.

Table 2. The PEPT-ML tracking parameters for the random walk paths using the interpolated times.

τ (ms)	$SS1$	MD (mm)	$TF1$	$SS2$	$OL2$	$TF2$
0.08	34	0.60	0.20	5	4	0.60
0.09	39	0.60	0.25	5	4	0.60
0.10	45	0.60	0.25	5	4	0.60
0.20	65	0.80	0.25	5	4	0.60
0.30	100	0.80	0.25	5	4	0.60
0.40	130	0.80	0.25	5	4	0.60
0.50	158	0.80	0.25	5	4	0.60
0.60	200	0.80	0.25	5	4	0.60
0.70	231	0.80	0.25	5	4	0.60
0.80	268	0.80	0.25	5	4	0.60
0.90	304	0.80	0.25	5	4	0.60
1.00	340	0.80	0.25	5	4	0.60

References

[1] Parker D, Broadbent C, Fowles P, Hawkesworth M and McNeil P 1993 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **326** 592–607

[2] Waters K, Rowson N, Fan X, Parker D and Cilliers J 2008 *Minerals Engineering* **21** 877–882

[3] Cole K, Barker D J, Brito-Parada P R, Buffler A, Hadler K, Mackay I, Mesa D, Morrison A J, Neethling S, Norori-McCormac A, Shean B and Cilliers J 2022 *MethodsX* **9** 101680 ISSN 2215-0161

[4] Boucher D, Jordens A, Sovechles J, Langlois R, Leadbeater T W, Rowson N A, Cilliers J J and Waters K E 2017 *Minerals Engineering* **100** 155–165

[5] Sarrut D, Bała M, Bardiès M, Bert J, Chauvin M, Chatzipapas K, Dupont M, Etxebeste A, Fanchon L M, Jan S *et al.* 2021 *Physics in Medicine & Biology* **66** 10TR03

[6] Herald M, Sykes J, Parker D, Seville J, Wheldon T and Windows-Yule C 2023 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **1047** 167831

[7] Herald M, Wheldon T and Windows-Yule C 2021 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **993** 165073 ISSN 0168-9002

[8] Perin R, Cole K, van Heerden M R, Buffler A, Lin Y Y, Zhang J, Brito-Parada P R, Shock J and Peterson S W 2023 *Applied Sciences* **13** 6690

[9] Cole K, Buffler A, Cilliers J, Govender I, Heng J, Liu C, Parker D, Shah U, van Heerden M and Fan X 2014 *Powder Technology* **263** 26–30 ISSN 0032-5910

[10] Leadbeater T, Buffler A, van Heerden M, Camroodien A and Steyn D 2023 *Nuclear Science and Engineering* 1–17

[11] Nicuşan A and Windows-Yule C 2020 *Review of Scientific Instruments* **91**

[12] Hoffmann A C, Dechsiri C, Van de Wiel F and Dehling H 2005 *Measurement Science and Technology* **16** 851

Emulated gamma spectroscopy with simulation and machine learning-based detection and analysis

C J Buckton[1], S M Wyngaardt[1], M Ngxande[2] and J J Van Zyl[1]

[1]The Department of Physics, Stellenbosch University, Private Bag X1, 7602 Matieland, South Africa

[2]The Computer Science Division, Dept of Mathematical Sciences, Stellenbosch University, Private Bag X1, 7602 Matieland, South Africa

E-mail: 19825706@sun.ac.za

Abstract. Modern student training in experimental nuclear physics relies heavily on the presence and operation of physical radioactive substances, whether for data capturing or analysis. Providing a smart system which can both simulate the radiation monitoring and source emission process, as well as the data analysis, would be useful as a safe alternative to basic training level nuclear experiments with numerous educational possibilities. The work presented here is the concept and design of a fully automated spectroscopy emulation system. It involves training and testing a deep learning-based intelligent detection model which will be used to detect, analyse and provide feedback on data collected from a simulated spectroscopy experiment (e.g. standard calibration and radioactive decay experiments for students). Emulation hardware will be combined with an efficient GEANT4-based simulation [1] and neural network as the detection model. Additionally, for the purpose of future physical or simulated experimentation, the idea is to also develop a large, comprehensive data repository using the simulation. Several factors and parameters including choice of simulated detector, source selection, geometry, etc. are chosen to represent a basic experimental setup with a compact scintillator, but which can also be updated in the future to simulate more complex experimental setups.

1. Introduction

With any practical application of high or low energy gamma spectroscopy in nuclear physics, comes the inherent danger of working with radioactive substances, and safety precautions which students in the field of nuclear physics may not be familiar with, particularly during the training phase of such experimentation. The idea to use emulation hardware with custom simulated gamma spectroscopy and a machine learning algorithm comes primarily from the goal of reducing risk and increasing safety through complete simulation of the source and detection process, removing the need for physical radioactive sources and shielding. The simulation of spectroscopy in this way is combined with an emulator, for providing a true signal based on the inputs of the simulator and also for handling specific processes and hardware not included in the simulation. Additionally, there is the option of including a deep-learning neural network for automated analysis and detection, depending on the nature of the experiment (e.g. for training students, having them perform analysis and then compare to the prediction of the neural network or evaluation of the supervisor). The complete emulation package concept can be seen in Figure 1:

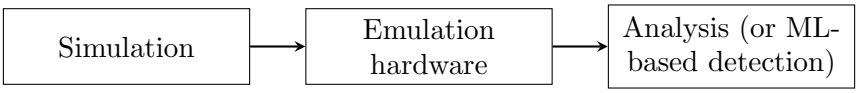


Figure 1. Processing steps from the start of the simulation to the detection/analysis.

The simulation will encompass all radioactive decay and source generation, as well as simulating the environment and core detector geometry (e.g. just the scintillation material in this case). Thereafter, the simulation will provide input to the emulation system to generate a signal, while emulating the interactions following scintillation, including the photo-multiplying tube and cathode. In a sense, the emulator is used to improve upon the simulation by accounting for physical processes not being simulated. The analysis is performed either individually or by using an automated machine-learning algorithm, which takes the resulting spectrum from the simulation and emulator and predicts which isotopes are present, through confidence scores.

2. Aims and Objectives

In general, the goal is to:

- Develop a comprehensive repository of real and synthetic data for use in remote or simulated experimental environments.
- Automate the process of detection using custom machine learning models.
- In conjunction with an spectroscopy emulator, use as a risk-free alternative to physical radiation monitoring and spectroscopy experiments.
- Provide a safe and secure environment for training students on related experiments.

Currently, we have developed a simulation which is being used to generate a database of synthetic spectral data, which will be used in future to train a neural network for automated detection. The goal is to eventually include the emulation system in the energy spectrum process, although the simulation itself is considered to be adequate for training and testing purposes of the network.

3. Methodology

3.1. Motivation for Deep-Learning

Convolutional neural networks are highly capable in object recognition and identification or classification problems, and are hence a prime candidate for choice of analysis network, although other models and network architectures are being investigated. The network will learn from a histogram input in a 'counts vs. energy' structure. Of course, it is not just the energy peaks in the spectrum, but also the associated backscatter, Compton scattering, pair production, collision interactions, etc. and possible secondary or tertiary reactions which will need to be picked up by the network. There is the additional issue of signal noise either inherited or due to background sources. There is also a possibility of few signal anomalies appearing in the data, which the model will need to know how to deal with. For this reason, a method will need to be developed to handle this behaviour. Kalman filters and noise reduction algorithms [2] can be developed to treat unexpected or unwanted behaviour in the system. It is also believed that the supervised training process of the network can include examples of noise, to assist in obtaining as clean a spectrum as possible, and to guide the network to the true reaction data.

Ideally, the network will learn to identify the isotopes by peaks and possibly backscatter patterns. This will prevent the inclusion of unanticipated sources confusing the network when making predictions. Of course, the learning process is largely randomised, and so it is often difficult to predict just how effective the approach to the training will be beforehand.

3.2. Synthetic Data Generation

Currently, the GEANT4 simulation toolkit is being considered to provide a large set of spectral data. A simple, but efficient GEANT4-based simulation is used for generation of synthetic spectra. The generated data will need to be standardised in a way that allows for compatibility during the learning process for the neural network. For the purposes of classification and detection, a viable data format would be a histogram of energies for a given source, or combination of sources. Note that since the simulation will also be providing a repository of sorts, the data it generates will need to be compared to real experimental data as a measure of confidence that the generated data can be trusted.

For the case of a Na(Tl) scintillator, a world geometry is defined to contain a cylindrical detector geometry where a volume of Na(Tl) is surrounded by a thin layer of aluminium oxide (Al_2O_3) and has an outer layer of Aluminium as seen in Figure 2.

A particle gun is used to initiate an event with 0 momentum. Naturally, more events means better statistics, but also larger computational cost. Hence, a value of 1 million events is chosen for efficiency, taking into account the fact that the neural network learns the spectral feature representation. That is, it cares more about the overall shape and features than the pure statistics, the former of which will not change significantly by increasing event numbers. The relevant features come from the physics classes included in the GEANT4 simulation, which include accounting for:

- Compton scattering, photoelectric process, backscatter and pair production.
- Optical physics for light ray interactions in the scintillation material.
- Radioactive decay physics for decay of the particle from particle generator.

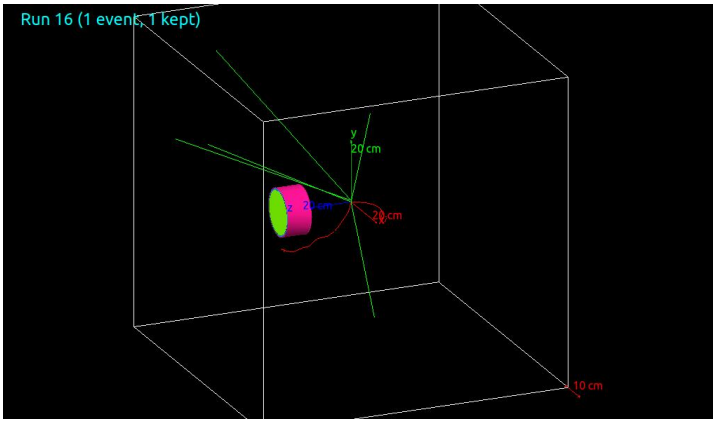


Figure 2. Experimental setup for simulation using an Na(Tl) scintillator (Na(Tl) in green, aluminium oxide in blue, pure aluminium in pink).

Furthermore, to account for the simulation using a scintillator, several other adjustments are made to correct for statistical accuracy:

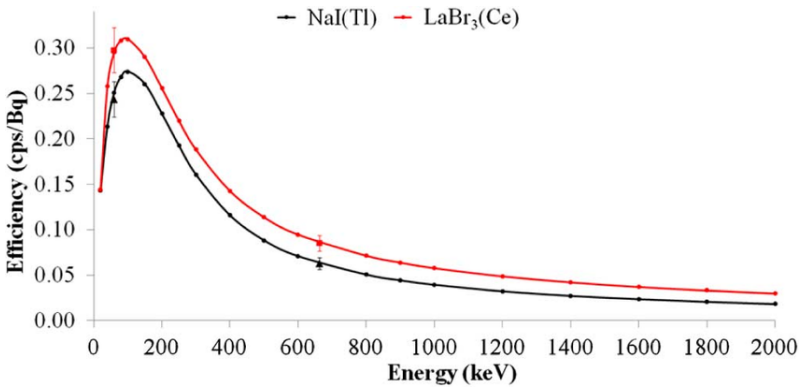


Figure 3. Efficiency curve used in simulation in black. Data points are estimated through inspection. Source: [3]

- An efficiency curve is manually implemented as input, so as to allow higher counts at lower energies, due to the detection efficiency per energy. The curve used specifically for this simulation is shown in Figure 3 above, while in practice the planned detector for the overall system will provide the values.
- To simulate the resolution of a detector, the final energy deposition should resemble a Gaussian curve at the relevant photopeaks, assuming a Gaussian distribution an appropriate detector response. Essentially, the simulated energy deposited `fEdep` is taken to be the center of a randomly sampled Gaussian distribution, with uncertainty

$$\sigma = \frac{R \times \texttt{fEdep}}{2\sqrt{2\ln 2}}. \tag{1}$$

  Here, R is the energy-dependent resolution, which like the efficiency curve can be determined on a per-detector basis. In this case, for the purposes of pure simulation the relationship between energy and detector resolution is taken from an experiment [4] as

$$R(\texttt{fEdep}) = -0.0038 \times \texttt{fEdep} + 9.5101. \tag{2}$$

- A suitable background threshold of 10 keV was chosen.

Finally, the resulting spectra are sorted into classes, one for each isotope, and the neural network will learn in a multi-classification structure, and provide a score per detected source. For consistency, all final training histograms and real data for testing are normalized by

$$\texttt{Relative\_Counts[E}_i\texttt{]}\,(\%) = \frac{\texttt{Counts[E}_i\texttt{]}}{\texttt{max(Counts)}} \times 100, \tag{3}$$

so that the energies are counted as a relative percentage of the maximum count.

4. The Emulator

The emulator itself consists of hardware, and various components necessary to convert the data fed into it by the simulation into a signal which is then made into a spectrum. In a sense, this replaces the physical hardware of a true experiment, such as multi-channel analyzer (MCA), amplifiers and pre-amplifiers, etc. Crucially, the simulation does not account for any physical processes beyond the core of the detector (scintillation material in this case). Processes such as photoelectron ejection through the photomultiplier tube, cathode, anode interactions, etc.

The whole package will be considered as a single emulation suite, which will be benchmarked and improved upon, and is of course subject to change. An added benefit of such a system will be the addition of remote access, allowing for remote controlled experimentation.

5. Testing and Benchmarking

Apart from the testing phase of the model through training, a physical test using a scintillator in a lab will be necessary to ascertain the true performance of the model. Using a NaI(Tl) scintillator and real experimental data, a test will be conducted firstly for the neural network's performance on real data. The overall error rate of the network will be the key factor in determining the viability of the final product. The full emulation suite will be benchmarked against a real experimental setup, and will consist of the simulation, neural network, and emulation hardware.

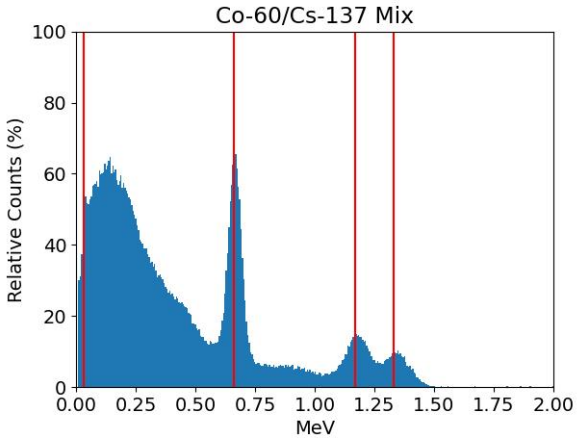


Figure 4. Example histogram of a mix of ^{60}Co and ^{137}Cs for 50 million events.

The simulation already shows promise in generating the necessary characteristics and features of various isotopes to be encoded into the neural network. Eventually the network will be expected to run autonomously and efficiently for real-time detection and prediction of any and all isotopes present in a given sample spectrum. Running the simulation for 50 million events, for example, would show that the parameters model the statistics after a scintillator to good accuracy. Figure 4 shows the overall shape of the data, where the red lines indicate key spectral features from left to right: x-ray fluorescence, the 662 keV ^{137}Cs photopeak, and the 1.17 MeV and 1.33 MeV ^{60}Co photopeaks. Note the significantly higher 662 keV photopeak, which is largely due to detector efficiency at the various photopeaks, seen in Figure 3.

6. Conclusion

Overall, the following ideas from the planned smart emulation system have been presented:

- GEANT4-based simulation with machine learning-based detection, analysis and emulator combined to form gamma spectroscopy emulation tool.
- The goal of testing the proposed smart system in simulation, recognition and ability to analyse.
- The idea to use the emulator to virtually generate full lab experiments for training purposes.
- Generation of large data (real + synthetic) for radioactive decay across many nuclei.
- Proposal of remote access to emulation system for remote experiments and related activities.

Some further remarks and considerations for improvement include, as mentioned before, to use experimentally determined values from the detector to be integrated into the emulation suite. Although convolutional neural networks are the primary focus as the detection and analysis system, alternative solutions and combinations of various models can be investigated for efficiency or performance improvements. The use of different detectors, changes in geometry and other simulation factors can be varied as the dataset increases, which may lead to better generalization of the network. Automating the process of data generation will also be a key time-saving component, for decreasing the run-to-run time and generate large batches of data in shorter amounts of time.

References

[1] Agostinelli S et al. (GEANT4) 2003 *Nucl. Instrum. Meth.* **A506** 250–303
[2] Ma'arif A, Iswanto, Nuryono A and Alfian R 2020 *Sig. Img. Proc. Lett.* **1** 11–22
[3] Casanovas R, Morant J and Salvad o M 2014 IEEE *Trans. Nucl. Sci.* **61**
[4] Akkurt I, Gunoglu K and Arda S S 2014 *Sci. Technol. Nucl. Install.* **2014** 186798 ISSN 1687-6075 URL https://doi.org/10.1155/2014/186798

DIVISION

C

PHOTONICS

Quantum random number generation using an on-chip linear plasmonic beamsplitter

C Strydom and M S Tame

Stellenbosch Photonics Institute, Department of Physics, SU, Matieland 7602, RSA

E-mail: conradstryd@gmail.com

Abstract. True random numbers are essential in cryptography, simulation and many other information processing tasks. Here, we experimentally demonstrate quantum random number generation with an on-chip linear plasmonic beamsplitter to generate true random numbers. The beamsplitter has a footprint of $2\,\mu\text{m} \times 25\,\mu\text{m}$ and is more compact than a previous demonstration, with a reduction in size by a factor of 2, thereby reducing the impact of loss. At the input grating of the beamsplitter, free-space single photons are converted into single surface plasmon polaritons, which propagate along one of two gold stripe waveguides to one of two output gratings where they are converted back into photons. The value of each random bit is determined by the output at which each photon is detected. In our experiment, we achieved a random number generation rate of 2.86 Mbits/s despite the presence of loss. By applying randomness extraction in the form of a deterministic shuffle followed by the recursive von Neumann algorithm to our raw bits, we obtained a sample of bits which passed the ENT and NIST Statistical Test Suites.

1. Introduction

Random numbers are ubiquitous in cryptography, simulation and coordination in computer networks, to name but a few [1]. However, they are very difficult to generate reliably using classical hardware, as the unpredictability relies on incomplete knowledge, which can introduce ordered features and compromise their utility. In contrast, the inherent randomness of quantum mechanics makes quantum hardware ideal for generating random numbers [1]. Studies have confirmed that in many important applications, such as molecular Monte Carlo simulations [2], results obtained with poor quality classical pseudorandom number generators can deviate significantly from those obtained with true quantum random number generators. While high-quality pseudorandom number generators exist, they are extremely resource intensive [3]. This has led to an increasing interest in the development of dedicated quantum hardware for random number generation. In particular, quantum random number generation has been successfully realised experimentally in a great variety of physical settings, ranging from cloud-based superconducting quantum computers [4] to photonic integrated circuits [5].

It was also recently shown that a truly random sample of bits can be generated using an on-chip plasmonic beamsplitter [6]. The main advantage of plasmonic hardware [7] is that it allows light to be confined below the diffraction limit, which enables a significant reduction in device footprints compared to the dielectric hardware conventionally used in photonic integrated circuits. In this work, we investigate quantum random number generation with an on-chip linear plasmonic beamsplitter. Our linear plasmonic beamsplitter is more compact than the cross-shaped plasmonic beamsplitter used in Ref. [6], with a reduction in size by a factor of 2. In

our experiment, free-space single photons focused onto the input grating of the on-chip linear plasmonic beamsplitter randomly couple into one of its two gold stripe waveguides, allowing us to obtain random bits. By applying standard randomness extraction protocols to raw bits generated at a rate of 2.86 Mbits/s, we obtained a sample of processed bits with a marginally reduced effective generation rate of 2.85 Mbits/s which passed industry standard tests. Our work successfully demonstrates the use of a highly compact on-chip plasmonic component for quantum random number generation and will be of interest to researchers developing plasmonic components for this and other important quantum information processing tasks [7].

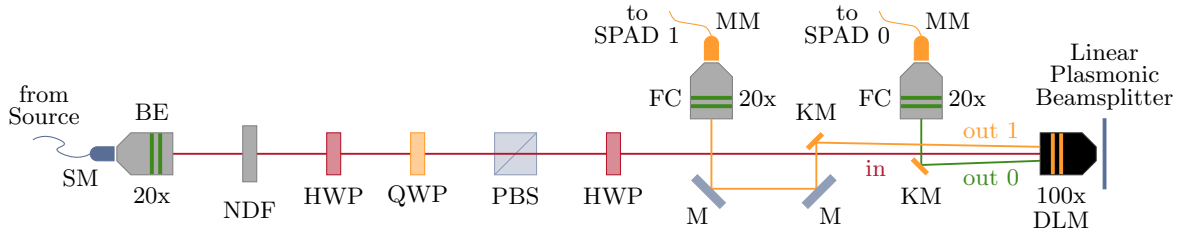
2. Experimental setup

The optical experimental setup for generating random numbers with an on-chip linear plasmonic beamsplitter is shown in figure 1(a) and a diagram of the beamsplitter used in our experiment is shown in figure 1(b). Our on-chip linear plasmonic beamsplitter has a footprint of just $2\,\mu\text{m} \times 25\,\mu\text{m}$, making it a highly compact device. It comprises a central 2-step input grating, with a $2\,\mu$m wide gold stripe waveguide and an 11-step output grating on either end. Each output grating has a period of 740 nm. The gold linear plasmonic beamsplitter is fabricated on a 0.17 mm thick silica glass substrate with a refractive index of $n = 1.5255$ using a combination of electron beam lithography and electron beam evaporation, as described in Ref. [6]. Figure 1(c) shows atomic force microscope (NT-MDT Smena) images of the final fabricated structure.

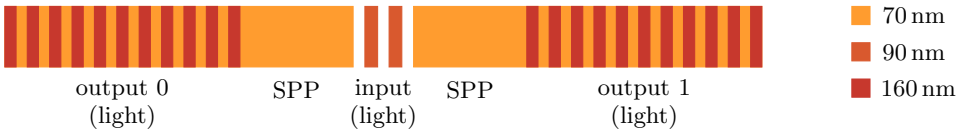
The light source used in our experiment is a continuous-wave laser (Thorlabs LPS-785-FC), with a vacuum wavelength of $\lambda_0 = 785$ nm and a frequency bandwidth of $\delta\nu = 5.36$ THz, operating above the lasing threshold. Polarised coherent laser light is injected into the optical setup in figure 1(a) via a beam expander (BE), which is connected to the continuous-wave laser by a polarisation-preserving single-mode optical fibre (SM). In our setup, the beam is first sent through a neutral density filter (NDF), which attenuates the coherent laser light down to the single-photon level, as will be explained later. The photons then pass through a half-wave plate (HWP), a quarter-wave plate (QWP), a polarising beamsplitter (PBS) and another HWP, which are used to refine and control their polarisation. A 100x diffraction limited microscope (DLM) objective lens is used to focus these free-space single photons onto the input grating of the on-chip linear plasmonic beamsplitter (a diffraction-limited spot of approximately $2\,\mu$m) where they are converted into single surface plasmon polaritons (SPPs). The single SPPs propagate along one of the two gold stripe waveguides of the linear beamsplitter to one of its two output gratings, where they are converted back into photons.

Out-coupled photons from the two output gratings of the beamsplitter are collected by the same DLM objective that was used to focus input photons onto its input grating, and are then reflected into fibre couplers (FCs) by knife-edge mirrors (KMs). Each FC is connected to a single-photon avalanche diode (SPAD) detector (Excelitas SPCM-AQRH-15), with a dead time of $T_d = 24$ ns, by a multi-mode optical fibre (MM). Each SPAD detector is in turn connected to a channel of a Picoquant TimeHarp 260, which can measure the arrival time of a photon at a detector to a precision of 25 ps. The arrival of a photon at one detector or the other results in the generation of a bit (either 0 or 1). Hence, it is the process by which a free-space single photon, focused onto the input grating of the linear plasmonic beamsplitter, randomly couples into one of its two gold stripe waveguides which ultimately allows us to obtain a random bit.

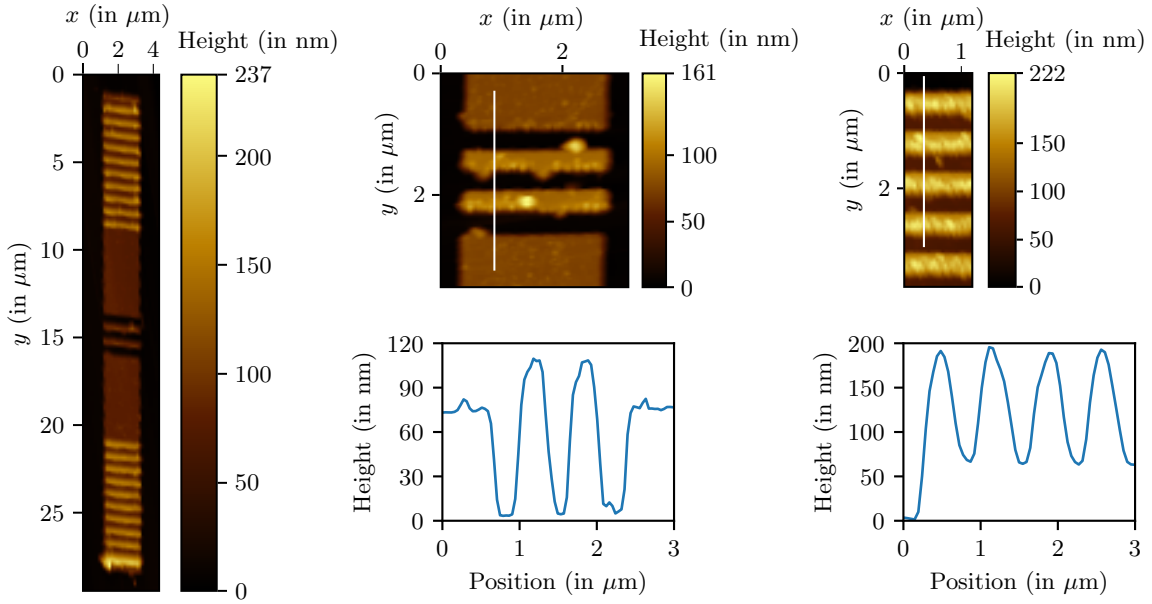
The attenuation of the coherent laser source by the NDF ensures that the linear plasmonic beamsplitter is operating in the single-excitation regime so that single-photon splitting is realised at its input grating. We can model the attenuated light entering the linear plasmonic beamsplitter as a weak coherent state $|\alpha\rangle = e^{-|\alpha|^2/2} \sum_n \frac{\alpha^n}{\sqrt{n!}} |n\rangle$, with mean excitation number $\langle \hat{n} \rangle = |\alpha|^2 \ll 1$ [6]. The photon number distribution is given by $p_n = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!}$, where p_0 and p_1 are the dominant components since $|\alpha|^2 \ll 1$. The use of a SPAD detector at each output of the



(*a*) Optical Setup



(*b*) Linear Plasmonic Beamsplitter



(*c*) Atomic Force Microscope Images

Figure 1. Quantum random number generation using an on-chip linear plasmonic beamsplitter. (*a*) Optical Setup shows the optical experimental setup for generating random numbers with an on-chip linear plasmonic beamsplitter. Legend for labels: single-mode optical fibre (SM), beam expander (BE), neutral density filter (NDF), half-wave plate (HWP), quarter-wave plate (QWP), polarising beamsplitter (PBS), diffraction-limited microscope (DLM), knife-edge mirror (KM), mirror (M), fibre coupler (FC), multi-mode optical fibre (MM), single-photon avalanche diode detector (SPAD). (*b*) Linear Plasmonic Beamsplitter shows a diagram of the on-chip linear plasmonic beamsplitter used in our experiment. (*c*) Atomic Force Microscope Images shows atomic force microscope images of the linear plasmonic beamsplitter (left), its input grating (top centre) and its top output grating (top right), as well as height profiles of its input grating (bottom centre) and its top output grating (bottom right).

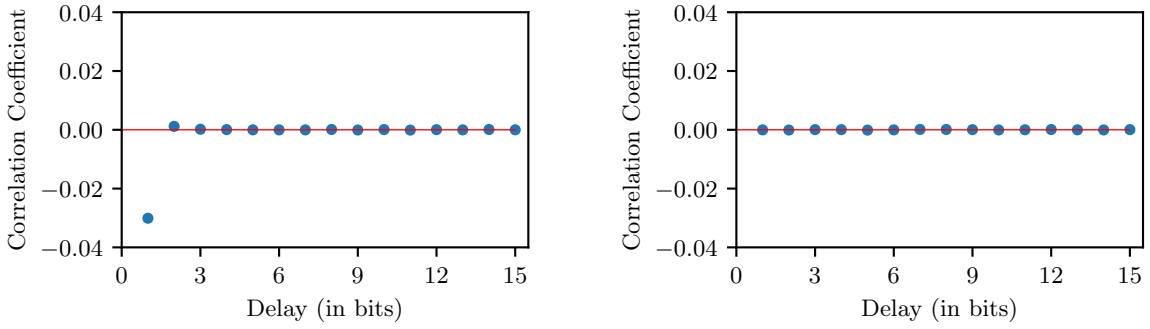
beamsplitter removes the vacuum component p_0 by post-selection so that only the single-photon component p_1 remains (with $p_i \approx 0$ for $i \geq 2$).

The random number generation rate is determined by the light intensity at which the setup is operated. While increasing the light intensity increases the generation rate, there are two important considerations which limit the light intensity [6]. Firstly, to ensure that the mean photon number $\langle \hat{n} \rangle \ll 1$ for light entering the linear plasmonic beamsplitter, the rate at which photons enter the beamsplitter, r, must be much less than the reciprocal of the coherence time, τ, of the source. For our continuous-wave laser, $\tau = \sqrt{2 \ln 2}/\pi\delta\nu = 7.00 \times 10^{-14}\,\mathrm{s}$, and using the method proposed in Ref. [6], we determined that $r = 1.20 \times 10^9\,\mathrm{s}^{-1}$ for our chosen light intensity. Hence $r \ll \frac{1}{\tau} = 1.43 \times 10^{13}\,\mathrm{s}^{-1}$, as required. Secondly, to ensure that the majority of photons arriving at each SPAD detector are detected, the photon detection rate, R, for each detector must be much less than the reciprocal of T_d. For our chosen light intensity, $R = 1.4 \times 10^6\,\mathrm{s}^{-1}$ for both detectors, so that $R \ll \frac{1}{T_d} = 4.2 \times 10^7\,\mathrm{s}^{-1}$.

3. Results

We were able to acquire 171,543,801 bits in 60 s, giving a random number generation rate of about 2.86 Mbits/s. In what follows, we will refer to the first 170 Mbits as the raw sample. The Pearson correlation coefficient [8] of the raw sample with 1-bit to 15-bit delays of itself is plotted in figure 2(a). The Pearson correlation coefficient ranges from -1 to 1, where a positive value indicates a positive correlation and a negative value indicates a negative correlation, while a value close to 0 suggests that no correlation is present. Hence we find that small negative correlations exist between adjacent bits in the raw sample, while short-ranged correlations between non-adjacent bits are negligible. The correlations between adjacent bits can likely be attributed to SPAD detector imperfections [6]. In the raw sample, the relative frequency of zeros is 0.49527 and the relative frequency of ones is 0.50473. Hence the raw sample shows a small bias towards one. This is likely a result of asymmetry in the linear plasmonic beamsplitter and collection optics [6].

We first deterministically rearranged or shuffled the 171,543,801 raw bits generated using our on-chip linear plasmonic beamsplitter in an attempt to remove the correlations between adjacent bits, and then applied the recursive von Neumann algorithm [9] to the shuffled bits in an attempt to remove the bias. In what follows, we will refer to the first 170 Mbits of the resulting sample of 170,925,124 processed bits as the processed sample. The von Neumann scheme constructs an unbiased sample of bits from a biased sample by mapping the bit-pairs 01 and 10, which occur with equal probability in a biased sample in which adjacent bits are uncorrelated, to



(a) raw sample (b) processed sample

Figure 2. Pearson correlation coefficient of the (a) raw sample and (b) processed sample with 1-bit to 15-bit delays of itself.

Table 1. ENT Statistical Test Suite results for the processed sample and expected results for a true random sample.

Statistical Test	Processed Sample	True Random Sample
Entropy	7.999992	8.000000
χ^2 Distribution	73.15%	10–90%
Arithmetic Mean	127.467	127.500
Monte Carlo value for π	3.14273000	3.14159265
Serial Correlation Coefficient	0.000114	0.000000

Table 2. NIST Statistical Test Suite results for the processed sample. 'Tested' shows the number of equal-length sequences into which the 170 Mbit processed sample was divided for a test. 'Threshold' shows the minimum number of sequences which must pass a test for the processed sample to pass. 'Passed' shows the number of sequences which passed a test. The p-value for a test quantifies the uniformity of the distribution of the test results obtained for the different sequences and must be greater than 0.0001 for the processed sample to pass. Medians of test results are given for tests which comprise more than five subtests (marked with *).

Statistical Test	Tested	Threshold	Passed	p-value
Frequency	1700	1670	1683	0.126572
Block Frequency	1700	1670	1686	0.955304
Cumulative Sums 1	1700	1670	1681	0.685984
Cumulative Sums 2	1700	1670	1683	0.969205
Runs	1700	1670	1678	0.873253
Longest Run of Ones	1700	1670	1682	0.334538
Binary Matrix Rank	340	331	340	0.594330
Discrete Fourier Transform	1700	1670	1677	0.366918
Non-overlapping Template*	1700	1670	1681.5	0.555022
Overlapping Template	170	164	169	0.149743
Universal Statistical	170	164	169	0.144842
Approximate Entropy	340	331	339	0.594330
Random Excursions*	170	104	109	0.709445
Random Excursions Variant*	170	104	109	0.367453
Serial 1	170	164	169	0.839406
Serial 2	170	164	168	0.849412
Linear Complexity	170	164	166	0.751633

the bits 0 and 1, respectively. In the standard von Neumann algorithm [10], the bit-pairs 00 and 11 are simply discarded, resulting in a substantial reduction in sample size and effective generation rate. In the recursive von Neumann algorithm [9] employed here however, these bit-pairs are used to construct additional biased samples to which the von Neumann scheme is recursively applied, and the reduction in the effective generation rate is kept to a minimum. We indeed find that randomness extraction only marginally reduced the effective generation rate to about 2.85 Mbits/s. Figure 2(b) confirms that our deterministic shuffle successfully removed the correlations between adjacent bits. The relative frequency of zeros and ones in the

processed sample is 0.50006 and 0.49994 respectively, which confirms that the von Neumann scheme successfully removed the bias.

The quality of the processed sample was analysed further using the ENT [11] and NIST [12] Statistical Test Suites. In the ENT Statistical Test Suite, five key statistical quantities are determined for the sample in question, and the values obtained are compared to the expected values for a true random sample. As can be seen in table 1, the values obtained for the processed sample show excellent agreement with the expected values for a true random sample. The NIST Statistical Test Suite comprises fifteen rigorous industry standard tests, which are mainly aimed at assessing a random number generator's suitability for use in cryptographic applications. Detailed NIST test results for the processed sample are presented in table 2. The processed sample passed the NIST Statistical Test Suite, which confirms that, with standard post-processing, our linear plasmonic beamsplitter can be used to generate random numbers of sufficient quality to meet the stringent requirements of cryptographic applications.

4. Conclusion

We successfully demonstrated quantum random number generation using an on-chip linear plasmonic beamsplitter. The linear plasmonic beamsplitter used in our experiment is more compact than the plasmonic beamsplitter used in previous work [6], with a reduction in size by a factor of 2, thereby reducing the impact of loss. In our setup, the value of each random bit is determined by the waveguide into which each free-space single photon couples at the input grating of the linear plasmonic beamsplitter. Raw bits, generated at a rate of 2.86 Mbits/s, initially showed small correlations between adjacent bits and a small bias towards one, due to detector imperfections and asymmetry in the beamsplitter and collection optics. However, by employing standard randomness extraction protocols, we obtained a sample of high-quality processed bits, with a marginally reduced effective generation rate of 2.85 Mbits/s, which passed the ENT and NIST Statistical Test Suites. Future work could involve integrating an on-chip source and on-chip detectors into our linear plasmonic beamsplitter to produce a highly compact, fully self-contained plasmonic quantum random number generator chip.

Acknowledgments

We acknowledge the use of the services of NTT-AT Japan for the fabrication of the on-chip linear plasmonic beamsplitter. This research was supported by the Harry Crossley Foundation, the University of Stellenbosch and the South African Quantum Technology Initiative (SAQuTI) through the Department of Science and Innovation of South Africa.

References

[1] Herrero-Collantes M and Garcia-Escartin J C 2017 *Rev. Mod. Phys.* **89** 015004
[2] Ghersi D, Parakh A and Mezei M 2017 *J. Comput. Chem.* **38** 2713–20
[3] Tian X and Benkrid K 2009 *Proc. of the 2009 NASA/ESA Conf. on Adaptive Hardware and Systems* (Washington, DC: IEEE Computer Society) pp 460–4
[4] Strydom C and Tame M S 2021 *The Proc. of SAIP2021, the 65th Annual Conf. of the South African Institute of Physics* (Pretoria: SAIP) pp 630–5
[5] Bai B, Huang J, Qiao G R, Nie Y Q, Tang W, Chu T, Zhang J and Pan J W 2021 *Appl. Phys. Lett.* **118** 264001
[6] Francis J T, Zhang X, Özdemir Ş K and Tame M S 2017 *Quantum Sci. Technol.* **2** 035004
[7] Tame M S, McEnery K R, Özdemir Ş K, Lee J, Maier S A and Kim M S 2013 *Nat. Phys.* **9** 329–40
[8] Edwards A L 1976 *An Introduction to Linear Regression and Correlation* (San Francisco, CA: Freeman) chapter 4 pp 33–46
[9] Peres Y 1992 *Ann. Stat.* **20** 590–7
[10] von Neumann J 1951 *Natl Bur. Stand. Appl. Math. Ser.* **12** 36–8
[11] Walker J 2008 https://www.fourmilab.ch/random/
[12] Rukhin A *et al.* 2010 https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-22r1a.pdf

Quantum Phase-based Plasmonic Biosensing for Enhanced COVID-19 Detection

Kelvin T. Mpofu [1*] and Patience Mthunzi-Kufa [1,2,3]

[1]Council for Scientific and Industrial Research (CSIR) , National Laser Centre , Pretoria , South Africa.
[2]School of Chemistry and Physics , University of KwaZulu-Natal , Durban 4041, South Africa.

[3]Orthopaedic Biomechanics Lab , Division of Biomedical Engineering , Department of Human Biology , Faculty of Health Sciences , University of Cape Town , Anzio Road , Observatory South Africa.

E-mail: kmpofu@csir.co.za*

Abstract. The recent COVID-19 pandemic has made clear the need for rapid and sensitive diagnostic tools to allow effective monitoring and control of the disease. In this study, we theoretically model an approach to COVID-19 detection by employing quantum phase-based surface plasmon resonance biosensing, with the intent to enhance the limit of detection measurement compared to its classical equivalent. We demonstrate a theoretical framework of a quantum plasmonic biosensor, designed to target the SARS-CoV-2 spike protein with high specificity. In this work we model and simulate the operation of the biosensor in an ideal noiseless setup as well as in a noisy setup which more realistically resembles the conditions in a lab. The modeled sensor explores the advantages of quantum phase sensitivity and surface plasmon resonance to achieve a precision level below the shot noise limit. The results show that a quantum plasmonic biosensor could potentially outperform its classical counterparts in terms of LOD in the absence of noise but does not perform as well when there is noise in the system, offering rapid and precise identification of viral presence at very low concentrations. The quantum state considered in this work is the NOON state, but this work opens up the potential to work with other quantum states such as squeezed states and Fock states. This work has the potential to lead to more precise optical diagnostic devices and pave the way for more effective public health strategies to combat future pandemics.

1. Introduction

The ongoing and continuous evolution of viruses and bacteria is a global concern, as it results in pandemics and epidemics [1]. Of late the world has been under attack by COVID-19. Infection with the COVID-19 coronavirus causes serious illness. This is mainly due to the response of the immune system of the patient or host, which results in the release of a large number of pro-inflammatory cytokines [2, 3, 4, 5]. This "cytokine storm" leads to intense inflammatory and immune reactions, this occurs mostly in the lungs, which results in acute respiratory failure or distress. Widespread vaccination has been instrumental in reducing the number of infections and hospitalizations, effectively alleviating the burden of COVID-19 [6]. The use of vaccines has played an essential role in mitigating infections, thus preventing disease-related deaths and hospitalizations, and is helping to control its spread. In addition to vaccination and protection strategies, early diagnosis of COVID-19 infected individuals is crucial for pandemic management.

Some examples of methods that are used to detect SARS-CoV-2 infection are molecular methods [7] and lateral flow-based methods [8]. Molecular biology based tests, for example polymerase chain reaction (PCR) tests, typically are very sensitive and specific in detecting viral RNA and are mostly recommended for symptomatic individuals and implementing public health measures [7]. Lateral flow-based antigen rapid detection assays detect viral proteins and, though less sensitive than molecular tests, are advantageous because they are affordable, fast and easily performed by individuals. These tests can be used to screen high-risk individuals, protect vulnerable populations, ensure safe travel and resume activities, and promote economic recovery. In order to enhance and extend regular disease detection and testing it is necessary to focus on the development of fast techniques, which require low-infrastructure tests or allow for self-testing with a sensitivity comparable to that of PCR testing. Diagnostic tests have played and will continue to play a crucial role in the transition from pandemic response to pandemic control. While the aforementioned methods have proven extremely helpful in the diagnosis of COVID-19, optical biosensors should also be taken into consideration. This is especially true given their ability to detect the virus quickly and accurately as well as their ability to improve these qualities by utilizing quantum resources. In light of the recent COVID-19 pandemic, the need for highly sensitive biosensors is highlighted. Optical biosensors are a good candidate because they typically offer rapid, highly sensitive, and precise detection. Plasmonic-based optical biosensors typically have high sensitivity and, consequently, low limit of detection (LOD) compared to other optical biosensors [9]. The incorporation of quantum optics can enhance their sensitivity and precision.

The plasmonic setup of interest in this work is known as the Kretschmann configuration [10, 11]. It has been incorporated in commercial products such as Biacore devices; however, these devices operate on an angular mechanism whose detection limit cannot be broken with the use of quantum states, unlike intensity and phase-based mechanisms [12, 13]. Plasmonic biosensors can work in an intensity-sensing approach, a wavelength-sensing approach, an angular-sensing-based approach, and a phase-sensing-based approach [14]. In phase-based plasmonic biosensors [15, 16, 17, 18, 19, 20], a phase shift is used to infer the presence of specific molecules in the sample. This work focuses on the simulation of phase-based plasmonic biosensors. In the phase-based surface plasmon resonance (SPR) biosensor, a Krestchmann configuration-based plasmonic biosensor is integrated into one of the Mach-Zender interferometer arms. Although classical phase-based SPR biosensors are highly sensitive, their limit of detection (LOD) is bound by the shot-noise limit (SNL) imposed on the coherent state of light. When quantum states of light are used as probe input states in the SPR biosensor, the biosensor is generally referred to as a quantum plasmonic biosensor. The LOD of phase-based optical biosensors in the context of the refractive index, n, can be calculated using the equation (the error-propagation formula, derived using a truncated taylor series expansion) Eq. 1 below,

$$\Delta n = \Delta\phi |\frac{\partial\phi}{\partial n}|^{-1} \tag{1}$$

where n is the refractive index on the biosensor surface, ϕ is the phase measured in the biosensing setup, Δn is the precision or LOD in the context of the refractive index and $\Delta\phi$ is the precision in the phase measurement or the LOD in the context of the phase. The derivative, $|\frac{\partial\phi}{\partial n}|$, measures the sensitivity of the biosensor, which is a classical property. In the error-propagtion formula we take the inverse of the derivative. The values for $\Delta\phi_{coh}$ which is the phase precision for the coherent state and $\Delta\phi_{\mathrm{NOON}}$ which is the phase precision for the coherent state, come from the literature [21, 22]. The optimized value for $\Delta\phi_{coh}$ is given by Eq. 2

$$\Delta\phi_{coh} = \frac{\sqrt{\eta_1} + \sqrt{\eta_2}}{2\sqrt{N\eta_1\eta_2}}. \tag{2}$$

Where η_1 and η_2 are losses introduced into the two modes (reference mode and signal mode respectively, shown in Figure. 2) of the Mach-Zender interferometer and N is the photon number. The optimized value for $\Delta\phi_{\mathrm{NOON}}$ is given by Eq. 10

$$\Delta\phi_{\mathrm{NOON}} = \frac{{\eta_1}^{\frac{N}{2}} + {\eta_2}^{\frac{N}{2}}}{2N\eta_1{\eta_2}^{\frac{N}{2}}}. \tag{3}$$

Where η_1 and η_2 are losses introduced into the two modes of the Mach-Zender interferometer and N is the photon number. Certain quantum states of light, such as the Fock state and squeezed states, have been shown to enhance detection below the SNL in an intensity-based SPR sensing setup [23, 24, 22, 25, 26, 27, 28, 29]. In this work, the effect of the NOON state on enhancing the detection precision below the SNL is considered. This work shows how the use of quantum states of light in SPR biosensing has the potential to give an enhancement over the use of classical states, and some limitations that may need to be overcome. Quantum SPR biosensors represent the future of optical biosensing.

In SPR-based biosensing the reflection coefficient given by r (shown in Eq. 4) [10].

$$r = \frac{e^{i2k_2d}r_{23} + r_{12}}{e^{i2k_2d}r_{23}r_{12} + 1}, \tag{4}$$

where

$$r_{uv} = \left(\frac{k_u}{\epsilon_u} - \frac{k_v}{\epsilon_v}\right) \Big/ \left(\frac{k_u}{\epsilon_u} + \frac{k_v}{\epsilon_v}\right). \tag{5}$$

Here, d is the thickness of the metal film, $k_i = \sqrt{\epsilon_i}(\omega/c)[1 - (\epsilon_1/\epsilon_i)\sin^2\theta_{in}]^{1/2}$ is the normal-to-surface component of the wave-vector in the i-th layer, ω is the angular frequency of the light in the signal mode and ϵ_i is the respective permittivity, where $\epsilon_1 = \epsilon_p$, $\epsilon_2 = \epsilon_m$ and $\epsilon_3 = \epsilon_a$ for the layers. The reflection coefficient, r, has an implicit time-dependance and can be expressed as $r(t)$, where t is time. The time component is actually embedded in the refractive index on the sensor surface, $n(t)$, which changes as binding interactions occur and is expressed as $n(t) = \sqrt{\epsilon_3(t)}$ (see Chapter 3 of [22] for details about the derivation of $n(t)$). The reflectance, R, of the light is expressed as $R(t) = |r(t)|^2$. The value of, r(t), is a unit-less ratio of electric field amplitudes hence, R(t) is also unit-less. The refractive index of the prism used in this work was $n_{prism} = 1.5107$, the wavelength $\lambda_0 = 810$ nm, and the gold film thickness, d = 50 nm.

In order to calculate Δn two steps are needed, firstly it is necessary to calculate $\Delta\phi$ and then the derivative of the phase ϕ respect to n_a, the refractive index of the analyte. In the case of the Mach-Zender interferometer the general expressions for $\Delta\phi$ are known or have previously been calculated in other works [22, 30]. In this work we use the result from the Appendix in [22] for the minimum value of $\Delta\phi$ which already includes the losses. To calculate the derivative of ϕ, it is necessary to use Eq. 7. This is because Eq. 7 allows us to rewrite the phase, ϕ, with respect to the reflectivity coefficient which itself has a dependence on the refractive index of the analyte, n_a. From the expression of the phase, including the refractive index of the analyte, we can numerically calculate the derivative of the phase with respect to the analyte. It should also be noted that a time-dependent value of the refractive index is used in the calculations; the derivations of this refractive index are shown in the work done here [22, 28]. To generate $R(t)$ as shown in Figure 1, we follow a transformation shown in [28] where the refractive index on the biosensor surface has a dependence on the angular sensorgram given by, $A(t)$.

$$A(t) = \begin{cases} A_c(1 - e^{-k_s t}) & 0 \le t < \tau \\ A_\tau e^{-k_d(t-\tau)} & t \ge \tau, \end{cases} \tag{6}$$

A_c and A_τ are functions of the concentrations of the sample [28]. The value of A_τ is given by $A_\tau = A_c(1 - e^{-k_s\tau})$. The affinity constant, k_s, is given by $k_s = k_a[L_0] + k_d$ where k_a is the association constant measured in $\mathrm{M^{-1}s^{-1}}$ (per molarity per second), k_d is the dissociation constant measured in $\mathrm{s^{-1}}$ and $[L_0]$ is the initial concentration of the ligand. The kinetic parameters in this work are $k_a = 1.2 \times 10^5 \mathrm{M^{-1}s^{-1}}$ and $k_d = 7.05 \times 10^{-4}\mathrm{s^{-1}}$ and an initial concentration L_0 of 50nM.

2. Sensing model

In this work the kinetic binding interaction between the S1 spike protein and a 20-base aptamer "CFA0688T" from BasePairBio together with the surface chemistry architecture is theoretically modelled. We start by studying the binding interaction between the S1 spike protein and the 20-base aptamer "CFA0688T", which is well documented in the work of Szunerits *et al.* [31]. The rate constants of ligand attachment (k_a) and dissociation (k_d) are the focus of binding kinetics, the affinity constant k_s measures the binding strength in kinetics. These parameters are critical in studying diesease diagnostics and even in drug discovery; for example, it has been discovered that a drug's association constant has a stronger correlation with its clinical efficacy, safety, duration of action, tolerability, indication, and therapeutic differentiation than its binding affinity [32]. An intensity-based surface plasmon resonance sensorgram derived from an angular sensorgram with kinetic parameters $k_a = 1.2 \times 10^5 \mathrm{M^{-1}s^{-1}}$ and $k_d = 7.05 \times 10^{-4}\mathrm{s^{-1}}$ is transformed to a phase-based sensorgram using the relationship shown in Eq. 7,

$$r(t) = |r|e^{i\phi(t)} \tag{7}$$

where $r(t) = \sqrt{\mathrm{R}(t)}$, is the reflectivity coefficient with respect to the horizontal polarization which is the polarization which causes plasmon resonance, $|r|$ is magnitude of the reflectivity coefficient over time, and $\phi(t)$ is the phase measured over time. The transformation allows us to modify and model the intensity-based binding reaction into a phase-based one, from which analysis can be conducted. The kinetic parameters appear in Eq. 7 through the refractive index term of the reflectivity coefficient. This follows from work which was done in [28], where we transform the angular sensorgram into an intensity-based sensorgram. In this work we perform a similar transformation then take a step further using Eq. 7 to transform the sensorgram from an intensity-based sensorgram into a phase-based sensorgram. This then allows the sensitivity derivative of Eq. 1 to be extracted. Figure 1 (a) shows the intensity-based figure binding sensorgram for the binding interaction between the S1 spike protein and a 20-base aptamer "CFA0688T", and Figure 1 (b) shows the phase based transformed equivalent which is used to calculate the LOD values. In this work, we consider a two-mode Mach-Zender intefereometer system that has a Krestchman configuration plasmonic biosensor integrated. By developing such a configuration, we can track the biological interactions on the biosensor surface as a function of a phase shift, as opposed to traditional angle-shift and intensity-change probes. The setup considered in this work is shown in Figure 2 and it shows how phase based plasmonic biosensing can be setup in the lab. The SPR angle considered for this work is an angle on the slope of the reflectivity versus incident angle plot that is close to the dip angle. This is because at the SPR dip angle a significant number of photons will be lost (absorbed), which means no phase measurement will be possible.

2.1. Quantum states considered

The typical states used in the Krestchmann configuration setup is the coherent state. As such we use it as the benchmark standard for comparison with quantum states. The two mode coherent state (TMC) is expressed mathematically as Eq. 8 below,

$$|\mathrm{TMC}\rangle = |\gamma\rangle_m |\sigma\rangle_n = \hat{D}_m(\gamma)\hat{D}_n(\sigma) |0\rangle_m |0\rangle_n , \tag{8}$$

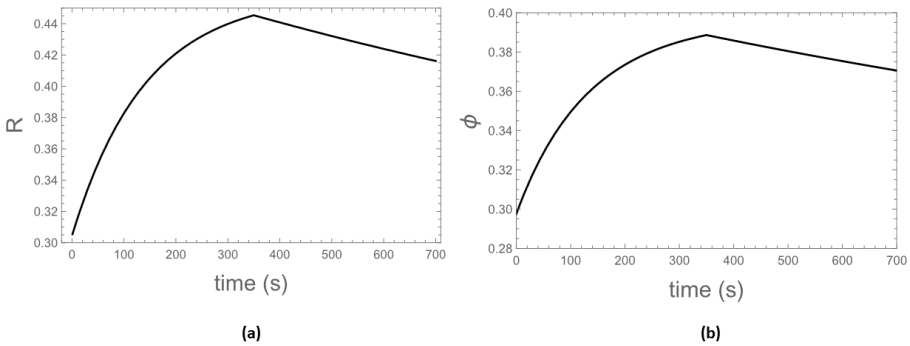


Figure 1 This is a plot which shows the sensorgrams of the kinetic binding interaction between the S1 spike protein and a 20-base aptamer "CFA0688T". (a) shows an intensity-based sensorgram derived from an angular sensorgram whose kinetic parameters are $k_a = 1.2 \times 10^5 \mathrm{M}^{-1}\mathrm{s}^{-1}$ and $k_d = 7.05 \times 10^{-4}\mathrm{s}^{-1}$. (b) Shows the phase (measured in radians) based system which is derived or transformed from the intensity based sensorgram using Eq. 7.

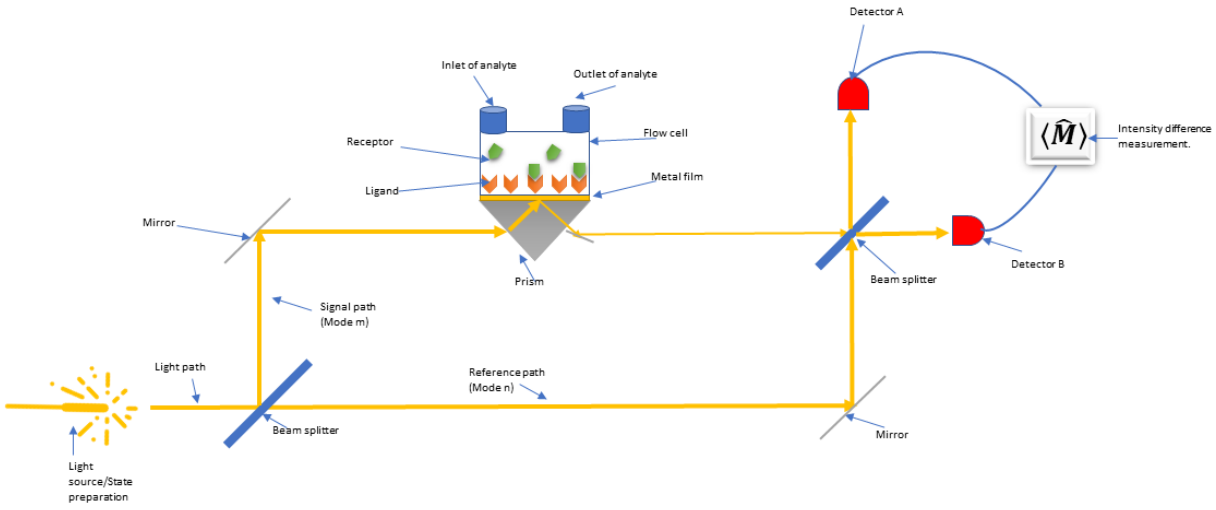


Figure 2 This is a picture of a surface plasmon resonance setup embedded in a Mach-Zender interferometer to allow for phase based biosensing. The setup shown here is applicable to the coherent state case and would have to be modified in the state preparation for the NOON state [33, 34], the figure does give a representation of how the experiment may be setup for the coherent state.

where $\hat{D}_m(\gamma) = e^{\gamma \hat{m}^\dagger - \gamma^* \hat{m}}$ is called the displacement operator for the mode m, which has a displacement parameter $\gamma \in \mathbb{C}$. $\hat{m}^\dagger$ and $\hat{m}$ are the creation and annihilation operators, respectively, for the mode m. The measurement operator, $\hat{M}$, used for the coherent state is the intensity difference measure,

$$\langle \hat{M} \rangle = \langle \hat{m}^\dagger \hat{m} - \hat{n}^\dagger \hat{n} \rangle. \tag{9}$$

The measurement $\hat{M}$ is implemented by taking the photon number difference or the intensity difference between the two modes using homodyne detection. In this work we look at the use of a quantum state of light known as the NOON state for comparison. The NOON state is an N-photon entangled state which can be generated using the Hong Ou Mandel effect [35] or via post processing of single photons generated via spontaneous parametric down conversion [36, 37] and circuit QED [38]. Mathematically, the NOON state can be expressed as shown in Eq. 10,

$$|\psi_{\mathrm{N00N}}\rangle = \frac{1}{\sqrt{2}}(|\mathrm{N}\rangle_m |0\rangle_n + |0\rangle_m |\mathrm{N}\rangle_n). \tag{10}$$

Where m and n are the modes of the states. The measurement operator used for the NOON state analysis in this work is written as in Eq. 11,

$$\hat{\mathrm{A}}_{\mathrm{NOON}} = |\mathrm{N}\rangle_m |0\rangle_n \langle 0|_m \langle \mathrm{N}|_n + |0\rangle_m |\mathrm{N}\rangle_n \langle \mathrm{N}|_m \langle 0|_n . \tag{11}$$

The measurement operator $\hat{\mathrm{A}}_{\mathrm{NOON}}$ can be realized experimentally via interference measurements at the output ports of a Mach-Zehnder interferometer [39].

2.2. Analysis

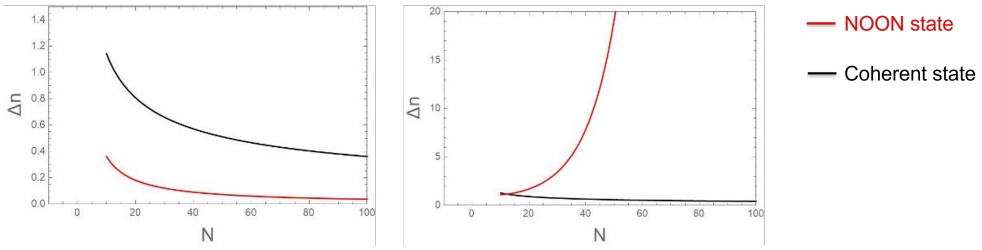


Figure 3 This is a plot which shows the LOD of the plasmonic biosensor in the context of the refractive index on the biosensor surface. Image (a) shows the detectable change in refractive index with photon number, N, in the ideal case where there are no losses to the environment. Image (b) shows the plasmonic sensing experiment but with the inclusion of a 20% loss in both modes of the eperiment. We see that the NOON state is not robust against any losses to the environment and the LOD quickly becomes unstable with increasing photon number, N.

From Figure 3 (a) it is clear that an ideal case of the simulation, i.e., when losses to the environment are not accounted for and are assumed to be zero then the NOON state outperforms the coherent state in limit of detection measures for varying photon numbers. However, in real lab experiments it is crucial to account for losses to the environment, hence in these simulations we account for environmental losses, when we add a 20% loss to both arms of the experiment, it is apparent and clear from Figure 3 (b) that the NOON state is not robust to environmental losses and the limit of detection measure grows exponentially fast.

3. Conclusion

In summary, this work undertakes a theoretical transformation of an SPR sensorgram from an angular shift picture to an intensity shift picture to a phase shift picture. In this work, it is shown that using the NOON state of light as a probe in the phase-based plasmonic biosensor we can enhance the LOD of measurement of the refractive index in the diagnosis of COVID-19.

The NOON state is unfortunately not robust against losses in the system, and when minor losses are introduced, it quickly fails to surpass the SNL and becomes worse with increasing number of photons. This is due to the fact that the $\Delta\phi$ term in the formula for Δn has a scaling factor $f(\phi, N, \eta) = \frac{\eta_1^{\frac{N}{2}} + \eta_2^{\frac{N}{2}}}{2(\eta_1\eta_2)^{\frac{N}{2}}}$ that depends on the photon number N, phase ϕ, and losses η. Hence, $\Delta\phi = \frac{1}{N} f(\phi, N, \eta)$ [30]. This scaling factor diverges or grows rapidly with increasing N when there are losses in the experiment, the divergence is described in detail in the work done in [30]. The scaling factor is also explicitly calculated in the work shown in the Appendix of work done in [22]. Despite this limitation, the knowledge that it is possible in theory to break the SNL is a great positive. The answer may lie in considering other quantum states of light, such as the squeezed states of light and or Fock basis states, which are known to be more robust against losses. This is a possible direction for future research, both theoretical and experimental. This work highlights the potential to use optical biosensors for disease detection and highlights how they can be enhanced to address future pandemics. At present, regarding the generation of NOON states, they are usually generated with a relatively small number of photons, often less than 10. This is because the difficulty of creating and maintaining entangled states increases exponentially with the number of particles involved. These are very low photon numbers for practical experiments which can handle up to milliwatt power intensities; hence, with the current state of NOON state technologies, it is better to work with the classical states. It is worth noting that quantum technologies are continually improving and there may come a time when it is easier to produce these quantum states at high photon numbers, in which case it will be worth using in these states for the precision enhancement they offer. Future work in this space will look at using squeezed states and Fock states, as these may be more robust against environmental losses and could result in a lower LOD and may be easier to produce in a practical setting.

Acknowledgements.— This research was supported by the Council for Scientific and Industrial Research (CSIR) and the Department of Science and Innovation (DSI) in South Africa.

Data Availability: The data sets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

[1] P. Singh, "Surface plasmon resonance: A boon for viral diagnostics," *Reference Module in Life Sciences*, 2017.

[2] M. P. Da Silveira, K. K. da Silva Fagundes, M. R. Bizuti, É. Starck, R. C. Rossi, and D. T. de Resende E Silva, "Physical exercise as a tool to help the immune system against covid-19: an integrative review of the current literature," *Clinical and experimental medicine*, vol. **21**, no. 1, pp. 15–28, 2021.

[3] S. Felsenstein, J. A. Herbert, P. S. McNamara, and C. M. Hedrich, "Covid-19: Immunology and treatment options," *Clinical immunology*, vol. **215**, p. 108448, 2020.

[4] E. Mortaz, P. Tabarsi, M. Varahram, G. Folkerts, and I. M. Adcock, "The immune response and immunopathology of covid-19," *Frontiers in immunology*, vol. **11**, p. 2037, 2020.

[5] D. Darif, I. Hammi, A. Kihel, I. E. I. Saik, F. Guessous, and K. Akarid, "The pro-inflammatory cytokines in covid-19 pathogenesis: What goes wrong?" *Microbial pathogenesis*, vol. **153**, p. 104799, 2021.

[6] J. P. Gómez Vázquez, Y. E. García, A. J. Schmidt, B. Martínez-López, and M. Nuño, "Testing and vaccination to reduce the impact of covid-19 in nursing homes: an agent-based approach," *BMC Infectious Diseases*, vol. **22**, no. 1, p. 477, 2022.

[7] Y. Yan, L. Chang, and L. Wang, "Laboratory testing of sars-cov, mers-cov, and sars-cov-2 (2019-ncov): Current status, challenges, and countermeasures," *Reviews in medical virology*, vol. **30**, no. 3, p. e2106, 2020.

[8] S. Agarwal, C. Warmt, J. Henkel, L. Schrick, A. Nitsche, and F. F. Bier, "Lateral flow–based nucleic acid detection of sars-cov-2 using enzymatic incorporation of biotin-labeled dutp for poct use," *Analytical and Bioanalytical Chemistry*, vol. **414**, no. 10, pp. 3177–3186, 2022.

[9] P. Damborskỳ, J. Švitel, and J. Katrlík, "Optical biosensors," *Essays in biochemistry*, vol. **60**, no. 1, pp. 91–100, 2016.

[10] E. Kretschmann and H. Raether, "Radiative decay of non radiative surface plasmons excited by light," *Zeitschrift für Naturforschung A*, vol. **23**, no. 12, pp. 2135–2136, 1968.
[11] M. S. Islam, A. Z. Kouzani, X. J. Dai, and W. P. Michalski, "Parameter sensitivity analysis of surface plasmon resonance biosensor through numerical simulation," in *IEEE/ICME International Conference on Complex Medical Engineering*. IEEE, 2010, pp. 171–176.
[12] L. Jason-Moller, M. Murphy, and J. Bruno, "Overview of biacore systems and their applications," *Current protocols in protein science*, vol. **45**, no. 1, pp. 19–13, 2006.
[13] M. Fivash, E. M. Towler, and R. J. Fisher, "Biacore for macromolecular interaction," *Current opinion in biotechnology*, vol. **9**, no. 1, pp. 97–101, 1998.
[14] J. Homola and M. Piliarik, *Surface plasmon resonance (SPR) sensors*. Springer, 2006.
[15] Y. Huang, H. Ho, S. Wu, and S. Kong, "Detecting phase shifts in surface plasmon resonance: a review," *Advances in Optical Technologies*, vol. **2012**, 2012.
[16] M. Kashif, A. A. A. Bakar, N. Arsad, and S. Shaari, "Development of phase detection schemes based on surface plasmon resonance using interferometry," *Sensors*, vol. **14**, no. 9, pp. 15 914–15 938, 2014.
[17] S. Zeng, X. Yu, W.-C. Law, Y. Zhang, R. Hu, X.-Q. Dinh, H.-P. Ho, and K.-T. Yong, "Size dependence of au np-enhanced surface plasmon resonance based on differential phase measurement," *Sensors and Actuators B: Chemical*, vol. **176**, pp. 1128–1133, 2013.
[18] Y. Huang, H. P. Ho, S. K. Kong, and A. V. Kabashin, "Phase-sensitive surface plasmon resonance biosensors: methodology, instrumentation and applications," *Annalen der Physik*, vol. **524**, no. 11, pp. 637–662, 2012.
[19] S. Mohammadzadeh-Asl, A. Keshtkar, J. E. N. Dolatabadi, and M. de la Guardia, "Nanomaterials and phase sensitive based signal enhancment in surface plasmon resonance," *Biosensors and Bioelectronics*, vol. **110**, pp. 118–131, 2018.
[20] X. Guo, "Surface plasmon resonance based biosensor technique: a review," *Journal of biophotonics*, vol. **5**, no. 7, pp. 483–501, 2012.
[21] J. Cooper and J. Dunningham, "Towards improved interferometric sensitivities in the presence of loss," *New Journal of Physics*, vol. **13**, no. 11, p. 115003, 2011.
[22] K. T. Mpofu, "Quantum plasmonic sensing with application to hiv research." Ph.D. dissertation, 2020.
[23] K. Mpofu, S. Ombinda-Lemboumba, P. Mthunzi-Kufa *et al.*, "Classical and quantum surface plasmon resonance biosensing," *International Journal of Optics*, vol. **2023**, 2023.
[24] C. Lee, B. Lawrie, R. Pooser, K.-G. Lee, C. Rockstuhl, and M. Tame, "Quantum plasmonic sensors," *Chemical Reviews*, vol. **121**, no. 8, pp. 4743–4804, 2021.
[25] C. Lee, F. Dieleman, J. Lee, C. Rockstuhl, S. A. Maier, and M. Tame, "Quantum plasmonic sensing: beyond the shot-noise and diffraction limit," *Acs Photonics*, vol. **3**, no. 6, pp. 992–999, 2016.
[26] K. Mpofu, C. Lee, G. Maguire, H. Kruger, and M. Tame, "Experimental measurement of kinetic parameters using quantum plasmonic sensing," *Journal of Applied Physics*, vol. **131**, no. 8, 2022.
[27] J.-S. Lee, S.-J. Yoon, H. Rah, M. Tame, C. Rockstuhl, S. H. Song, C. Lee, and K.-G. Lee, "Quantum plasmonic sensing using single photons," *Optics express*, vol. **26**, no. 22, pp. 29 272–29 282, 2018.
[28] K. Mpofu, C. Lee, G. Maguire, H. Kruger, and M. Tame, "Measuring kinetic parameters using quantum plasmonic sensing," *Physical Review A*, vol. **105**, no. 3, p. 032619, 2022.
[29] K. Mpofu and P. Mthunzi-Kufa, "Enhanced signal-to-noise ratio in quantum plasmonic image sensing including loss and varying photon number," *Physica Scripta*, 2023.
[30] M. Bohmann, J. Sperling, and W. Vogel, "Entanglement and phase properties of noisy noon states," *Physical Review A*, vol. **91**, no. 4, p. 042332, 2015.
[31] S. Szunerits, H. Saada, Q. Pagneux, and R. Boukherroub, "Plasmonic approaches for the detection of sars-cov-2 viral particles," *Biosensors*, vol. **12**, no. 7, p. 548, 2022.
[32] Y. Fang, "2.17—label-free screening technologies," *Comprehensive Medicinal Chemistry*, vol. **3**, pp. 416–433, 2017.
[33] L. Zhang and K. W. C. Chan, "Scalable generation of multi-mode noon states for quantum multiple-phase estimation," *Scientific reports*, vol. **8**, no. 1, p. 11440, 2018.
[34] L. Zhang, "Theoretical studies of multi-mode noon states for applications in quantum metrology and proposals of experimental setups for their generation," 2018.
[35] C.-K. Hong, Z.-Y. Ou, and L. Mandel, "Measurement of subpicosecond time intervals between two photons by interference," *Physical review letters*, vol. **59**, no. 18, p. 2044, 1987.
[36] P. Walther, J.-W. Pan, M. Aspelmeyer, R. Ursin, S. Gasparoni, and A. Zeilinger, "De broglie wavelength of a non-local four-photon state," *Nature*, vol. **429**, no. 6988, pp. 158–161, 2004.
[37] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, "Super-resolving phase measurements with a multiphoton entangled state," *Nature*, vol. **429**, no. 6988, pp. 161–164, 2004.
[38] Q.-P. Su, C.-P. Yang, and S.-B. Zheng, "Fast and simple scheme for generating noon states of photons in circuit qed," *Scientific reports*, vol. **4**, no. 1, p. 3898, 2014.

[39] I. Afek, O. Ambar, and Y. Silberberg, "High-noon states by mixing quantum and classical light," *Science*, vol. **328**, no. 5980, pp. 879–881, 2010.

Majority voting algorithm for TB detection: Machine learning approach

Nkgaphe Tsebesebe[1,2], Kelvin Mpofu[1], Sphumelele Ndlovu[1], Sudesh Sivarasu[2] and Patience Mthunzi-Kufa[1,2,3]

[1]Council for Scientific and Industrial Research, National Laser Centre, P.O Box 395, Building 46A, Pretoria 0001, South Africa.
[2]Department of Human Biology, Division of Biomedical Engineering, University of Cape Town, Cape Town 7935, South Africa
[3]School of Chemistry and Physics, University of KwaZulu-Natal, Durban 4001, South Africa

E-mail: ntsebesebe@csir.co.za

Abstract. An accurate examination of chest radiographs remains a challenge for physicians. As part of the solution, most researchers proposed the use of algorithms to perform automatic classification of chest radiographs. The current work continues with the effort to develop machine learning algorithms to perform the classification of chest radiograph images as healthy and TB infected. In this work, a combination of machine learning algorithms (i.e., a artificial neural network, a support vector machine, and decision tree algorithms) use a majority voting technique to classify chest X-ray images into healthy and TB-infected. The performance of individual algorithms is optimized by hyperparameter techniques such as grid search. In this work we develop a voting system where the three machine learning algorithms collectively classify each image and the class with the majority number of votes becomes the new label. The voting system of this work achieves an accuracy of 99% in classifying 800 chest radiograph images. The system classifies TB-infected images by 98% and noninfected images by 100%, providing high precision and recall. The voting system outperforms the individual algorithms' performances. The voting algorithm proposed in this work can be beneficial in medical facilities in providing second opinions to medical professionals by reducing the difficult task of classifying abnormalities in chest radiographs.

1. Introduction

Respiratory or chest diseases are one of the primary causes of mortality and disability worldwide [1]. Chest infections are common diseases which include among others tuberculosis (TB) [1]. Chronic obstructive pulmonary diseases (COPDs) are considered the third highest reason for death worldwide [1]. In 2021, TB caused by Mycobacterium tuberculosis (Mtb) with a single bacterial agent claimed approximately 1.6 million lives worldwide [2]. On the other hand, approximately 10 million people were reported to have developed active TB in 2020 [2]. The death toll for 2021 was an increase from the 1.4 million TB-related deaths worldwide reported in 2019 [2, 3]. To ensure proper treatment, control, and complete recovery, early detection of TB becomes a crucial step [1]. Although there are numerous methods used to detect TB up to date, chest radiography is one of the most widely used examination methods to perform TB diagnosis due to its relatively low cost and ease of operation.

The chest radiograph examines the thorax (chest) and identifies acute and chronic cardiopulmonary conditions to allow medical professionals to make the diagnosis [1]. In X-ray diagnosis, the partly transmitted and absorbed radiation is used to produce an X-ray image [4]. The X-ray images show variations in transmission caused by structures in the object of varying thickness, density or atomic composition [4]. The chest radiographs used in this work contain detailed information on the health of patients, hence accurate examination of the information is important to produce medical reports of patients [2, 4]. Conventional methods of diagnosis involve physicians analyzing the images and making conclusions. However, the accurate examination of the information in detail remains a challenge to physicians [1]. Arshia Rehman et. al. [1] reported that it is a very difficult task for clinical radiologists to classify abnormalities on chest radiographs. As part of the solution, many researchers proposed the use of algorithms to perform automatic classification of chest radiograph images [1]. The approach offers multiple advantages, such as reduced diagnosis errors, identification of complex patterns, and overcoming human bias [1].

The current work continues with the effort to develop machine learning algorithms to perform the classification of chest radiograph images as healthy and infected with TB. The approach presented in this paper makes use of the majority voting technique where three machine learning models (Support Vector Machine (SVM), decision tree, and artificial neural network) are combined and a prediction of each model on the same x-ray image is considered a vote. The majority vote count was considered to be a final prediction on a particular image under examination. The aim of this technique was to improve the overall classification performance of the algorithms. Collective voting has shown to outperform individual algorithms [5].

2. Methodology

2.1. Machine Intelligence Library

The traditional models, i.e., the Support Vector Machine (SVM) and the decision tree of the current study were implemented from scikit-learn [6]; the deep learning model, i.e., the artificial neural network was implemented from Keras [7] with Google TensorFlow [8]. All the models operated with the aid of some scientific computing libraries, including NumPy [9] and Matplotlib [10] for mathematical operations and data visualization.

2.2. Dataset and data pre-processing

This study used publicly available chest radiograph images from the Kaggle TB database [11]. In figure 1 are samples of chest X-ray images from the database. The images are labeled; hence the current study follows the supervised machine learning approach. That is, the labeled data sets were used to train the algorithms to make accurate predictions. The current work use total of 2006 chest radiographs, where 1206 chest radiographs, 800 infected with TB and 406 healthy were used for training and testing; 800 are reserved for validating the voting algorithm. All the images before training and testing the algorithms are resized into 200 x 200 pixels, then flattened into a 1-dimensional array with 40000 pixels. The pixel values were divided by a maximum value that a pixel can take for an 8-bit image (255). As such, each pixel value was scaled to have minimum value of 0 and maximum of 1, where 0 and 1 represented a black and white color in a gray scale images, respectively.

Principal Component Analysis (PCA) was utilized to perform feature selection considering a 98% variance ratio to detect sensitive information from the images while minimizing information loss. The data was split so that 80% of the chest X-ray images were for training the machine learning models and 20% of the images were for testing the model. Lastly, machine learning models were trained on the data set.

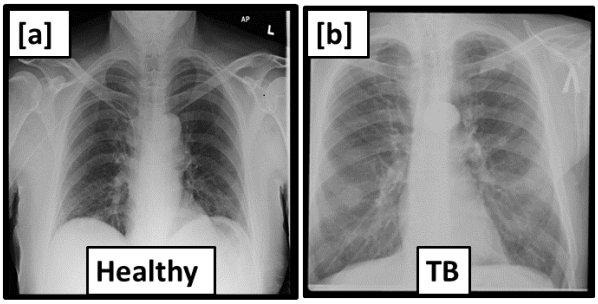


Figure 1. The samples of chest X-ray images used in this work.

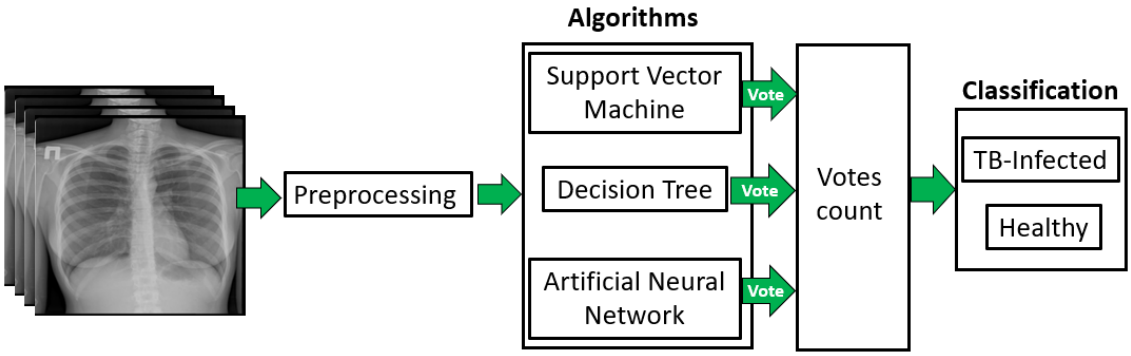


Figure 2. A graphic representation of the voting algorithm.

2.3. Majority Voting Algorithms

This study considers three machine learning models to cast predictions as votes, where the majority votes determine the prediction of each X-ray image. A representation of the voting algorithm is displayed in Figure 2.

2.4. Support vector machine

Support vector machines belong to a family of generalized linear classification [12]. They are a set of supervised learning methods that can solve pattern classification and regression problems [12]. The SVM work by minimizing the classification and maximizing the geometric margin. This happens by mapping the input vector to a higher-dimensional space from which the maximal separating hyperplane is generated [12]. Conversely, parallel hyperplanes are generated on the side of the hyperplane that separates the data. In the current work, the kernel function which transforms the data into a higher-dimensional space to perform separation of the data was a linear function. The hyperplane took the form of:

$$m \cdot x + c = 0 \tag{1}$$

where c is a scalar and m is a number of images, and x are features. Parallel hyperplanes are defined as:

$$m \cdot x + c = 1 \tag{2}$$

$$m \cdot x + c = -1 \tag{3}$$

Since the data in this work are linearly separable, the hyperplanes were selected so that there were no points between them. The distance between the hyperplanes was determined as:

$$\frac{2}{|m|} \tag{4}$$

The support vector machine in this case minimized $|m|$ from

$$m \cdot x_i - c \geq 1 \; or \; m \cdot x_i - c \leq 1, \; 1 \leq i \leq n \tag{5}$$

2.5. Decision tree

A decision tree is a type of supervised learning model that adopts a flow chart tree structure, where every internal node represents a test on an attribute, the branches denote the prediction of the test and the leaf nodes represent labels of the classes [12, 13]. This work implemented a decision tree algorithm with an information gain function to facilitate data splitting. The function considered class probability and log_2 of class probability as follows:

$$information \; gain \; or \; Entropy = \sum_{n=2}^{C} -p_i \times log_2(p_i) \tag{6}$$

Where p represents the probability of samples in a class, C represents a number of classes and i is a number of samples. The decision tree algorithm of this work used 8 minimum number of chest X-ray images to perform the splitting of the internal note to a maximum tree depth of 17.

2.6. Artificial neural network

This study implemented a feedforward neural network with two different activation functions: the rectified linear unit (ReLU) and the sigmoid in the hidden layer, and the classification layer, respectively. The activation function ReLU possesses biological and mathematical information that helps improve deep neural network training [14]. The functionality of the ReLU activation function depends on the thresholding values at 0. It computes the values from the following function:

$$f(x) = max(0, x) \tag{7}$$

The function output 0 when $x < 0$ implying the neuron is not activated in the case and output linear function when $x \geq 0$, neurons are activated in this case. Sigmoid activation on the classification layer was defined as:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{8}$$

Where e is the exponential function and x are the features. The function translates the inputs range from $(-\infty; +\infty)$ to a range in $[0; 1]$ [15]. This enabled a prediction of either TB-infected or healthy. The artificial neural network of this work was implemented sequentially with densely connected layers. The model had only one hidden layer consisting of 64 neurons and one neuron at the classification layer to facilitate a binary classification. The output of the hidden layer was normalized from a normalization batch. The model was optimized by Adam optimizer with a learning rate of 0.19 and a loss function of categorical crossentropy. The model was trained with 100 chest X-ray images at the time, for 80 training cycles.

3. Results and Discussion

3.1. Performance matrices

The performance of the majority-voting algorithm depends mainly on the performance of each algorithm casting votes in the process. Models in this current paper are evaluated for over-fitting

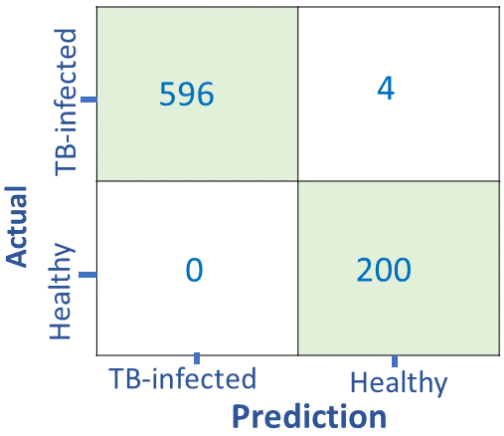


Figure 3. Performance metrics of the voting algorithm

and under-fitting before they can be implemented in the majority voting. Table 1 shows the performance scores of the models during training and testing. All the algorithms of this study have less that 5% percentage differences in accuracy between training and testing, implying good results. The accuracy scores were calculated at the machine level using the following formula [16]:

$$Accuracy = \frac{True\ positives + True\ negatives}{True\ positives + True\ negatives + False\ positives + False\ negatives} * 100 \quad (9)$$

The support vector machine and artificial neural network both achieve 97.93% accuracy during testing, while the decision tree falls behind by 4.13% to 93.8%. After passing the testing stage, the models were saved to the machine and they were grouped to cast votes, where the majority votes on each individual chest radiograph image was considered a prediction of the voting algorithm. This approach was used avoid the effect of not-normalized output from deep learning models. A decision boundary of 0.5 was used to round the output values from the ANN model into a whole number; 0 for all values less than 0.5 and 1 for all output values equal or greater than 0.5. The performance of the voting algorithm was evaluated from the ability of the algorithm to classify 800 images that were initially reserved for testing, as TB-infected or healthy. The algorithm correctly classified 596 images as TB-infected and 200 images as healthy; it incorrectly classified 4 images as healthy while they were TB-infected. This performance is shown from the confusion matrix in Figure 3.

Table 1. Performance scores of the models during testing.

Algorithms	Training Score (%)	Testing Score (%)
`Support Vector Machine`	99.45	97.93
`Decision Tree`	98.26	93.80
`Artificial Neural Network`	99.87	97.93

3.2. Clinical indicators

Table 2 and Table 3 are clinical indicators of the voting algorithm in the chest X-ray images. A detailed evaluation of the majority algorithm was conducted from the clinical indicators. The voting algorithm correctly predicted chest radiograph images by 99.50% as reflected by the accuracy in table 1. The accuracy measure in this work denotes a total number of correctly identified images amongst all the images. The algorithm identified TB-infected images by 100% (positive predictive value (PPV)) and healthy images by 98.04% (negative predictive value (VPV)) denoted by a precision value, this was dependent on the ability of the algorithm to distinguish features from the image. The principal component analysis with 98% variance ratio helped the algorithm differentiate features by 94.83% (sensitivity) which resulted in 100% ratio between correctly predicted images and all predicted images. The balance between healthy and TB-infected images was obtained by 95.15% (F1 score) from the mean accuracy and the sensitivity of the algorithm [17]. The accuracy of the algorithm prediction depends mainly on the quality of the input images [16]. Thus, the high accuracy of the algorithm is due to the linear nature of the data set. The linear kernel of the support vector machine produced a maximum margin between the two classes. On the other hand, the artificial neural network model was able to handle the data distribution by facilitating a lower learning rate. That is, during the training process, the weights associated with neuron connections of the model were updated at a lower rate after forward passes of data through the network. Another advantage is that the model had only one hidden layer; this reduced the overfitting by increasing the generalizability of the model on the test data. However, this was the major factor in the long prediction time of 92.98 seconds. The overall performance of the majority algorithm was confirmed by the high percentage (95.15%) of the F1 score. The F1 score is a machine learning evaluation metric that measures a weighted precision and sensitivity average [18].

In comparing of this work with previous work, S.I. Nafisah and G. Muhamad [17] developed an automatic TB detection from chest radiographs using deep learning models. Their best-performing convolutional neural network model (EfficientNetB3) achieved an accuracy of 99.10%. As such, the current work offers a slight improvement in accuracy. Although the works used different approaches, the slight improvement of the current work can be associated with the collaborative functionality of the algorithms as reflected in the F1 score. This collaborative functionality mainly depends on the number of models used in the voting process. This work used an odd number of models to avoid ties and overrule a vote of a model which may misclassify the input data.

Table 2. Clinical indicators of the voting algorithm on chest radiographs

Algorithm	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)
`SVM+DT+ANN`	99.50	100.00	94.83	100.00

Table 3. Continuation of clinical indicators of the voting algorithm on chest radiographs

Algorithm	PPV (%)	NPV (%)	F1 Score (%)	Time(s)
`SVM+DT+ANN`	100.00	98.04	95.15	92.98

4. Conclusions

Classification algorithms play an important role in the early screening of TB. This paper proposed a majority voting algorithm to classify chest radiograph images as healthy and infected with TB. The collaborative functionality of the models (support vector machine, decision tree, and artificial neural network) provides an accuracy of 99.50%. The voting algorithm offers an advantage to overrule missed predictions coming from one model. Another contribution to the accuracy of the algorithm was offered by the PCA, which enabled each individual model to detect sensitive information from the images while minimizing information loss. The use of the voting algorithm proposed in this work can be beneficial in medical facilities in providing second opinion to medical professionals by reducing the difficult task for clinical radiologists to classify abnormalities on chest radiograph images.

5. Acknowledgments

We acknowledge the Department of Science and Innovation (DSI) for funding this research.

6. References

[1] Rehman A, Khan A, Fatima G, Naz S, Razzak I Review on chest pathogies detection systems using deep learning techniques. Artificial Intelligence Review 2023 **56**(11):12607–53. Available from: http://dx.doi.org/10.1007/s10462-023-10457-9.

[2] Ndong Sima CAA, Smith D, Petersen DC, Schurz H, Uren C, Möller M. The immunogenetics of tuberculosis (TB) susceptibility. Immunogenetics 2022, **75**(3):215–30. Available from: http://dx.doi.org/10.1007/s00251-022-01290-5.

[3] Chakaya J, Petersen E, Nantanda R, Mungai BN, Migliori GB, Amanullah F, et al. The WHO Global Tuberculosis 2021 Report – not so good news and turning the tide back to End TB. International Journal of Infectious Diseases 2022, **124**:S26–9. Available from: http://dx.doi.org/10.1016/j.ijid.2022.03.011.

[4] Carlsson C.A. and Alm Carlsson G.A., 1973. Basic physics of X-ray imaging. Linkoping University Electronic Press. Available from: https://www.diva-portal.org/smash/record.jsf?pid=diva2\%3A276160\&dswid=-7479.

[5] Salim A, Yousif M, Bader S, Sagheer AM. Proposed combination method for ensemble classification technique based on majority voting. AIP Conference Proceedings 2023; Available from: http://dx.doi.org/10.1063/5.0154218.

[6] Pajankar A, Joshi A. Introduction to Machine Learning with Scikit-learn. Hands-on Machine Learning with Python. 2022 65–77. Available from: http://dx.doi.org/10.1007/978-1-4842-7921-2_5.

[7] Moolayil J. Learn Keras for Deep Neural Networks. Apress; 2019. Available from: http://dx.doi.org/10.1007/978-1-4842-4240-7

[8] Manaswi NK. Understanding and Working with Keras. Deep Learning with Applications Using Python 2018;31–43. Available from: http://dx.doi.org/10.1007/978-1-4842-3516-4_.

[9] van der Walt S, Colbert SC, Varoquaux G. The NumPy Array: A Structure for Efficient Numerical Computation. Computing in Science & Engineering 2011 **13**(2):22–30. Available from: http://dx.doi.org/10.1109/mcse.2011.37/.

[10] de Andrade E. SciPy Tools Plenary on Matplotlib. Proceedings of the Python in Science Conference. SciPy; 2020 Available from: http://dx.doi.org/10.25080/majora-342d178e-029.

[11] "Kaggle," 2023. [Online]. Available from: https://www.kaggle.com/datasets/tawsifurrahman/tuberculosistb-chest-xray-dataset.

[12] Kieu STH, Bade A, Hijazi MHA, Kolivand H. A Survey of Deep Learning for Lung Disease Detection on Medical Images: State-of-the-Art, Taxonomy, Issues and Future Directions. Journal of Imaging 2020 **6**(12):131. Available from: http://dx.doi.org/10.3390/jimaging6120131.

[13] Bhavsar H. and Panchal M.H., 2012. A review on support vector machine for data classification. International Journal of Advanced Research in Computer Engineering and Technology (IJARCET), **1**(10):185-189. Available from: https://citeseerx.ist.psu.edudocument?repid=rep1&type=pdf&doi=d683a971524a0d76382ce335321b4b8189bc8299.

[14] Agarap A.F., 2018. Deep learning using rectified linear units (relu). arXiv preprint arXiv:1803.08375. Available from: https://arxiv.org/abs/1803.08375.

[15] Rasamoelina AD, Adjailia F, Sincak P. A Review of Activation Function for Artificial Neural Network. 2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI) 2020; Available from: http://dx.doi.org/10.1109/sami48414.2020.9108717.

[16] An L, Peng K, Yang X, Huang P, Luo Y, Feng P, et al. E-TBNet: Light Deep Neural Network for Automatic Detection of Tuberculosis with X-ray DR Imaging. Sensors 2022 **22**(3):821. Available from: http://dx.doi.org/10.3390/s22030821.

[17] Nafisah SI, Muhammad G. Tuberculosis detection in chest radiograph using convolutional neural network architecture and explainable artificial intelligence. Neural Computing and Applications 2022: 1-21; Available from: http://dx.doi.org/10.1007/s00521-022-07258-6.

[18] Shashikant R, Chetankumar P. Predictive model of cardiac arrest in smokers using machine learning technique based on Heart Rate Variability parameter. Applied Computing and Informatics. 2020 **19**(3/4):174–85. Available from: http://dx.doi.org/10.1016/j.aci.2019.06.002.

The combined effect of hypericin-mediated photodynamic therapy and *Punica granatum* on MCF-7 breast cancer cells

Nosipho Fakudze, Paromita Sarbadhikary, Blassan P. George* and Heidi Abrahamse

Laser Research Centre, Faculty of Health Sciences, University of Johannesburg, P.O. Box 1701, Doornfontein 2028, South Africa.

E-mail: **blassang@uj.ac.za**

Abstract. As per cancer statistics, breast cancer is the most prevalent cancer in South African women, accounting for 22.6% of all cancers in women and 16% of cancer deaths in women. Phytochemicals/secondary metabolites are compounds found in plants that have antitumor, antiproliferative, and antioxidant effects used in traditional medicine. Current systematic studies on the use of phytochemicals for various diseases have proved to have significant success with reduced side effects. Photodynamic therapy (PDT) is a cancer therapeutic modality, which has been approved for treating various cancer types. Although being promising for cancer treatment, both phytomedicine and PDT suffers from their own limitations. Combination therapy is one of the avenues being explored to overcome those limitations with enhanced therapeutic outcome. This study aims to investigate the combined anticancer effect of IC_{50} dose of hypericin mediated PDT with 1 J/cm^2 laser treatment at 594 nm to enhance the anticancer potential of *P. granatum* ethyl acetate extract against MCF-7 cells. Cytotoxicity in MCF-7 cells treated with extract, PDT alone and in combination was evaluated with 3-(4,5-dimethyl-2-thiazyl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) assay. Morphological changes were investigated using brightfield microscopy and cell death analysis using Annexin V/PI flow cytometry assay. The results showed enhancement in cytotoxicity (~90%) in combination treatment at a dose of 400 μg/mL of *P. granatum* aril ethyl acetate extract and 0.09 μM hypericin-PDT in comparison to single treatment. In conclusion, *P. granatum* and hypericin-mediated PDT in combination treatment for breast cancer is a viable treatment modality for further investigation.

1. Introduction

In the coming decades, regardless of the availability of resources, cancer, which is already the leading cause of death in many high-income countries, is expected to overtake other diseases as the leading cause of morbidity and mortality worldwide. The condition is worrisome particularly in low- and middle-income countries, which now bear 80% of the global burden of this illnesses [1]. Research and strategic frameworks have been developed to assist to mitigate and treat this disease. Breast cancer is one of the most common causes of death in women and has the leading incident rate [2]. Breast cancer mortality rates continue to increase every year despite the increased and advanced screening and treatment modalities [3]. Treatment modalities for breast cancer include surgery, hormone therapy, radiation, and chemotherapy [4]. These treatments have the benefit of cancer killing effects but also induces unwanted adverse effects. This has brought about the need for improved treatment remedies with little to no adverse effects.

Phytochemicals are natural products found in plants that in recent years gained interest in cancer treatment modalities due to their antioxidant, antitumor, and antiproliferative potentials with little to no toxic effects [5]. *P. granatum* is a medicinal plant with anticancer, antibacterial, antimicrobial, and antiproliferative properties and have been used in traditional medicine throughout the Asian region. In recent years extracts of this plant has been explored in cancer treatment in with positive treatment outcomes [6].

PDT is form a of minimally invasive cancer treatment which have gained interest as an alternative and palliative treatment modality recently [7]. The benefits of PDT include low invasiveness and low cost of treatment but has drawbacks such as low penetration depth and photosensitivity [8]. Hypericin is a naturally occurring compound isolated from *Hypericum perforatum* with phototoxic properties used in several PDT studies with significant cancer cell death induction [5]. Combination therapy is one of the avenues explored in cancer therapy to combat treatment limitations and results in a synergistic effect by inducing cancer cell apoptosis, inhibition of malignant growth, and cell cycle arrest [9]. This study aims to combine the therapeutic effects from hypericin-PDT and *P. granatum* extract on breast cancer cells.

2. Methodology

2.1. Cell culture, experimental groups and treatment

The MCF-7 breast cancer cell line (ATTC® HTB-22) was used in this study. It was grown in a monolayer in a T75 tissue culture flask in Dulbecco's Modified Eagles's Medium (DMEM) and supplemented with 10% fetal bovine serum, 1% penicillin-streptomycin, and 1% amphotericin B. Cells were incubated under the conditions of 95% humidified incubator operating at 37 °C and 5% carbon dioxide (CO_2). Seeding density of control and experimental groups was 7000 cells per well in 96 well plates. After 24 h and 48 h treatment of cells with various concentrations (50, 100, 200, and 400 µg/mL) of *P. granatum* aril ethyl acetate extract alone and PDT with 0.0125, 0.025, 0.1 and 0.2 µM concentrations of hypericin for 3 h were carried out. The highest cytotoxic concentration of *P. granatum* (400 µg/mL) treated for 24 h and IC_{50} dose (0.09 µM) of hypericin groups were used for combined treatment. Laser treatment was carried out using semiconductor diode laser at 594 nm wavelength, at an optimized fluency of 1 J/cm^2. Cytotoxic studies were done after completion of all treatments.

2.2. Morphology

The cellular morphological alterations were evaluated for each group using Olympus CKX 41 inverted microscope following the treatment.

2.3. Cell viability

Roche cell proliferation kit (11465007001) was used to determine cell viability after treatment. MTT reagent (10 µL) was added to each well in 90 µL of culture media containing cell and incubated for 3 h. Thereafter 100 µL of solubilization buffer was added and plate was incubated overnight. Absorbance was measured using multiplate reader at 570 nm.

2.4. Apoptosis assay

FITC Annexin V/PI (Becton Dickinson, 556570 FITC Annexin V Apoptosis Detection Kit II) assay was used to evaluate and quantify percentage of apoptotic and necrotic cells. After treatment, cells were prepared as per manufactures instructions and Annexin V protein complex was added, followed by addition of propidium iodide (PI) as vital dye. Measurement was carried out using flow cytometer (BD Accuri™ C6) BD C Sampler analysis software.

2.5. Statistical analysis

In this study each experiment was performed in triplicate and repeated thrice independently. The experimental groups were compared to control MCF-7 cells and data expressed as mean and standard error. The student's *t*-test was used as an analysis tool for determining the difference in control and

experimental groups with 95% confidence interval. Data was investigated using Sigma Plot v.14.0 software. Statistically significant values were expressed as p <0.05 (*), p < 0.01 (**) or p < 0.001 (***).

3. Results

3.1. Absorption spectra

Absorption spectra of *P. granatum* aril ethyl acetate extract and hypericin were recorded in DMSO. Results showed that the extract doesn't represent any absorption peak in the visible region between 400-800 nm but at 300 nm as represented in Figure 1 (a) thus, cannot be useful in PDT. Therefore, the cytotoxicity effects of the extract were combined with the phototoxic effect of hypericin which has a strong absorption peak at 600 nm as shown in Figure 1 (b).

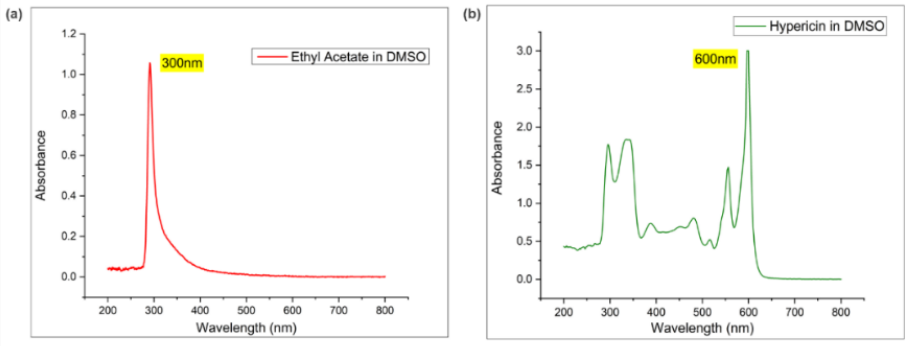


Figure 1. Absorption spectra of (a). *P. granatum* aril ethyl acetate extract and (b) hypericin recorded in DMSO as solvent.

3.2. Cell viability assay

3.2.1. *Single treatment*

Cell viability assay showed cytotoxicity at 400 µg/ml concentration of *P. granatum* extract at 24 and 48 h represented in Figure 2 (a). However, as shown in Figure 2 (b), the induced cytotoxicity by extract was less than 50%, thus IC_{50} dose cannot be calculated.

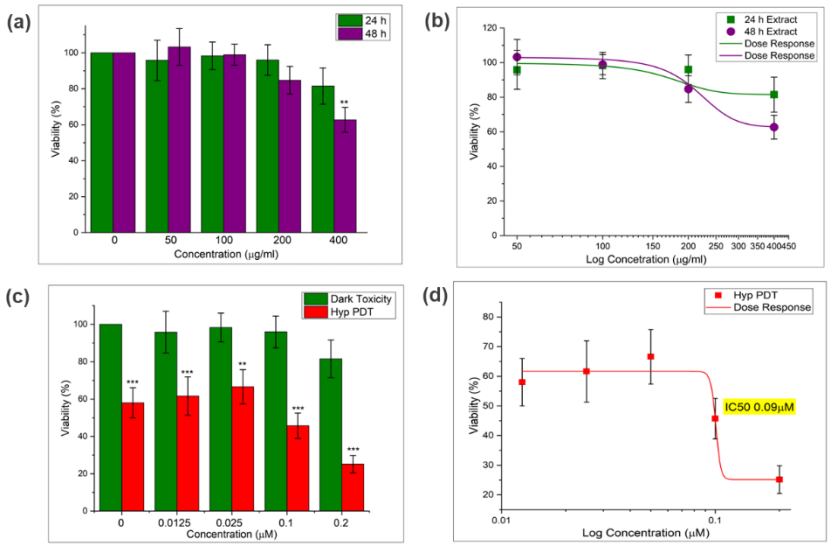


Figure 2. Effect of single treatment on cell viability of MCF-7 cells. (a) *P. granatum* aril ethyl acetate extract treatment in concentration range of 50-400 µg/ml, (b) dose response curve of *P. granatum* aril ethyl acetate extract induced cytotoxicity, (c) phototoxicity induced by different concentration of hypericin (0.0125- 0.2 µM) irradiated at 1 J/cm^2 and (d) dose response curve of hypericin induced phytotoxicity. (p <0.05 (*), p < 0.01 (**) or p < 0.001 (***)).

Cells treated with hypericin followed by irradiation after 3 h, showed a significant decrease in cell viability with all concentrations (0.001-0.2 μM) as shown in Figure 2 (c). The dark toxicity group included hypericin treatment of cells for 3 h without any light irradiation, showing no significant dark toxicity in cells. The IC_{50} of hypericin was determined to be at 0.09 μM concentration (Figure 2 (d)). The effect of DMSO as solvent on cells was eliminated through vehicle control which showed no toxicity in cells.

3.2.2. Combined treatment

Combined treatment showed a significant decrease in cell viability as represented in Figure 3 (a). MCF-7 cells were treated with ethyl acetate extract (400 μg/mL) for 24 h followed by hypericin-PDT at 0.001-0.2 μM concentrations. An increase in the concentration of hypericin caused a significant decrease in cell viability. As shown in Figure 3 (b) the calculated IC_{50} dose of 0.06 μM for combination treatment is lower than effective concentration of hypericin PDT (IC_{50} ~ 0.09 μM) and extract treatment alone.

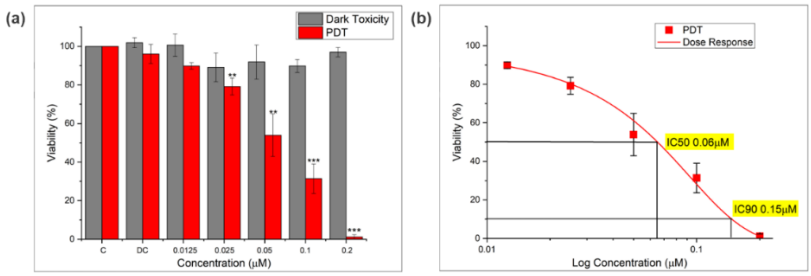


Figure 3. Effect of combination treatment. (a) Cell viability percent of MCF-7 cells treated with 400 μg/ml *P. granatum* aril ethyl acetate extract for 24 h and different hypericin concentrations (0.001-0.2 μM) irradiated at 1 J/cm^2. (b) Dose response curve of induced cytotoxicity by combination treatment. ($p < 0.05$ (*), $p < 0.01$ (**) or $p < 0.001$ (***)).

3.3. Morphological alterations

Morphological changes of untreated and treated cells were observed for signs of cell death and representative images for each group is shown in Figure 4.

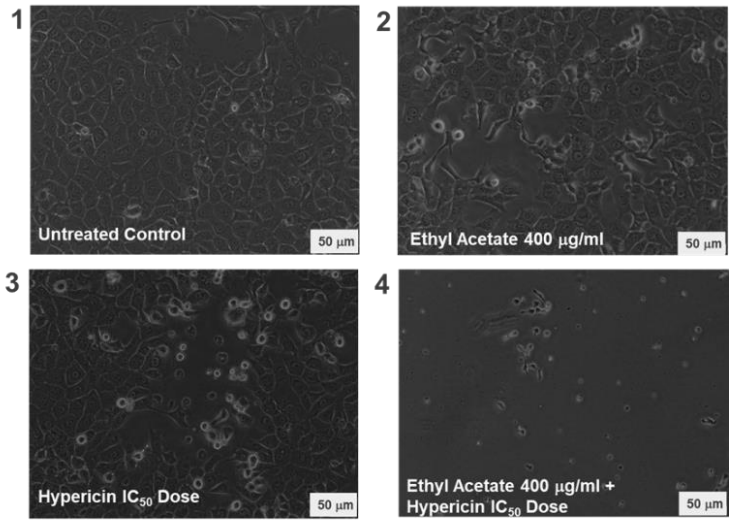


Figure 4. Morphological alterations of MCF-7 cells after treatment. Representative images of (a) Untreated cells, (b) *P. granatum* aril ethyl acetate extract treatment at 400 μg/ml for 24 h, (c) hypericin-PDT IC_{50} at 0.09 μM and (d) Combination treatment with *P. granatum* aril ethyl acetate extract at 400 μg with hypericin PDT IC_{50}. (200X Magnification, Scale bar, 50 μm).

Alone treatment with *P. granatum* aril ethyl acetate extract at 400 µg/mL (figure 4 (b)) on MCF-7 cells showed modest induction in cell damage while hypericin-mediated PDT at IC_{50} of 0.09 µM (figure 4 (c)) showed relatively more cells with damaged and rounded up morphology when compared to the untreated cells (Figure 4 (a)). Combination treatment showed significant number of cells with altered and damaged morphology along with a significant decrease in cells' population (Figure 4 (d)).

3.4. FITC Annexin V/PI assay

FITC Annexin V/PI assay was carried out to show the percentage of live, early apoptotic, late apoptotic, and necrotic cells following the different treatment. A representation of cell death for each treatment group is shown in Figure 5. As shown in Figure 5 (b), untreated MCF-7 cells showed ~ 98% of live cells while single treatment with *P. granatum* aril ethyl acetate extract at 400 µg/mL for 24 h showed a significant decrease in live cells (~22%) with a marked increase in cells in early apoptosis (~20%). Combination treatment showed only ~56% of live cells with increase in percentage of late apoptotic (~18%) and necrotic (~22%) cells with ~ 4% of early apoptosis cells.

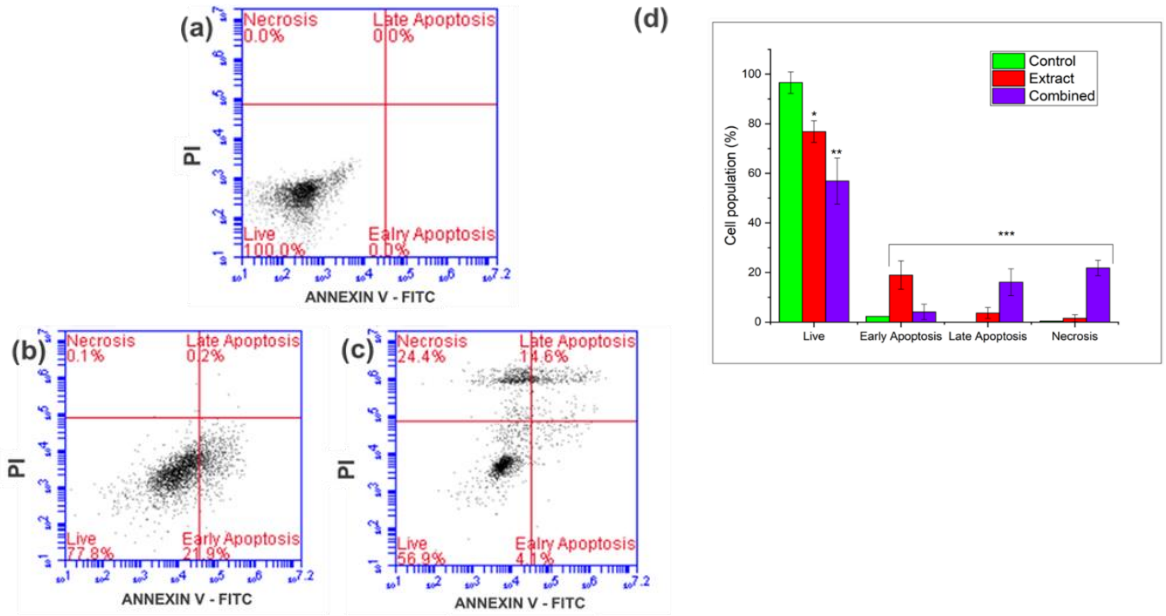


Figure 5. FITC Annexin V/PI assay. Representative histogram of (a) untreated MCF-7 cells (b) *P. granatum* aril ethyl acetate extract (400 µg/ml) treatment for 24 h (c) combination treatment (*P. granatum* aril ethyl acetate extract at 400 µg/ml + hypericin-PDT at IC_{50}) and (d) percentage of live, early apoptotic, late apoptotic and necrotic cells populations for each experimental group. ($p < 0.05$ (*), $p < 0.01$ (**) or $p < 0.001$ (***)).

4. Discussion

P. granatum extracts have shown the therapeutic effects against different types of cancer, including breast cancer [10]. In this study the ethyl acetate extract was found to be moderately cytotoxic at a concentration of 400 µg/ml at 48 h treatment (Figure 2(a)). Our study is the first report on evaluation of cytotoxic potential of South African cultivar of *P. granatum*. Cytotoxicity induced by extracted fraction from *P. granatum* arils can be explained due the presence of bioactive phytochemicals in the fruit. Others studies with *P. granatum* extracts on MCF-7 cells have reported, 320 µg/ml of methanol extract of peel treated for 48 h induced ~65% loss of cell viability [11], while IC_{50} dose of 377.88 µg/ml of ethanol extract of peel and 49.08 µg/ml of arils juice extract from P7 -Izmir 1513 Turkish cultivar was determined at 24 h [13]. However, the results on investigation on combination treatment utilizing extracted *P. granatum* arils phytochemicals with ethyl acetate fraction and hypericin-mediated PDT

against breast cancer cells showed strong cytotoxic effects, compared to single treatment with extract and hypericin-PDT (Figure 3(a)). Cell viability and morphological results (Figure 4) clearly showed hypericin-PDT significantly enhanced the cytotoxicity induced by extract alone, almost reaching a cell death of ~90% with 0.15 μM of hypericin-PDT in combination with 400 μg/mL extract treatment. Further, the FITC Annexin V/PI assay also showed increase in cell death percentage with almost ~40% of cells undergoing either apoptosis or necrosis in combined treatment group. This showed the combination therapy can induce both apoptosis and necrosis in MCF-7 cells, which can be due to additive effect of both the treatment. Although no study has been reported for combination of *P. granatum* extract with PDT, while one study have reported the combination effect of pomegranate extracts (40 μg/mL) and genistein (~20 μg/mL) on MCF-7 cells which induced apoptotic cell death in ~50% cell population, which was more effective than single treatments [14]. Although no direct comparison can be made from this study, but it can be theorised that effect of hypericin-PDT causes extreme cell damage which promotes the cells to undergo necrosis as compare to programmed apoptotic pathway induced by extract alone.

5. Conclusion

In conclusion, combination treatment with plant based non-phototoxic and phototoxic phytochemicals can be potential strategy for treatment of cancer with the advantages of less drug delivery doses, better therapeutic outcome and minimally induced unwanted toxicity.

Acknowledgements

The authors sincerely thank the South African Research Chairs initiative of the Department of Science and Technology and the National Research Foundation (NRF) of South Africa (Grant No. 98337), South African Medical Research Council (Grant No. SAMRC EIP007/2021), South African Development Community (SADC), as well as grants received from the NRF Research Development Grants for Y-Rated Researchers (Grant No. 137788), Council for Scientific Industrial Research (CSIR)-National Laser Centre (NLC), African Laser Centre (ALC), and University Research Committee (URC).

References

[1] Bray F, Jemal A, Grey N and Ferlay 2012 *Lancet Oncol.* **13** 790–801
[2] Sung H, Ferlay J, Siegel R L, Laversanne M, Soerjomataram I, Jemal A and Bray F 2021 *CA Cancer J. Clin.* **71** 209–249
[3] Azamjah N, Soltan-Zadeh Y and Zayeri F 2019 *Asian Pac. J. Cancer Prev.* **20** 2015–2020
[4] Waks A G and Winer E P 2019 *JAMA* **321** 316
[5] de Andrade G P, de Souza T F M, Cerchiaro G, Pinhal M A, Ribeiro A O and Girão M J 2021 *Photodiagnosis Photodyn. Ther.* **35** 102343
[6] Ahmadiankia N 2019 *Iran. J. Basic Med. Sci*. **22** 977–988
[7] Montaseri H, Nkune N W and Abrahamse H 2022 *J. Photochem. Photobiol. A* **11** 100136
[8] Gunaydin G, Gedik M E and Ayan S 2021 *Front. Chem.* **9**
[9] Mokhtari R B, Homayouni T S, Baluch N, Morgatskaya E, Kumar S, Das B and Yeger H 2017 *Oncotarget* **8** 38022–38043
[10] Moga M A, Dimienescu O G, Bălan A, Dima L, Toma S I, Bîgiu N F and Blidaru A 2021 *Mol*. **26** 1054
[11] Sreeja S, Santhosh Kumar T R, Lakshmi B S and Sreeja S 2012 *J. Nutr. Biochem.* **23** 725–732
[12] Manasathien J and Indrapichate K 2017 *J. Sci. Technol.* **24** 63-74
[13] Eroglu Ozkan E, Seyhan M F, Kurt Sirin O, Yilmaz- Ozden T, Ersoy E, Hatipoglu Cakmar S D, Goren A C, Yilmaz Aydogan H and Ozturk O 2021 *J. Food Biochem.* **45**
[14] Jeune M A L, Kumi-Diaka J and Brown J 2005 *J. Med. Food* **8** 469–475

Curcumin an emerging natural photosensitiser for lung cancer photodynamic therapy

G Kah, R Chandran* and H Abrahamse

Laser Research Centre, Faculty of Health Sciences, University of Johannesburg, P.O. Box 17011, Doornfontein 2028, South Africa

Email: rahulc@uj.ac.za

Abstract. Worldwide, lung cancer remains the main cause of malignant tumours and it contributes to a high mortality rate in cancer cases. Conventional therapeutic approaches for lung cancer are characterised as ineffective since resistance to therapy and toxic after-effects do occur in patients. Photodynamic therapy (PDT) is a promising therapeutic approach that requires a specific light source to activate a photosensitiser to induce cytotoxic cell death via the generation of reactive oxygen species. Curcumin is a natural compound from *Curcuma longa* that has been confirmed as a photosensitiser. This study aims to investigate the in vitro effects of curcumin-mediated PDT on lung cancer. Lung cancer cell lines (A549) were grown in complete media in an incubator maintained at 85% humidity, 5% CO_2, and a temperature of 37 ℃. Treatment of cells was achieved by exploring different concentrations of curcumin, and the cells were irradiated with a 470 nm diode laser at 10 J/cm^2 fluency. Post-PDT tests, including microscopic evaluation of treated cells and biochemical analysis, were performed to determine the cellular response of the cells treated. The results revealed morphological alterations and decreased viability in treated cells, which signifies cytotoxic damage. Thus, the increased cytotoxic effect established in this study suggests that curcumin can be an effective natural photosensitiser (PS) over synthetic ones for mediating lung cancer PDT. Further, in vivo studies may be needed to evaluate how curcumin can be utilized for lung cancer PDT in clinical settings.

1. Introduction

Lung cancer is a leading cause of human mortality worldwide, and both males and females are affected. These mortality rates are predominantly contributed by aggressive local invasion and distant metastasis in patients. Smoking highly predisposes people to the risk of lung cancer [1]. The conventional therapy for lung cancer includes chemotherapy, surgery, and radiotherapy. Hormones and immunotherapy are exploited as secondary treatment modalities. However, these therapies are noted to be very expensive and non-target-specific, leading to unsolicited toxicities, severe side effects, and general deterioration of the patients' effects [2].

PDT is stated as a low-cost and safe approach that is very promising for cancer treatment. PDT methods utilise light with an appropriate wave number to activate the photosensitiser. This will trigger reactive oxygen species (ROS) production, leading to apoptosis in tumour cells [3, 4].

Natural compounds are now greatly considered excellent candidates for novel drug discovery, and curcumin is mentioned as such. Curcumin is a yellow biological compound from the *Curcuma longa Linn* (turmeric) plant that has been largely used as a spice, food colourant, cosmetics, and drug [4]. It can be produced in affordably gram-scale quantities while requiring simple storage and handling.

Biologically curcumin is safe even at elevated doses of 12 g/kg per day, and the Food and Drug Administration (FDA, USA) affirms that it is generally safe [5]. It presents a wide spectrum of pharmacological and therapeutic efficacy, including antibacterial, antiviral, antioxidation, chemotherapeutic, antiamyloid, chemopreventive, anti-inflammatory, anticancer, antiarthritic, and thrombo-suppressive, with minimal or no aftereffects [4]. Curcumin can enhance different biochemical reactions by nucleophilic reactions (irreversible and reversible reactions) and donating hydrogen molecules to trigger oxidation. Moreover, curcumin has photokilling, photobiology, and photochemical potentials and is also noted to display a wide absorption spectrum (300–500 nm) [5, 6]. This indicates that curcumin may satisfactorily meet the prerequisites of an efficient photosensitiser. Nonetheless, it seems no clinical application of curcumin in cancer PDT has been approved to date. This study, therefore, explores curcumin as photosensitiser in lung cancer PDT.

2. Methodology

2.1. UV-visible spectroscopy

Curcumin (PS) can be photo-excited by blue light of different wavelengths depending on the solvent in which it is prepared [4]. Curcumin PS (C1386, Sigma-Aldrich) was freshly prepared in PBS according to the manufacturer's instructions, and the different concentrations (100, 50, 25, and 10 μM) from the stock preparation were subjected to analysis. The Uv-visible spectrophotometer (Genway, Lasec) was set at 200 nm to 800 nm, and the absorption of the PS at various concentrations was recorded.

2.2. Curcumin PS subcellular localisation in A549 cells

A549 cells were seeded on glass coverslips placed in 3.4 cm^2 culture plates (Corning Inc.) with 3 mL culture media and incubated for 24 hours to let the cells attach. The media was removed, and the cells were rinsed in PBS. Curcumin PS (16 μM) prepared in 3 mL of prewarmed media was pipetted into the culture plates, and the cells were incubated for 24 hours. The media was removed, and the cells were washed three times in PBS to get rid of the unabsorbed traces of the PS. The cells were maintained for 15 mins at room temp in 1 mL paraformaldehyde (Sigma-Aldrich) at 4% in PBS (v/v). One mL of 0.5% TrixtonX-100 prepared in PBS (v/v) was used to permeabilise the cells for 15 mins and the cells were washed three times with PBS. A 200 uL of 4-6-diamidino-2-phenylindole (DAPI, Invitrogen) at 300 nM concentration was used to counterstain the cell for 5 mins. The cells were washed, and the coverslips with stained cells were removed from culture plates and inverted onto slides, then sealed. A drop of the mounting medium was added, and the slides were examined using the fluorescent microscopic instruments (Carl Zeiss Axio Z1 microscope).

2.3. Cell culture and analysis

Lung cancer cell lines for this study were purchased from ATCC (A549, ATCC® CCL-185). The cells were seeded in culture flasks containing RPMI-1640 (Roswell Park Memorial Institute 1640) complete media and incubated at 85% humidity, 5% CO_2, and a temperature of 37°C to grow. Full-grown cells or monolayers of confluent cells were detached using TryplExpress and washed in Hanks Balances Salt Solution (HBSS) by centrifugation. Cells obtained were then seeded at 1 x 10^4 cells per well in 96-well plates for biochemical assay and 2.5 x 10^5 cells in 3.4 cm cultured plates for morphological and localisation analysis and incubated for 24 hours to enable the cells to attach.

2.3.1. Curcumin dark cytotoxicity

A549 cells treated in the dark for 24 h with various concentrations of curcumin (1, 2, 4, 8, 16, 32, and 64 μM) were washed with PBS to remove unabsorbed curcumin. Fresh media was added to the cells, and the cells were incubated for another 24 h. Viability analyses were assayed using an MTT assay (Proliferation Kit I (MTT), Sigma-Aldrich).

2.3.2. Irradiation treatment and post-analysis

A549 cells were seeded for biochemical and morphological analysis to attach. The attached cells were subjected to freshly prepared curcumin PS of various concentrations (1, 2, 4, 8, and 16 μM) for 4 h. PBS was used to wash the cells, and a final volume was pipetted into the culture for irradiation. The irradiation of cells was performed for 125 seconds, at 10 J/cm^2 fluency using a 470 nm blue light laser (Laser provided by the National Laser Centre of South Africa). After irradiation, PBS was removed and replaced with prewarmed fresh medium, and cells were then maintained for 24 h in the incubator. Post-irradiation analysis including, cell viability studies was done using the MTT assay (Cell Proliferation Kit I, Sigma), while the cytotoxic effect of the treatment was evaluated using the lactate dehydrogenase (LDH) cytotoxicity assay (CytoTox96®, G1780-Promega). Also, morphological studies were then performed using the light microscope.

2.4. Statistical analysis

All analyses were performed in triplicate, and the GraphPad Prism software was used to determine the IC_{50}, mean, and standard errors for the biochemical assays. The significant levels ($p < 0.05$*****, $p < 0.01$****,** or $p < 0.001$*******) between the control and tested variables were determined by exploring Dunnett's analysis of variances (one-way ANOVA) method.

3. Results and Discussion

3.1. *Uv-visible spectroscopy*

Figure 1 reveals the obtained spectra for the PS. The maximum absorption for curcumin was recorded at λ = 470.3 nm. A steady linear increase in absorption peak intensity was also noted as the concentrations of curcumin increased. This established wavelength for curcumin PS was then explored during the irradiation treatment of A549 cells.

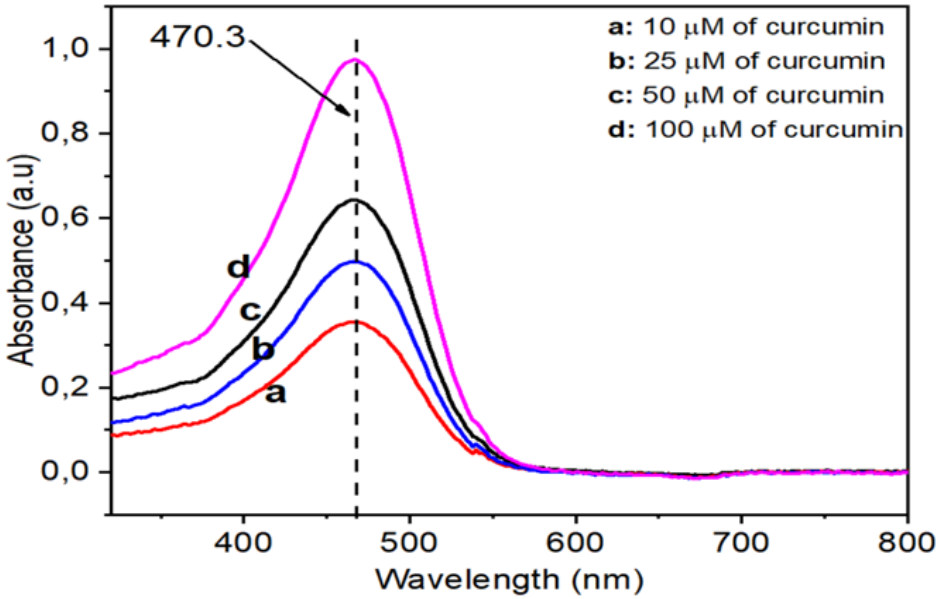


Figure 1. Curcumin PS absorption spectra of concentrations.

3.2. Curcumin PS subcellular uptake in A549 cells

The images (Figure 2) indicate a considerable cellular uptake of curcumin PS by A549 cells. The DAPI stain (control) counterstained the cell nuclei. The merged image shows an overlapping between the green fluoresces from curcumin PS and the blue fluoresce from the nuclei, which is suggestive of an efficient internalization of curcumin PS in A549 cells. Potential internalization of PS in the cell is noted as an important prerequisite for an effective photodynamic effect [7].

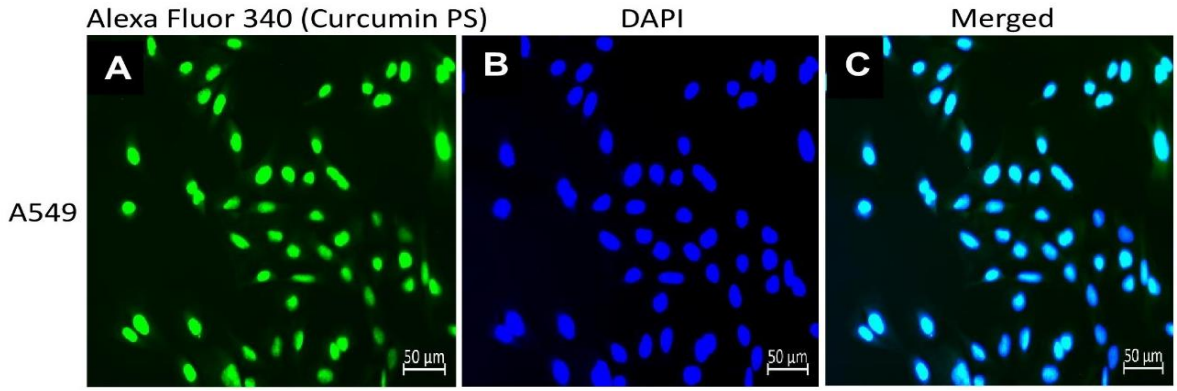


Figure 2. Fluorescence image with (A) fluorescence of curcumin (PS) uptake by A549 cells, (B) Blue fluorescence of DAPI from the cell nuclei (control), (C) light blue fluorescence of merged images of A and B.

3.3. Biochemical assays

3.3.1. Curcumin PS dark cytotoxicity

The dark cytotoxicity effect was assayed for non-irradiated curcumin-treated cells. Most of the concentrations of curcumin utilised did not produce significant toxicity when compared to the control (cell without treatment) (Figure 3A). However, significant decreases in cell proliferation were noted with cells treated with 32 and 64 µM of curcumin. The findings suggest curcumin PS at optimised concentrations of 1, 2, 4, 8, and 16 µM may be more suitable in lung cancer PDT since no dark cytotoxicity may be induced at these concentrations.

3.3.2. LDH cytotoxicity evaluation

The release of lactate dehydrogenase (LDH) by injured cells can be linked to cytotoxic effects associated with cell membrane damage or cell death [8]. No change in LDH was noted with cells treated with 1 and 2 µM of curcumin PS and laser light irradiation when compared with the untreated cells (control). However, a significant increase in LDH levels (***$p < 0.001$) was recorded for cells treated with light irradiation and curcumin PS at 4, 8, and 16 µM (Figure 3B). This indicates cytotoxic damage was induced via curcumin-mediated PDT. Also, the cytotoxic impact of curcumin in the mediation of PDT was noted to be highly significant with an increase in its concentration.

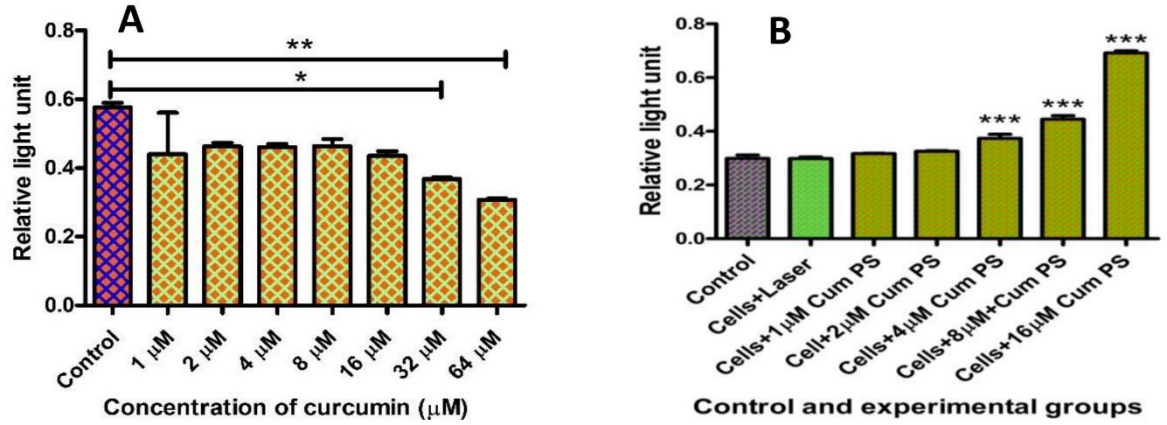


Figure 3. Biochemical assays with (A) Dark cytotoxicity after cell treatment with curcumin only, (B) LDH cytotoxicity post-PDT.

3.3.3. MTT cell viability evaluation

The viability of treated cells treated cells was determined using the MTT assay. Mitochondrial dehydrogenase can be produced by living cells. This enzyme can convert MTT to formazan, a purple crystal [9]. Levels of significance between the untreated cells (control group) and PDT groups at *p < 0.05 and ***p < 0.001 were noted for treated groups that utilised 4, 8, and 16 µM of curcumin PS. No significant differences were established for the treated groups that utilised 1 and 2 µM of curcumin (Figure 4C). The findings showed that curcumin PS and laser light treatments can substantially induced a decrease cell viability, depending on the concentration of the PS.

3.3.4. IC_{50} value determination

The lowest inhibitory concentration (IC_{50}) was determined using the data obtained from the MTT viability assay. The IC_{50} value, as indicated on the graph, is 3.5 µM. (Figure 4D). This implies that the optimal concentration for curcumin PS in lung cancer-mediated PDT is 3.542 µM, and this indicates a 50% induced cytotoxic effect.

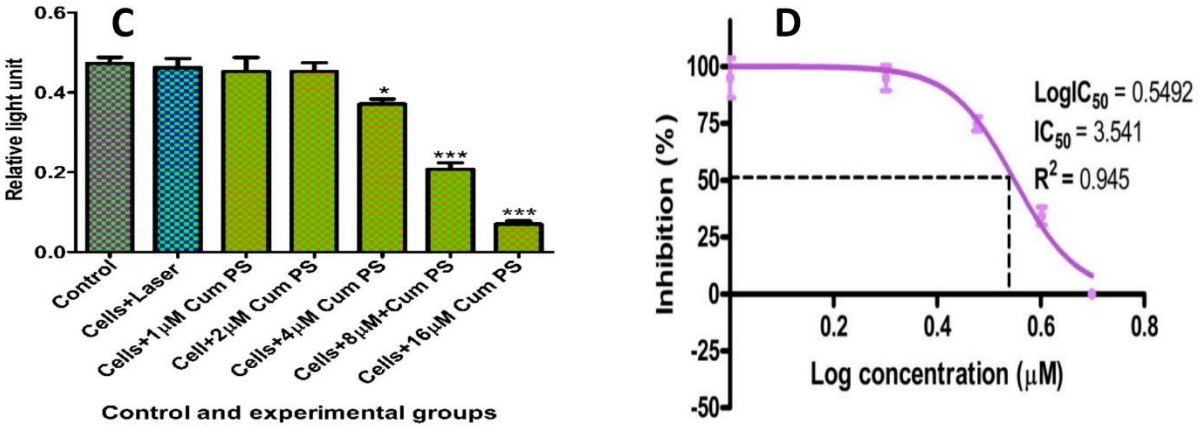


Figure 4. Biochemical assays with (C) MTT cell viability post-PDT, D) IC_{50} value estimation.

3.4. Morphological elevation of A549 cell post-PDT

The morphology of lung cancer cells without treatment (control) showed well-attached cells with normal morphology. Nonetheless, morphological changes were observed 24 h post-PDT included the shrinkage of cells, cell membrane damage, unattached cells, floating cells, and dead cells (Figure 5).

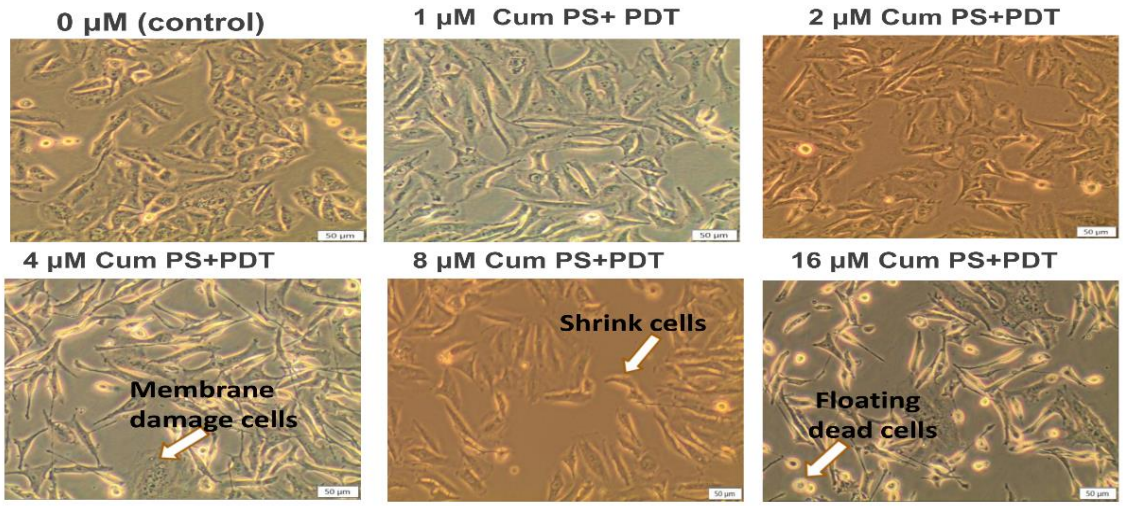


Figure 5. Microscopic image of A549 cells post-PDT

4. Conclusion

The effectiveness of curcumin-mediated PDT on A549 cells was demonstrated via cellular uptake, biochemical, and morphological studies. Effective uptake of curcumin by A549 cells was seen. Also, low viability and increased cytotoxicity were established following curcumin and light irradiation treatments of A549 cells. The curcumin-mediated PDT also triggered morphological alterations and cell death. This promising finding on curcumin in lung cancer PDT shows that curcumin can be a good natural photosensitizer to be exploited in cancer therapy.

Acknowledgments

The authors thank the Laser Research Center and the University of Johannesburg for their facilities. Appreciation also goes to the South African Research Chairs Initiative of the Department of Science and Technology (Grant No. 98337) and the University of Johannesburg GES 4.0 for funding this project.

References

[1] Testa U, Castelli G, Pelosi E 2018 Lung Cancers: Molecular characterisation, clonal heterogeneity and evolution, and cancer stem cells *Cancers (Basel)* **10** (8) 248

[2] Ailioaie L M, Ailioaie C, Litscher G 2021 Latest innovations and nanotechnologies with curcumin as a nature-inspired photosensitiser applied in the photodynamic therapy of cancer *Pharmaceutics* **13** (10) 1562

[3] Hempstead J, Jones D P, Ziouche A, Cramer G M, Rizvi I, Arnason S, Hasan T, Celli J P 2015 Low-cost photodynamic therapy devices for global health settings: characterisation of battery-powered LED performance and smartphone imaging in 3D tumor models *Sci Rep* **5** 10093

[4] Kah G, Chandran R, Abrahamse H 2023 Curcumin a natural phenol and its therapeutic role in cancer and photodynamic therapy: a review *Pharmaceutics* **15** (2) 639

[5] Kazantzis KT, Koutsonikoli K, Mavroidi B, Zachariadis M, Alexiou P, Pelecanou M, Politopoulos K, Alexandratou E, Sagnou M (2020 Curcumin derivatives as photosensitisers in photodynamic therapy: photophysical properties and in vitro studies with prostate cancer cells *Photochem Photobiol Sci* **19** (2) 193–206

[6] Dias L D, Blanco K C, Mfouo-Tynga I S, Inada N M, Bagnato V S 2020 Curcumin as a photosensitiser: From molecular structure to recent advances in antimicrobial photodynamic therapy *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* **45** 100384

[7] Halevas E, Arvanitidou M, Mavroidi B, Hatzidimitriou A G, Politopoulos K, Alexandratou E, Pelecanou M, Sagnou M 2021 A novel curcumin gallium complex as photosensitiser in photodynamic therapy: Synthesis, structural and physicochemical characterisation, photophysical properties and in vitro studies against breast cancer cells *Journal of Molecular Structure* **1240** 130485

[8] Forkasiewicz A, Dorociak M, Stach K, Szelachowski P, Tabola R, Augoff K 2020 The usefulness of lactate dehydrogenase measurements in current oncological practice *Cellular & Molecular Biology Letters* **25** (1) 35

[9] Kielbik A, Wawryka P, Przystupski D, Rossowska J, Szewczyk A, Saczko J, Kulbacka J, Chwiłkowska A 2019 Effects of photosensitisation of curcumin in human glioblastoma multiforme cells *In Vivo* **33** (6) 1857–1864

Photonic crystal-based biosensing for TB detection

Charles Maphanga [1, 2*], Saturnin Ombinda-Lemboumba [1], Yaseera Ismail [2], Patience Mthunzi-Kufa [1, 2, 3]

[1]Council for Scientific and Industrial Research, National Laser Centre, P O BOX 395, Pretoria, 0001, South Africa
[2]University of KwaZulu-Natal, School of Chemistry and Physics, Westville Campus, University Road, Durban, South Africa
[3]Department of Human Biology, Division of Biomedical Engineering, University of Cape Town, Cape Town 7935, South Africa

E-mail: CMaphanga@csir.co.za

Abstract. Over the last three decades, biosensors based on photonic crystals (PhCs) have been developed and continue to receive significant recognition owing to their distinctive electromagnetic properties and broad applications. Studies that have used refractive index as a parameter to design optical biosensing devices based on PhCs have resulted in optical devices that are sensitive with quick response time for small variations in samples. In this study, a 1-dimensional (I-D) PhC biosensor chip was developed for the diagnosis of TB. A pathogen-specific mycolic acid (MA) TB biomarker was studied based on the detection of refractive index changes on functionalised PhC biosensing surface. The MA biomarker was used as the biorecognition element to capture anti-mycobacterium tuberculosis antibodies, and a custom-built optical biosensing setup was used for optical biosensing to monitor biomolecular interactions between the antigen and antibody. Functionalised and successfully characterised gold nanoparticles (AuNPs) were introduced on the biosensing surface to enhance the detection signal. The biosensing surface was further characterised using atomic force microscopy (AFM). Analysis of biomolecular binding events on the biosensing surface was achieved using the optical biosensing setup by measuring transmitted light through the biosensor chip, and successfully distinguishing differences between the experiment and control samples. From our findings, it was realised that mycolic acid antigen can be used as a biomarker for active TB and can be successfully immobilised on a biosensing surface to capture anti-mycobacterium tuberculosis antibodies. It was evident that the PhC-based optical biosensing technique was successful in detecting small refractive index changes on the biosensing surface for the diagnosis of TB. These results pave the way toward the development of a PhC-based point-of-care (POC) diagnostic device for TB.

1. Introduction

The PhCs are periodic nanostructures that manipulate the flow of light and have garnered significant interest in the field of biosensing with their potential application in the development of point-of-care (POC) diagnostic devices. PhC-based biosensing plays a crucial role in the development of point-of-care diagnostics. They can be engineered to exhibit strong light-matter interactions, resulting in enhanced sensitivity for detecting analytes. By tailoring the structural properties of PhCs, such as the periodicity and refractive index contrast, the interaction of light with biomolecules can be optimised, leading to increased sensitivity in detecting molecular binding events [1–3]. PhC-based biosensors offer

label-free detection, eliminating the need for fluorescent or radioactive labels [4–6]. This simplifies the assay procedure and reduces costs. The interaction of analytes with the PhC structure causes changes in the optical properties, such as the shift in the wavelength or change in intensity, enabling direct detection without the use of additional labels or probes [6].

One of the advantages of PhC-based biosensors is their capability for real-time monitoring of biomolecular interactions. The changes in the optical properties of the PhC structure occur instantaneously upon analyte binding, allowing the detection and analysis of binding kinetics and affinity [6]. Real-time monitoring enables dynamic measurements, enhancing the understanding of biomolecular interactions and enabling timely diagnostic decisions. In this work, a 1-D PhC surface was modified and used as the biosensor chip. A PhC biosensor chip was developed by immobilising the mycolic acid TB antigen surface of the PhC. Mycolic acid was used as a biorecognition element to specifically bind the anti-*mycobacterium tuberculosis* antibody as the analyte. Furthermore, gold nanoparticles (AuNPs) were covalently conjugated with a secondary antibody and introduced on the biosensing surface. The introduction of functionalised AuNPs to the biosensing surface plays a crucial role in changing the material refractive index. AuNPs have unique optical properties that can be easily manipulated compared to bulk gold. These special characteristics provide a plethora of advantages, especially in biosensing and diagnostic imaging. One of the major attributes is the interchange of light with electrons on its exterior [7]. The resonance wavelength and bandwidth of the AuNPs are dependent on the shape, particle size, refractive index of the surrounding media, and temperature [7]. Herein, Covalent conjugation chemistry was applied for the bioconjugation of AuNPs with the secondary antibody. EDC and NHS chemistry were used to bind goat anti-rabbit IgG H&L secondary antibody to PEG. Bioconjugation occurs upon successful binding of the antibody to the AuNPs surface, thus resulting in a wavelength shift in absorption spectrum by a few nanometres when analysed using UV-vis absorption spectroscopy [8].

The AuNPs were characterised before and after bioconjugation using ultraviolet-visible (UV-vis) spectroscopy before introducing them on the biosensing surface. The biosensing surface was also characterised layer-by-layer during the development process of the biosensor chip using atomic force microscopy (AFM). The biosensor chip was analysed using a custom-built PhC setup, transmitted light intensity was measured, and a shift in the wavelength due to refractive index changes was detected using a spectrometer. The acquired data were analysed using OriginPro software to visualise the transmission spectra between the neat PhC, experiment, and control samples.

2. Materials and Methods

2.1. Atomic force microscopy

The AFM was used to perform layer-by-layer characterisation of the biosensing surface for surface morphology studies of the PhC biosensor chips at different stages of the functionalisation process. Herein, the 1-D PhC biosensor chip was washed with absolute ethanol and blow-dried with nitrogen (N_2) gas. The chip was subsequently subjected to oxygen plasma treatment. The PhC was immersed in (3-Aminopropyl) triethoxysilane (APTES) (1%, v/v) (Sigma-Aldrich, A3648) at room temperature overnight. Post incubation, the PhC was washed thoroughly by sonication in absolute ethanol for 5 minutes and followed by sonication in ultrapure water (resistivity of 18 MΩ) for 5 minutes. The clean PhC was blow-dried with N_2 gas and placed in the incubator at 90°C for 1 hour. The silanised PhC surface was functionalised with stearic acid (STA) (141 mM) (Sigma-Aldrich, S4745). Mycolic acid (MA) from *mycobacterium tuberculosis* (bovine strain) (0.5 mg/mL) (Sigma-Aldrich, M4537) was immobilised on the STA monolayer by incubating the chip in a solution of MA overnight at room temperature. Post incubation, the chip was rinsed thoroughly with 1x phosphate buffered saline (PBS) (Sigma-Aldrich, D1283). A solution of bovine serum albumin (BSA) (1%, v/v) (Roche Diagnostics, 10711454001) was added to the chip for an hour. Anti-*mycobacterium tuberculosis* primary antibody (8.8 mg/ml) (Abcam, ab905) added to the MA monolayer and incubated overnight. A goat anti-rabbit IgG H&L secondary antibody (2.1 mg/ml) (Abcam, ab6702) was conjugated to gold nanoparticles

(AuNPs) (Sigma-Aldrich, 797707) and herein called the bioconjugate. (AuNPs-IgG) and introduced on the surface containing primary antibody. In the specifity control sample, the primary antibody was not introduced on the surface; however, the TB bioconjugate was introduced on the surface containing MA. AFM images were captured with a Vecco AFM system (Digital Instruments, USA).

2.2. Transmission electron microscopy

The morphological analysis of the AuNPs used in this study was performed using a High-Resolution Transmission Electron Microscope (HRTEM) (JEOL JEM-2100, Japan) operated at a voltage of 200 Kilovolt (KV). A copper grid was used for sample preparation and dipped into a sample of AuNPs suspended in PBS. The HRTEM operates at an imaging mode with high magnification, thus enabling high-resolution imaging of the crystallographic structure of nanomaterials at the atomic scale level. In this analysis, the HRTEM analysis was used for analysing the morphology, particle size and shape to confirm the structural properties of the AuNPs for subsequent studies.

2.3. Ultraviolet-visible (UV-vis) absorption spectroscopy

UV-vis absorption spectroscopy was performed to characterise the gold nanoparticles (AuNPs) before and after bioconjugation. Covalent and direct bioconjugation strategies were explored. To perform covalent conjugation, N-hydroxysuccinimide (NHS) (13 mM) (Sigma-Aldrich, 56480) and 1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide) (EDC) (5 mM) (Sigma-Aldrich, E7750) with similar molar ratio were added to heterobifunctional polyethylene glycol (PEG) (SH-PEG-COOH) (Sigma-Aldrich, 757845) and shaken for an hour on an orbital shaker. Subsequently, the anti-*mycobacterium tuberculosis* antibody was introduced to the reaction mixture and incubated for 2 hours. The mixture was complex was added to 200 µL of AuNPs and incubated for an hour on an orbital shaker. To perform physical conjugation, the anti-*mycobacterium tuberculosis* antibody was added directly to 200 µL AuNPs. The mixture was incubated on an orbital shaker for 30 min. Additionally, neat AuNPs were used as a reference. Bioconjugation was confirmed by measuring the samples using NanoDrop 8000 spectrophotometer (Thermo Fisher Scientific, United States).

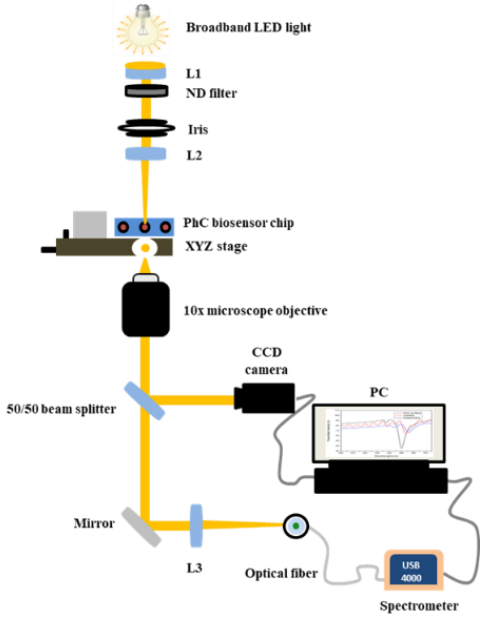


Figure 1: A custom-built photonic crystal biosensing setup used to measure transmitted light. The setup consists of a broadband LED light source, collimating lenses (L1 and 2), XYZ stage, a 10x microscope objective, and an imaging system consisting of a CCD camera and computer.

2.4. Photonic crystal biosensing

A custom-built optical biosensing setup with a transmission spectroscopy conformation was used to perform PhC biosensing in air (Figure 1). The setup was assembled using a broadband light-emitting diode (LED) light source. The light was collimated using a combination of collimating lenses and focused on the sample stage. Light transmitted through the PhC biosensor chip on the stage was collected using a 10x microscope objective. A 50/50 beam splitter directed the light to the imaging system and the spectrometer using a fiber. Transmitted light was focused into the fiber using a focussing lens and directed to the USB 4000 portable spectrometer (Ocean Optics Inc, United States) connected to the computer. The imaging system enabled the visualisation of biosensing activities on the sample stage and was formed by a charge-coupled device (CCD) camera connected to a computer. SpectraSuit software was used to perform spectroscopic analysis of the different samples. The samples were prepared in triplicates under the same conditions and the analysis was repeated three times ($n=3$). During sample analysis, background measurements were taken before each recording. The acquired data was further analysed using OriginPro8 software to plot wavelength-shift differences between the samples.

3. Results

3.1. Atomic force microscopy

Atomic force microscopy was used in this study to perform layer-by-layer characterisation of the PhC biosensor chip surface during the chip functionalisation process. The samples included neat PhC, control, and experiment samples. The control sample consisted of mycolic acid immobilised on the biosensing surface as the biorecognition element and the goat anti-rabbit IgG H&L secondary antibody bioconjugated to AuNPs (TB-bioconjugate). There was no primary antibody on the surface of the control sample. Mycolic acid used in this study can only bind to the anti-*mycobacterium tuberculosis* antibody, and the secondary antibody is designed to bind to the primary antibody. The experiment sample consisted of mycolic acid, an anti-*mycobacterium tuberculosis* primary antibody, and the TB bioconjugate. The images taken using AFM are shown in (Figure 2), and each of the scans represents a 5µm X 5µm lateral area that was scanned.

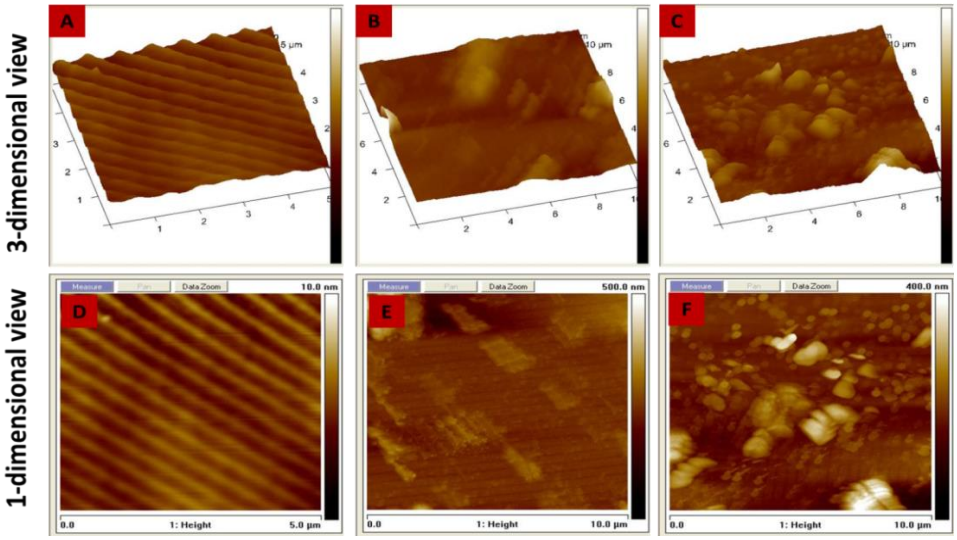


Figure 2: Atomic force microscopy micrographs of the PhC, control, and experiment group displayed in 3-dimensional and 1-dimensional view projection. Figures A and D show the 1-D PhC displayed in 3-D and 1-D view. Figures B and E show the control sample in 3-D and 1-D view projection. The experiment sample displayed much more surface roughness compared to the control sample.

Figure 2A and D shows the neat 1-D PhC surface depicting the periodic arrangement of the grating structure. The consistent crest and trough pattern of the PhC periodicity is represented in Figure 2A. The surface is uniform, with no signs of irregularities, defects, or contaminants. Prominent surface roughness was observed in the experiment sample (Figure 2C and F) when compared to the control sample (Figure 2B and E). The Z-range increased from 150 nm in the neat 1-D PhC sample used as a reference to 344 nm in the control sample. The experiment sample containing immobilised bioconjugated AuNPs displayed a z-range of 473 nm.

3.2. Transmission electron microscopy

The morphology and size of AuNPs were evaluated. Figure 3 shows the TEM image of bare AuNPs, specifically Au NanoUrchins used in this study. Under TEM analysis, the AuNPs appeared multi-branched with spiky, uneven surfaces. The AuNPs had an average size of 60 nm.

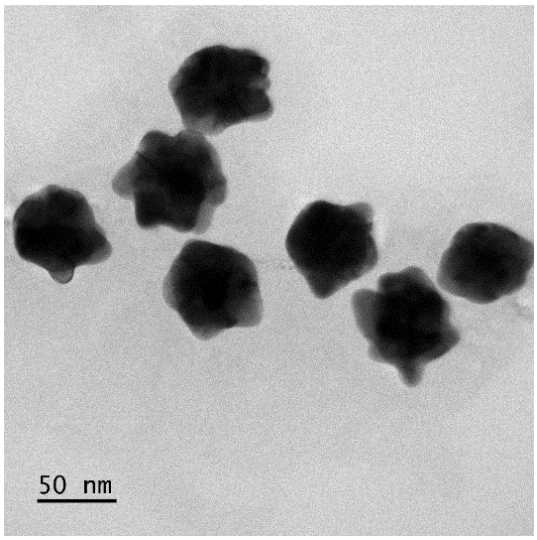


Figure 3: Transmission electron micrograph of gold NanoUrchins. The nanoparticles displayed a multibranched, spiky, uneven surface morphology with an average size of 60nm. The length of the scale bar is 50nm.

3.3. UV-vis absorption spectroscopy

To characterise the AuNPs before and after bioconjugation using covalent conjugation chemistry and physical conjugation, UV-vis absorption spectroscopy was performed (Figure 4). At 250 nm – 350 nm, there is no absorption peak for the AuNPs. At 250 nm – 350 nm, the high absorption intensity of antibodies physically conjugated (TB-Bioconjugate-P) to the AuNPs was detected, and lower absorption intensity was recorded for the covalently bioconjugated antibodies (TB-Bioconjugate-C). At 500 nm – 750 nm, approximately 645 nm absorbance peak for the AuNPs was detected, and a redshift in both the covalent and physically conjugated AuNPs was detected. In this region, a notable difference in absorbance intensity was also recorded between the AuNPs, TB-Bioconjugate-P, and TB-Bioconjugate-C.

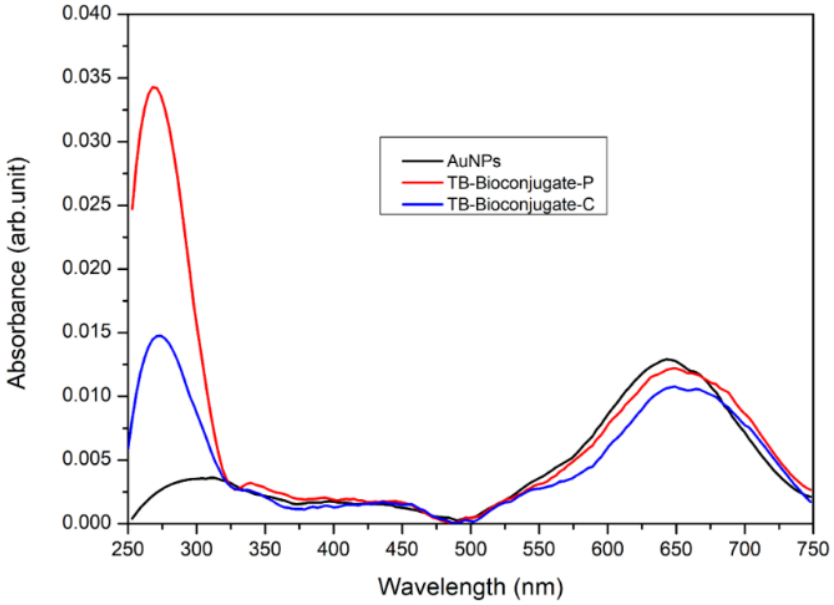


Figure 4: Absorbance spectra of neat AuNPs, covalent conjugated AuNPs (TB-Bioconjugate-C), and physical bioconjugated AuNPs (TB-Bioconjugate-P). There are clear differences in absorbance intensity of the antibody (250 nm – 350 nm) and the AuNPs (500 nm – 750 nm) when comparing the two bioconjugation techniques.

3.4. Photonic crystal biosensing

A custom-built PhC biosensing setup was used to differentiate and detect differences between the neat 1-D PhC, MA, MA with immobilised primary antibody (PAb), MA with immobilised PAb and bioconjugated secondary antibody (SAb), and MA with immobilised bioconjugated SAb (Figure 5). Transmitted light intensity was measured between the five samples. There were clear resonance wavelength-shift differences between the samples analysed. The neat PhC reference sample (herein referred to as the reference sample) was detected at a resonance wavelength of 650.08 nm, which corresponds to the manufacturer's recommendations. The MA sample was detected at a wavelength of 653.28 nm, MA with immobilised PAb was detected at a wavelength of 660.27 nm, MA with immobilised PAb and bioconjugated SAb was detected at 664.45 nm, and MA with immobilised SAb (herein referred to as the control sample for specificity) was detected at 653.28 nm. A redshift of 3.2 nm was detected when MA was compared to the reference sample, a redshift of 14.37 nm was detected when the sample containing immobilised PAb (herein referred to as the experiment sample) was compared to the reference sample, a redshift of 10.19 nm was detected when the sample containing MA/PAb and AuNPs bioconjugated SAb (herein referred to as the signal enhancement sample) was compared to the reference, and a redshift of 6.8 nm was detected when the sample containing AuNPs bioconjugated SAb (herein referred to as the specificity control sample) was compared to the reference sample. Additionally, a blueshift of 11.17 nm was detected between the specificity control sample and the experiment sample.

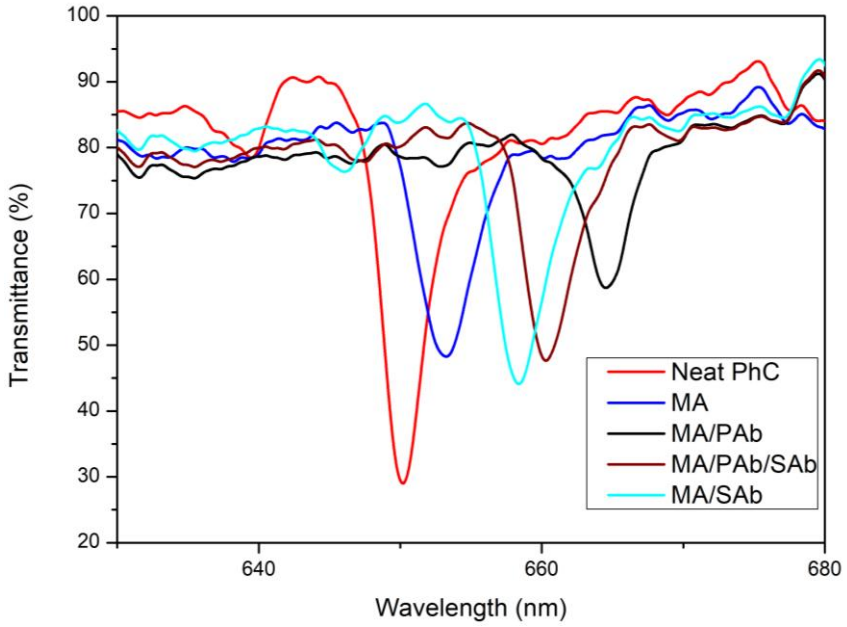


Figure 5: Photonic crystal spectrograph for the detection of anti-*mycobacterium tuberculosis* antibodies. The resonance wavelength of neat 1-D photonic crystal (PhC), mycolic acid (MA) mycolic acid with immobilised primary antibody (PAb), MA with immobilised PAb and bioconjugated secondary antibody (SAb), and MA with immobilised bioconjugated SAb was obtained using the custom-built PhC biosensing setup. There were clear detectable resonance wavelength differences between the five samples.

4. Discussion

The PhC-based biosensing surface was characterised layer-by-layer using AFM during the surface functionalisation process to confirm the binding of molecules on the surface. Immobilisation of molecules on the biosensing surface was achieved through covalent attachment and the formation of self-assembled monolayers. Samples, including the neat PhC, control, and experiment, were studied. The neat PhC was used as the reference sample. A biorecognition element (MA) was immobilised on both the control and experiment samples. However, a complementary PAb was not introduced on the surface of the control sample, but rather a bioconjugated SAb was introduced to both the control and experiment samples. In the control, the bioconjugate was used to test for the specificity of the biorecognition element. In the absence of a primary antibody on the surface of the control sample, the prominent surface roughness observed in the experiment sample (Figure 2C and F) was not detected (Figure 2B and E). Various structures resembling AuNP-antibody complexes were visualised in Figures C2 and F of the experiment sample (Figure 2). This also supports findings by Lee et al., 2015 who demonstrated the possibility of visualising AuNP-antibody complexes as a result of the presence of immobilised AuNPs. The increased surface roughness in the experiment sample could be attributed to the immobilisation of the AuNPs contained in the bioconjugate bound to the experiment sample. This is also supported by the Z-range values of 150 nm in the neat 1-D PhC sample, 344 nm in the control sample, and 473 nm in the experiment sample. This can be an indication that the immobilisation of the SAb to the AuNPs was successful, and further attached to the PAb in the experiment sample.

On TEM, the AuNPs appeared multibranched with a spiky uneven surface. Multibranched and spiky AuNPs have a highly porous and rough surface structure, providing a significantly larger surface area compared to their spherical counterparts [10]. This increased surface area allows efficient immobilisation of biomolecules, leading to enhanced sensitivity in biosensing applications. The larger surface area also results in improved binding affinity, thus improving the detection limits of the

biosensor [10]. The unique surface morphology of multibranched and spiky AuNPs can lead to signal amplification in biosensing. The presence of multiple branches or spikes on the nanoparticles creates numerous nanogaps, which can trap and amplify the signal produced during the sensing process. This amplification effect enhances the detection sensitivity and improves the signal-to-noise ratio of the biosensor [10–12].

The AuNPs were further functionalised with a signal or secondary antibody using both the covalent and physical bioconjugation techniques. Bioconjugation occurs upon successful binding of the antibodies to the AuNPs surface, thus resulting in LSPR spectrum redshift which is an increase in wavelength by a few nanometres [13]. From the findings, it was realised that covalent conjugation resulted in high binding efficiency of antibodies to the surface of the AuNPs indicated by the sharp decline in the absorption intensity of the antibodies at 250 nm – 350 nm and a redshift in the wavelength at 500 nm – 750 nm (Figure 4). The physical conjugation strategy resulted in lower binding efficiency of antibodies to AuNPs indicated by the high absorbance intensity of the antibodies at 250 nm – 350 nm, and a less prominent redshift in the wavelength at 500 nm – 750 nm (Figure 4). The AuNPs do not absorb at 250 nm – 350 nm; hence no absorption peak was detected.

After the characterisation of the biosensing surface, PhC-based biosensing of the chip in air was performed using the custom-built optical biosensing setup. During analysis, wavelength shift between the samples was studies. There were distinct differences in wavelength redshift between the five samples analysed. The redshift detection is due to changes in the refractive index between samples [14]. The reference sample (neat PhC) exhibited a resonance wavelength of 650.08 nm, consistent with the manufacturer's specifications. Notably, the MA sample demonstrated a redshift of 3.2 nm compared to the reference, while the introduction of immobilised PAb led to a larger redshift of 14.37 nm, marking the experiment sample and a positive TB sample in diagnostics. Further modifications involving the incorporation of AuNPs bioconjugated SAb onto the MA/PAb structure resulted in a redshift of 10.19 nm when compared to the reference sample. However, this shift was not greater than the one detected in the absence of the bioconjugate (14.37 nm). This indicates the possibility of detecting anti-*mycobacterium tuberculosis* antibodies without the introduction of nanomaterials. Additionally, the specificity control sample involving AuNPs bioconjugated SAb showed a redshift of 6.8 nm relative to the reference, serving as a benchmark for specificity. Moreover, a notable blueshift of 11.17 nm emerged when comparing the specificity control sample (MA/SAb) to the experiment sample (MA/PAb). Overall, the differences in resonance wavelength shifts across the various samples provide valuable information and a diagnostic foundation that can be used to distinguish between negative and positive TB samples, offering potential implications for sensor development for TB diagnostics. These findings suggest a possible diagnosis of TB using a label-free PhC optical biosensing platform, and further exploration in this direction could lead to advancements in bioanalytical techniques and sensor design.

5. Conclusion

In conclusion, the study successfully detected anti-*mycobacterium tuberculosis* antibodies using a 1D photonic crystal-based biosensor chip. Mycolic acid was effectively immobilised on the biosensor chip to facilitate specific binding to anti-*mycobacterium tuberculosis* antibodies. To rule out non-specific binding, a bioconjugated secondary antibody was introduced, which did not bind to the biorecognition element, and this was confirmed using atomic force microscopy (AFM). AFM also revealed distinct differences in height and surface roughness among the samples studied. The bioconjugated secondary antibody was successfully bound to the primary antibody on the biosensing surface, inducing a change in surface refractive index. Additionally, the developed custom-built PhC optical biosensing setup enabled the detection of refractive index changes through wavelength shifts between the five samples studied. The observed redshifts in resonance wavelengths across various samples were indicative of specific binding interactions. The distinct spectral responses and shifts provide a significant diagnostic foundation for the detection of negative and positive TB specimens. These findings collectively suggest the promising prospect of diagnosing tuberculosis using the label-free PhC optical biosensing platform.

Subsequent studies will focus on the reproducibility of the assay for label-free detection and the limit of detection studies.

References

[1] Wang H, Zhang K.Q, *Sensors (Switzerland)*. **13** (2013) 4192–4213. https://doi.org/10.3390/s130404192.

[2] Konopsky V.N, Karakouz T, Alieva E.V, Vicario C, Sekatskii S.K, Dietler G, *Sensors (Switzerland)*. **13** (2013) 2566–2578. https://doi.org/10.3390/s130202566.

[3] Tokel O, Inci F, Demirci U, *Chem. Rev*. **114** (2014) 5728–5752. https://doi.org/10.1021/cr4000623.

[4] Luan E, Shoman H, Ratner D.M, Cheung K.C, Chrostowski L, *Sensors (Switzerland)*. **18** (2018) 1–42. https://doi.org/10.3390/s18103519.

[5] Cunningham B.T, Zhang M, Zhuo Y, Kwon L, Race C, *IEEE Sens. J*. **16** (2016) 3349–3366. https://doi.org/10.1109/JSEN.2015.2429738.

[6] Dorfner D, Zabel T, Hürlimann T, Hauke N, Frandsen L, Rant U, Abstreiter G, Finley J, *Biosens. Bioelectron*. **24** (2009) 3688–3692. https://doi.org/10.1016/j.bios.2009.05.014.

[7] Shafiqa A.R, Abdul Aziz A, Mehrdel B, *J. Phys. Conf. Ser*. **1083** (2018). https://doi.org/10.1088/1742-6596/1083/1/012040.

[8] Zhang X, *Nanophotonics*. **10** (2021) 1329–1335. https://doi.org/10.1515/nanoph-2020-0603.

[9] Lee J.H, Kim B.C, Oh B.K, Choi J.W, *J. Nanosci. Nanotechnol*. **15** (2015) 1117–1122. https://doi.org/10.1166/jnn.2015.9336.

[10] Stassi S, Cauda V, Canavese G, Manfredi D, Roppolo I, Martino P, 2014, *Handbook of Nanomaterials Properties*. Berlin Heidelberg, Springer. https://doi.org/10.1007/978-3-642-31107-9.

[11] Zou Y, Chen H, Li Y, Yuan X, Zhao X, Chen W, Cao F, Cai N, Huang X, Yang F, Liu W, *J. Saudi Chem. Soc*. **23** (2019) 378–383. https://doi.org/10.1016/j.jscs.2018.11.005.

[12] He M, Cao B, Gao X, Liu B, Yang J, *J. Mater. Res*. **34** (2019) 2928–2934. https://doi.org/10.1557/jmr.2018.503.

[13] Oliveira J.P, Prado A.R, Keijok W.J, Wagnner P, Antunes P, Yapuchura E.R, Cesar M, Guimarães C, *Sci. Rep*. **9** (2019) 1–8. https://doi.org/10.1038/s41598-019-50424-5.

[14] Maphanga C, Manoto S, Ombinda-lemboumba S, Ismail Y, *Sens. Bio-Sensing Res*. **39** (2023) 100545. https://doi.org/10.1016/j.sbsr.2022.100545.

DIVISION

D

ASTROPHYSICS AND SPACE SCIENCE

Kepler and *Gaia* DR2 views of open cluster NGC 6819

G.Mekonnen[1], O.P Abedigamba[2]

[1]Department of Physics, University of Zululand, South Africa
[2]Department of Physics, Kyambogo University, Uganda.

Abstract. In this work we derive the distance modulus from *Gaia* parallaxes for *Kepler* Input Catalogue (KIC) stars located within 10 arcmin in the field of NGC 6819. The mean distance modulus is used to determine probable new KIC members of NGC 6819 cluster. We find 7 probable KIC new astrometric distance members which have been previously overlooked in the literature as revealed by *Gaia* Data Release 2 (DR2). Of these, one is a solar-like star (SOLR), one γ Dor star (GDOR), 2 rotational variables, 1 eclipsing binary and 2 unknown types. There is no signifant difference in the mean distance modulus when probable new astrometric members are included in deriving the mean distance modulus.

1. Introduction

Space photometric and astrometric observations obtained with missions such as *Kepler* and *Gaia* respectively, have been useful in the study of stellar variability. The *Gaia* space craft, a European Space Agency mission, was launched in 2013 to provide provide the astrometry of stellar sources and precise 3D map of our galaxy. The second data release of *Gaia* contains parallaxes and proper motions for a large sample of stars (*Gaia* DR2; [13]). *Gaia* DR2 parallaxes have been useful in obtaining precise distance and luminosities of pulsating stars [2] used to refine the Hertzsprung–Russell diagram diagram. Pulsating stars are located in both the field and open clusters. The stellar clusters belonging to the disk of the Galaxy are refered to as open clusters and constitute ideal laboratories to study stellar formation and evolution [9]. Open clusters are groups of stars with the same age and same initial chemical composition, formed at a single event from the same gas cloud. [9] used *Gaia* DR2 to obtain a list of members and mean cluster parameters such as distances and proper motions for 1229 clusters. The open cluster NGC 6819 was among the clusters studied by [9].

NGC6819 centered at, RA = 19 : 41 : 18 , DEC= 40 : 11 : 12, is one of the open clusters located in the *Kepler* and *Gaia* field of view. It is a moderately old cluster (≈ 2.5 Gyr; [4]) with near-solar or slightly super-solar metallicity ([Fe/H]= $+0.09 \pm 0.03$; [7]). Using photometry from the *Kepler* Input Catalogue, [3] determined an unreddened distance modulus of 12.20±0.06 mag. [10] found the distance modulus to be 12.22 mag in BV photometry. [3] found eight δ Scuti, 39 new solar-like candidates, three γ Dor variables and several eclipsing binaries in the field with some being photometric distance members. We studied NGC 6819 because it is special in a sense that it has dozen stars in the range between 10^{th} and 12^{th} magnitude in $5'$ circle. Moreover, in *kepler* field of view in the years between $2009 - 2013$, the center of the galaxy was observed in 200×200 pixel and provided a unique high precision long time series data set.

[1] mengistieg@unizulu.ac.za
[2] oyigamba@gmail.com

Determination of membership of stars in open clusters is not an easy process; one needs to take all available information into account. All stars in a cluster tend to have the same motion in three dimensions. The proper motion in Right Ascension (pmRA) and proper motion in Declination (pmDec) can be measured and members tend to have the same proper motions. A star is probably a member of the cluster if it has the same RV or pmRA and pmDec (or both). In addition, one can use distance modulus to deduce membership. The mean proper motion, radial velocity and distance modulus for all cluster members in NGC 6819 were found to be pmRA = -2.6±0.5, pmDec = -4.2±0.5 mas/yr [23], 11.747±0.086 mag [1].

In this paper we study the field containing the open cluster NGC 6819 using *Gaia* DR2 parallaxes & proper motions and time series photometry obtained using the *Kepler* space mission. The aim is to study the probable new members in detail with Asteroseismolgy.

2. Data

Part of the data in this study were obtained with NASA's *Kepler* telescope. We used the publicly available *Kepler* data (Q0 - Q17) in this study, where Q stands for quarters referring to the interval in which the data are downloaded on certain time interval, i.e., Q0 is a 10-d commissioning run. The long cadence (LC) data were used. The light-curve files of the stars contain simple aperture photometry (SAP) flux and a more processed version of SAP with artefact mitigation included, i.e., presearch data conditioning (PDC) flux. Jumps between quarters were calculated and resulting periodograms were obtained. For the *Kepler* Input Catalogue (KIC) stars, revised effective temperatures were used. Parallaxes and proper motions were obtained from the second data release of Gaia.

3. Method and analysis

We selected all *Kepler* Input Catalogue (KIC) stars measured by *Gaia* within a radius of 10 arcmin of the cluster centre and with *Kepler* magnitude Kp $<$ 15 mag. In the selected region the field contains 233 KIC stars. Fig. 1 shows the luminosity-effective temperature (HR) diagram with luminosities derived from *Gaia* parallaxes.

To determine the luminosities from Gaia DR2 parallaxes, one requires knowledge of the apparent magnitudes, interstellar extinctions, and bolometric corrections. A correction for interstellar extinction needs to be applied to the r magnitude using r_0 = r - 0.874A_V [21]. The value of A_V listed in the *Kepler* Input Catologue (KIC) is from a simple reddening model which depends only on galactic latitude and distance. In the case of the cluster NGC 6819, we take uniform A_V as was previously done [3, 1]. From the *Gaia* DR2 parallax (π), the absolute magnitude is calculated using $r_{abs} = r_0 + 5(\log10\ \pi + 1)$. The absolute bolometric magnitude is given by $M_{bol} = r_{abs}$ + BC_r - $M_{bol\odot}$ with the solar absolute bolometric magnitude $M_{bol\odot}$ = 4.74 and BC_r is the bolometric correction.

The luminosity relative to the Sun is found using log $L/L_\odot$ = -0.4M_{bol} and the resulting values for each of the 233 selected stars are shown in Fig. 1 and Fig. 2 as filled circles. Probable cluster members are shown as filled circles (black) in Fig. 1. These were matched with the KIC stars within 10 arcmin of the cluster centre. We used the well known KIC cluster members to obtain distance modulus (DM) to the cluster (DM = 12.30±0.12 mag). New members were then chosen from those with a DM close to the mean value that we have found, together with the *Gaia* mean proper motions (pmRA = -2.920±0.035 mas/yr, pmDec = -3.862±0.038 mas/yr) of the known KIC members. Stars which were clearly non-members (outliers) based on their distances and proper motions were removed.

4. Results and Discussion

The list of the new KIC members with their classification type, distances, and proper motions are shown in Table 2. The mean cluster distance and proper motion were obtained by including

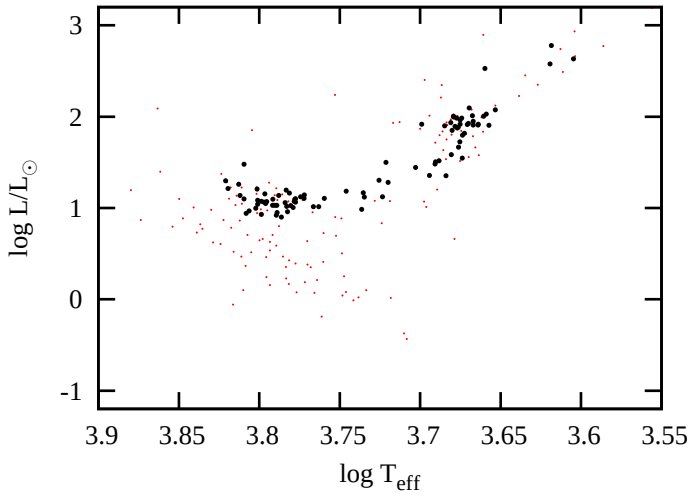


Figure 1. The large filled circles (black) denote stars considered as members while the small filled circles (red) denote all KIC stars within 10 arcmin of the cluster centre.

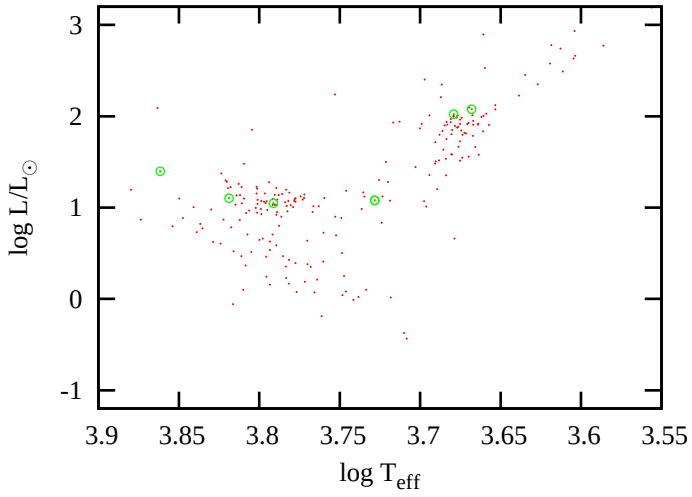


Figure 2. Small filled circles (red) denote all KIC stars within 10 arcmin of the cluster centre. New KIC members according to *Gaia* proper motion and distance modulus (parallax) are shown as (green) open circles.

KIC new members to the well known members. Our focus was to improve the mean distance of the cluster. We found mean distance DM = 12.30±0.12 mag. There is no significant difference in the mean distance modulus when probable new KIC members are included in deriving the average distance modulus. Fig. 2 is the HR diagram showing the location of the new KIC members according to *Gaia* proper motion and distance modulus. The new KIC members are shown as green open circles (Fig. 2).

In Fig. 3 we show periodograms of five stars of known classification as indicated in each panel. In the survey of the field of NGC 6819, [3] classified and measured rotation periods of 129 stars. The presence of harmonics in the frequency spectra or periodogram are good indicators of starspots. To classify a star as a rotational variable, one needs to examine the

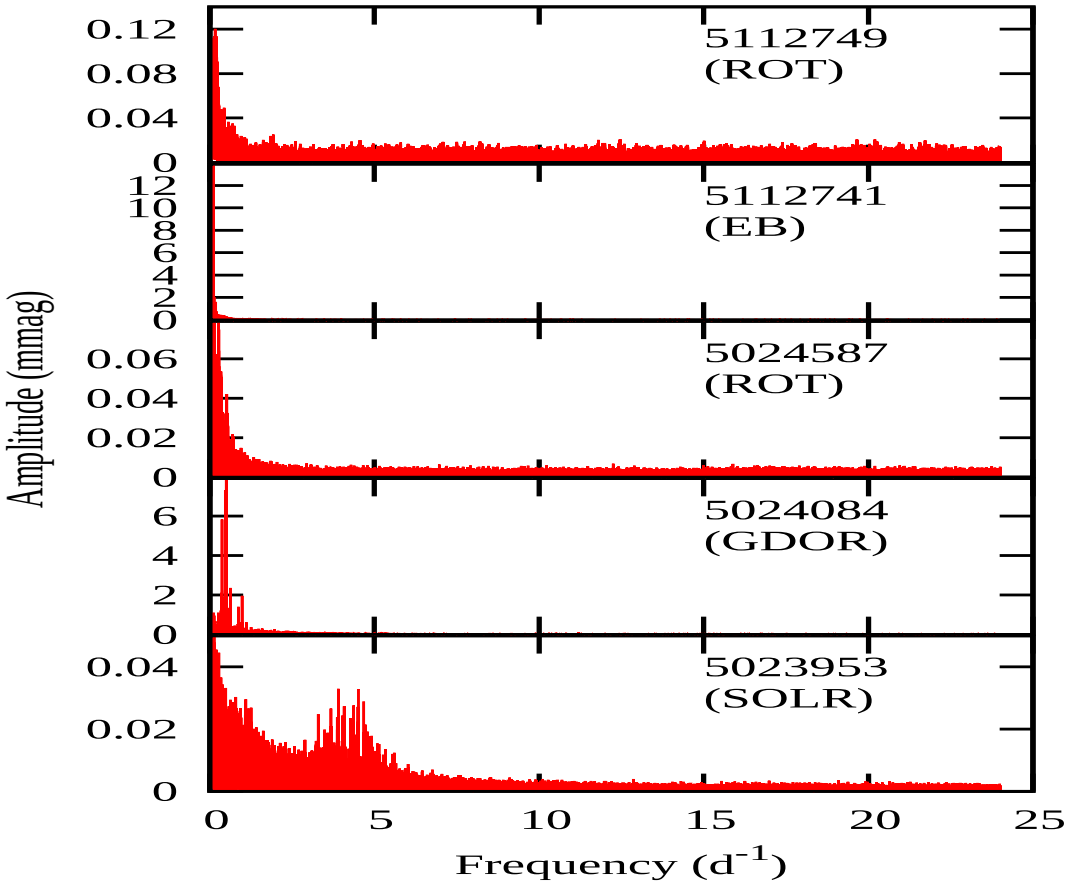


Figure 3. Periodograms of five stars of known classification.

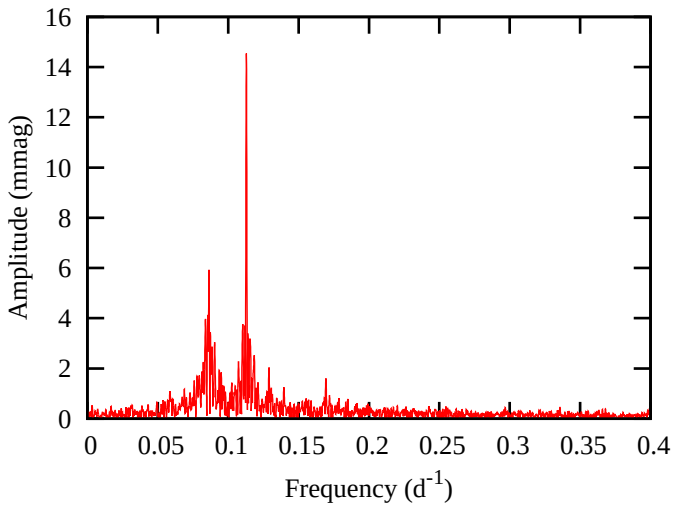


Figure 4. Periodogram of KIC 5112741 showing two dominant peaks.

Table 1. A list of 7 stars with proper motion and distance modulus consistent with that of NGC 6819 open cluster. The columns are KIC - *Kepler* Input Catalogue; pmRA - proper motion in Right Ascension; epmRA - error in proper motion in Right Ascension; pmDE - proper motion in Declination in milli-second of arc per year; epmDE - error in proper motion in Declination; Plx - parallax in milli-second of arc; ePlx - error in parallax; Kp - *Kepler* magnitude; the values of effective temperatures $\log(T_{\rm eff})$, luminosities, $\log L/L_{\odot}$; Type - variations seen in the periodograms; DM - distance modulus in mag; eDM - error in distance modulus in mag.

KIC	pmRA (mas/yr)	epmRA (mas/yr)	pmDE (mas/yr)	epmDE (mas/yr)	Plx (mas)	ePlx (mas)
5023953	-2.728	0.035	-3.770	0.037	0.3292	0.0203
5024084	-3.070	0.033	-3.822	0.034	0.3395	0.0181
5024429	-2.937	0.050	-3.745	0.046	0.3282	0.0238
5024587	-2.953	0.029	-3.987	0.033	0.3493	0.0164
5112585	-2.928	0.035	-3.899	0.037	0.3485	0.0195
5112741	-2.861	0.037	-3.789	0.040	0.3411	0.0207
5112749	-2.764	0.039	-3.762	0.040	0.3632	0.0223

Table 2. Table continued from column 8 Kp - *Kepler* magnitude; the values of effective temperatures $\log(T_{\rm eff})$, luminosities, $\log L/L_{\odot}$; Type - variations seen in the periodograms; DM - distance modulus in mag; eDM - error in distance modulus in mag.

Kp	$\log(T_{\rm eff})$	$\log L/L_{\odot}$	Type	DM (mag)	eDM (mag)
12.612	3.6792	2.0232	SOLR	12.4127	0.1338
14.718	3.8187	1.1013	GDOR	12.3458	0.1156
14.943	3.7913	1.0463	-	12.4193	0.1573
13.952	3.8616	1.3971	ROT	12.2840	0.1018
14.750	3.7283	1.0776	-	12.2890	0.1214
12.449	3.6680	2.0770	EB	12.3356	0.1316
14.953	3.8013	0.9446	ROT	12.1993	0.1332

periodogram and see if there are any peaks in the low frequency range (peak at frequency less than about $1\,{\rm d}^{-1}$) which can be attributed to rotation. Two stars were found to have peaks in the low frequency range in their periodograms: KIC 5024587 (peak at 0.097 mmag, frequency = 0.061 $\rm d^{-1}$) and KIC 5112749 (peak at 0.115 mmag, frequency = 0.154 $\rm d^{-1}$). We have identified and classified them as rotational variables. Harmonics is clearly present in the periodogram of KIC 5024587, at $f_3 = 2f_5$. Where $f_3 = 0.257\,{\rm d}^{-1}$ and $f_5 = 0.128\,{\rm d}^{-1}$. Thus, a starspot is present in KIC 5024587. Further more, there is no harmonics in the periodogram of KIC 5112749, hence there is no starspot. We have measured the rotation period ($\rm P_{rot}$) of the two stars showing rotational modulation as KIC 5024587 ($\rm P_{rot}$ = 3.8662 d) and KIC 5112749 ($\rm P_{rot}$ = 6.4975 d).

All cool stars have convective envelopes. In the periodogram, solar-like oscillations are easily identified because of the localized comb-like structure with amplitudes which decrease sharply from a central maximum. The periodogram of KIC 5023953 in Fig. 3 (bottom panel) shows comb-like structure typical of a solar-like oscillator. [15] assigned KIC 5023953 (WOCS 3011) a radial velocity binary likely member (BLM). [19] classified KIC 5023953 (WOCS 3011) as single-lined binary (SB1) member with 93% membership probability and radial velocity, $v_{\rm r}$ = +3.01±0.11 km/s. KIC 5023953 is to the left of the red giant branch where one expect to find clump stars, hence confirming the result reported in the literature.

Variability is detected in the light curve and periodogram of KIC 5112741 (Fig.4). The periodogram of KIC 5112741 shows two dominant periods of 8.854 and 11.576 d. The low-frequency peak might be because of binarity. We therefore, interprete the cause of variability as contact binary (eclipsing binary). It is interesting to note that KIC 5112741 which we have classified as an eclipsing binary occupies the position where one expect to find red giant showing solar-like oscillation.

Multi-periodic pulsating stars with frequencies in the range 0.5–5 $\rm d^{-1}$ are generally considered

to be γ Dor variables [3]. Based on this typical characteristic, only one candidate KIC 5024084 was found to exhibit such frequencies. The highest amplitude (9.903 mmag) is at $f = 0.490\,\mathrm{d}^{-1}$, but there are other peaks as well with no harmonics. The fact that harmonics are absent means there is no starspot. Assuming the frequency of the highest peak is the rotational frequency (F_{rot}), this implies $P_{\mathrm{rot}} = 2.041$ d. This is a γ Dor showing rotation modulation.

5. Conclusion

Stellar variability of the probable new KIC members were studied in detail using *Kepler* time series photometry and we have found a mean distance modulus DM = 12.30±0.12 mag. We also found a γ Dor variable with no harmonics present in the periodogram, thus no starspot. One star showing solar-like oscillation in its periodogram was found. We were able to interprete variability in one of the probable new KIC members as due to contact binary. Two stars could not be classified into any well known variable based on the visual inspection of their periodograms. Possible future work should look at estimating the basic parameters of the eclipsing variable (KIC 5112741).

5.1. Acknowledgments

This work has made use of data from Gaia and the Kepler mission.

References

[1] O. P. Abedigamba, L. A. Balona, and R. Medupe. Distance moduli of open cluster NGC 6819 from Red Giant Clump stars. , 46:90–93, July 2016.

[2] L. A. Balona. Gaia luminosities of pulsating A-F stars in the Kepler field. , June 2018.

[3] L. A. Balona, T. Medupe, O. P. Abedigamba, G. Ayane, L. Keeley, M. Matsididi, G. Mekonnen, M. D. Nhlapo, and N. Sithole. Kepler observations of the open cluster NGC 6819. , 430:3472–3482, April 2013.

[4] S. Basu, F. Grundahl, D. Stello, T. Kallinger, S. Hekker, B. Mosser, R. A. García, S. Mathur, K. Brogaard, H. Bruntt, W. J. Chaplin, N. Gai, Y. Elsworth, L. Esch, J. Ballot, T. R. Bedding, M. Gruberbauer, D. Huber, A. Miglio, M. Yildiz, H. Kjeldsen, J. Christensen-Dalsgaard, R. L. Gilliland, M. M. Fanelli, K. A. Ibrahim, and J. C. Smith. Sounding Open Clusters: Asteroseismic Constraints from Kepler on the Properties of NGC 6791 and NGC 6819. , 729:L10, March 2011.

[5] W. Borucki, D. Koch, N. Batalha, D. Caldwell, J. Christensen-Dalsgaard, W. D. Cochran, E. Dunham, T. N. Gautier, J. Geary, R. Gilliland, J. Jenkins, H. Kjeldsen, J. J. Lissauer, and J. Rowe. KEPLER: Search for Earth-Size Planets in the Habitable Zone. In F. Pont, D. Sasselov, and M. J. Holman, editors, *Transiting Planets*, volume 253 of *IAU Symposium*, pages 289–299, February 2009.

[6] D. Bossini, A. Vallenari, A. Bragaglia, T. Cantat-Gaudin, R. Sordo, L. Balaguer-Núñez, C. Jordi, A. Moitinho, C. Soubiran, L. Casamiquela, R. Carrera, and U. Heiter. Age determination for 269 Gaia DR2 open clusters. , 623:A108, March 2019.

[7] A. Bragaglia, E. Carretta, R. G. Gratton, M. Tosi, G. Bonanno, P. Bruno, A. Calì, R. Claudi, R. Cosentino, S. Desidera, G. Farisato, M. Rebeschini, and S. Scuderi. Metal Abundances of Red Clump Stars in Open Clusters. I. NGC 6819. , 121:327–336, January 2001.

[8] T. M. Brown, D. W. Latham, M. E. Everett, and G. A. Esquerdo. Kepler Input Catalog: Photometric Calibration and Stellar Classification. , 142:112, October 2011.

[9] T. Cantat-Gaudin, C. Jordi, A. Vallenari, A. Bragaglia, L. Balaguer-Núñez, C. Soubiran, D. Bossini, A. Moitinho, A. Castro-Ginard, A. Krone-Martins, L. Casamiquela, R. Sordo, and R. Carrera. A Gaia DR2 view of the open cluster population in the Milky Way. , 618:A93, October 2018.

[10] J. Choi, C. Conroy, Y.-S. Ting, P. A. Cargile, A. Dotter, and B. D. Johnson. Star Cluster Ages in the Gaia Era. , 863:65, August 2018.

[11] Isabel L. Colman, Timothy R. Bedding, Daniel Huber, and Hans Kjeldsen. The Kepler IRIS Catalog: Image Subtraction Light Curves for 9150 Stars in and around the Open Clusters NGC 6791 and NGC 6819. , 258(2):39, February 2022.

[12] E. Corsaro, D. Stello, D. Huber, T. R. Bedding, A. Bonanno, K. Brogaard, T. Kallinger, O. Benomar, T. R. White, B. Mosser, S. Basu, W. J. Chaplin, J. Christensen-Dalsgaard, Y. P. Elsworth, R. A. García, S. Hekker, H. Kjeldsen, S. Mathur, S. Meibom, J. R. Hall, K. A. Ibrahim, and T. C. Klaus. Asteroseismology of the open clusters NGC 6791, NGC 6811, and NGC 6819 from nineteen months of Kepler photometry. *ArXiv e-prints*, May 2012.

[13] Gaia Collaboration, A. G. A. Brown, A. Vallenari, T. Prusti, J. H. J. de Bruijne, C. Babusiaux, and C. A. L. Bailer-Jones. Gaia Data Release 2. Summary of the contents and survey properties. *ArXiv e-prints*, April 2018.

[14] X.-H. Gao, S.-K. Xu, and L. Chen. 3D cluster members and near-infrared distance of open cluster NGC 6819. *Research in Astronomy and Astrophysics*, 15:2193, December 2015.

[15] K. T. Hole, A. M. Geller, R. D. Mathieu, I. Platais, S. Meibom, and D. W. Latham. WIYN Open Cluster Study. XXIV. Stellar Radial-Velocity Measurements in NGC 6819. , 138:159–168, July 2009.

[16] D. Huber, V. Silva Aguirre, J. M. Matthews, M. H. Pinsonneault, E. Gaidos, R. A. García, S. Hekker, S. Mathur, B. Mosser, G. Torres, F. A. Bastien, S. Basu, T. R. Bedding, W. J. Chaplin, B.-O. Demory, S. W. Fleming, Z. Guo, A. W. Mann, J. F. Rowe, A. M. Serenelli, M. A. Smith, and D. Stello. Revised Stellar Properties of Kepler Targets for the Quarter 1-16 Transit Detection Run. , 211:2, March 2014.

[17] J. S. Kalirai, H. B. Richer, G. G. Fahlman, J.-C. Cuillandre, P. Ventura, F. D'Antona, E. Bertin, G. Marconi, and P. R. Durrell. The CFHT Open Star Cluster Survey. II. Deep CCD Photometry of the Old Open Star Cluster NGC 6819. , 122:266–282, July 2001.

[18] E. Komucyeya, O. P. Abedigamba, E. Jurua, and S. K. Anguma. New red giant star in the Kepler open cluster NGC 6819. , 61:65–68, May 2018.

[19] K. E. Milliman, R. D. Mathieu, A. M. Geller, N. M. Gosnell, S. Meibom, and I. Platais. WIYN Open Cluster Study. LX. Spectroscopic Binary Orbits in NGC 6819. , 148:38, August 2014.

[20] A. Pietrinferni, S. Cassisi, M. Salaris, and F. Castelli. A Large Stellar Evolution Database for Population Synthesis Studies. I. Scaled Solar Models and Isochrones. , 612:168–190, September 2004.

[21] M. H. Pinsonneault, D. An, J. Molenda-Zakowicz, W. J. Chaplin, T. S. Metcalfe, and H. Bruntt. Effective temperature scale for KIC stars (Pinsonneault+, 2012). *VizieR Online Data Catalog*, 219:90030, June 2012.

[22] M. H. Pinsonneault, Y. P. Elsworth, J. Tayar, A. Serenelli, D. Stello, J. Zinn, S. Mathur, R. A. García, J. A. Johnson, S. Hekker, D. Huber, T. Kallinger, S. Mészáros, B. Mosser, K. Stassun, L. Girardi, T. S. Rodrigues, V. Silva Aguirre, D. An, S. Basu, W. J. Chaplin, E. Corsaro, K. Cunha, D. A. García-Hernández, J. Holtzman, H. Jönsson, M. Shetrone, V. V. Smith, J. S. Sobeck, G. S. Stringfellow, O. Zamora, T. C. Beers, J. G. Fernández-Trincado, P. M. Frinchaboy, F. R. Hearty, and C. Nitschelm. The Second APOKASC Catalog: The Empirical Approach. , 239:32, December 2018.

[23] I. Platais, N. M. Gosnell, S. Meibom, V. Kozhurina-Platais, A. Bellini, C. Veillet, and M. S. Burkhead. WIYN Open Cluster Study. LV. Astrometry and Membership in NGC 6819. , 146:43, August 2013.

[24] T. Reinhold, A. Reiners, and G. Basri. Rotation and differential rotation of active Kepler stars. , 560:A4, December 2013.

[25] J. M. Rosvick and D. A. Vandenberg. BV Photometry for the ˜2.5 Gyr Open Cluster NGC 6819: More Evidence for Convective Core Overshooting on the Main Sequence. , 115:1516–1523, April 1998.

[26] D. Stello, D. Huber, T. R. Bedding, O. Benomar, L. Bildsten, Y. P. Elsworth, R. L. Gilliland, B. Mosser, B. Paxton, and T. R. White. Asteroseismic Classification of Stellar Populations among 13,000 Red Giants Observed by Kepler. , 765:L41, March 2013.

[27] D. Stello, S. Meibom, R. L. Gilliland, F. Grundahl, S. Hekker, B. Mosser, T. Kallinger, S. Mathur, R. A. García, D. Huber, S. Basu, T. R. Bedding, K. Brogaard, W. J. Chaplin, Y. P. Elsworth, J. Molenda-Żakowicz, R. Szabó, M. Still, J. M. Jenkins, J. Christensen-Dalsgaard, H. Kjeldsen, A. M. Serenelli, and B. Wohler. An Asteroseismic Membership Study of the Red Giants in Three Open Clusters Observed by Kepler: NGC 6791, NGC 6819, and NGC 6811. , 739:13, September 2011.

[28] M. Vrard, T. Kallinger, B. Mosser, C. Barban, F. Baudin, K. Belkacem, and M. S. Cunha. Amplitude and lifetime of radial modes in red giant star spectra observed by Kepler. , 616:A94, August 2018.

[29] C. L. Watson, A. A. Henden, and A. Price. The International Variable Star Index (VSX). *Society for Astronomical Sciences Annual Symposium*, 25:47, May 2006.

Solar energetic particle transport between Earth and Mars

A. M. Dumont[1], **R. D. Strauss**[1], **A. Posner**[2], **J. T. Stevens**[1]

[1]Centre for Space Research, North-West University, South Africa
[2]NASA/HQ, Washington, DC, USA

E-mail: dumont.aimee@gmail.com

Abstract. Observing and modelling solar energetic particles (SEPs) can aid in establishing an early warning system to prevent any hazardous impact on humans and technology in space, due to high levels of ionising radiation. During the Hohmann-type transfer from Earth to Mars of the MAVEN spacecraft, the Hohmann-Parker effect allows for the application of one-dimensional SEP transport. Soft X-ray data from the GOES satellite is used to approximate the acceleration and subsequent injection of the SEPs in the model due to solar flares. The transport of SEPs are then modelled along the Parker heliospheric magnetic field, and compared to electron intensity observations from the MAVEN and WIND spacecraft. Understanding SEP behaviour along Earth-Mars transfer trajectories will be necessary, in particular so that the Hohmann-Parker effect can be fully exploited for the safety of future crewed missions.

1. Introduction

Solar energetic particles (SEPs) are highly energetic particles accelerated near the Sun which consist of electrons with energies ranging from 1 keV - 10 MeV, and ions ranging between 50 keV - 10 GeV [1; 2]. Understanding how SEPs are produced and transported is fundamental in the protection of future crewed space missions, as these particles contribute to radiation exposure to astronauts and could cause acute radiation sickness, cancer, as well as other hazardous health conditions [3; 4; 5]. These particles are created by solar events, mostly solar flares and coronal mass ejections, where after the particles travel predominantly along the Parker [6] heliospheric magnetic field (HMF) [7]. The Parker HMF is assumed to be attached to and rotating with the photosphere of the Sun [2]. The magnetic field produced by the Sun is frozen into the solar wind (SW) plasma without any force exerted on it, which causes the field to curve into an Archimedean spiral in the solar equatorial plane, where it curves more at lower SW speeds and curves less at higher SW speeds [8]. The HMF can be understood by considering ideal SW conditions that have a steady radial outflow of constant speed, which is independent of radial and latitudinal position [8; 3].

As most planetary missions use the Hohmann [9] transfer manoeuvre, the transfer of a spacecraft into an elliptical trajectory orbit between two planets, studying SEPs in the HMF can help establish a system to warn organisations that have human activity in space [10; 11]. The Hohmann-Parker effect uses ideal SW conditions to show that a spaceship using the Hohmann transfer manoeuvre to transfer between two planets will experience a small angular magnetic connection to either one or the other planet [3]. During a transfer between Earth and Mars, the spacecraft will initially be connected to Earth's magnetic field line which is shown in Figure 1(a), and later connects to the Mars' field line.

The Neupert [12] effect, which refers to a relationship between the soft X-ray (SXR) photon flux and the flux of SEP electrons accelerated during a solar flare, can be used to develop a model that approximates the injection of SEPs into the interplanetary space along the HMF [3; 5; 13]. This paper

uses data from the geostationary operational environmental satellite (GOES), global geospace science wind satellite (WIND), and the Mars atmosphere and volatile evolution (MAVEN) spacecraft to observe how SEP electrons are transported from solar events towards Mars to determine if it could be used as a sufficient early warning system. Data from GOES[1] is used to approximate SEP injections along the HMF and the resulting modelled SEP intensity is compared to WIND and MAVEN electron observations. For this work an one-dimensional SEP transport model is used, as motivated by the good magnetic connectivity during the Hohmann-type transfer trajectory, which means that perpendicular transport due to random walking magnetic field lines [14] are neglected.

2. The modulation model

The distribution function,

$$f(\mathbf{x}, \mathbf{v}, t), \tag{1}$$

gives the particle number density in the general six-dimensional $(\mathbf{x}, \mathbf{v})$ space at time t where the time evolution of f is described by the Boltzmann equation [15]

$$\underbrace{\frac{\partial f}{\partial t}}_{\text{A}} + \underbrace{\mathbf{v} \cdot \boldsymbol{\nabla} f}_{\text{B}} + \underbrace{\frac{\mathbf{F}}{m} \cdot \frac{\partial f}{\partial \mathbf{v}}}_{\text{C}} = \underbrace{\left(\frac{\partial f}{\partial t}\right)_c}_{\text{D}}. \tag{2}$$

Here, term A is the time-dependent change and term B describes the change in the guiding centre position around the magnetic field which drifts parallel to the HMF [16]. For the spatially one-dimensional case, this reduces to $\mathbf{v} \cdot \boldsymbol{\nabla} f = v_{\parallel} \frac{\partial f}{\partial s} = \mu v \frac{\partial f}{\partial s}$. Term C is the pitch-cosine angle change due to particles moving to different regions in the HMF, which for the one-dimensional case reduces to $\frac{\mathbf{F}}{m} \cdot \frac{\partial f}{\partial \mathbf{v}} = \frac{v}{2L}\left(1-\mu^2\right)\frac{\partial f}{\partial \mu}$, which uses the mirroring condition with the focusing length, L [17]. Term D is the time rate change of the pitch-angle dependence due to collisions, $\left(\frac{\partial f}{\partial t}\right)_c = \frac{\partial}{\partial \mu}\left(D_{\mu\mu}\frac{\partial f}{\partial \mu}\right)$ where $D_{\mu\mu}$ is the pitch-angle diffusion coefficient [18].

The one-dimensional focused transport equation derived here is also referred to as the Roelof [19] equation which demonstrates how SEPs transport in the HMF and can be written as

$$\frac{\partial f}{\partial t} + \mu v \frac{\partial f}{\partial z} + \frac{v}{2L}\left(1-\mu^2\right)\frac{\partial f}{\partial \mu} = \frac{\partial}{\partial \mu}\left(D_{\mu\mu}\frac{\partial f}{\partial \mu}\right), \tag{3}$$

which give the evolution of the distribution function for the one-dimensional scenario, $f(s, \mu, t)$, where s is the line element in the $\hat{z}$ direction and μ is the pitch-angle of the SEPs along the HMF [20]. From equation 3 an omnidirectional distribution function can be calculated once the distribution function (equation 1) is obtained. The omnidirectional intensity is $j = p^2 f$, where p^2 is the particle momentum which is assumed to be constant for these mono-energetic simulations so that

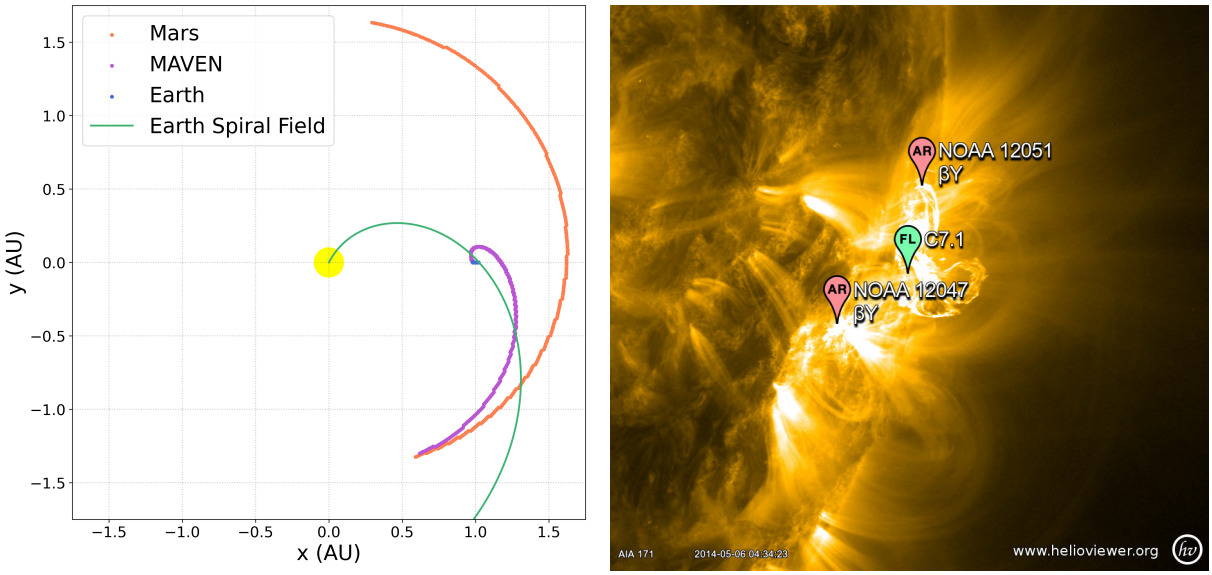
$$j \propto \frac{\int_{-1}^{1} f(s, \mu, t)\, d\mu}{\int_{-1}^{1} d\mu} = \frac{1}{2}\int_{-1}^{1} f(s, \mu, t)\, d\mu, \tag{4}$$

where j is measured in particles/cm^2/s/sr/MeV [20]. The mean free path is calculated from the pitch-angle diffusion coefficient $D_{\mu\mu}$, as,

$$\lambda_{\parallel} = \frac{3v}{8}\int_{-1}^{1} d\mu \frac{\left(1-\mu^2\right)^2}{D_{\mu\mu}(\mu)}, \tag{5}$$

[1] https://www.swpc.noaa.gov/products/goes-x-ray-flux

[2] https://student.helioviewer.org



(a) (b)

Figure 1: Panel (a) shows the trajectory of the MAVEN spacecraft's voyage from Earth to Mars, in purple, in a system co-rotating with the Earth and the Sun at the origin to demonstrate the Hohmann-Parker effect. The Parker HMF spirals represent the magnetic foot point of the Earth at Earth's initial and final coordinates during the MAVEN voyage. Panel (b) is an image taken from Helioviewer[2]which shows the active region NOAA 12051 that was active on the surface of the Sun between 121 to 128 DOY before it disappeared behind the Sun's western limb.

which is the average distance particles would move between collisions that are parallel to the guiding centre position around the HMF [21].

The Parker [6] HMF on which the SEPs are assumed to travel, is given as

$$\vec{B}_{HMF} = AB_0 \left(\frac{r_0}{r}\right)^2 \left[\hat{r} - \tan\psi\hat{\phi}\right], \tag{6}$$

where $\hat{\phi}$ is the unit vector in the azimuthal direction, A is the magnetic field polarity at Earth (which is at 1 AU), at $r = r_0$, r is the radial distance, ψ is the winding angle of the Parker HMF, and B_0 is the normalization value [17]. The line integral of the magnetic field, $d\vec{l}$, in spherical coordinates can be written as

$$d\vec{l} = dr\,\hat{r} + rd\theta\,\hat{\theta} + r\sin\theta d\phi\,\hat{\phi}, \tag{7}$$

where, for the Parker HMF, $d\phi = \frac{-\Omega}{V_{sw}} dr$ and $d\theta = 0$ [2]. This line element integral is used to trace Parker HMF lines in Figure 1(a).

3. MAVEN spacecraft

MAVEN was launched into a Hohmann-type orbit in November 2013 and arrived at Mars in September 2014 [3; 22]. The SEP instrument on board MAVEN measured the energy and angular distributions of electrons and ions [22]. In Figure 1(a) the coordinates for Earth, Mars, and MAVEN were obtained from NASA's heliocentric trajectories website [3] between the 323 day of year (DOY) in 2013 and the 246 DOY in 2014. These coordinates were plotted using a Python3 code in a coordinate system co-rotating with Earth with equation 6 overlayed to illustrate the Parker spirals at ideal conditions at Earth and Mars on

[3] https://omniweb.gsfc.nasa.gov/coho/helios/heli.html

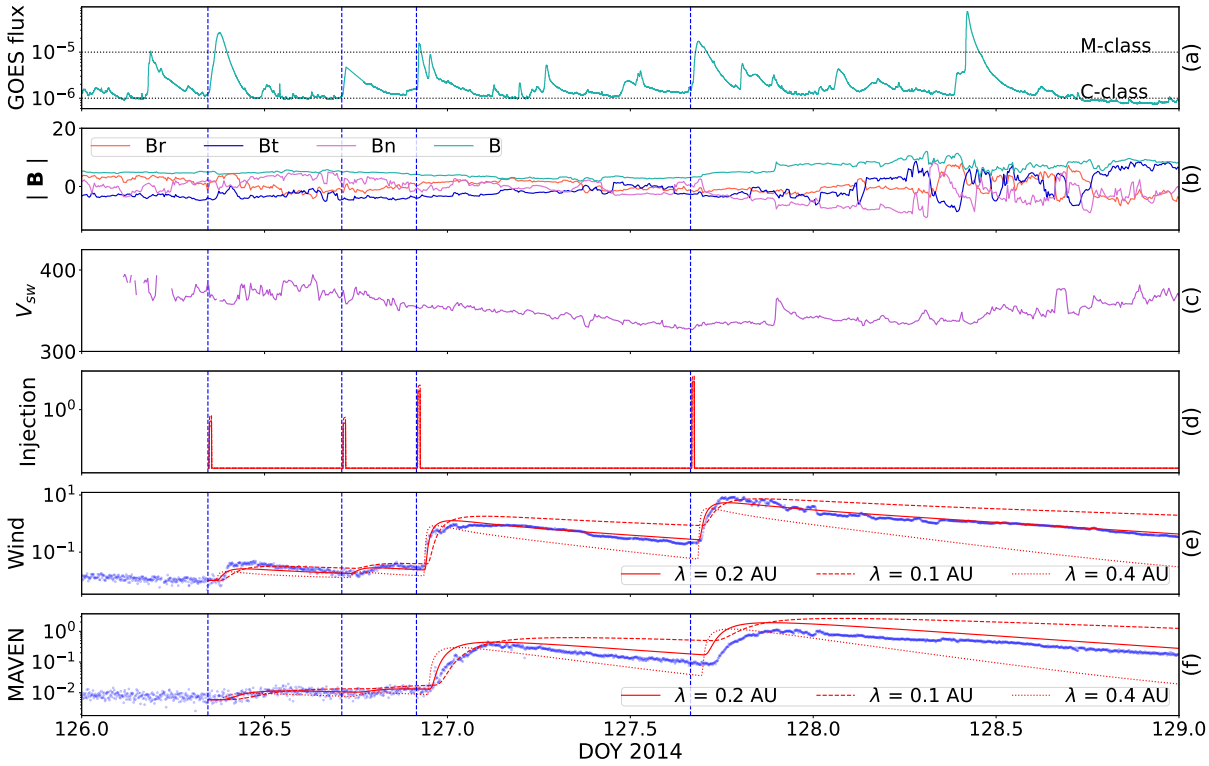


Figure 2: From top to bottom, panels (a)-(f), are the measurements of the SEP event between 126-129 DOY. Panel (a) is the SXR flux time profile measured from GOES in W/m^2. Panel (b) is the total magnetic flux as well as the magnetic components of the HMF which is measured in nT. Panel (c) is the SW speed observed from the ACE spacecraft, which is measured in km/s. Panel (d) is the simulated SEP injections in the HMF near the Sun. Panel (e) is the SEP electron flux observed by the WIND spacecraft, which measures the electrons in particles/cm^2/s/sr/MeV, with an overlay of modelled solutions for different values at $\lambda_{\parallel}$. Panel (f) is the SEP electron flux observed from the MAVEN spacecraft, which measures the electrons in particles/cm^2/s/sr/MeV, with an overlay of modelled solutions for different values at $\lambda_{\parallel}$. During this time interval MAVEN was located at radial distances between 1.31 AU on DOY 126 and 1.32 AU on DOY 129. The time interval is from 6 May (DOY 126) to 9 May 2014 (DOY 129).

323 DOY. Figure 1(a) demonstrates that MAVEN is initially well connected to the magnetic field line of Earth and gradually connected to the magnetic field line of Mars as the spacecraft neared the end of its voyage. It is apparent that the Hohmann elliptical trajectory is curved similarly to the shape of the Parker HMF spiral. Therefore, the SEPs follow approximately the same magnetic connection between Earth and Mars for this time interval and an one-dimensional SEP transport model can be used.

4. Results

The model used was a modified version of the one-dimensional SEP transport FORTRAN90 code to simulate the propagation of SEPs along a single magnetic field line, developed by Strauss & Fichtner [23] [4]. The model solves the focused transport equation (equation 3) to obtain the generalised one-dimensional distribution function (equation 1), where $f(\mathbf{x}, \mathbf{u}, t) = f(s, \mu, t)$, that can then be compared to SEP observations. The model used an ideal condition where the speed of the SW remains constant at $V_{sw} = 350$ km/s. In the model, $\lambda_{\parallel}$ was adjusted to correspond to the data and thus treated as a free

[4] `https://github.com/RDStrauss/SEP_propagator`

parameter.

The time interval in Figure 2 ranged between 126 to 129 DOY in 2014. GOES observed six major flares, of which the SEPs of only four flares were observed at the MAVEN and WIND spacecraft, the other two flares were not located near the magnetic foot point of the Earth, therefore SEPs from these flares were observed by neither WIND or MAVEN. Figure 1(b) shows the active region NOAA 12051 that produced the solar flares that occurred between 126 to 129 DOY in 2014. The most prominent flares observed on the electron SEP counters in GOES were C-class, M-class and X-class flares. An M-class flare is 10 times stronger than a C-class flare and an X-class is 10 times stronger than an M-class [24]. The magnetic foot point of the Earth is the region of the Sun that is magnetically connected to Earth, which is located approximately at W60 [25]. The NOAA 12051 flare occurred at approximately S13W86 on the Sun's surface, which is close to the magnetic foot point of the Earth.

From Figure 2 there was a quiet magnetic field observed up until 128 DOY, after which the magnetic field appears to be more perturbed. The SW speed ranged between 325-400 km/s, which is within ideal wind speeds. The GOES data was used to determine when injections take place while the magnitude of the SEP injection for each flare was adjusted manually to fit the data. There were four injections placed in the model, which represents the flares from the GOES data. $\lambda_{\|}$ was adjusted in the one-dimensional model until a reasonable fit was obtained to fit the data. The best fit of the mean free path was when $\lambda_{\|} = 0.2$ AU and the background was set to 0.01 particles/cm^2/s/sr/MeV, which shifted the model's default line to the data.

5. Conclusion

The GOES data demonstrated that there was a connection between the SEPs released in solar flares and the magnetic foot point of the Earth; the majority of solar flares that occurred near the magnetic foot point of the Earth were observed along the Parker spiral towards Earth and Mars. The different flare classes affected the number of particles observed in the data. The mean free path of the particles differed during each of the solar flares in Figure 2, indicating that the SEPs interacted differently with the turbulence along the HMF. The MAVEN data is given in Figure 1(a) demonstrated that MAVEN followed the Hohmann bi-tangent orbit which mostly correlated with the Hohmann-Parker effect, therefore the SEPs ejected from solar flares follow approximately the same trajectory between Earth and Mars. A later section of the MAVEN mission does not follow the Parker spiral that connects Earth with the Sun, which could lead to an ineffective Earth-based (Earth-Sun L1) warning system if transporting astronauts between Earth and Mars. Here, a Mars-Sun L1 observing system would be able to provide accurate warnings. SEP warnings from Mars L1 would also apply for much of the return journey [5]. Furthermore, an Earth-Sun L4 mission would support most SEP warning schemes as described in [26] with X-ray and extreme ultraviolet observations (and others) of SEP events originating just behind the western limb of the Sun as viewed from Earth. These SEP events can be hazardous to astronauts as shown by [27].

Acknowledgments

[1] The Helioviewer was created by the following scientific teams: SDO (NASA), SOHO (ESA/NASA), STEREO (NASA), PROBA2 (ESA), Yohkoh (JAXA/NASA/PPARC), Hinode (JAXA/NASA/PPARC), GSFC, Royal Observatory of Belgium, LMSAL, SDAC, Stanford University, Harvard-Smithsonian Astrophysical Observatory, MSU/SDO-FFT.©2018 WGBH Educational Foundation. All Rights Reserved.

[2] The trajectory data was produced by NASA Heliocentric Trajectories for Selected Spacecraft, Planets, and Comets.

[3] The GOES 8-15 X-ray/magnetic field/particle data are produced in real-time by the NOAA Space Weather Prediction Center (SWPC) and are distributed by the NOAA National Geophysical Data Center (NGDC).

References

[1] Reames D V 2013 *Space Science Reviews* **175** 53–92

[2] Steyn P J 2022 *A multi-wavelength approach to solar energetic particle transport using remote sensing and in-situ observations.* Ph.D. thesis North-West University (South Africa)

[3] Posner A, Odstrĉil D, MacNeice P, Rastaetter L, Zeitlin C, Heber B, Elliott H, Frahm R, Hayes J, von Rosenvinge T *et al.* 2013 *Planetary and Space Science* **89** 127 – 139

[4] Moloto K D, Engelbrecht N E and Burger R A 2018 *The Astrophysical Journal* **859** 107

[5] Posner A and Strauss R D 2020 *Space Weather* **18**

[6] Parker E N 1958 *The Astrophysical Journal* **128** 664

[7] Corti C, Whitman K, Desai R, Rankin J, Strauss R D, Nitta N, Turner D and Y Chen T 2022 *arXiv e-prints*

[8] Owens M J and Forsyth R J 2013 *Living Reviews in Solar Physics* **10** 5

[9] Hohmann W and United States National Aeronautics and Space Administration 1960 *The Attainability of Heavenly Bodies* NASA technical translation [National Aeronautics and Space Administration]

[10] Loto'aniu P, Rodriguez J, Redmon R, Machol J, Kress B, Seaton D, Darnel J, Rowland W, Tilton M, Denig W *et al.* 2017 *EGU General Assembly Conference Abstracts* p 9663

[11] Miele A, Ciarcià M and Mathwig J 2004 *Journal of Optimization Theory and Applications* **123** 233 – 253

[12] Neupert W M 1968 *The Astrophysical Journal Letters* **153** L59

[13] Steyn R, Strauss R D, Effenberger F and Pacheco D 2020 *Journal of Space Weather and Space Climate* **10** 64

[14] Giacalone J and Jokipii J R 1999 **520** 204

[15] Choudhuri A R 2004 *The Physics of Fluids and Plasmas: An introduction for astrophysicists* [Cambridge University Press]

[16] Bian N H and Li G 2022 *The Astrophysical Journal* **924** 120

[17] Van den Berg J P 2019 *Modelling of solar energetic particles by stochastic processes.* Master's thesis North-West University (South Africa)

[18] Chen F F 1984 *Introduction To Plasma Physics And Controlled Fusion* 2nd ed [Plenum Press]

[19] Roelof E C 1969 *Lectures in High-Energy Astrophysics* ed Ögelman H and Wayland J R pp 111–135

[20] Strauss R D, van den Berg J P and Rankin J S 2022 *The Astrophysical Journal* **928** 22

[21] van den Berg J, Strauss R D and Effenberger F 2020 *Space Science Reviews* **216** 146

[22] Jakosky B M, Lin R P, Grebowsky J M, Luhmann J G, Mitchell D F, Beutelschies G, Priser T, Acuna M, Andersson L, Baird D *et al.* 2015 *Space Science Reviews* **195** 3–48

[23] Strauss R D and Fichtner H 2015 *The Astrophysical Journal* **801** 29

[24] Muranushi T, Shibayama T, Muranushi Y H, Isobe H, Nemoto S, Komazaki K and Shibata K 2015 *Space Weather* **13** 778–796

[25] Palmerio E, Kilpua E K J, Witasse O, Barnes D, Sánchez-Cano B, Weiss A J, Nieves-Chinchilla T, Möstl C, Jian L K, Mierla M *et al.* 2021 *Space Weather* **19**

[26] Whitman K, Egeland R, Richardson I G, Allison C, Quinn P, Barzilla J, Kitiashvili I, Sadykov V, Bain H M, Dierckxsens M *et al.* 2022 *Advances in Space Research* ISSN 0273-1177

[27] Posner A, Arge C N, Staub J, StCyr O C, Folta D, Solanki S K, Strauss R D T, Effenberger F, Gandorfer A, Heber B *et al.* 2021 *Space Weather* **19** e02777

Modelling galactic cosmic ray drifts in the presence of a Fisk-type heliospheric magnetic field

J.S. Troskie, N.E. Engelbrecht and P.J. Steyn

Centre for Space Research, North-West University, 11 Hoffman Street, Potchefstroom 2531

E-mail: jst99960@gmail.com

Abstract. Drift effects due to gradients and curvature in the heliospheric magnetic field (HMF) have long been known to affect cosmic ray (CR) modulation. Thus far, only drift effects due to the geometry of the HMF proposed by Parker [1] have been studied. However, several other HMF models exist, including the model proposed by Fisk [2], which, notably, has a latitudinal component. The Fisk-Parker hybrid HMF proposed by Burger & Hitge [3] assumes that the HMF is Parkerian at the poles and in the ecliptic plane, and Fisk-like at mid-latitudes. In this proceeding, the effects of the Fisk-Parker hybrid HMF on CR drifts are discussed; these results show, for the first time, that a Fisk-type HMF results in a CR drift velocity profile which differs significantly from what has first been described for the Parker HMF by Jokipii & Thomas [4]. Furthermore, different approaches to modelling drift due to the heliospheric current sheet are investigated; these include the models proposed by Burger [5] and Engelbrecht et al. [6].

1. Introduction

Previous studies on galactic cosmic ray (GCR) modulation (e.g., Burger et al. [7] and Sternal et al. [8]) have shown that Fisk-type heliospheric magnetic fields (HMFs) potentially play a role in GCR transport. However, no previous study has self-consistently modelled the influence of Fisk-type fields on GCR drift velocities. The present study aims to investigate this.

Given a spatially inhomogeneous magnetic field $\vec{B}$, a charged particle (such as a cosmic ray) will drift (an effect detailed in introductory plasma physics texts, e.g., Chen [9]). Drift effects have been taken into consideration in cosmic ray modulation studies for several decades, proving instrumental in explaining, for example, the 22-year GCR intensity cycle observed by neutron monitors (see, e.g., Engelbrecht et al. [10] and references therein). Isenberg and Jokipii [11] show that, for a nearly isotropic distribution of cosmic rays, the average drift velocity due to spatial variation in $\vec{B}$ is given by $\vec{v}_d = \nabla \times \kappa_A \hat{e}_B$, where $\hat{e}_B = \vec{B}/B$ is the unit vector in the direction of the magnetic field. The quantity κ_A is the drift coefficient (Forman et al. [12]), given by $\kappa_A = pv/(3qB) = vR_L/3$, where q is the particle magnitude, p is the particle momentum, v is the particle speed perpendicular to the magnetic field and R_L is the Larmor radius. Note that this drift coefficient only holds when the influence of turbulence on particle drifts is negligible (e.g., Engelbrecht et al. [13]).

2. The Parker heliospheric magnetic field

The heliospheric magnetic field, or HMF, is most commonly assumed to be described by the model proposed by Parker [1], given by the equation

$$\vec{B} = A\left(\frac{r_0}{r}\right)^2 \hat{r} - \frac{Ar_0^2}{rV_{SW}}\Omega \sin\theta\hat{\phi}. \tag{1}$$

The quantity A represents the magnetic field strength at a heliocentric distance r_0 from the Sun, r is the heliocentric distance, V_{SW} is the solar wind speed, Ω is the solar equatorial rotation rate and θ is the colatitude, measured from the solar North pole. Furthermore, A has a sign. If it is positive, $\vec{B}$ is directed away from the Sun in the Northern hemisphere and directed towards the Sun in the Southern hemisphere. If it is negative, the direction of $\vec{B}$ is reversed.

Fig. 1 shows field lines for the Parker HMF. In the ecliptic plane, a given field line will form an Archimedean spiral about the Sun. At other colatitudes, field lines will lie on cones centered around the rotation axis.

3. The heliospheric current sheet

The heliospheric current sheet, or HCS, is the structure separating regions of opposing polarity in the dipolar HMF, where B drops to zero. The HCS is known to play a role in cosmic ray drift; for $A > 0$, protons drift from over the polar regions of the Parker HMF towards the HCS, along which they drift to the edge of the heliosphere (see Jokipii & Thomas [4]), with drift directions being reversed for $A < 0$.

A number of HCS drift models (e.g., the models of Burger [5] and Engelbrecht et al. [6]) have been developed, as the expression for the gradient and curvature drift velocity (given in the introduction) has to be modified in order to model drift along the HCS. Burger [5] employs a hyperbolic tangent function to do so, such that

$$\vec{v}_d = \nabla \times \kappa_A \tanh\left[k\left(\theta_{ns} - \theta\right)\cos\nu\right]\hat{e}_B, \tag{2}$$

where the HCS is given by Kóta & Jokipii [14] as $\theta_{ns} = \pi/2 - \arctan\left(\tan\alpha \sin\phi^*\right)$, with $\phi^* = \phi + \phi_0 + r\left(\Omega/V_{SW}\right)$ and

$$\cos\nu = \left[1 + \left(r\frac{\partial\theta_{ns}}{\partial r}\right)^2 + \left(\frac{1}{\sin\theta}\frac{\partial\theta_{ns}}{\partial\phi}\right)^2\right]^{-1/2}. \tag{3}$$

The factor k determines the width of the region over which current sheet drift occurs, and α is the heliospheric tilt angle.

Engelbrecht et al. [6] transform the drift coefficient in a similar manner:

$$\vec{v}_d = \nabla \times \kappa_A f\left(\delta\right)\hat{e}_B, \tag{4}$$

where $f\left(\delta\right) = 2S\left(\delta\right) - 1$, with S being a higher-order smoothstep function and $\delta = (1/2) + r\left(\theta_{ns} - \theta\right)/\left(4R_L\right)$. When $\delta > 1$, $S\left(\delta\right) = 1$, while $S\left(\delta\right) = 0$ when $\delta < 0$. Furthermore, the first and second derivatives of $S\left(\delta\right)$ are zero when $\delta \geq 1$ or $\delta \leq 0$. This ensures that HCS drift effects only impact particles within 2 Larmor radii of the current sheet.

4. Fisk-type HMF models

In the HMF model proposed by Fisk [2], a polar coronal hole (PCH) is assumed to be centered around the Sun's magnetic axis, which rotates at the equatorial rate. At the PCH, V_{SW} is high compared to V_{SW} elsewhere. Furthermore, the differential rotation rate of the Sun's photosphere (e.g., Snodgrass [15]), into which HMF lines are rooted, is taken into account. Field lines expand nonradially with the fast solar wind from the PCH. At the heliographic North pole, where no differential rotation occurs, a single field line connects the photosphere to the source surface, beyond which the solar wind expands radially. The angle between this line and the rotation axis is the Fisk angle, denoted by β; see Zurbuchen et al. [16].

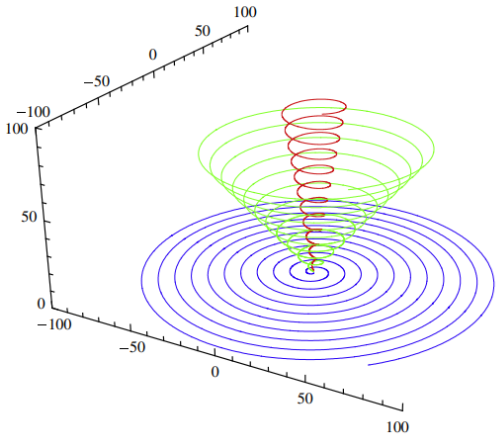


Figure 1. Field lines calculated for a Parker field at three different latitudes. Figure taken from Engelbrecht & Burger [17].

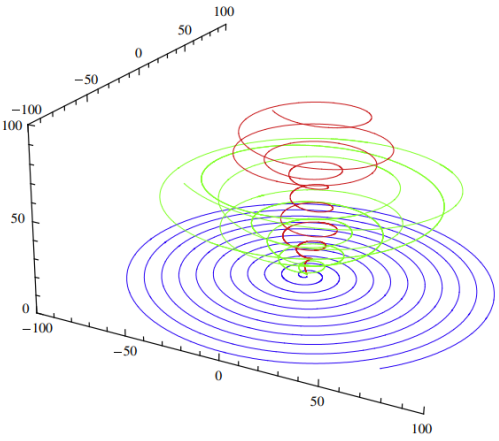


Figure 2. Field lines calculated for a Fisk-Parker hybrid field at three different latitudes. Figure taken from Engelbrecht & Burger [17].

The Fisk-Parker hybrid field proposed by Burger & Hitge [3] is Parkerian at the poles and the equator, and Fisk-like at mid-latitudes. The components of the refined hybrid model are given by Burger et al. [7]:

$$B_r = A\left(\frac{r_0}{r}\right)^2 \tag{5}$$

$$B_\theta = \frac{Ar_0^2}{rV_{SW}}\omega^* \sin\beta^* \sin\phi^* \tag{6}$$

$$B_\phi = \frac{Ar_0^2}{rV_{SW}}\left[\omega^* \cos\beta^* \sin\theta + \frac{d}{d\theta}\left(\omega^* \sin\beta^* \sin\theta\right)\cos\phi^* - \Omega \sin\theta\right], \tag{7}$$

where ω is the differential rotation rate, and $\omega^* = \omega F_S$ and $\beta^* = \beta F_S$ vary with the latitude-dependent transition function F_S. When $F_S = 1$, the components reduce to those of a pure Fisk field, and when $F_S = 0$, the components reduce to those of the Parker HMF.

In the ecliptic plane, a Fisk-Parker hybrid field line (Fig. 2) will form the Archimedean spiral seen for a Parker field line (Fig. 1). However, in contrast to the Parker model, field lines originating at intermediate colatitudes display considerable latitudinal excursions and a significant azimuthal dependence.

5. The Parker drift velocity field

In figures to follow, drift effects for protons are shown, and a relatively high value of $k = 60$ is assumed when the Burger [5] HCS drift model is used, as higher values of k result in a transition function more closely resembling the Heaviside step function (e.g., Mohlolo et al. [18]). Fig. 3 shows drift velocity profiles for the Parker HMF, calculated from the Burger [5] HCS drift model using Eq. 2 (brighter regions are associated with higher drift speeds). A solar wind speed of 400 km/s, a proton energy of 2 GeV and the HCS expression used by Jokipii & Thomas [4] have been assumed for comparative purposes. In the left panel, for which $A > 0$, protons are seen to drift towards the Sun from over the polar regions, towards the HCS away from the

poles and outwards along the HCS. In the right panel, for which $A < 0$, drift directions are reversed with respect to the $A > 0$ field. For negatively charged particles, such as electrons, drift velocities are reversed with respect to the profiles shown in Fig. 3. Away from the HCS, the Parker drift velocity field does not show azimuthal variation, as the Parker HMF (Eq. 1) has no ϕ-dependence.

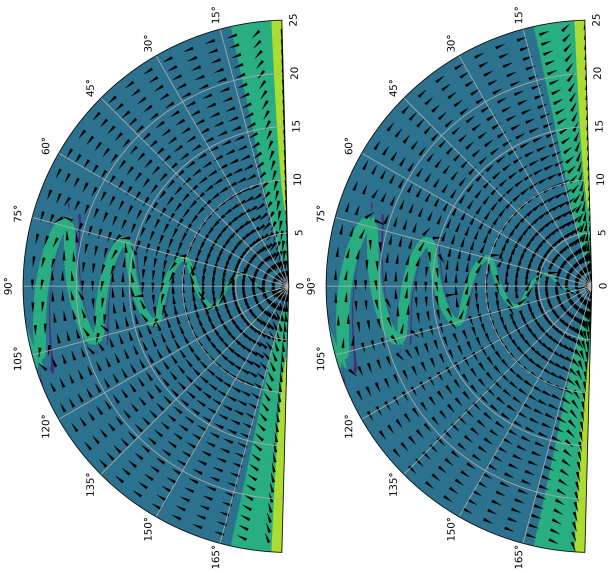


Figure 3. $A > 0$ (left) and $A < 0$ (right) drift velocity profiles for the Parker HMF. A proton energy of 2 GeV has been assumed.

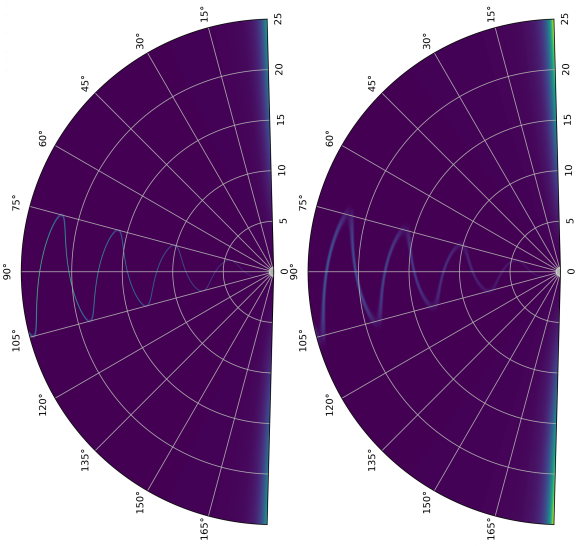


Figure 4. Drift speed profiles for 1 GeV (left panel) and 10 GeV protons (right panel) calculated from the Engelbrecht et al. [6] drift model.

In Fig. 4, the total (3D) drift speed calculated from the Engelbrecht et al. [6] model is shown for 1 GeV (left panel) and 10 GeV protons (right panel). Because of the R_L-dependence of the transition function employed by these authors, the HCS drift region increases in width as the proton energy increases. The log-scaled colouring implemented in Fig. 3 has been removed in order to show greater detail.

6. The Fisk-Parker drift velocity field

In figures to follow, the Engelbrecht et al. [6] drift model will be implemented, while the HCS expression of Kóta & Jokipii [14] will be used. In deriving the Fisk-Parker hybrid HMF, a constant solar wind speed is assumed (see Burger & Hitge [3]); an observationally-motivated average value of 600 km/s is chosen (McComas et al. [19]).

Fisk-type fields show drift velocity profiles which differ significantly from those calculated for the Parker HMF. In Fig. 5, drift velocity profiles calculated from Fisk-type fields are compared; $\vec{v}_d$ for the Fisk-Parker hybrid HMF is shown on the left, while the drift velocity calculated from a pure Fisk HMF, as derived by Zurbuchen et al. [16], is shown on the right. For comparative purposes, the HCS has been removed for both fields to emphasise gradient and curvature drift effects due to the different geometries.

Fig. 5 shows that Fisk-type drift velocity fields (both hybrid and pure) feature high-$|\vec{v}_d|$ structures extending from over the Sun's poles. The structures are associated with directional changes in the drift velocity. Other "true" Fisk-type drift features (i.e., features not resulting from the transition function used in the hybrid field) include mid-latitude structures associated with reduced drift speeds and directional changes (e.g., the structures seen between 20 and 25

AU and around $135° \leq \theta \leq 150°$ in both panels of Fig. 5). A feature which is unique to the hybrid field is a series of whorl-like mid-latitude structures associated with reduced drift speeds and directional changes - see $60° \leq \theta \leq 75°$ and $105° \leq \theta \leq 120°$ in the left panel (hybrid field) of Fig. 5.

In Fig. 6, calculated using the Engelbrecht et al. [6] HCS drift model for 1 GeV protons, it can be seen that a change in the sign of $\vec{B}$ (i.e., a switch from $A > 0$, shown in the left panel, to $A < 0$, shown in the right panel) again results in a reversal of the direction of $\vec{v}_d$.

Azimuthal variations in the Fisk-Parker hybrid drift velocity profile are shown in Fig. 7, which assumes the Engelbrecht et al. [6] drift model and a proton energy of 1 GeV. The high-speed protrusions seen over the poles are seen to form large-scale spiralling structures in 3D, and the mid-latitude reduced-drift structures seen at a given azimuth are projections of reduced-$|\vec{v}_d|$ tubes associated with directional changes throughout the heliosphere.

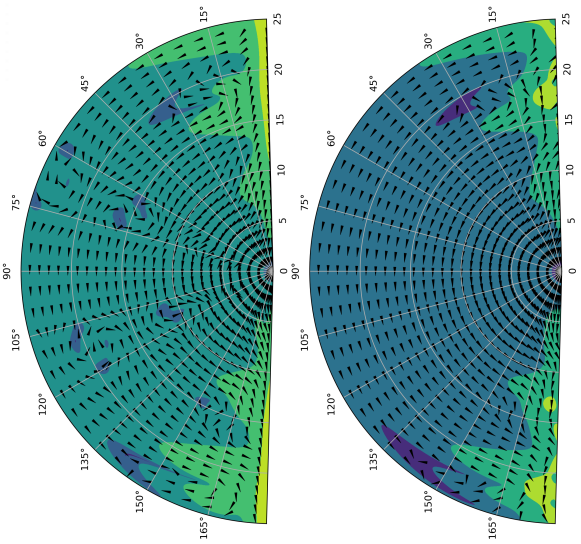


Figure 5. $A > 0$ drift velocity profiles for the Fisk-Parker hybrid (left panel) and pure Fisk (right panel) HMF models.

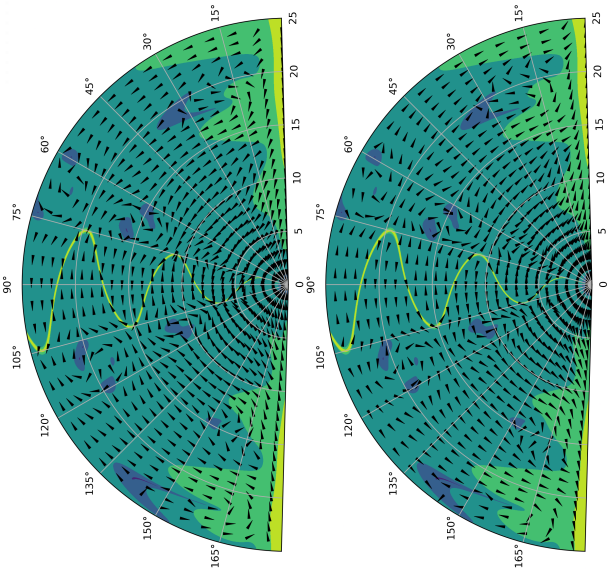


Figure 6. As for the Parker model, drift velocity directions are reversed when $A < 0$ (right panel).

7. Conclusions

Drift velocity profiles calculated from Fisk-type HMF models differ significantly from those calculated using the Parker model. Differences arise from the fact that Fisk-type fields, in contrast to the Parker HMF, have both a strong latitudinal component and a significant azimuthal dependence. Because drift velocities associated with Fisk-type fields have not yet been self-consistently implemented in modulation studies investigating the effects of the HMF models on GCR intensities (e.g. Sternal et al. [8]), it remains to be seen how the implementation of the results of this study would alter previously-drawn conclusions as to the influence of Fisk-type fields on GCR modulation. This will be the subject of a future investigation.

Future work will also include looking into the effect of turbulence, known to reduce drift effects (e.g., Engelbrecht et al. [13]), on drift velocity profiles associated with a Fisk-type HMF. Furthermore, drift velocity profiles associated with other Fisk-type HMF models (e.g., Schwadron & McComas [20] and Hitge & Burger [21]) will be considered.

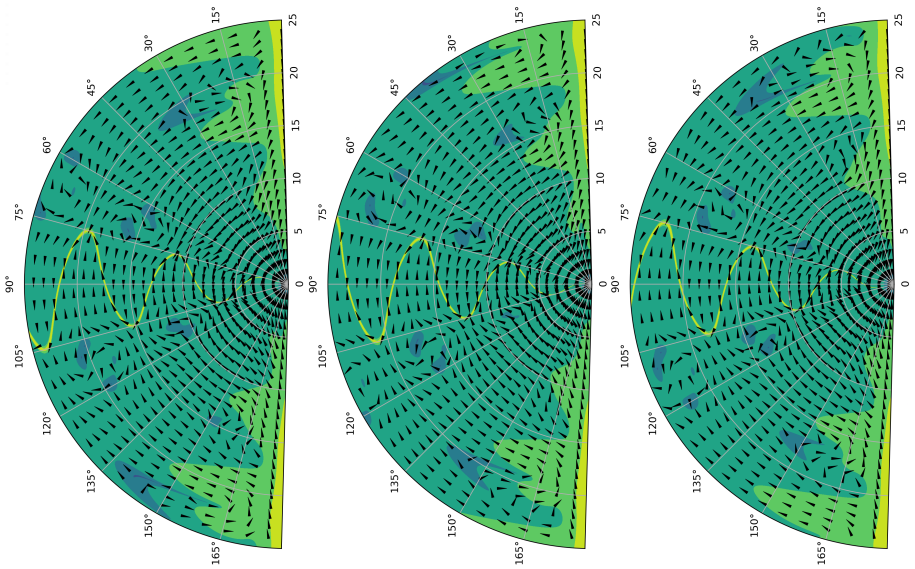


Figure 7. The $A > 0$ Fisk-Parker hybrid HMF drift velocity profile for 1 GeV protons at $\phi = 0°$, $\phi = 120°$ and $\phi = 240°$.

8. Acknowledgements

This work is based wholly on research supported by the National Research Foundation of South Africa (NRF grant numbers 142146 and 137793). Opinions expressed and conclusions arrived at are those of the authors and are not necessarily to be attributed to the NRF.

References

[1] Parker E N 1958 *Astrophysical Journal* **128** 664
[2] Fisk L A 1996 *Journal of Geophysical Research: Space Physics* **101** 15547–15553
[3] Burger R A and Hitge M 2004 *Astrophysical Journal* **617** L73
[4] Jokipii J R and Thomas B 1981 *Astrophysical Journal* **243** 1115–1122
[5] Burger R A 2012 *Astrophysical Journal* **760** 60
[6] Engelbrecht N E, Mohlolo S T and Ferreira S E S 2019 *Astrophysical Journal Letters* **884** L54
[7] Burger R A, Krüger T P J, Hitge M and Engelbrecht N E 2008 *Astrophysical Journal* **674** 511
[8] Sternal O, Engelbrecht N E, Burger R A, Ferreira S E S, Fichtner H, Heber B, Kopp A, Potgieter M S and Scherer K 2011 *Astrophysical Journal* **741** 23
[9] Chen F F 1974 *Introduction to Controlled Fusion* (Springer)
[10] Engelbrecht N E, Effenberger F, Florinski V, Potgieter M S, Ruffolo D, Chhiber R, Usmanov A V, Rankin J S and Els P L 2022 *Space Science Reviews* **218** 33
[11] Isenberg P A and Jokipii J R 1979 *Astrophysical Journal* **234** 746–752
[12] Forman M A, Jokipii J R and Owens A J 1974 *Astrophysical Journal* **192** 535–540
[13] Engelbrecht N E, Strauss R D, Le Roux J A and Burger R A 2017 *Astrophysical Journal* **841** 107
[14] Kóta J and Jokipii J R 1983 *Astrophysical Journal* **265** 573–581
[15] Snodgrass H B 1983 *Astrophysical Journal* **270** 288–299
[16] Zurbuchen T H, Schwadron N A and Fisk L A 1997 *Journal of Geophysical Research: Space Physics* **102** 24175–24181
[17] Engelbrecht N E and Burger R A 2010 *Advances in Space Research* **45** 1015–1025
[18] Mohlolo S T, Engelbrecht N E and Ferreira S E S 2022 *Advances in Space Research* **69** 2574–2588
[19] McComas D J, Barraclough B L, Funsten H O, Gosling J T, Santiago-Muñoz E, Skoug R M, Goldstein B E, Neugebauer M, Riley P and Balogh A 2000 *Journal of Geophysical Research: Space Physics* **105** 10419–10433
[20] Schwadron N A and McComas D J 2003 *Geophysical Research Letters* **30**
[21] Hitge M and Burger R A 2010 *Advances in Space Research* **45** 18–27

On the $R_h = ct$ universe

Edmund Kyazze and Amare Abebe
Centre for Space Research, North-West University, Potchefstroom, 2520, South Africa
E-mail: edmundkyazze@gmail.com

Abstract. Current astronomical observations show that the universe is undergoing an accelerated expansion and, although the standard model of cosmology is the closest we have to a standard theory of gravitation, it fails to adequately describe our Universe without the ad hoc introduction of dark energy and dark matter to late-time cosmology and inflation to early-universe cosmology. This has certainly created dilemmas for the wider astronomical community and several alternative cosmological models are being considered. Most current work is limited to the study of background dynamics so a brief overview of the equations that govern the evolution of cosmological perturbations in the $R_h = ct$ model is presented. This model is consistent with observations and has received particular attention in the last few years. Like the ΛCDM model, the $R_h = ct$ model is based on FLRW cosmology with the total energy density ρ and pressure p of the cosmic fluid satisfying the condition of the vanishing total gravitational mass condition: $\rho + 3p = 0$. The results derived will be compared to those from the standard model and analysed to see if the model allows for the formation of structure in the Universe.

Introduction

In modern cosmology the Cosmological Principle says that, on large enough scales, the Universe is homogeneous and isotropic. Furthermore, current observations suggest that it is expanding [1]. This expansion should satisfy the Einstein Field Equations (EFEs) for a universe whose total matter content may be assumed to be a perfect fluid. In particular, there should be some time dependence in the metric describing this expansion. To preserve the homogeneity and isotropy of space and incorporate time evolution a time-dependent expansion factor $a(t)$ is introduced to describe the time evolution of a homogeneous, isotropic universe [14]. Thus, to explore what sort of cosmological expansion is allowed by the theory, it is necessary to figure out what kind of cosmological expansion solves the EFEs.

Mathematically, there are several geometries that obey the Cosmological Principle but only three are of considerable cosmological interest: flat space, spherically-curved space and hyperbolically-curved space – where the latter two are a bit less obvious. However, the 4-dimensional space-time metrics that describe these spatial geometries are diagonal, which makes things much easier. Additionally, uniform scale expansion of the pure space hyper-surface can be incorporated into each of these metrics by multiplying each of the three pure space components by the scale factor. One is then free to insert these metrics into the EFEs and constrain the scale factor in the form of differential equations.

However, it is not actually necessary to insert three different metrics into the EFEs. Changing to co-moving coordinates allows all three metrics to be combined into one to obtain the standard

Friedmann-Lemaitre-Robertson-Walker (FLRW) metric with the mostly positive signature:

$$ds^2 = -c^2dt^2 + a^2(t)\left[\frac{1}{1-kr^2}dr^2 + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right)\right]. \tag{1}$$

The results are substituted into the EFEs and collapsed down to the two Friedmann equations:

$$\left(\frac{\dot{a}}{a}\right)^2 \equiv H^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2}, \tag{2}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) \tag{3}$$

where c and G are defined in the usual way. H is the Hubble parameter and the curvature parameter $k = \{-1, 0, 1\}$ for hyperbolically-curved, flat and spherically-curved space. The quantities ρ and $p = w\rho c^2$ are the mass density and pressure, respectively. w is the equation of state (EOS) parameter.

Birkhoff's Theorem

The imaginary surface beyond which all null geodesics recede from an observer is called the gravitational horizon. The radius at which this happens is the Schwarzschild radius. To understand this radius in a cosmological setting, imagine an observer situated at the centre of a spherical cavity of proper radius R_C and then surround them with a spherically-symmetric mass with a surface of proper radius $R_S < R_C$. The space-time outside this gravitating, nonrotating, spherical mass must then be given by the Schwarzschild metric [2, 15].

The density $\rho = \rho(t)$ is a function of cosmic time only so if R_S is allowed to increase, eventually a threshold will be reached of enclosed mass – at which point R_S becomes the gravitational horizon,

$$R_h = \frac{2GM}{c^2}, \tag{4}$$

where

$$M(R_h) = \frac{4\pi R_h^3}{3}\rho \tag{5}$$

is the proper mass enclosed within a sphere of proper radius R_h. Thus, for $k = 0$

$$R_h = \frac{c}{H} \tag{6}$$

The $R_h = ct$ Model

According to Weyl's postulate, any proper radius $R(t)$ used in an FLRW cosmology to express – not the co-moving distance r between two points but, rather – the changing distance that increases as the Universe expands must be expressible as the product of the constant co-moving distance and a universal function of time: $R(t) = a(t)r$ [14].

Therefore, $R_h = c/H(t)$ means that $aH = c/r$ is a constant and the Universe is expanding but with zero acceleration. This corresponds to either an empty universe (in which $\rho = p = 0$) or one characterised by an EOS parameter $w = -1/3$. John (2019) refers to this as the non-empty 'always coasting' model. Furthermore, this also means that $a \propto t$ and $H(z) = H_0(1 + z)$. H_0 is the current value of the Hubble parameter and z is the redshift related to the scale factor in the usual way. This, of course, leads directly to $R_h = ct$.

Cosmological Perturbations

The Cosmological Principle breaks down on smaller scales where there is clear evidence that the Universe is not homogeneous. The large-scale structures observed today are believed to have formed through small density perturbations whose growth rate depends on the dynamical expansion of the Universe [10].

To study these perturbations, it is convenient to define the dimensionless density contrast $\delta = \delta\rho/\rho_0$ in terms of the density fluctuation $\delta\rho(t, \vec{x})$ and the unperturbed density of the homogeneous background $\rho_0(t)$ [7]. In units where $c = 1$, the resulting differential equation describing the growth of density perturbations is

$$\frac{d^2\delta}{dt^2} + (2 - 3w)\, H \frac{d\delta}{dt} = \frac{3}{2} H^2 (1 + 3w)(1 - w)\, \delta + \frac{c_s^2 \nabla^2 \delta}{a^2}, \tag{7}$$

where $c_s^2 = d_\rho p$ is the sound speed squared. The second term on the LHS is due to cosmic expansion and always suppresses the growth of the perturbations. The combined term on the RHS relates gravity and pressure. One is free to describe a flat comoving geometry using Fourier decomposition

$$\delta(t, \vec{x}) = \frac{V}{(2\pi)^3} \int_k \delta_k(t) e^{\iota k \cdot x}\, d^3k,$$

where the complex Fourier coefficients δ_k only depend on the cosmic time and k and x are the comoving wavevector and radius, respectively [7]. For $3w = -1$, the dynamical equation is

$$\ddot{\delta}_k + \frac{3}{t}\dot{\delta}_k = \frac{\kappa^2 c^2}{3H_0^2 t^2}\delta_k. \tag{8}$$

For $p < 0$ the pressure term drives the growth of perturbations. The solution to this equation is

$$\delta_k(t) = C_1 t^{-1+\sqrt{1+k^2c^2/3H_0^2}} - C_1 t^{1-\sqrt{1+k^2c^2/3H_0^2}}, \tag{9}$$

where the mass conservation condition $\delta_{k=0} = 0$ was used. For real $\delta(x)$, $\delta(x)$=$\delta^*(x)$ implies $\delta_k^* = \delta_{-k}$.

It is often convenient to assume the perturbation is a Gaussian random field so that the waves in the Fourier decomposition have random phases. This way, the field may be specified entirely by its power spectrum [10], written as

$$P(k) = \langle |\delta_k(t)|^2 \rangle,$$

where the averaging is done over all space.

The solution $\delta_k(t)$ has an explicit time dependence that also appears in its power spectrum, so, although the general shape of the graph remains the same, its turning point slowly tends towards higher values of k as the Universe ages. In contrast, the power spectrum in the ΛCDM model is time-independent so the peak does not change as the Universe ages [12].

Conclusion

$R_h = ct$ hints at a great simplification of the Universe. Where the standard model predicts a decelerating expansion in the early radiation-dominated Universe (without inflation) and an accelerated late-time dark energy-dominated Universe, the $R_h = ct$ model suggests that, as per

the Cosmological Principle and Weyl's postulate, the mean effect of deceleration and acceleration is that the Universe has been expanding at the same rate $a(t) \propto t$ since the Big Bang. This simplification agrees with several late-time results: it correctly gives the age of the Universe as well as several galaxy clusters and agrees quite well with supernovae data. [1, 11].

Furthermore, $R_h = ct$ makes no ad hoc assumptions of early-time inflation and late-time acceleration due to dark energy. But despite these successes, there is the problem of the non-physical equation of state. Further work is being done to investigate this model in alternative and/or modified theories of gravity to resolve this issue. In particular, why does it agree with late-time cosmology so well?

References

[1] Bikwa, O; Melia, F; Shevchuk, A. 2012. *Photon Geodesics In Friedmann–Robertson–Walker cosmologies.* Monthly Notices of the Royal Astronomical Society. **421**. 3356 – 3361

[2] Bilicki, M; Seikel, M. 2012. *We Do Not Live In the $R_h = ct$ universe.* Monthly Notices of the Royal Astronomical Society. **425**. 1664 – 1668

[3] Harrison, E. 1991. *Hubble Spheres and Particle Horizons.* The Astrophysical Journal. **383**. 60 – 65

[4] Efstathiou, G; Hobson, M.P; Lasenby, A. N. 2006. *General Relativity: An Introduction for Physicists.* Cambridge University Press.

[5] John, M. V. 2019. *$R_h = ct$ and the Eternal Coasting Cosmological Model.* Monthly Notices of the Royal Astronomical Society. **484 (1)**. *L*35

[6] Sultana, J; Kazanas, D. 2017. *The $R_h = ct$ Universe in Alternative Theories of Gravity.* Monthly Notices of the Royal Astronomical Society. **472**. 2583 – 2589

[7] Melia, F; Shevchuk, A.S.H. 2012. *The $R_h = ct$ Universe.* Monthly Notices of the Royal Astronomical Society. **419**. 2579 – 2586

[8] Melia, F. 2013. *The $R_h = ct$ Universe Without Inflation.* Astronomy and Astrophysics. **553**. *A*76

[9] Melia, F. 2014. *On Recent Claims Concerning the $R_h = ct$ Universe.* Monthly Notices of the Royal Astronomical Society. **446**. 1191 – 1194

[10] Melia, F. 2017. *The Linear Growth of Structure in the $R_h = ct$ Universe.* Monthly Notices of the Royal Astronomical Society. **464**. 1966 – 1976

[11] Melia, F. 2018. *The Apparent (Gravitational) Horizon In Cosmology.* American Journal of Physics. **86**. 585 – 593

[12] Melia, F; Yennapureddy, M. K. 2020. *Structure Formation and the Matter Power-Spectrum in the $R_h = ct$ Universe.* Physics of the Dark Universe. **31**. 100752

[13] Misner, C. W; Thorne, K. S; Wheeler, J. A. 1973. *Gravitation.* W. H. Freeman and Company.

[14] Mukhanov, V. 2014. *Physical Foundations of Cosmology.* Cambridge University Press.

[15] Narlikar, J. 2002. *An Introduction to Cosmology: Third Edition.* Cambridge University Press.

[16] Riess, A. G et al. 2004. *Type Ia Supernova Discoveries at $z > 1$ From the Hubble Space Telescope: Evidence for Past Deceleration and Constraints On Dark Energy Evolution.* The Astrophysical Journal. **607**. 665 – 687

[17] Sultana, J. 2016. *The $R_h = ct$ Universe and Quintessence.* Monthly Notices of the Royal Astronomical Society. **457**. 212 – 216

[18] Weinberg, S. 2008. *Cosmology.* Oxford University Press.

[19] Wang, F. Y; Yu, H. 2014. *Reconciling the Cosmic Age Problem in the $R_h = ct$ Universe.* The European Physical Journal C. **74**. 3090

Combining large radio and optical astronomical surveys: exploring the MeerKAT Galaxy Cluster Legacy Survey

S.J.J. Nel and S.I. Loubser

Centre for Space Research, North-West University, 11 Hoffman St, Potchefstroom, 2520, South Africa

E-mail: 30022525@mynwu.ac.za, Ilani.Loubser@nwu.ac.za

Abstract. In this paper, we contrast and test the Broos method and the Likelihood Ratio method to evaluate the reliability and completeness of the cross-matching method, and the resulting catalogues, to determine which method is better suited to optimally combine large radio and optical surveys. The radio and optical surveys used for the cross-matching was the MeerKAT Galaxy Cluster Legacy Survey (MGCLS), and the Dark Energy Camera Legacy Survey (DECaLS) respectively. The test was carried out by selecting nine clusters with varying degrees of completeness. The selected clusters were then used to perform the cross-matching where the corresponding statistical properties were calculated by both methods. We found that there is no change in the completeness between the two methods; however, there does seem to be some difference in the reliabilities. The Broos method produced consistently larger reliabilities at lower and intermediate completeness levels, while the likelihood ratio method stayed consistent throughout. We decided to incorporate the Likelihood Ratio method into our cross-matching procedure because of its stability in reliability across all levels of completeness.

1. Introduction

The Southern Hemisphere finds itself in an era of large radio and optical astronomical surveys. One challenge is to effectively combine and cross-match source catalogues from surveys in different wavelengths, thereby maximizing the science impact of surveys. We used the Likelihood Ratio method to cross-match compact source catalogues extracted from MGCLS with optical sources detected in DECaLS DR8. These matched catalogues can then be further explored for various science applications and to possibly improve the cross-matching method for future surveys.

One of the aspects of the cross-matching method that we set out to study is the manner by which the statistical properties of the cross-matched catalogues are calculated. It is important to know what the limitations of the different methods are so that the appropriate method can be implemented depending on the data that are available. The two main methods used in our cross-matching code to calculate these properties are the Broos method [1] and the Likelihood Ratio method [2]. The reason we are considering these two methods is because of the differences in the way they calculate the necessary properties. The Likelihood Ratio method is a natural choice for these calculations since we are already using the method to perform the cross-matching between the radio and optical surveys. This technique was tested and adapted for cross-matching by

[2] and was later also used as a tool to identify optical and infrared counterparts by [3]. The Likelihood Ratio method calculates false and true positives by determining the probability of each match that was made [3].

The Broos method developed by [1] uses a technique in which they simulate the behavior of their matching algorithm to detect all false positive and true negative matches; they do this by creating two populations, an associated population and an isolated population. These two populations are used to divide the data into two separate cases, one in which the source has a real counterpart (associated population) and one in which there is no counterpart (isolated population) [1]. This method was initially formulated to eliminate false positive matches, which arises from the fact that there were orders of magnitude more source entries in their IR catalogues than what was present in their X-ray catalogues [1]. These two methods will be compared by their ability to formulate matches between sources (completeness) and by the accuracy of those matches (reliability).

In Section 2 we will discuss the procedure that was used to compare and test the Broos and Likelihood Ratio methods. The results of the comparison will be given in Section 3, and they will be discussed and concluded in Section 4.

2. Method of comparison

The comparison procedure started by selecting nine galaxy clusters from the surveys that had a variety of completeness levels, which has been produced in the cross-matching procedure. To do so, we first had to determine the redshift range of the galaxy clusters in our surveys. This is an important step in the selection process since our optical surveys (DECaLS) contain only photometric redshifts, which are known to be less reliable than their spectroscopic counterparts. The photometric redshifts were derived from fitting spectral energy distributions to broad-band photometric measurements as described in [4]. We use the boundaries determined by [5] to exclude clusters that are not found in the photometric redshift range $0.15 < z_{ph} < 0.40$. These boundaries were determined by [5] to be the point of separation between photometric and spectroscopic redshifts. This separation is caused by broad band measurements that could not take place due to the lack of u-band filters in DECaLS [5]. Figure 1 shows the distribution of galaxy cluster redshifts used to identify acceptable galaxy clusters in the MGCLS survey to use for this comparison.

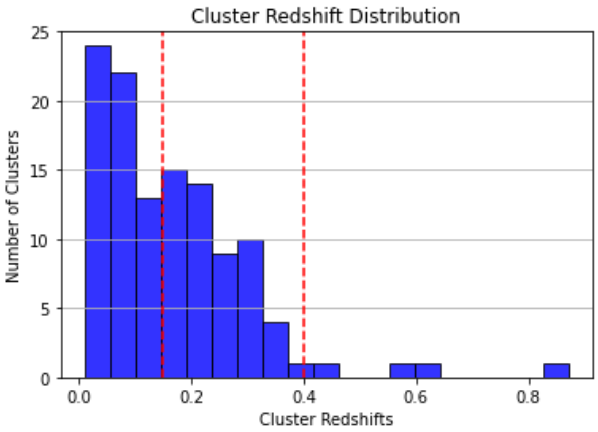


Figure 1. Distribution of MGCLS galaxy cluster redshifts. The red dashed vertical lines show the redshift range that can be used for accurate photometric redshift studies.

Galaxy clusters identified to have redshifts in the range where photometric redshifts are

accurate can be used for the comparison, we can choose three clusters that have a low, intermediate, and high completeness level. Growth in the completeness as a function of Likelihood Ratio (LR) is used so that we can determine whether the methods can make accurate matches, which is independent of the amount of matches found. With this in mind, we selected the nine clusters from table 1 and 2 from [6]. The selected galaxy clusters can be seen in our Table 1. To test and compare the methods, we ran the cross-matching code as described in [7], and [6] for all nine clusters using both methods. We then plotted their completeness and reliability outputs against LR. In the cross-matching method, the LR cutoff at 0.5 is used as the probability above which a match can be considered reliable. The cutoff is also used to find a balance between completeness and reliability in LR. In general, the higher the LR, the more reliable but less complete the matched sample. These plots can be seen in figure 2 in the results section.

3. Results

Table 1. The selected galaxy clusters. This table is divided into three main groups Low, Intermediate, and High, indicating the relative level of completeness found in that group. The individual galaxy clusters and their corresponding completeness values [6] are also indicated.

Low		**Intermediate**		**High**	
Cluster Name	Completeness	Cluster Name	Completeness	Cluster Name	Completeness
J0014.3-6604	0.185	J0449.9-4440	0.397	Abell 370	0.483
Abell 33	0.328	J0232.2-4420	0.411	Abell 2744	0.492
J0416.7-5525	0.356	J0177.8-5455	0.420	Abell 2813	0.502

We present the nine selected MGCLS clusters that were used to generate the statistical properties of both methods, as well as the graphs of the output of those methods which are used for the comparison of their performance. Figure 2 shows the comparison between the Broos and Likelihood Ratio method for three of the selected clusters. We only show three clusters here - one for each level of completeness, as all three clusters per level of completeness show the same trend. We decided to only show the graphs for the cluster with the lowest completeness (J0014.3-6604), the lowest intermediate completeness (J0449.9-4440), and the highest completeness (Abell 2813). When comparing the two methods, we see that in each of the three cases, the completeness (blue line) is the same for both methods. This indicates that both methods found the same number of matches during the cross-matching procedure. If we now look at the reliability (red dashed line) in the $0 < LR < 0.50$ region, we see that for all three clusters, the Broos method has achieved maximum reliability quicker at lower LR than the Likelihood Ratio method. This means that even the first few matches that are made by the Broos method can be considered to be accurate matches, whereas the Likelihood Ratio method could give more false positives at lower LR.

Looking at the reliability at an LR of 2 shows that there exists a discrepancy between the two methods. The difference between the maximum reliability of the two methods is greater in the lower completeness level and smaller at larger completeness levels. However, there does seem to be a trend where the Broos method's reliability decreases with increasing completeness. But the reliability of the Likelihood Ratio method is not very sensitive to changes in varying completeness levels. However, little emphasis will be placed on these trends, as it is hardly ever necessary to use LR values as large as this.

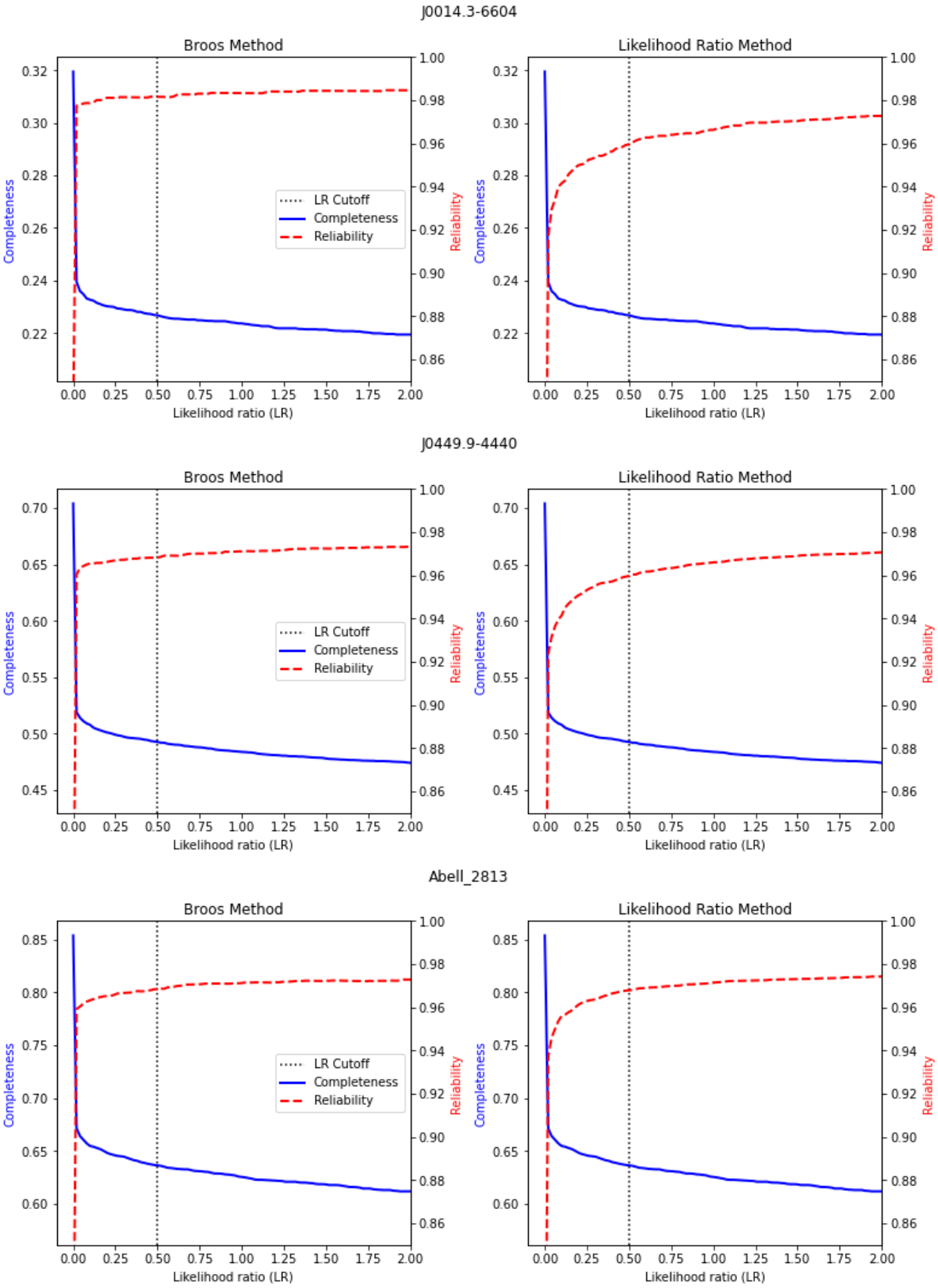


Figure 2. This figure shows the completeness (blue solid line) and the reliability (red dashed line) of the Broos method and the Likelihood Ratio method plotted against the probability of finding a match. The cluster completeness increases from top to bottom. The line at 0.50 (black vertical dotted line) is the LR cutoff which is used to provide an estimate of the spurious identification rate [7].

4. Conclusion

We have set out to investigate the comparison of the Broos and Likelihood Ratio methods, with a special emphasis placed on their ability to calculate statistical properties from the cross-matching procedure. We achieved this by using data from MGCLS and DECaLS to perform the cross-matching, which we then used to calculate the properties of the two methods. To ensure that our comparison is accurate, we excluded galaxy clusters from the test based on their redshifts according to the findings of [5]. They showed that photometric redsifts are not consistent with spectroscopic redshifts at low and high redshifts. For the sake of a complete comparison, we also selected clusters based on their completeness values to test the methods under various circumstances. The test was performed by completing the cross-matching for each of the selected clusters with the two methods calculating the statistical properties during the procedure. The comparison was then made by plotting the completeness and reliability determined by both methods for each selected cluster. We found that the two methods reached the same completeness in every case, but varied in their determination of reliability. We also found that the Broos method outperformed the Likelihood Ratio method in the region $0 < LR < 0.50$, as its reliability reaches a maximum much faster at lower LR in most circumstances. This indicates that when the uncertainty in the matching is larger (low LR), the certainty that a match is a true match is greater with the Broos method than with the Likelihood Ratio method.

From this comparison, we found that the Broos method decreased in reliability as the completeness levels increased, which is in stark contrast to the stability shown by the Likelihood Ratio method. It is also important to note that the Broos method was born out of necessity to deal with surveys that differed wildly in number of source entries [1]. Thus, the ability of the Broos method to produce accurate matches is independent of the number of source entries in the surveys used for the cross-matching procedure. From this comparison we find that the methods' usability is dependent on the data used, and on whether the researcher chooses to prioritize completeness or reliability. After incorporating the most suitable method, future work will involve determining the cluster membership for all clusters found in MGCLS.

References

[1] Broos P S, Feigelson E D, Townsley L K, Getman K V, Wang J, Garmire G P, Jiang Z and Tsuboi Y 2007 *The Astrophysical Journal Supplement Series* **169** 353

[2] Sutherland W and Saunders W 1992 *Monthly Notices of the Royal Astronomical Society* **259** 413–420

[3] McAlpine K, Smith D, Jarvis M, Bonfield D and Fleuren S 2012 *Monthly Notices of the Royal Astronomical Society* **423** 132–140

[4] Hilton M, Hasselfield M, Sifón C, Battaglia N, Aiola S, Bharadwaj V, Bond J R, Choi S K, Crichton D, Datta R *et al.* 2018 *The Astrophysical Journal Supplement Series* **235** 20

[5] Kesebonye K C, Hilton M, Knowles K, Cotton W D, Clarke T, Loubser S I, Moodley K and Sikhosana S P 2023 *Monthly Notices of the Royal Astronomical Society* **518** 3004–3016

[6] Loubser I 2023 *North-West University Research Repository* https://doi.org/10.25388/nwu.21975542.v1

[7] Knowles K, Cotton W, Rudnick L, Camilo F, Goedhart S, Deane R, Ramatsoku M, Bietenholz M, Brüggen M, Button C *et al.* 2022 *Astronomy & Astrophysics* **657** A56

Investigating the solar differential rotation rate by sunspot tracking using terrestrial solar observations

Calmay Lee, P.J. Steyn
North-West University, Center for space research
E-mail: calmay.lee9@gmail.com

Abstract. The Centre for Space Research at the North-West University, South Africa, commissioned a solar telescope in 2021. The first science application of the solar telescope was investigating the differential rotation rate of the Sun. The solar differential rotation rate is known to be latitudinally dependent. In this study, the photospheric rotation rate is experimentally determined by tracking multiple active regions over successive days using a dedicated hydrogen-alpha solar telescope. The observational campaign runs over a period of 46 days between 15 August and 14 October 2022. A total of 29 active regions are studied. The results of this study are compared to three published models. On average, a faster rotation rate is observed when compared to accepted model predictions. Factors influencing the observational data are investigated, such as the influence of the number of successive observations of each active region, the area of active regions (containing multiple sunspots), as well as the location of each active region with respect to the solar limb.

1. Introduction

The rotation of the Sun is not uniform; instead it exhibits differential rotation where its angular rotation rate is greater at the solar equator compared to the solar poles [1]. The solar differential rotation rate has been extensively studied during the past 50 years, including its connection to the tachocline which is the transition region between the radiative zone and the differentially rotating convective zone (see e.g. [2], [3], [4], [5]).

Studies such as [6], [7], [8] and [9] have investigated this phenomenon and have engaged in real-time solar observations spanning several years. In general, there are two approaches to determine the photospheric rotation rate. The first method utilizes spectrographic observations of Doppler displacements in specific spectral lines near the solar limb [10]. The second method involves analyzing the longitudinal movements of semi-permanent features on the solar disk, such as sunspots, faculae, and magnetic field patterns [10]. This study makes use of the latter method, following sunspot movement across the solar disk. The main objective of the study is observing sunspots and analyzing their latitude dependent differential rotation rate. Furthermore, this analysis involves comparisons between observational data with three existing models documented in literature.

2. Solar rotation periods

There are two distinct solar rotation periods, namely the synodic- and the sidereal period [11]. The sidereal period is measured from a solar-centric reference frame and refers to the time

required for a solar feature to complete one revolution around the Sun [12]. The synodic period is an Earth-based reference frame and represents the apparent rotation period as observed from Earth. Due to the orbit of the Earth around the Sun, this synodic perspective influences the apparent rotation rate of a sunspot. Taking this orbital factor into account, the synodic period at the equator is derived as approximately 26 days. While observations are conducted in synodic time, the conversion to the sidereal period is necessary since historical rotation rates are universally reported in terms of the sidereal period.

Equation (1) shows the conversion of the synodic period to the sidereal period

$$P = \frac{S \times E}{S + E}, \tag{1}$$

where P represents the sidereal period and S the synodic period [13]. E is the orbital period of the Earth (365.25 days).

3. Differential rotation models

Several researchers have investigated the solar differential rotation rate with the aim to establish a theoretical equation that accurately characterizes this phenomenon. The approach involved continuous solar observations spanning multiple years to derive an estimation of the latitude-dependent solar differential rotation rate. The models considered within this investigation all have approximately the same form, with only the coefficients differing between models.

The rotation profile of the Sun as determined by [8] using a periodic standard functional fit given by

$$\omega(\theta) = 2.902 - 0.464 \sin^2\theta - 0.328 \sin^4\theta \quad \mu\text{rad/sec}, \tag{2}$$

shows the rotation rate ω in micro radians per second, where θ is the solar latitude. The first term refers to the solar equatorial rotation rate (when θ is equal to 0°), followed by the second and third periodic terms which denotes the rotation rate of the photosphere at different solar latitudes.

In the study by [7] they determined the fit to be

$$\omega(\theta) = 2.779 - 0.351 \sin^2\theta - 0.442 \sin^4\theta \quad \mu\text{rad/sec}, \tag{3}$$

where θ also represents the solar latitude. Finally, [6] obtained the functional fit

$$\omega(\theta) = 2.904 - 0.492 \sin^2\theta \quad \mu\text{rad/sec}, \tag{4}$$

where θ is again the solar latitude.

4. Instrumentation and imaging

This study made use of the recently commissioned North-West University (NWU) Solar Observatory in Potchefstroom, South Africa. The observational instruments include a Lunt® 50 mm aperture hydrogen-alpha (Hα) telescope mounted on an iOptron® CEM 60 equatorial mount. A monochromatic charged-coupled device (CCD) camera is used for imaging. The resolution of the of CCD camera is 1360 x 1024 pixels and each pixel has a size of 4.65 x 4.65 μm. The focal length of the telescope is 350 mm and therefore provides a resolution of 2.74 arc-seconds/pixel. For solar observations, this equates to approximately 2.1×10^3 km/pixel on the solar disk. Although sunspots vary considerably in size (≈ 10 - 160×10^3 km), the resolution provided by this setup is enough to resolve sunspots on the solar surface. Daily solar images were captured with the Hα telescope throughout the two-month campaign. It should be mentioned that, in principle, this study could also be performed with a standard white-light solar telescope.

However, the Hα telescope (restricted to the 656.2 nm emission line) was used since it offers more detail regarding the chromospheric conditions within and around the active regions (ARs), such as magnetic field properties, plages and spicules. However, the study focuses solely on determining the locations (heliolatitude and heliolongitude coordinates) of either single sunspots or sunspot groups within ARs and its change over time. The daily-captured solar images are vertically flipped since the NWU Solar Observatory is located in the southern hemisphere. This enables easy comparison with international data networks. A Stonyhurst diagram, a convenient spherical coordinate system [14], is superimposed onto the solar images. Next, the heliolatitude and heliolongitude coordinates of each sunspot (AR) are recorded. For cases where ARs contain multiple sunspots (sunspot groups), the most prominently identifiable sunspot (the largest) is traced. Finally, the results are converted from the synodic period to the sidereal period using equation (1). As an example, this process is illustrated in figure A1 in Appendix A where two successive full-disk solar images are shown between 19 June and 21 June 2023 while tracking AR 3335. This AR is located in the southern hemisphere ($-16.5°$), the two observations were done 46.2 hours apart, the change in heliogrpahic longitude was determined to be 25.3°, which resulted in a differential rotation rate of 2.7 μ rad/sec.

5. Observational results

Figure 1 shows the sidereal rotation rates of the 29 tracked ARs during the observational campaign. While the presumption holds that sunspots follow a constant heliographic latitude, minor deviations were noticed, typically measuring less than 1°. For completeness sake, the small deviations in heliographic latitude are averaged. These are expressed in radians, while the averaged change in heliographic longitude over time is expressed in micro radians per second (μ rad/sec), which equates to the sidereal differential rotation rate. The standard deviation of the rotation rate for each AR is calculated based on ARs sharing the same latitude or falling within a range of ± 0.1 radians. Negative heliographic latitudes refer to the southern hemisphere, while positive values indicate the northern hemisphere. The results are categorized into two distinct groups. The first group shows faster-than-anticipated periods, while the second group aligns closely with the models, within the bounds of standard deviation error margins. Figure 1 shows several sunspots rotating at a faster rate than predicted by the models of [6], [7], and [8]. The scattering of raw data is expected, as discussed in the paper by [9], in which the equatorial rotation rates were examined for the years 1921 through 1982, spanning a 62 year period and effectively illustrates the level of variability one can expect while tracing equatorial tracers for the purpose of determining the equatorial rotation rate. The sidereal rotation rates recorded during this observational campaign show a stronger alignment with the models proposed by [8] and [6]. While some ARs adhere to the projected model trends, our primary focus is on understanding the reasons behind the deviations in other ARs and there are multiple influential factors that contribute to these deviations.

The number of successive days that an AR is observed plays an important role in determining the differential rotation rate (more observations lead to a more accurate rotation rate). Therefore, terrestrial weather conditions play an influential role and may limit the amount of observation days. In an ideal scenario, an extended observation period coupled with a larger count of tracked ARs would increase the reliability of the results. However, this project was restricted to a two-month observation period. The study of [9] also highlights sunspot size as a significant contributor to the scatter of data. Sunspots and sunspot groups of greater size exhibit a comparatively slower rotation rate than smaller- and single sunspots. This suggests that sunspots at the same latitude could potentially display differing rotation rates due to their size. The influence of sunspot size on its differential rotation rate is a important focus area for future work.

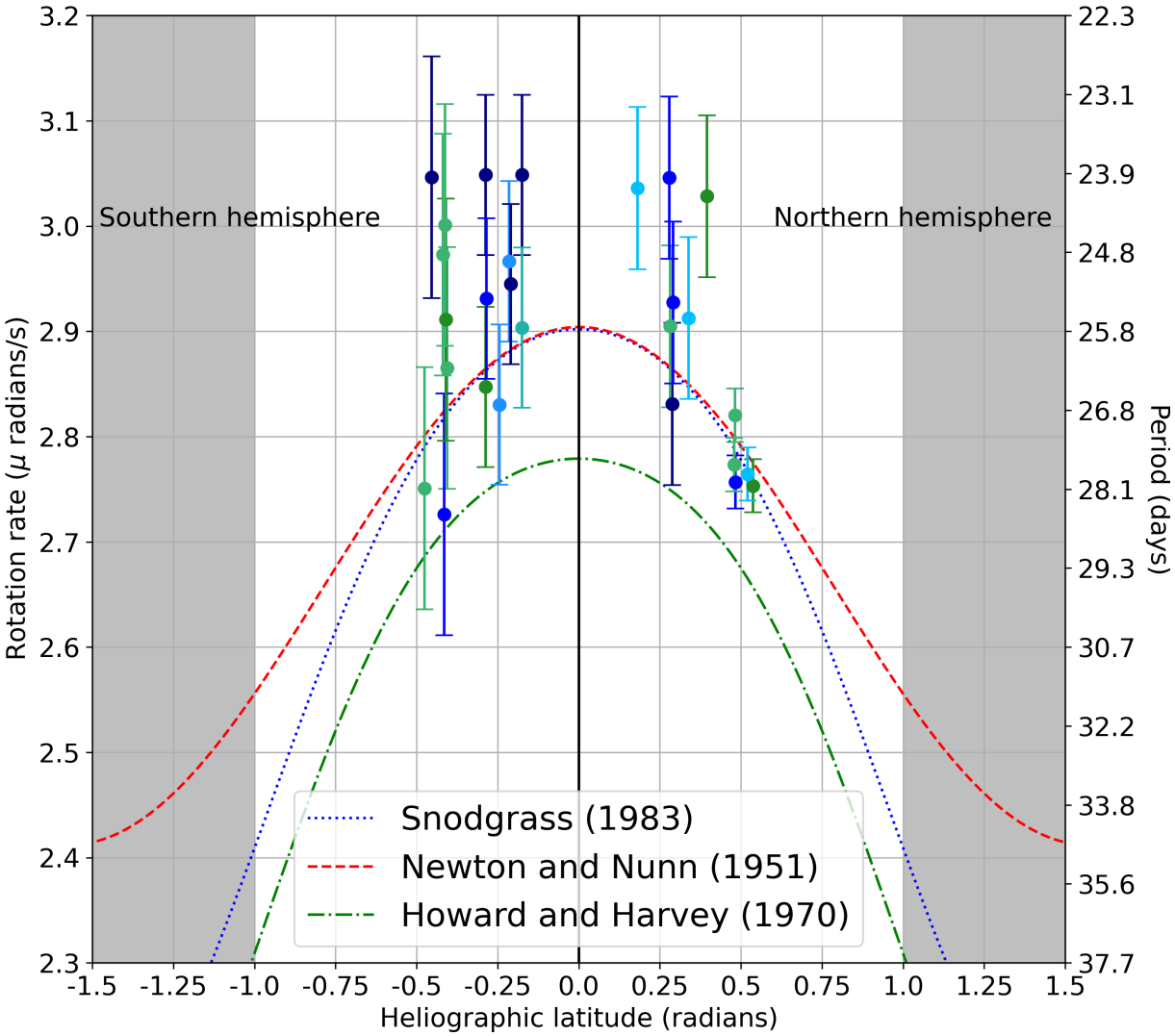


Figure 1. Sidereal differential rotation rates of 29 ARs as determined by the observational campaign from 15 August to 14 October 2022 at the NWU Solar Observatory. The left y-axis shows the rotation rate in μrad/s. The right y-axis shows the period of each AR in days. The three models are shown in the legend.

6. Discussion and conclusion

Six influencing factors were recognized for the observed deviation of the observational data from the model patterns in this study. Improving these factors will substantially increase the precision of future research. While some ARs conform to the anticipated model trends within an acceptable error margin, the primary emphasis remains on the exploration of ARs that deviate from the model profiles.

Firstly, the number of successive days that an AR is observed influences the final result. The credibility of data decreases with a lower number of successive observations, and particularly affect ARs monitored over only a two-day period. Secondly, the sizes of sunspots has an influence on the rotation rate as shown by [9], where smaller sunspots travel at a faster rate compared to larger sunspots or sunspot groups. Several additional factors also contribute to the observed variability in the collected data. For instance, a 1° discrepancy in the recorded longitude leads

to a 0.2 μrad/s change in the final rotation rate outcome. The error introduced by manually operating the telescope and imagining processes contribute to inaccuracies. Minor inaccuracies at this stage of the data processing can progress into an error in the results. Future software and hardware upgrades of the solar telescope will enable full automation of telescope and imaging processes. These upgrades will mitigate errors that arise due to manual operation. Moreover, complications arising from projection near the solar limb play a role. ARs close to the solar limb present a challenge, as a sunspot in this area might seem to have a smaller $\Delta\phi$ (change in heliographic longitude), when, in fact, the sunspot is moving away from the observer along the curvature of the Sun. Furthermore, the time span between observations, if exceeding 24 hours, can alter the data and impacts the overall average which gives rise to outliers. Lastly, the duration of this observational campaign spanned 46 days. In contrast, [8] observed sunspots and sunspot groups for over 15 years, compared to [6] which spanned a decade. This disparity is reflected in the equatorial rotation rate variation observed by [7], who analyzed sunspots for only a year. Consequently, extended observation periods are advantageous in accumulating a more substantial sunspot sample size. In conclusion, this study pursued three key objectives: imaging sunspots and sunspot groups (within ARs) and determining their heliographic latitude-dependent differential rotation, comparing data from the NWU Solar Observatory with the discussed models, and contributing to a larger data collection for theoretical comparison. The results of this study shows a faster-than-expected rotation rate at most heliographic latitudes, especially close to the solar equator. The possible reasons for this are discussed in the text.

Acknowledgments
This work is based on the research supported in part by the National Research Foundation of South Africa (Ref Numbers TTK2204183656).

References

[1] Bruzek A and Durrant C J 1977 *Illustrated glossary for solar and solar-terrestrial physics* vol 69 (Taylor & Francis)
[2] Schroeter E H 1985 **100** 141
[3] Thompson M J, Toomre J, Anderson E R, Antia H M, Berthomieu G, Burtonclay D, Chitre S M, Christensen-Dalsgaard J, Corbard T, De Rosa M, Genovese C R, Gough D O, Haber D A, Harvey J W, Hill F, Howe R, Korzennik S G, Kosovichev A G, Leibacher J W, Pijpers F P, Provost J, Rhodes E J J, Schou J, Sekii T, Stark P B and Wilson P R 1996 *Science* **272** 1300–1305
[4] Beck J G 2000 **191** 47–70
[5] Chandra S, Vats H O and Iyer K N 2010 **407** 1108–1115 (*Preprint* 1005.0490)
[6] Newton H and Nunn M 1951 *Monthly Notices of the Royal Astronomical Society* **111** 413–421
[7] Howard R and Harvey J 1970 *Solar Physics* **12** 23–51
[8] Snodgrass H B 1983 *Astrophysical Journal* **270** 288–299
[9] Howard R 1984 *Annual review of astronomy and astrophysics* **22** 131–155
[10] Howard R 1977 *Illustrated glossary for solar and solar-terrestrial physics* (Springer) pp 7–12
[11] Karttunen H, Kröger P, Oja H, Poutanen M and Donner K J 2007 *Fundamental astronomy* (Springer)
[12] Bartlett A K 1904 *Popular Astronomy, vol. 12, pp. 649-651* **12** 649–651
[13] Freedman R, Geller R and Kaufmann W 2010 *Universe* (W. H. Freeman) ISBN 9781429231534 URL https://books.google.co.za/books?id=XfR8QgAACAAJ
[14] Çakmak H 2014 *Experimental Astronomy* **38** 77–89

Appendix A.

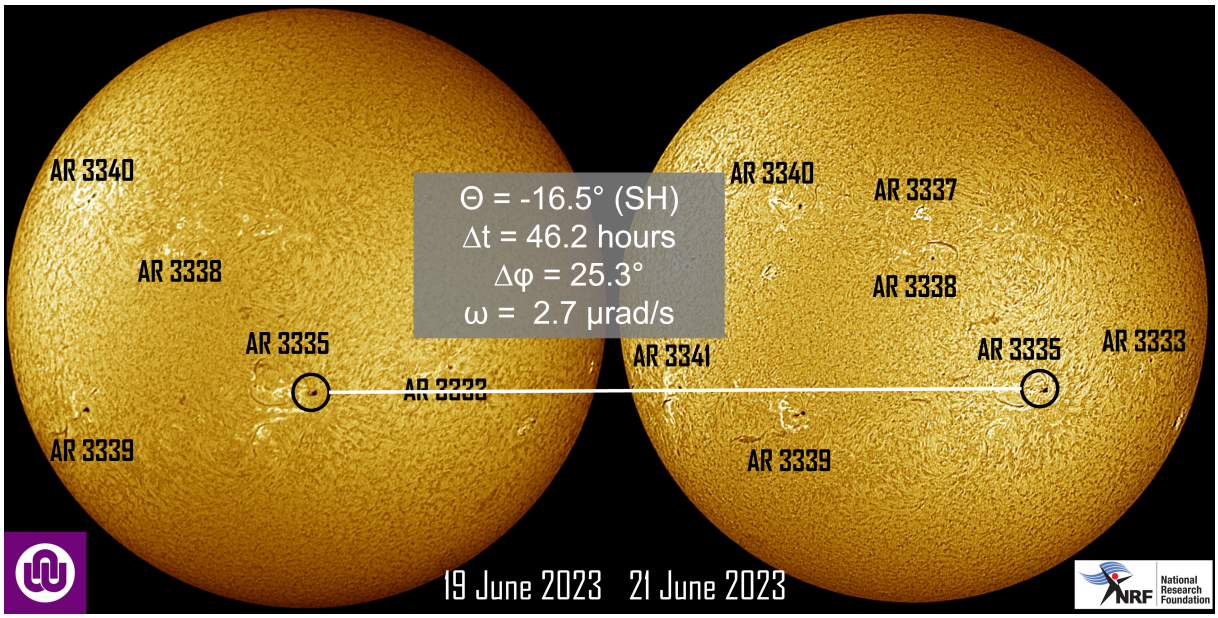


Figure A1. Colourised full−disk examples of daily solar images captured by the 50 mm Hα telescope from the NWU Solar Observatory to determine the differential solar rotation rate. AR 3335 is used as an example between 19 and 21 June 2023. Other ARs are also shown. The details are discussed in the text.

Commensal searching for radio transients/variables in data from the MHONGOOSE Large Survey Project (LSP).

VM Tshilengo[1], NE Maluta[1,2], PA Woudt[1,3]

1. University of Venda, Department of physics, P/Bag X 5050, Thohoyandou, 0950, South Africa

2. The National Institute for Theoretical and Computational Sciences (NITheCS), Gauteng, South Africa

3. Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa

E-mail: tshilengov@gmail.com , Eric.Maluta@univen.ac.za , pwoudt@ast.uct.ac.za

Abstract. The MeerKAT radio telescope is the newly built South African precursor radio telescope array for the mid-frequency component of the Square Kilometre Array (SKA1-Mid). Its excellent sensitivity and wide field of view, combined with commensal access to a large number of MeerKAT surveys, allows astronomers to gain new insights in time-domain radio astronomy, particularly in uncovering the population of faint radio transients and variables. In this study, we use the South African Radio Astronomy (SARAO) Science Data Processing (SDP) images from the MHONGOOSE Large Survey Project (LSP) to search for new radio transients. We used the Transient Pipeline (TraP) on the cloud compute infrastructure of the Inter-University Institute for Data Intensive Astronomy (IDIA) to characterise the light curves of point sources in the field using statistical variability parameters. Here we show the first results of our search for radio transients/variables in the MHONGOOSE data set of NGC1566 showing the preliminary detection four variable sources which are likely to be AGNs.

Keywords: Transients; Variables; Radio Astronomy

1. Introduction

Commensal searches for transients and variables are proving to be an effective technique to explore the radio sky, especially when using facilities with a wide field of view. These searches can increase our understanding of the populations of sources in the radio sky by constraining transient rates and interesting sources (e.g.,[1]). Transient radio astronomy focuses on a wide range of observing frequencies (30 MHz – 150 GHz, e.g.,[2]), in this work we used observations at L-band (covering 900–1670 MHz), that is, the mid-frequencies of the MeerKAT [3] searchingfor dynamic and explosive events happening out in space using the MeerKAT radio telescope. MeerKAT is an ideal telescope for searching for dynamic and explosive events taking place in the universe. These events can be triggered by different events such as stellar flares, supernovae, etc. MHONGOOSE is targeting 30 different galaxies (pointings) observed over 10 epochs per pointing [4].

In this work, we present transient searches from the MHONGOOSE survey, specifically from the observation of the field of NGC 1566 obtained from 21 October 2020 to 02 November 2020. NGC1566 is an active (Seyfert 1), spiral galaxy in the southern hemisphere at a distance of 21 Mpc.

2. Methodology

For this project we use the Science Data Product (SDP) 4k continuum images, which are automatically generated by the South African Radio Astronomy Observatory (SARAO) data pipeline. Each one of the 10 images represents a $\sim$5 hr-long observation. SDP images are generally used for quality assurance and are not subject to robust calibrations. However, the advantage is that they are available immediately after the observation is concluded and potentially allow for the discovery of image-domain transients in near-to real time. In this project we are testing the use of the SDP image for transient detection and benchmark this against transient searches in the calibrated continuum images from the MHONGOOSE project derived from the simultaneous narrow-band continuum data. In this paper we focus on early results from the SDP images.

The analysis of the data and search for radio transients was performed using software developed by the LOFAR Transients Key Science Project team, named the Transients Pipeline (TraP, [5]). TraP is built to search for transients in the image plane, whilst also storing light curves and variability statistics of all detected sources. A full and detailed overview of the TraP can be found in [5]. Table 1 shows some of the input parameters used on TraP. When running the images through the pipeline, a detection threshold of 8σ was used, which is the threshold for blind detection of a source, along with an analysis threshold of 3σ, which is the threshold used for analysing information about the source (such as position and flux, and uncertainties in those quantities). We used the Cube Analysis and Rendering Tool for Astronomy (CARTA, [6]), an image visualization and analysis tool, to inspect the positions of the new sources in the original images from one epoch to another. We also inspected the light curves of these new sources and check the position for counterparts from other survey catalogues.

Table 1. The parameters used within the TraP.

detection threshold	8σ
extraction_radius_pix	1500 pixels
force beam	True
new_source_sigma_margin	1
elliptical_x	2.0
beamwidths_limit	1.0
analysis_threshold	3σ

3. Results and Discussion

We analysed the first results in terms of the distribution of the statistical parameter of all the point sources in the NGC 1566 field where we plotted V against η in Figure 1. For each detected sources the dimensionless variability parameter (η and V, see equation 35 and 36 of [5]) were calculated by the TraP. Any sources with variability parameters exceeding 2σ for one or both values of η and V are considered as potentially variable. Variable sources are typically identified using the two key variability parameters calculated by the TraP: the reduced weighted $\chi 2$ of a fit assuming a constant brightness, η, and the coefficient of variation, $V = s/<I>$ (where s is the standard deviation of the flux density measurements and $<I>$ is the average flux density of the source, as defined in [5].

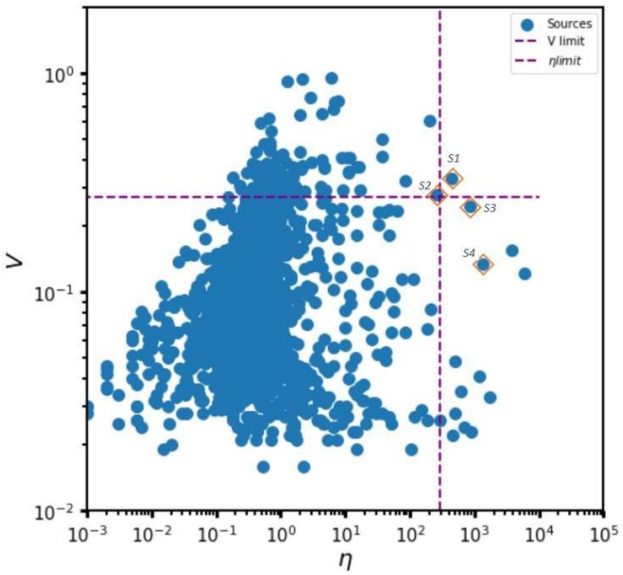


Figure 1. The distribution of the variability parameters V, a coefficient of variation, η, the significance of the variability for each light curve detected. The dotted line represents the V limit of 0.27 and η limit of 288.2. Our sources are blue in colour where four sources are highlighted (S1-S4) and discussed in the text. They are marked with an orange rhombus shape.

In the following discussion we highlight four different sources that stand out in terms of their η and V variability statistics (S1-S4 in Figure 1, and Table 2). Figure 2 shows the individual light curves of the four sources labelled in Figure 1. From the first light curve (S1 = SRC29675: top left), one can notice that the flux starts around 0.34 mJy then it jumps up where we haveour peak flux of 0.85 mJy it drops and goes up again. For SRC28339 (= S2, top right) the flux starts at 0.48 mJy then itjumps up where we have our peak flux of 0.87 mJy it drops and goes up again. As for the third light curve (S3 = SRC28214: bottom left) it starts with a flux around 0.9 mJy then went up and drop the pattern repeat and ithas a peak flux of 1.7 mJy. The last light curve (S4 = SRC29725: bottom right) starts at a flux of 3.2 mJy then went up again and then drop down which has a peak flux of 3.78 mJy before it drops down; it is one of the brightest sources that has been observed in our data.

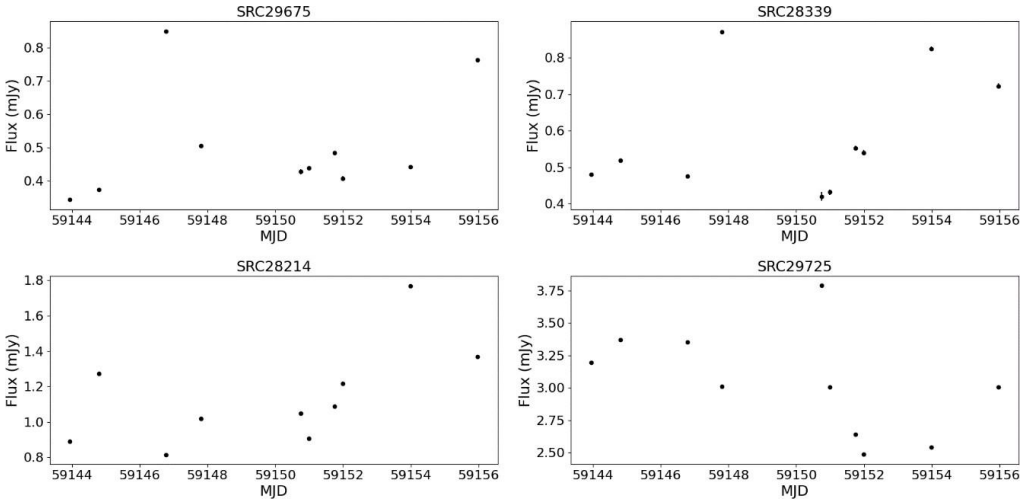


Figure 2. The observed radio light curve of SRC29675, SRC28339, SRC28214 and SRC29725 showing the dates and peak fluxes. Typical uncertainties in the flux density measurements are in the range of 0.04-0.07 mJy and are too small to be visible in the light curve.

Table 2. Properties of the variable sources identified in the field of NGC 1566.

	SourceID	RA (2000.0)	DEC (2000.0)	Peak flux (mJy)	Peak flux error (mJy)	R(mag)	V	η
S1	29675	04:17:27.2	-55:13:07.8	0.8486	0.0042	20.6055	0.332	438.1
S2	28339	04:20:28.6	-55:10:47.2	0.8713	0.0048	19.0882	0.280	259.8
S3	28214	04:20:44.7	-55:24:34.8	1.7697	0.0059	24.3780	0.248	867.6
S4	29725	04:17:15.4	-54:45:16.9	3.7872	0.0067	18.8728	0.134	1344.8

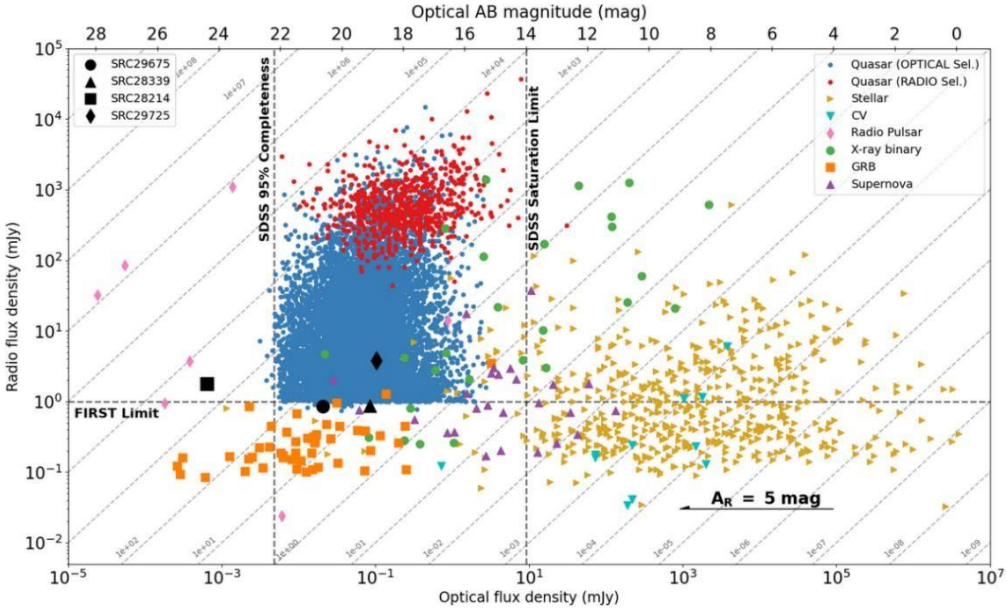


Figure 3. Plot showing radio versus optical flux density of our sample of objects from different classes, Adapted from Figure 1 of [7]. Four variable sources identified are shown with black symbols.

The light curves shown in Figure 2 do not show any obvious correlated variability and seems to suggest that radio transient behaviour – at least to first first order – are not due to calibration issues. This will need to be investigated further with the aid of the independently reduced narrow-band continuum data for this field from the MHONGOOSE survey.

3.1. SRC29675 (=S1)

For SRC29675 we searched Simbad for counterparts, given the coordinates of the source and found no known association for this target and in PanSTARRS no counterpart was found. However, the Dark Energy Survey (DES) DR1 catalogue indicates a possible association with J041727.13-551307.7 and the NED catalogue shows a counterpart WISEA J041727.12-551307.6 (far infrared). In Figure 3, SRC29675 is shown as a black circle, is an area close to the optically-selected Quasar sample (blue dots) near the FIRST sensitivity limit. It is likely an AGN.

3.2. SRC28339 (= S2)

For SRC28339 we searched for the existing multi-wavelength catalogues by querying the location in Simbad and no associated source was found. However, in Gaia the source was associated with EDR3 4778660198581705344 and on DES catalog as J042028.54-551047.1. It was confirmed on NED that is a galaxy WISEA J042028.55-551047.2, with magnitude of 19.08 mag in the optical R band (Table 2). Looking at Figure 3, this source is represented by black triangle, in a region dominated by AGN. Therefore, it is safe to say this source is likely to be an AGN.

3.3. SRC28214 (= S3)

For SRC28214 we searched Simbad for a counterpart given the coordinates of the source and found no known association for this target. In PanSTARRS there is no counterpart. However, in DES DR1 there is a very faint counterpart, DES J042044.62-552434.7 at an optical R-band magnitude of 24.4 mag [7]. In the NED catalogue there is a source associated with WISEA J042044.62-5524350 [8] within an 0.2 arcmin search radius. From Figure 3, we mark this source as a black square. It lies somewhat removed from the optically-selected quasar sample towards fainter magnitudes. It is likely an Active Galactic Nuclei (AGN) as it can be traced via the diagonal line (representation of the change in optical and radio flux as a function of distance) to the sample of AGN. The vertical cut-off of the optically-selected quasar distribution is a selection effect representing the completeness limit of the SDSS sample.

3.4. SRC29725 (= S4)

We searched the existing multi-wavelength catalogues by querying the location in Simbad and there was no counterpart. However, in DES DR1 there is a counterpart, J041715.28-5445116.9, with a R-band magnitude of 18.9 mag. This is the brightest optical counterpart to a variable radio source in our sample. From Figure 3, it lies in clearly in the Quasar dominated area.

4. Conclusion

We searched for transients and variable sources in MeerKAT observations of the NGC 1566 fields. We identified four variable sources in this field, however no notable transients were detected. We were able to considerably enhance the known population of variable AGNs in this MeerKAT field and gain a better knowledge of these sources by recognizing the variability of these sources through imaging surveys. Studies of AGNs can establish whether the variability is due to external factors like scintillation or intrinsic factors [9].

Acknowledgements

The authors would like to thank Zwidofhela Khangale for his helpful comments in this paper. We thank Alex Andersson for assisting with the use of TraP. V.M Tshilengo acknowledges support from SARAO postgraduate bursary. This research made use of the Inter-University Institute for Data Intensive Astronomy (IDIA) research cloud for data processing. IDIA is a South African university partnership involving the University of Cape Town, the University of Pretoria, and the University of the Western Cape. The MeerKAT telescope is operated by the South African Radio Astronomy Observatory (SARAO), which is a facility of the National Research Foundation, an agency of the Department of Science and Innovation of South Africa.

References

[1] Carbone, D., Van Der Horst, A. J., Wijers, R. A. M. J., Swinbank, J. D., Rowlinson, A., Broderick, J. W., Cendes, Y. N., Stewart, A. J., Bell, M. E., Breton, R. P., Corbel, S., Eislöffel, J., Fender, R. P., Grießmeier, J.-M., Hessels, J. W. T., Jonker, P., Kramer, M., Law, C. J., Miller-Jones, J. C. A., … Zarka, P. (2016). New methods to constrain the radio transient rate: Results from a survey of four fields with LOFAR. *Monthly Notices of the Royal Astronomical Society*, *459*(3), 3161–3174. https://doi.org/10.1093/mnras/stw539

[2] Varghese, S. S., Obenberger, K. S., Dowell, J., & Taylor, G. B. (2019). Detection of a Low-frequency Cosmic Radio Transient Using Two LWA Stations. *The Astrophysical Journal*, *874*(2), 151. https://doi.org/10.3847/1538-4357/ab07c6

[3] Camilo, F., Scholz, P., Serylak, M., Buchner, S., Merryfield, M., Kaspi, V. M., Archibald, R. F., Bailes, M., Jameson, A., Straten, W. V., Sarkissian, J., Reynolds, J. E., Johnston, S., Hobbs, G., Abbott, T. D., Adam, R. M., Adams, G. B., Alberts, T., Andreas, R., … Zitha, S. (2018). Revival of the Magnetar PSR J1622–4950: Observations with MeerKAT, Parkes, *XMM-Newton* , *Swift* , *Chandra* , and *NuSTAR*. *The Astrophysical Journal*, *856*(2), 180. https://doi.org/10.3847/1538-4357/aab35a

[4] Blok, W., Adams, E., Amram, P., Athanassoula, E., Bagetakos, I., Balkowski, C., Bershady, M., Beswick, R., Bigiel, F., Blyth, S., Bosma, A., Booth, R., Bouchard, A., Brinks, E., Carignan, C., Chemin, L., Combes, F., Conway, J., Elson, E., & Zijlstra, A. (2017). An Overview of the MHONGOOSE Survey: Observing Nearby Galaxies with MeerKAT.

[5] Swinbank, J. D., Staley, T. D., Molenaar, G. J., Rol, E., Rowlinson, A., Scheers, B., Spreeuw, H., Bell, M. E., Broderick, J. W., Carbone, D., Garsden, H., Van Der Horst, A. J., Law, C. J., Wise, M., Breton, R. P., Cendes, Y., Corbel, S., Eislöffel, J., Falcke, H., … Zarka, P. (2015). The LOFAR Transients Pipeline. *Astronomy and Computing*, *11*, 25–48. https://doi.org/10.1016/j.ascom.2015.03.002

[6] Comrie, A., Kuo-Song Wang, Shou-Chieh Hsu, Moraghan, A., Harris, P., Pang, Q., Pińska, A., Cheng-Chin Chiang, Tien-Hao Chang, Hwang, Y.-H., Hengtai Jan, Lin, M.-Y., & Simmonds, R. (2021). CARTA: The Cube Analysis and Rendering Tool for Astronomy (2.0.0) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.3377984

[7] Stewart, A. J., Muñoz-Darias, T., Fender, R. P., & Pietka, M. (2018). On the optical counterparts of radio transients and variables. *Monthly Notices of the Royal Astronomical Society*, *479*(2), 2481–2504.

[8] Cutri, R. M., Wright, E. L., Conrow, T., Fowler, J. W., Eisenhardt, P. R. M., Grillmair, C., Kirkpatrick, J. D., Masci, F., McCallon, H. L., Wheelock, S. L., Fajardo-Acosta, S., Yan, L., Benford, D., Harbut, M., Jarrett, T., Lake, S., Leisawitz, D., Ressler, M. E., Stanford, S. A., … Wittman, M. (2013). Explanatory Supplement to the AllWISE Data Release Products. In *Explanatory Supplement to the AllWISE Data Release Products* (p. 1). https://ui.adsabs.harvard.edu/abs/2013wise.rept. 1C

[9] Reeves, S. N., Sadler, E. M., Allison, J. R., Koribalski, B. S., Curran, S. J., Pracy, M. B., Phillips, C. J., Bignall, H. E., & Reynolds, C. (2016). H i emission and absorption in nearby, gas-rich galaxies – II. Sample completion and detection of intervening absorption in NGC 5156. *Monthly Notices of the Royal Astronomical Society*, *457*(3), 2613–2641. https://doi.org/10.1093/mnras/stv3011

[10] Rowlinson, A., Meijn, J., Bright, J., van der Horst, A. J., Chastain, S., Fijma, S., Fender, R., Heywood, I., Wijers, R. A. M. J., Woudt, P. A., Andersson, A., Sivakoff, G. R., Tremou, E., & Driessen, L. N. (2022). Search and identification of transient and variable radio sources using MeerKAT observations: A case study on the MAXI J1820+070 field. *Monthly Notices of the Royal Astronomical Society*, *517*(2), 2894–2911. https://doi.org/10.1093/mnras/stac2460

Application of Data-Driven Deep Learning Hybrid Models for Forecasting of Atmospheric Tides Measured by a SuperDARN HF Radar

Nkanyiso Mbatha[1], Zanele Bhengu[1], and Venkataraman Sivakumar [2,3]

[1]University of Zululand, KwaDlangezwa, 3886, South Africa

[2]University of KwaZulu Natal, Discipline of Physics, Private Bag x 54001, Westville Campus, Durban, 4000

[3]National Institute for Theoretical and Computational Sciences, University of KwaZulu Natal, Westville Campus, Durban, 4000

E-mail: **mbathanb@unizulu.ac.za**

Abstract. In this research, mesosphere-lower thermosphere (MLT) tidal waves (Diurnal Tides (DT) and Semidiurnal Tides (SDT)) were extracted from the SuperDARN HF radar observations of meteor echoes at Kerguelen (49.22°S, 70.14°W), Antarctica using the Short-Time Fourier Transform (STFT) for the year 2008 data. Here, we have developed a robust data-driven hybrid time series-forecasting model (named the Empirical Wavelet Transform (EWT) long short-term memory (LSTM) network (EWT-LSTM) and Empirical Mode Decomposition LSTM (EMD-LSTM)). These models are compared to a standalone LSTM and a Wavelet Denoising WD-LSTM. The extracted tidal wave signals from HF Radar data were divided into 70% for training model and remaining 30% to test the simulations. In terms of the power spectra, significant peaks of both the DT and SDT were observed. The performance of the model was assessed using the testing data set after models had been trained using training data sets. Results showed that the hybrid LSTM model based on the EWT outperforms other models and is followed by the EMD-LSTM. This study demonstrates that the performance of the deep learning LSTM model is improved by preprocessing the signal using signal decomposition techniques like EWT and EMD before pushing it into the LSTM.

1. Introduction

The Kerguelen HF radar (49.22°S, 70.14°W), is part of the global Super Dual Auroral Radar Network (SuperDARN) [1]. The initial design of these HF radars was for the detection of plasma irregularities in the E and F regions of the ionosphere, and hence the analysis of the behavior of plasma in the auroral oval and the polar cap. However, these HF radars also indicated a capability to detect coherent echoes at the closest range gates which were termed “grainy-near-range echoes” [2]. The meridional wind is found to be dominated by the semidiurnal tide most of the year, with maxima in spring and autumn. The study by Hall et al., [2] also reported a significant presence of 24-hour (diurnal) and 8-hour tides in the meridional wind component of the SuperDARN radar. Thus, the SuperDARN radar meridional wind is able to capture these tidal waves, but it is primarily known to be dominated by semidiurnal tides most of the year [2,3].

The strategic location of SuperDARN radars makes them susceptible to intense storms that can cause heavy icing on antennas and hence the malfunctioning of the HF radar. Thus, with the new developments in the field of deep learning for time series forecasting, there is a need to explore new pattern-finding techniques to better estimate missing data on MLT velocities and wave dynamics. Thus, the study aims to utilize the meridional neutral wind data measured by the Kerguelen SuperDARN HF to extract diurnal and semidiurnal tidal wave time series. Data-driven hybrid time series forecasting models that are based on the decomposition of time series are thereafter used to forecast the tidal wave signal.

2. Method and Data

The data used in this study is from the Kerguelen HF radar, which is part of the SuperDARN HF radar network (http://psddb.nerc-as.ac.uk/data/access/coverage.php?menu=4&source=1&class=20&script=1) located both in the Southern and Northern Hemispheres (See Figure 1). This network consists of more than 30 low-power HF radars that are designed to face up into the Earth's upper atmosphere. The geographical locations of these radars begin at mid-latitudes and extend into the polar regions. SuperDARN HF radars operate continuously and observe the motion of charged particles (plasma) in the ionosphere and other effects that provide important information about the Earth's space environment. In terms of their operating principles, SuperDARN radars are coherent scatter radars where the ionosphere's signal is analogous to the Bragg scatter of X-rays in crystals [1,4]. Although it is possible for the transmitted signal to be dispersed in different directions, in this design, the signal of primary importance is that which is returned along the same path as the transmitted signal, allowing radars to detect direct backscatter. The scale size of the irregularities from which the signal is scattered is closely correlated with the half-wavelength of the transmitted signal, primarily due to the nature of Bragg scatter.

In general, SuperDARN radars are designed to operate in the HF band of the radio spectrum between 8 and 20 MHz [1] however, most radars can be tuned to operate over a narrower range of frequencies, typically between 10 and 14 MHz [4]. Of importance to this research is that the ionization that is created in the atmosphere by meteors can also lead to backscattered signals in SuperDARN data [2,3]. Since the meteor trail height normally runs between 85 and 120 km, this backscattered signal is typically detected at close ranges [2,5]. The detection of these coherent echoes at the closest range gates which were termed "grainy-near-range echoes" [2] makes it possible to study both the zonal wind and meridional winds with a mean height of ~94 km in the Mesosphere-Lower Thermosphere region of the atmosphere. Thus, in this study, tidal wave activity is extracted from the meridional wind component of what is known to be SuperDARN radar neutral wind data.

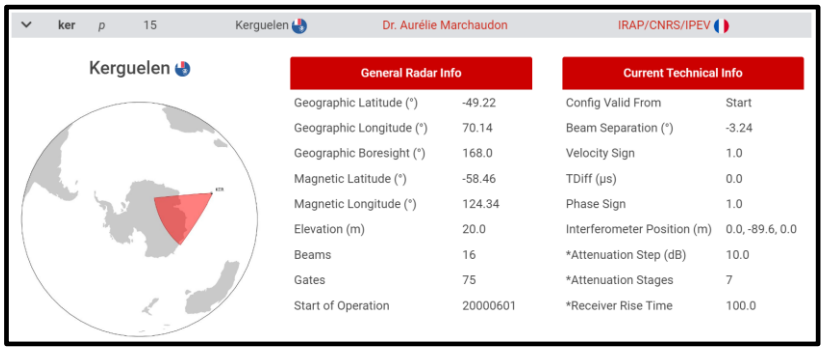
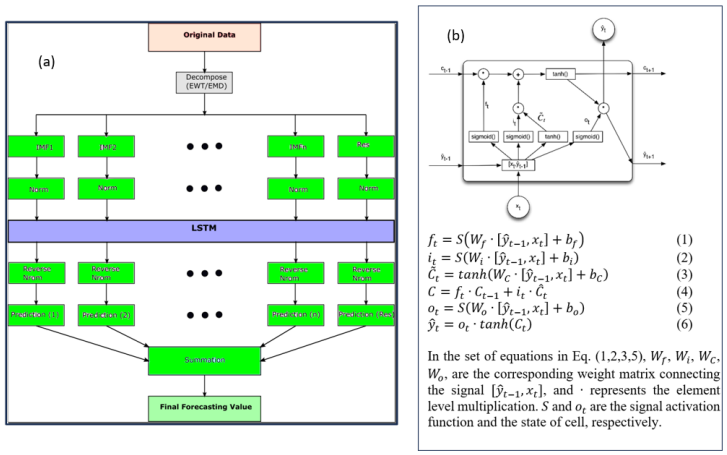


Figure 1. The technical specification of SuperDARN radar at Kerguelen (49.22 S, 70.14 E) (https://superdarn.ca/pydarn).

The tidal wave (semidiurnal and diurnal tides) activity was extracted by transforming the meridional wind time series to a frequency domain using the Fourier Transform (FT) method. Because of the need to retrieve both the frequency and time-space of these tides, the Short-Time Fourier Transform (STFT) method is found to be a better option to use in this study. The STFT method is based on windowing the time series into smaller segments that are assumed to be stationary. The FT is then applied to the partitioned segments while systematically advancing the window in time along the full length of the time series. To extract tides, the wind velocity used was a 4-day equivalent of hourly meridional wind data, which is advanced by 1-day (24 hours of data points) at a time. A Fourier transform is then applied to each of the 4-day stationary windows, creating what is called a dynamic Fourier spectrum. In this dynamic Fourier spectra, a 12-hour cycle and a 24-hour cycle are selected, representing the semidiurnal and diurnal tides, respectively.

In this study, four models are developed and used for forecasting the tidal wave signal. One is a standalone LSTM model (LSTM), and the other three models are hybrid LSTM models. These models are said to be hybrid because they use signal preprocessing techniques such as the empirical wavelet transform (EWT) (EWT-LSTM model), the empirical mode decomposition (EMD) (EMD-LSTM model) decomposition step, and the wavelet denoising (WD-LSTM). After preprocessing the tidal wave signal, the STFT power spectra of tidal waves time series are then divided into 70% training and 30% testing data sets, where the training data set is used as training input data for the data-driven forecasting models based on long short-term memory networks (LSTM). A block diagram showing the architecture of the model design is shown in Figure 2(a).

The LSTM is a type of recurrent neural network (RNN) architecture that is designed to address the limitations of traditional RNNs when dealing with long-term dependencies in sequential data. It was introduced by Sepp Hochreiter and J¨urgen Schmidhuber in 1997 [6]. It is a family of RNNs that is often used with deep neural networks. A single cell structure of the LSTM is made up of three gates: (1) a forget gate, which controls if/when the context is forgotten; (2) an input gate, which controls if/when a value should be remembered by the context; and (3) an output gate, which controls if/when the remembered value is allowed to pass from the unit. During the training process, this LSTM's unique structure is capable of successfully addressing the issues of gradient loss and gradient explosion. More details about the LSTM can be found in a study by Hochreiter et al., [6]. The structure of a single LSTM cell is shown in Figure 2(b). Figure 2b shows the single cells of the LSTM and all its 3 gates which include the input gate (i_t), forget gate (f_t) and the output gate (o_i). Equations of the gates, cell state, candidate cell states and the final output are also shown. A detailed operation procedure of each cell of the LSTM is given in a study by Hochreiter [6].



$$f_t = S(W_f \cdot [\hat{y}_{t-1}, x_t] + b_f) \quad (1)$$
$$i_t = S(W_i \cdot [\hat{y}_{t-1}, x_t] + b_i) \quad (2)$$
$$\hat{C}_t = tanh(W_C \cdot [\hat{y}_{t-1}, x_t] + b_C) \quad (3)$$
$$C = f_t \cdot C_{t-1} + i_t \cdot \hat{C}_t \quad (4)$$
$$o_t = S(W_o \cdot [\hat{y}_{t-1}, x_t] + b_o) \quad (5)$$
$$\hat{y}_t = o_t \cdot tanh(C_t) \quad (6)$$

In the set of equations in Eq. (1,2,3,5), W_f, W_i, W_C, W_o, are the corresponding weight matrix connecting the signal $[\hat{y}_{t-1}, x_t]$, and $\cdot$ represents the element level multiplication. S and o_t are the signal activation function and the state of cell, respectively.

Figure 2. (a) The architecture flow of the model and (b) The structural representation of the long short-term memory (LSTM) unit.

In terms of the signal decomposition methods used here, the Empirical Wavelet Transform (EWT), which was introduced by Gilles [7], decomposes a signal or an image on wavelet-tight frames that are built adaptively. In 1D space, like the time series that is used in this study, the procedure consists of detecting the supports of some intrinsic mode functions (IMFs) in the Fourier spectrum and then using these supports to build Littlewood-Paley-type wavelets [7]. On the other hand, empirical mode decomposition (EMD) is an adaptive, intuitive, direct, and highly efficient decomposition method that has been described in detail by [8]. In this method, the signal is decomposed into a nearly orthogonal combination of simple time series [9]. These elements are also known as residuals and IMFs. The time series are analyzed using the EMD methodology while still being in the time domain.

The predictive performance models designed in this study are tested using model performance statistics. A Taylor diagram [10] and statistical evaluation metrics such as mean Absolute error (MAE), correlation coefficient (R), root means squad error (RMSE), and standard deviation (STD) will be utilized.

3. Results and Discussion

As mentioned above, the time series of both the SDT and DT was obtained by taking the STFT of the Kerguelen HF meridional wind component for the period January to December 2008. The dynamic Fourier spectra are shown in Figure 3(a), and the time series of the DT and SDT, respectively, are shown in Figure 3(b). The power spectrum was normalized by σ^2/N, here N is the number of data points and σ^2 is the variance of the time series. In Figure 2(b), the horizontal red line indicates the 95% confidence level using a chi-square test, assuming white noise as the background spectrum [11]. In terms of tidal wave activity, there are significant peaks in both the DT and the SDT all through the year. Similar to what was reported by Hibbins et al., [12] a maxima in the SDT component is observed in both spring and autumn. Moreover, the power spectra also indicate a burst of activity for quasi-two-day waves (QTDW) at the beginning of the year.

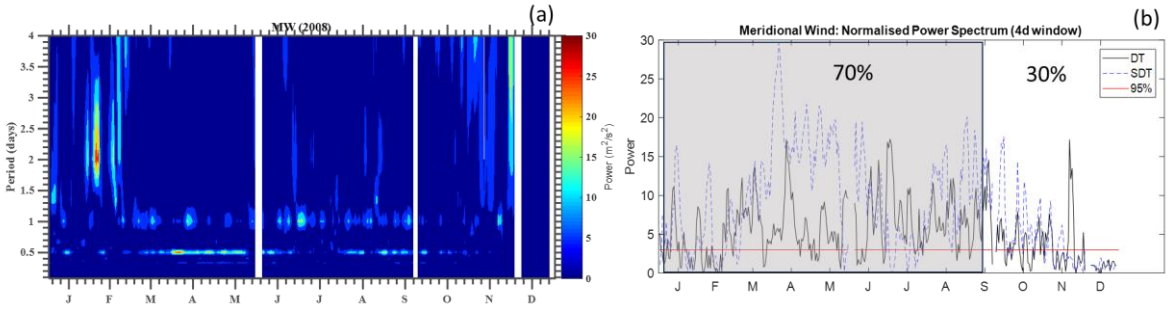


Figure 3. Normalised power spectra (b) of hourly meridional wind for year 2002, and (b) the time series of the extracted diurnal tide (DT) and semidiurnal tide (STD) power spectra.

In this research, we employ four different forecasting models on the DT and SDT time series shown in Figure 3(b). All these models are based on the LSTM, and the LSTM system used here is a Python TensorFlow (https://www.tensorflow.org/learn) LSTM system. The other model is the WD-LSTM, which is based on the signal denoising using the Daubechies (usually termed as db8) wavelet family in the PyWavelets (https://pywavelets.readthedocs.io/en/latest/) wavelet transform software for Python. The final two models are based on denoising the signal using EMD (EMD-LSTM) and EWT (EWT-LSTM), respectively. The architecture of the data-driven LSTM models used here is similar to that of Mbatha and Bencherif [13]. As shown in Figure 3(b), for the training and testing of models, the DT and SDT signals were divided into 70% training (grey shaded area) and 30% testing (white area).

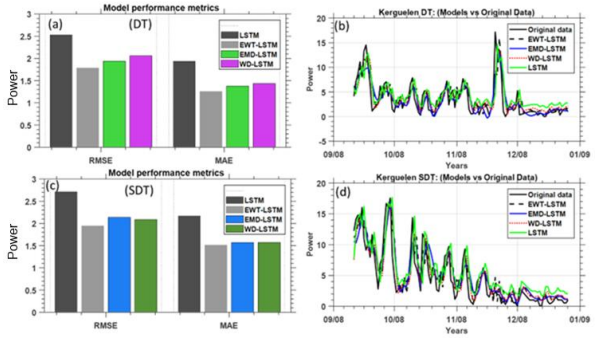


Figure 4. Model Performance statistic and the testing data sets and model forecasts for DT and SDT.

Figure 4 (b and d) depicts the actual testing datasets (black line) and their forecasts (green line: LSTM, black dashed line: LSTM, red dotted line: WD-LSTM, blue line: EMD-LSTM, and black dashed line: EWT-LSTM) for DT (b) and SDT (d) data. The corresponding RMSE and MAE statistics are shown in Figure 4 (a and c). In general, it is notable that the four models can significantly track the original data. But the performance statistics shown in Figure 4 (a and c) indicate that EWT-LSTM outperforms other models, with relatively low RMSE and MAE values for DT and SDT signals, respectively. For a more comprehensive presentation, the four data-driven predictive models were examined using a graphical demonstration of the Taylor diagram, as displayed in Figure 5 for DT (a) and SDT (b), respectively. It should be noted that there seems to be a prediction delay that is observed in both Figure 4(b) and (d). This is associated with the random nature of the signal introduced by the fact that the RNN LSTM always need to update the new information by using the previous or old information that it stored.

The Taylor diagram indicates two statistical metrics (i.e., correlation coefficients and standard deviations of each model), which are used to quantify the comparability between models and the actual data. The distance from the reference point (observation) is a measure of the RMSE. It is noted that only the non-negative correlation (which is above 0.8 for all models) is shown, which indicates that all

models' patterns follow those of the original data. Based on the Taylor diagram, it is found that the point of the EWT-LSTM is closest to the observed point (shaded red dot), followed by the EMD-LSTM and the WD-LSTM. This indicates the higher ability of LSTM models, which are based on the signal decomposition step. The EWT method of signal decomposition seems to outperform the EMD presumably [7,13].

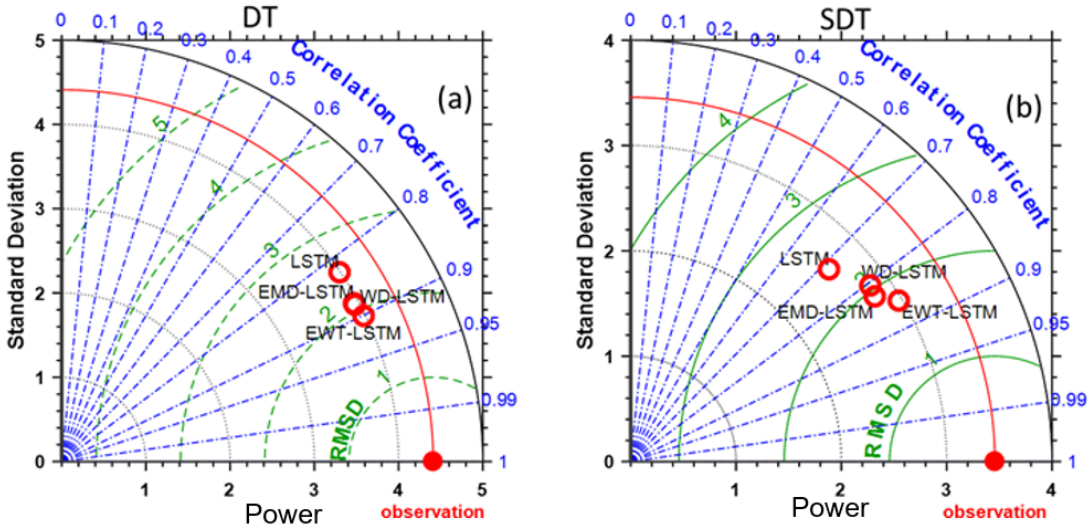


Figure 5. Taylor diagram graphical representation of 4 predictive models developed for forecasting DT (a) and STD (b) time series.

4. Conclusion

The main objectives of this study were to extract the time series of power spectra of the DT and the SDT and apply the relatively new deep learning model (LSTM) that is based on signal processing methods like WD, EWT, and EMD to simulate daily observations of these tidal waves in the MLT region. The performance of these models is evaluated using several statistical metrics namely the correlation coefficient, root mean squared error (RMSE), mean squared error (MAE), and the standard deviation. In general, the LSTM model shows a good memory ability in forecasting the atmospheric tides variability. The sensitivity of the LSTM is improved significantly by preprocessing the signal using WD, EWT, and EMD methods. The EWT method seems to produce the best hybrid LSTM model. A study by Gilles [7] compared the performance of the EMD and EWT signal decomposition methods. This study reported that EMD is a useful decomposition method even though it lacks mathematical theory. Thus, this study introduced the EWT as a better approach for signal decomposition because of its well-defined mathematical theory. EWT method uses Fourier transforms theory to produce EWT frequency domain segmentation of the time series. On the other hand, the EMD is a simplified method that extract intrinsic mode functions through identifying local extrema between two successive zero crossings of the signal and joining all maxima by cubic spline line considered as upper envelop and then calculating the mean of the upper and lower envelope for successive IMFs all the way to the residue.

References

[1] Greenwald, R. A., et al., DARN/SuperDARN: A global view of the dynamics of high-latitude convection, *Space Sci. Rev*., 71, 761–793, 1995.

[2] Hall, G. E., J. W. MacDougall, D. R. Moorcroft, J.-P. St.-Maurice, A. H. Manson, C. E. Meek, Super Dual Auroral Radar Network observations of meteor echoes, *J. Geophys. Res*., 102, 14,603–14,614, 1997

[3] Jenkins B, Jarvis MJ, Forbes DM. Mesospheric wind observations derived from Super Dual Auroral Radar Network (SuperDARN) HF radar meteor echoes at Halley, Antarctica: preliminary results. *Radio Science*. 1998 Jul;33(4):957-65

[4] Nishitani, N., Ruohoniemi, J.M., Lester, M., Baker, J.B.H., Koustov, A.V., Shepherd, S.G., Chisham, G., Hori, T., Thomas, E.G., Makarevich, R.A. and Marchaudon, A., 2019. Review of the accomplishments of mid-latitude Super Dual Auroral Radar Network (SuperDARN) HF radars. *Progress in Earth and Planetary Science*, 6(1), pp.1-57.

[5] Chisham, G. and Freeman, M.P., 2013. A reassessment of SuperDARN meteor echoes from the upper mesosphere and lower thermosphere. *Journal of Atmospheric and Solar-Terrestrial Physics*, *102*, pp.207-221.

[6] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780

[7] Gilles, J. Empirical wavelet transform. *IEEE Trans. Signal. Process*. 2013, 61, 3999–4010

[8] Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C. and Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London. Series A: *mathematical, physical and engineering sciences*, 454(1971), pp.903-995.

[9] Moore, K.J., Kurt, M., Eriten, M., McFarland, D.M., Bergman, L.A. and Vakakis, A.F., 2018. Wavelet-bounded empirical mode decomposition for measured time series analysis. *Mechanical Systems and Signal Processing*, 99, pp.14-29

[10] Taylor, K.E. Summarizing multiple aspects of model performance in a single diagram. J. Geophys. Res: Atmos. 2001, 106, 7183–7192.

[11] Torrence, C. and Compo, G.P., 1998. A practical guide to wavelet analysis. *Bulletin of the American Meteorological society*, 79(1), pp.61-78.

[12] Hibbins, R.E., Espy, P.J., Orsolini, Y.J., Limpasuvan, V. and Barnes, R.J., 2019. SuperDARN observations of semidiurnal tidal variability in the MLT and the response to sudden stratospheric warming events. *Journal of Geophysical Research: Atmospheres*, *124*(9), pp.4862-4872.

[13] Mbatha, N. and Bencherif, H., 2020. Time series analysis and forecasting using a novel hybrid LSTM data-driven model based on empirical wavelet transform applied to total column of ozone at buenos aires, argentina (1966–2017). *Atmosphere*, *11*(5), p.457.

DIVISION

E

PHYSICS FOR DEVELOPMENT, EDUCATION, AND OUTREACH

Evaluating essential skills

D B Fish

Unizulu Science Centre, University of Zululand, KwaDlangezwa, South Africa

E-mail: **thefish@iafrica.com**

Abstract. Operating from the University of Zululand, Unizulu Science Centre (USC) had been running face to face matric workshops for 25 years, presenting practicals and sharpening skills for over 200 000 matric science students. The 2020 lockdown presented a dilemma: matrics needed assistance more than ever, but schools were closed and large gatherings impossible. USC piloted a unique offline, digital video project (Essential Skills for Matric Science – ES) which was reported on at the 2021 SAIP conference. USC worked to convert a 4 hour contact workshop into first 8, then later 16, one-hour videos, highlighting the essential skills for Matric Physical Sciences Paper 1 – the physics paper. While these videos were made available on the internet (via the SAIP website) for download or streaming, they were physically distributed on memory sticks to teachers, along with an accompanying workbook. Local industry funding initially and then further SAIP funding (with support from Allan Gray) saw 100 000 booklets printed and 2000 memory sticks manufactured over the next three years. These have been distributed to schools in all nine provinces and also used as the basis for teacher training, especially by the SAIP. The framework for ES has already been used to generate a similar tool for Life Science and Chemistry. Videos and booklets for Mathematics are under development. Looking back over three years, it is essential to try to measure the impact of this project and how it can be made more effective for the future. A simple evaluation tool was used with key stakeholders who had been very involved in ES during this time, to gather their inputs on the ES project, how well it had worked and what could be improved. A synthesis of their responses will be presented to inform an evaluation of what has been achieved and what can be achieved in future. In particular the transition of ES from a distance only model (during lockdown) to a hybrid one (after lockdown) will be discussed.

1. Introduction

Unizulu Science Centre (USC) was established in November 1986 on the main Unizulu campus. It moved to its current location in Richards Bay in 1996 and serves the three (out of 12) northern educational districts of KwaZulu-Natal, comprising about 550 secondary schools. One of its main outreach programmes has been matric workshops which have been presented face to face to over 200 000 pupils over the last 25 years. The 2020 lockdown presented a dilemma: matrics needed assistance more than ever, but schools were closed and large gatherings impossible; so USC piloted a unique offline, digital video project (Essential Skills for Matric Science – ES) which was reported on at the 2021 SAIP conference. A 4 hour contact workshop was converted into first 8, then later 16, one-hour videos, highlighting the essential skills for Matric Physical Sciences Paper 1 – the physics paper. A firm partnership with the SAIP saw these videos made available on the internet for download or streaming, but they were mostly physically distributed on memory sticks to teachers, along with an accompanying workbook. Local industry funding initially (from South32) and then further SAIP funding (with support from Allan Gray) saw 100 000 booklets printed and 2000 memory sticks manufactured and distributed over the next three years.

During the COVID lockdown the production and distribution of resources was prioritised as matric needs were critical. Looking back now over these three years (2020 - 2022), it is essential to try to measure the impact of this project and how it can be made more effective for the future.

2. Review of Research into Free Choice (or Informal) Learning Evaluation

The "Physics Education" track has changed over the years and now covers "Physics for Development, Education and Outreach." While Physics Education Research (PER) in a formal setting has been developed over many years and has been discussed extensively in this group, development and outreach take place outside the classroom and cannot be evaluated in the same way. ES is a good example of an outreach project which, although covering the formal curriculum, happens in an out of school setting thus requiring unique evaluation tools.

Falk and Dierking [1] stress that: "Learning in Museums (Science Centers and outreach programmes) is not the same as the learning expected and assessed in schools. Measuring learning in and from museums requires different tools and different assumptions." Brody et al [2] suggest that: "If we take learning to include outcomes such as knowledge, understanding, attitudes, beliefs, values as well as interpersonal and social skills, then the opportunities afforded for assessment are substantial." Thus free choice learning evaluation can include more than just an evaluation of learning gains and offers rich opportunities for the promotion of physics education.

Johnson, and Majewska [3] further elucidate the differences between formal, informal and non-formal learning and discuss the research methods appropriate for each sector. With free-choice or informal learning, it is easier to take a constructivist [4] approach, acknowledging pupils' prior knowledge and socio-economic backgrounds [5] which can also be tested in the evaluation more easily than in formal education. Further useful material is available in two National Research Council studies from 2009 [6] and 2012 [7].

ES makes extensive use of PhET simulations and PASCO experimental equipment. Many studies have been done around the effectiveness of PhET simulations in instruction [8] and [9] in different contexts and locations. There have also been studies done around the use of PASCO interfaces, equipment and software [10], but these have mostly been at university, rather than at school, level. They have also been conducted mostly in more developed countries. It is hoped that this study will be able to contribute further to this literature and especially comment on the effectiveness of incorporating both PhET and PASCO resources into instructional materials – especially at secondary school level and in an African context.

Finally, Stocklmayer and Gilbert [11] argue for: "greater complementarity between formal school science education and the opportunities offered for science learning that are available outside of school, in what we will call the 'informal sector'." suggesting that ES forms a good bridge between the formal and informal education sectors.

3. Methodology

An evaluation tool was developed to be used with key stakeholders who had been very involved in ES during this time, to gather their inputs on the ES project, how well it had worked and what could be improved. A simple, one-page questionnaire (attached as an appendix) with a combination of qualitative and quantitative questions was sent via email (as the author was in the USA) to 40 stakeholders in ES in May 2023. Responses were received from 20 people. A synthesis of their responses is presented below to inform an evaluation of what has been achieved and what can be achieved in future. With the huge distribution of ES to all 9 provinces and 100 000 pupils it was not feasible to try to do a detailed follow up. The results from this limited set of stakeholders are intended to be indicative, rather than prescriptive.

4. Results

The questionnaire is attached as an appendix or can be requested from the author. Responses will be summarised and grouped in logical sections.

Question 1 asked in what role the respondent had encountered ES. Of the 20 responses, most (9) came from facilitators of some kind and 4 from South African Institute of Physics (SAIP) facilitators, with 3 from Science Centre staff and 2 from teachers. The remainder was made up of a translator (into Afrikaans) a science advisor with the Education Department and an industry representative. While there was a broad spectrum of respondents, it would have been pleasing to have more responses from teachers who actually employed the materials, rather than just from facilitators who mostly distributed them.

In response to question 2: "in which years did you use the resources?" the following results emerged:

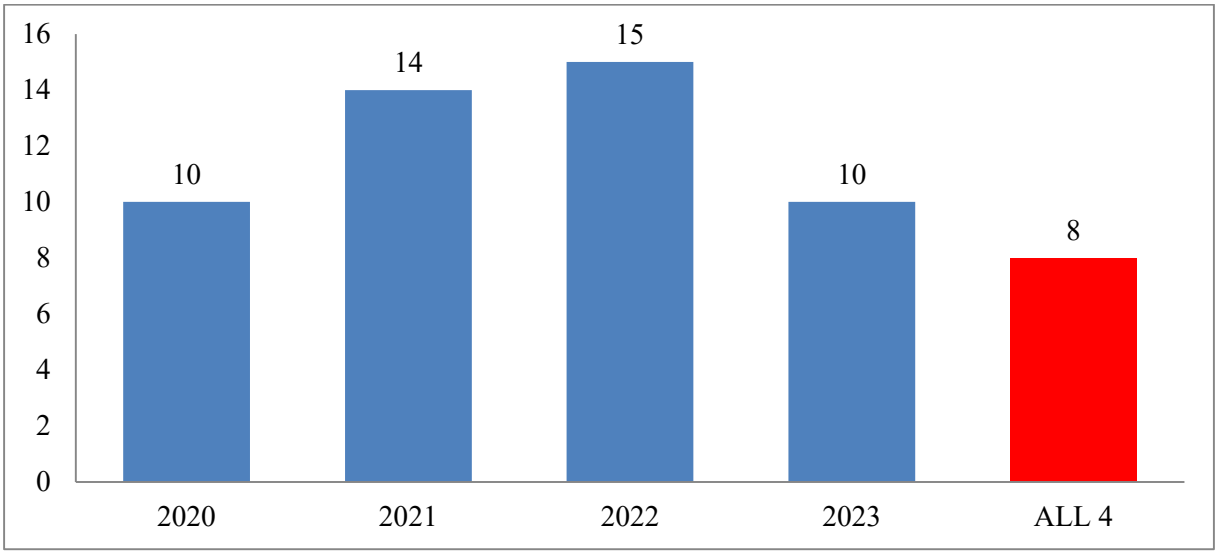


Figure 1: Graph of number of respondents using ES in a particular year.

While it is understood that some respondents only learnt of the resources after the first year, it is pleasing that some had used them for all 4 years, during which time the resources increased and improved.

Question 4 asked respondents how useful they had found the ES materials on a Likert Scale from 1 (not helpful) to 9 (extremely helpful). All answers were either 8's or 9's so the general response (understanding that Likert Scale responses should not be simply averaged) would have fallen between 8 and 9. This is very pleasing especially when taken with the more detailed responses to question 10.

Subsequent questions interrogated the contents of ES. Question 7 asked which of the four essential skills presented (graphs, free-body diagrams, equations and multiple choice questions) were most useful, generating the results in Fig. 2.

Before this workshop took an "essential skills" approach, it started as a "graphs" workshop, so it was pleasing to see that this skill stood out as being considered the most important. Most respondents commented that all sections were extremely useful, especially as skills rather than content was emphasised, addressing pervasive gaps. A sample comment follows:

> "Each school/teachers have their different strong or weak sections of the physics, I would always ask teachers where do they lack and use ES to help them to teach their specific problem section."

On being asked if any of the four skills were considered less important, 19 respondents said "no" and one said perhaps free-body diagrams and multiple choice questions. The survey asked for suggestions of other important skills, which was mostly left unanswered but a few useful suggestions were made, including: converting units, dimensional analysis and data analysis.

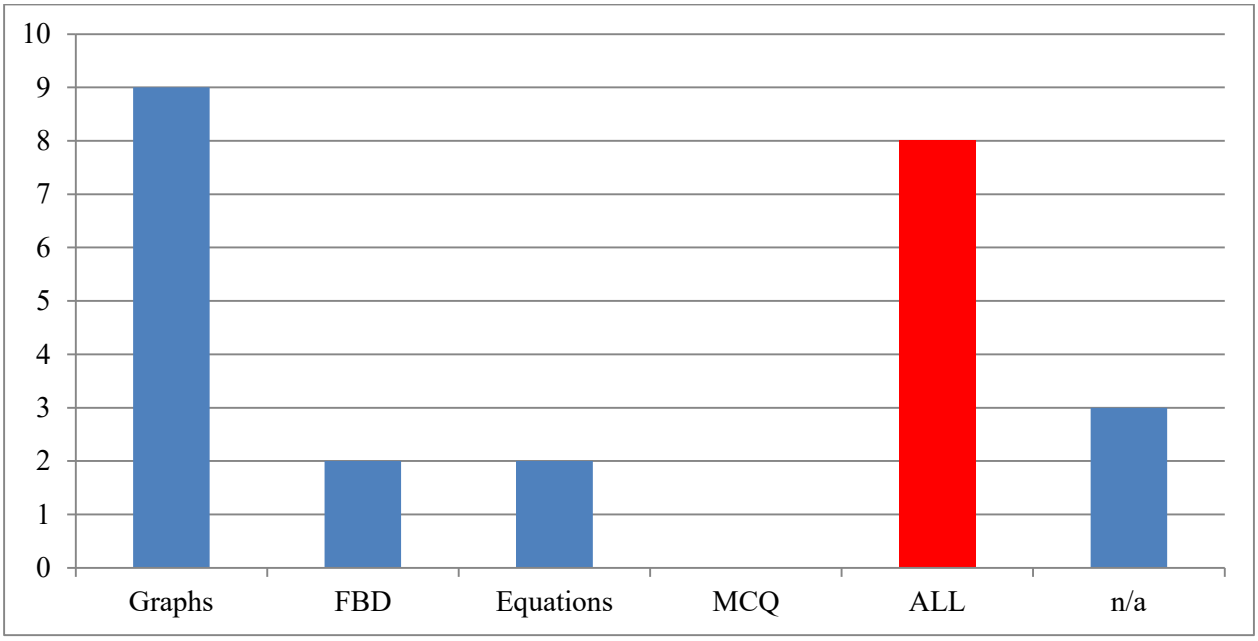


Figure 2: Graph of number of respondents choosing each skill as most useful

Each instructional video ended with an optional short, 5 minute, interview with various people in physics and industry. Question 6 asked if these had been used, which they mostly had, but more pleasing was the comments about these videos:

- the career videos inspired learners, who can "see themselves" (or their "mirror self") in the interviews,
- videos are extremely informative and there is no career guidance at schools;
- [pupils] were encouraged to take science as a career and change to taking science stream
- I have used them mostly with the grade 10 early in the year to pursue them to take science and pure maths.

It was encouraging to see that these interviews appear to have succeeded in adding a human dimension to the science and making it more relevant to the lives of pupils.

Finally the survey asked in question 8: "In what practical ways could ES be improved or made more useful?" Most respondents answered: "none", but some useful suggestions were offered, many of which have already been implemented:

- send out earlier in the year (implemented in 2023)
- make a Grade 10 resource (in progress)
- get feedback from teachers and students (like this survey)
- make a section for teacher development (SAIP is assisting with this)
- more simulations and experiments (we are working with PhET and PASCO on this)

On being asked in question 9 which other subjects would benefit from the ES approach, respondents answered maths (8), chemistry (7) and biology (4). Geography and accounting were each mentioned once. This confirms our development work which has already completed ES programmes for biology (2021) and chemistry (2023) and maths is in process.

A pressing question (10) is whether ES is still needed now that the COVID lockdown is over and face to face teaching has been resumed. All respondents answered "yes" to this question and gave reasons which were encouraging. They mentioned that there are still not enough teachers or that teachers are missing. Furthermore that we need to supplement teaching and provide digital resources for the 4th Industrial Revolution. That ES addresses pervasive knowledge gaps (for pupils and teachers) and puts particular emphasis on skills – which get higher cognitive learning. Finally the comment that students still ask for the resources. There are advantages and disadvantages to face to face, distance and hybrid

teaching models which have become even more apparent during the years of COVID. It was encouraging to see that ES could work well in multiple settings, ensuring its prolonged use after COVID was over.

5. Limitations and future study

We are well aware that this study was limited in many ways: in terms of geographical area, scope of work (only ES for physics) and number of respondents. A larger scale survey, especially one including more teachers and some pupils would yield new and valuable insights into ES.

ES provides many opportunities for future study and there is already a huge dataset available spread throughout South Africa. While many studies have been done on digital or online materials overseas there are very few in Africa, suggesting possibilities for exploring new areas. Further research together with PhET (simulations) and PASCO (experimental interfaces and data-logging) would provide valuable feedback to their efforts as well.

6. Conclusion

The survey results helped to confirm that ES has been a success. It has assisted many thousands of pupils and teachers throughout SA. Indeed, a project which South32 helped us to start in 2020 in order to serve just the 5000 matrics of the King Cetshawyo District in KwaZulu-Natal has endured for four years and now been distributed to over 100 000 pupils and their teachers in all 9 provinces. What began as a crisis response to COVID continues to be relevant and helpful after the pandemic has ended. The ES programme has been widely reported on all over SA and at at least 2 international conferences. The initial Physics materials have now been expanded covering Biology and Chemistry and work has begun on Mathematics resources. The Physics booklet has been translated into Afrikaans and work has begun on an isiZulu translation.

Furthermore, questions asking for suggestions for further development mostly confirmed the direction which we have already taken. Besides this confirmation, the survey yielded valuable suggestions and insights which can only improve and expand ES in the future.

References

[1] Falk J H and Dierking L D 2016 *The museum experience revisited*. Routledge.
[2] Brody M, Bangert A and Dillon J 2008 Assessing Learning in Informal Science Contexts: *Commissioned paper by the National Research Council for Science Learning in Informal Environments Committee*
[3] Johnson M and Majewska D 2022 Formal, non-formal, and informal learning: What are they, What are they, and how can we research them? *Cambridge University Press & Assessment Research Report*
[4] Driver R 1989 *International Journal of Science Education* **11** 481
[5] Fish D, Allie S, Pelaez N and Anderson T 2017 *Public Understanding of Science* **26(7)** 806
[6] Learning Science in Informal Settings 2009: https://nap.nationalacademies.org/catalog/12190/learning-science-in-informal-environments-people-places-and-pursuits
[7] Identifying and Supporting Productive STEM Programs in Out-of-School Settings 2012 https://nap.nationalacademies.org/catalog/21740/identifying-and-supporting-productive-stem-programs-in-out-of-school-settings
[8] Haryadi R and Pujiastuti H 2020 *Journal of Physics: Conference Series* **1521** 022017
[9] Wieman, C E, Adams W K and Perkins K K 2008 *Science* **322(5902)** 682
[10] Shakur A and Connor R 2018 *The Physics Teacher* **56(3)** 152
[11] Stocklmayer S M, Rennie L J and Gilbert J K 2010 *Studies in Science Education* **46:1** 1

Appendix: Questionnaire

SAIP QUESTIONS ABOUT ESSENTIAL SKILLS (ES – the programme of videos and booklets for matric science)

Three short functional/ technical questions:
1) In what role did you encounter ES? - Teacher, student, SAIP facilitator, advisor, other . . . .?
2) In which year/s did you use the resources? (2020, 2021, 2022 and/or 2023)
3) Describe how you made use of the resources - ie. Showed videos on data projector, TV, laptop? Sound?

Longer questions:
4) How helpful did you find the ES resources: (Please circle a number from 1 – 9 below)
NOT HELPFUL: 1 2 3 4 5 (AVE.) 6 7 8 9 EXTREMELY HELPFUL
5) The ES materials comprise 16 separate videos, each about 1 hour long. How many of these did you complete with a single class/ in a single year? (Or just state NOT APPLICABLE if you just distributed ES resources)
6) Did you show the short interviews/ careers videos at the end of each section? What was the response to that?
7) The course highlights four Essential Skills for Physics: Graphs, Free Body Diagrams, Equations and MCQ
- Which of these sections/ skills did you find *most* useful/ relevant? Why?
- Do you think some of these which are included are less important?
- Are there others (not included) which you feel are important?

8) In what practical ways could ES be improved or made more useful for you?
9) In which other subjects would you like to have access to an ES programme?
10) Is there still a need for ES now that we can teach face to face after COVID? Why?

Surreptitiously feeding education theory to physics students

Jennifer Williams
Rhodes University, Makhanda, South Africa
E-mail: Jennifer.Williams@ru.ac.za

Abstract. Physics is acknowledged to be one of the academic disciplines that students find the hardest, certainly to master but even simply to succeed in. Reasons proposed for this are varied but often incorporate the necessity of working in both the abstract (mathematical) and concrete (physical) realms and the need to transfer knowledge between these realms, while, at the same time, expending great cognitive effort to make sense of the real, physical world in many, often seemingly different and unrelated, contexts. I have found the education theory of LCT (Legitimation Code Theory) and its concept of semantic waves to be a powerful explanatory device for this process. I briefly introduce this theory and then give examples of the ways in which I make the translations between these realms explicit to physics students during lectures.

1. Introduction

This journey began for me when a student, I will call him "Anthony", who was in the Waves 201 class that I was lecturing, commented to me: "The problem with this course is that what the physics and the equations are telling me is different to what I think ought to be happening, my common-sense view!" This comment encapsulates a range of reasons why students find physics so hard. I have found the education theory of LCT (Legitimation Code Theory) to be a powerful way of explaining the difficulties students have with reconciling these different representations that Anthony referred to. As I have become more aware of the potential pitfalls involved with learning physics, I have been able to spot when my students are falling into them and I have become more explicit about pointing them out and trying to equip students to navigate them successfully.

I will briefly outline the LCT concepts involved and how they relate to the inherent difficulties of learning physics. I then give examples of the areas of difficulty that students have and the ways in which I make the transitions between representations explicit in my teaching.

2. Legitimation Code Theory

Legitimation Code Theory (LCT) is an education theory that seeks to understand what is valued ("legitimated") in various aspects of teaching and learning activities. The aspect of LCT that is relevant in this paper is that of "semantics". Firstly, "semantic gravity" is the extent to which a concept is contextualised (corresponding to high semantic gravity) or abstracted (corresponding to low semantic gravity). Secondly, "semantic density" is the amount of meaning that is condensed into a concept, word, or symbol. High semantic density means that there is

a large amount of meaning contained in the notation and low semantic density means that the language is more 'everyday', with less specific meaning attached to the notation [1].

Concepts can thus be mapped onto two perpendicular axes, showing the degrees of semantic gravity and semantic density in the concept under consideration. Semantic gravity (or, less commonly, semantic density) can also be plotted versus time. For instance, this is often used to track the levels of semantic gravity through a lecture or some other learning activity (see, for example, [2]).

3. LCT and physics

Physics is the academic subject that attempts to make sense of the real physical world (that is, aspects of the world that are non-living). It does this by isolating and abstracting relevant phenomena from the world by observation and experiment. From these observations physical laws and principles are formulated. These laws and principles are usually expressed and manipulated in symbolic mathematical equations. From the mathematical manipulations of physical laws and principles, predictions are made that are tested by experiments. The results of these experiments lead to the refining, extension or reformulation of the physical laws and principles, improving our ability to make sense of and understand the real physical world.

Thus physics involves transitions from the real world to an abstracted, but still contextual, experimental situation, through physical laws and concepts to generalised and abstract mathematical representations and then all the way back to the real world. It is impossible to do physics well without being proficient in these transitions and this is one reason why physics is generally acknowledged to be such a difficult subject to study [3, 4]. The necessity for a physicist to transition from the real world through concrete representations to physical and mathematical representations means that the process of learning physics lends itself to being understood in terms of the semantics dimensions of LCT.

The real world, the concrete nature of experiments, and the contextual problem statements of typical assessment questions correspond to high semantic gravity (rooted in context) and low semantic density (everyday use of language). As we transition towards the realm of physical laws and principles, semantic gravity decreases (the concepts are more abstract) and semantic density increases (the meaning contained in words becomes more specific and technical). Finally, in the mathematical realm, we find the lowest semantic gravity (most abstracted concepts) and highest semantic density (a large amount of meaning contained in a single symbol). Generally, in physics, we find that low semantic gravity tracks with high semantic density, although this is not exclusively so, for instance, it is necessary for students to watch out for meaningful words (with high semantic density) in otherwise concrete and contextualised (low semantic gravity) problem statements. For example, phrases like "rough surface" or "light string" convey meanings that are crucial to the successful solution of the problem.

In a typical physics lecture, time is spent in all these different realms. A lecturer may start with a concrete, contextual example and move up through the physical ideas involved to a mathematical formulation. Then perhaps they come back to interpret what a result means in the real world. When these different semantic gravities are plotted versus time a wave-like path is produced. These are called "semantic waves".

Referring back to Anthony's remark that I opened with, he was, in fact, expressing the difficulties he was having in translating between the different representations (the concrete, the principles and the mathematics) and making sure they were all 'telling the same story'. In LCT terminology, he was struggling to 'ride the semantic waves'. Anthony is not alone in this confusion and struggle and I will give some examples of the difficulties that I have noticed that students have in making these transitions and the ways that I have attempted to help them in my lecturing practices.

4. From "making sense" to physics and back

4.1. Cognitive Dissonance

I am sure that the reader is familiar with introductory questions similar to this one which has been adapted from Paul Hewitt's "Conceptual Physics" textbook [5]:

> A car travelling at 100 km/h strikes an unfortunate bug and splatters it. The force of the collision
>
> A is greater on the bug.
> B is greater on the car.
> C is the same for both.
> D cannot be compared without more information.

In a typical first-year class all four answers are given! When I remind my students about Newton's Third Law, they are happy to change their answer to option C, but they clearly don't really believe it in their hearts. This is an example of what is known as "cognitive dissonance" [6] in education theory and it is a sign that there are discordant beliefs that need to be reconciled in order for learning to happen [7].

When I observe an example of this in my teaching, I use it as an opportunity to tell my students "If ever you find yourself thinking 'I don't understand this, so I must *just learn it*', it is a sign that something is wrong and you need to wrestle with it until it makes sense." Sometimes, it is the student's common-sense view that is mistaken, and at others, their understanding of the physics concepts is incomplete. I use the above example to open the door to a fruitful class discussion that serves to both reduce misconceptions about Newton's Third Law and also to model the process of wrestling with conflicting understandings in order to achieve cognitive consistency [7].

4.2. When intuition fails

There are times when our intuition, or common-sense view, fails us. When this happens, there are valuable lessons to be learnt. One such example of where this happens that I use repeatedly with introductory physics students are conceptual physics questions such as the "Next Time Questions" compiled by Hewitt [8]. If these questions are answered too quickly, it is easy to get them wrong. This is an opportunity to point out to my students the value of wrong answers, as they give us a starting point for addressing misconceptions. If the right answer is given too early, I assume that everyone understands and move on, leaving a large portion of the class still struggling with cognitive dissonance.

There are also times when intuition fails in later years. In special relativity, we just have no intuition at all of what happens when we travel at speeds comparable to the speed of light. Trying to answer questions based on our low-speed intuition invariably leads to confusion. In this case, I use analogy to help, repeatedly reminding students that they are used to similar effects when spatial co-ordinate axes are rotated [9, 10]. I also explicitly remind them of what it is reasonable for them to expect to make intuitive sense. That is, that when considered inertial frame by inertial frame, the solution should make sense, and that there should be no causal inconsistencies between frames.

A colleague says about teaching quantum mechanics, "I only ask them questions that I know they'll get wrong!" In this case, it is not to teach the value of wrong answers, but for students to learn not to attempt to rely on their intuition in quantum mechanics. Can we ever get to the point of being able to trust our intuition in quantum mechanics? Is there even any intuition to rely on? But those are discussions for another time.

4.3. The dangers of familiarity, part 1

There is a danger for lecturers of taking the transitions between the physics representations and the concrete, contextual representations of a concept for granted. I received some extremely valuable feedback in this regard from an experienced colleague who peer reviewed my lecturing. He said "demonstrations ... may well show features that are obvious to you, but which may not be so to many students." It is important to remember where our students are coming from – that they do not have the prior knowledge or familiarity with a concept that we have; and that it is therefore valuable to them if we make the transitions between the concrete and the physics representations explicit. We cannot assume that our students can "see" the concrete representation of the physics concept unaided, especially when a topic is new and unfamiliar to them.

5. From physics to maths and back

5.1. Making connections

Regrettably, most students see the different subjects, and even topics, they study as being separate and unconnected. I was taken aback when my son asked of high-school maths, "why do we have to learn about parabolas?" Of course, he has not got to the stage of learning about projectile motion in physical science yet. Especially at first, it is necessary to emphasise to our students the connections to their prior knowledge and to other topics and subjects. For example, when teaching first-year students about significant figures, I tell them, "This is why they taught you about 'hundreds, tens and units' at primary school"; or when I introduce the concept of instantaneous velocity as a derivative, which they usually have not seen yet in maths, I say "Remember, you saw it here first!"

Eventually, students start making these connections for themselves. It is rewarding when honours classical mechanics students come to class excited about the numerical methods lectures they are attending in the maths department because they can see the applicability to the classical mechanics course.

5.2. Students are often more comfortable with maths

In my experience, students are often more comfortable with mathematical representations that they can manipulate procedurally and are less eager to return to the physics and sense-making realms in which they are more likely to experience cognitive dissonance. I consider that students therefore rely too heavily on the maths, using it as a crutch to bolster their shaky understanding of the physics. This is evidenced by a tutor in physics who also tutors in chemistry when he confessed to me "I don't know how to explain the spdf orbitals to the first years without using maths!"

In order to try and help them to make these translations to physics and sense making, I encourage students to form study groups so that they can explain concepts to each other, as it is often in teaching others that we gain a deeper understanding ourselves [11].

5.3. Physics answers versus maths answers

This point is best illustrated with an example. When teaching magnetic fields to first years, I ask a question such as:

> What happens to the radius of the circle the charge is moving in when the magnetic field is increased?

They correctly answer that it decreases. But when they are probed as to why, they will give the reason as "because the magnetic field is on the denominator." While this is true, it is the 'maths answer' instead of the 'physics answer'. It usually takes extensive probing in a class discussion to elicit "because the magnetic field is now stronger so the force is greater which pulls the charge

in more" as the physics answer. We then explicitly talk about the difference between 'physics answers' and 'maths answers' and what is required in a physics course.

I also point out that the maths and physics realms can be used to inform each other. For example, we often first notice that the magnetic field is on the denominator before we try explain in physical terms why that must be so. Another example of the maths informing the physics is when an answer comes out with a negative sign – that negative sign tells us something about the physics, that the quantity has decreased, or it is in the negative direction, or that an event happened before time $t = 0$. Conversely, our choice of co-ordinate system when setting up a problem in the physics realm informs us what signs quantities must take in the mathematical representation, for example in a conservation of momentum problem.

5.4. The dangers of familiarity, part 2

I received some very helpful feedback from a student regarding the difficulty that students have in translating between the maths and physics realms. She reminded me that, as a lecturer, I was very familiar with the physical concepts and the mathematical symbols representing them, but that students have not yet attained this same familiarity. She pointed out that when I wrote a symbol on the board, I would say the algebraic name of the symbol, for example, just "v", instead of "velocity". She said that they could read that for themselves, but that then they had to add an extra layer of translation, saying to themselves "v is velocity" which added unnecessarily to the cognitive load they were already experiencing, especially for more complicated symbols or more conceptually difficult topics. Explicitly making the link to the physics realm when talking or writing in the maths realm helps students to move between the realms more easily.

6. Telling the whole story coherently

Every so often, it is helpful to encourage our students to take a step back in order to remind them of the 'big picture' of the topic being studied; and to enable them to make the connections from the maths back to the physics, the concrete situation and the real world. As part of this process, they should be convinced that the different representations are telling the same story coherently and that the story they are all telling corresponds with the real world which we are seeking to understand.

I do this with the first-year students as a tutorial exercise using sets of small cards. Each card tells a story in a different way – using words, diagrams, graphs or equations. The students then work to group together the cards which are telling the same story. At the end of the first semester, also in a tutorial, I tell them to "produce a summary of Physics 101", with the hope that they will construct the 'big picture'; see the threads and connections between the topics; and realise which concepts are foundational principles and which are derivative special cases.

7. Conclusion

Since first becoming aware, through Anthony's frustrated comment, of the difficulties my students were facing in learning physics, I have found that 'the more I look, the more I see'. For example, even since preparing this paper, I have become more conscious of the pitfalls which await students due to the semantic density of the words used in physics. This is particularly relevant in my context in which almost all of my students are second-language English speakers. I have become aware that as I introduce a new concept with its corresponding name, I need to be more explicit about the ways in which this word is used differently in physics to its everyday meaning. I plan to suggest that students make a glossary for themselves of these words which have a special, semantically dense meaning in physics.

It is my hope that this paper prompts the reader to look for instances in their own lecturing where they can more explicitly help their students to surf the semantic waves with confidence.

References

[1] Winberg C, McKenna S and Wilmot K 2020 *'Nothing so practical as good theory': Legitimation Code Theory in higher education* (Routledge) pp 1–15

[2] Conana H, Marshall D and Case J 2020 *A Semantics analysis of first-year physics teaching* (Routledge London) pp 162–179

[3] Ornek F, Robinson W R and Haugan M R 2007 *Science Education International* **18**(3) pp 165–172 ISSN 1450-104X

[4] Bray A and Williams J 2020 *Journal of Physics: Conference Series* **1512**(1) ISSN 17426596

[5] Hewitt P G 2002 *Conceptual Physics* (Pearson Education) ISBN 0321051858

[6] Festinger L 1962 *Scientific American* **207**(4) pp 93–106 ISSN 0036-8733

[7] Gawronski B and Brannon S M 2019 *What is cognitive consistency, and why does it matter?* (American Psychological Association) pp 91–116 ISBN 1433830108

[8] Hewitt P Conceptual physics next time questions URL `http://www.arborsci.com/Labs/CP_NTQ.aspx`

[9] Taylor E F and Wheeler J A 1992 *Spacetime Physics* (Macmillan) ISBN 0716723271

[10] Moore T A 1995 *A Traveler's Guide to Spacetime: An Introduction to the Special Theory of Relativity* (McGraw-Hill) ISBN 9780070430273

[11] Redish E F 1994 *American Journal of Physics* **62**(9) pp 796–803 ISSN 0002-9505

Analysis of the understanding of vectors in Physics for first-year university entering students

V Nekhubvi, R Randela

Department of Physics, Faculty of Science, Engineering, and Agriculture, University of Venda

E-mail: **Vhutshilo.Nekhubvi@univen.ac.za**

Abstract. Physics education covers various topics, including classical mechanics, electromagnetism, thermodynamics, quantum mechanics, and relativity. In most topics in physics, vectors are essential because they represent physical quantities such as displacement, velocity, acceleration, force, and momentum. These quantities have both a magnitude and a direction, and vectors provide a convenient way to represent them mathematically. Teaching about vectors in the first year includes theoretical and practical aspects. It aims to equip learners with the mathematical and scientific skills necessary for higher levels and career paths. This paper investigates how first-year physics students deal with vector problems in physics. The study formally assessed 212 students, focusing on vector concepts to explore the challenges students face in answering questions, intending to equip first-year students with the right skills to solve vector problems. The study aimed to investigate the conceptual difficulties that the First-year entering students have in understanding vector concepts. The results show that most students need help understanding vectors in physics. The study has highlighted the areas where these students need help understanding vectors. In summary, vectors are essential in mathematics, science, engineering, and technology. They provide a way to express and manipulate complex physical and mathematical quantities and are used in various applications.

1. Introduction

Regardless of sufficient resources, physics education remains a realm of knowledge that eludes easy access. The Department of Physics of the University of Venda was surveyed in 2022 and 2023 to determine whether first-year students knew complex physics concepts before entering the class. Amongst the concepts, the vector concept was found to be the most challenging in both student categories (2021 and 2022). The survey results motivated the department to explore further research on understanding vectors. The following section provides literature on students' challenges when learning about vector concepts.

Shaffer P S and McDermott L CA [1] conducted a long-term, large-scale investigation into student understanding of motion to assess students' ability to qualitatively determine the magnitude and direction of instantaneous velocity and acceleration based on knowledge of an object's trajectory. The notable points are the mention of relevant prior research and the emphasis on student learning as the driving force behind their work. Their motivation was the identification of difficulties and persistence of challenges with the concept of velocity and acceleration. Prior research outcomes highlighted poor performance and persisting challenges when it came to understanding acceleration, which drove the research focus on improving the application of operational definitions of velocity and acceleration as vectors to real-world motion analysis. The study aimed to enhance the teaching and learning of

kinematics by developing and implementing tutorials, assessing their effectiveness, and gathering insights from many participants across multiple educational institutions.

Pablo and Genaro [2] studied students' understanding of vector concepts in problems without physical context. The discussed issue relates to the availability of multiple-choice testing instruments that evaluate students' understanding of vector concepts, as such tests are lacking. The second issue is the need for a large-population study at the university level to analyse students' understanding of vector concepts after completing their introductory physics courses.

In their study, Wutchana et al. [3] highlighted the significance of addressing students' misconceptions about vectors to enhance their learning experience in physics. Their focus was on high-school students. In their methodology, they developed a worksheet treated as the experimental intervention and the student's responses to a diagnostic quiz served as the dependent variable. The pre-and post-test scores were used to evaluate the effectiveness of the worksheet.

Hisham et al. [4] discussed the importance of understanding vector concepts for physics students at all levels. Their focus was on first-year medical students. The research aimed to improve students' understanding of vectors and, subsequently, their comprehension of physics concepts since it was found that students need help with simple graphical representations and basic vector operations. To achieve that, they employed a test of understanding of vectors (TUV), a multiple-choice test designed to assess students' understanding of vector concepts, to 240 students at the beginning of the semester before they were exposed to any vector-related instruction (pre-test). They further discussed the assessment of the test's reliability and discriminatory index through the analysis of difficulty and discriminative indices for each question. Their findings highlighted the poor performance of students in understanding vector concepts, particularly in terms of graphical properties and procedures. They recommended a clear improvement in teaching and learning strategies to address these misconceptions and enhance students' understanding.

The literature reveals that authors have investigated various aspects and approaches to tackle challenges. They have found similarities in their objectives, methodologies, and the recognition of the necessity for enhancing teaching and learning strategies. One of the methods employed to assess any difficulties students may face with vector concepts was the utilization of standardized tests.

2. Methodology

2.1. Method and material

A formative assessment in the form of a test paper with six questions that require understanding vectors in a physical context was developed. The method ensured that the questions were aligned with the learning objectives of vectors as a topic. Instructions were provided to the participants regarding the time limit and other applicable guidelines. The test was administered to 212 first-year university students. The test facilitated monitoring progress, identifying areas for improvement, and differentiating between students who have mastered competencies and those who have not, as alluded to by Friatma et al. [5]. This methodology provides a framework for studying students' understanding of vectors in physics education. The methodology also includes data collection and analysis of the administered test paper

2.2. Data Collection and Analysis

Statistical analysis was run to evaluate test performance, discriminatory index, and difficulty index. Participants' responses to the test paper were collected, ensuring their confidentiality and anonymity. Lastly, challenges were identified and discussed. Point biserial correlations between each test question and the total score on the test were used to assess the ability of individual questions to discriminate between high and low-performing students. The closer this value is to 1, the better the question distinguishes the learners who get a high score from those who get a low score [6]. Question difficulty is the fraction of learners who correctly answered a question, ranging from 0.0 to 1.0. The closer the difficulty of a question approaches zero, the more difficult that question is [6].

3. Results

The results are based on the written test, of which details are presented in Table 1. Firstly, the table contains information about the individual questions, their marks, and their performance represented as a percentage of at least 50%. This information can be used to identify areas of improvement for the individual and to track their progress over time. It can also give tailored feedback and support, enabling students to focus on their areas of concern and build upon them.

Table 1. Information about the individual questions, their marks, and their performance

#Q	Question information	Marks	No: of students Scored >=50%	% of students Scored >=50%
Q1	Give the names of all physical quantities, their SI units, and the symbols for the dimension used in this scenario.	6	153	72
Q2	If the vector $\vec{r}(t) = (2t + 1)\vec{\imath} + 2t^2\vec{\jmath}$ represents the movement of the jet. Find; (a) $\vec{r_i}$ when $t_i = 0$ hr. (b) $\vec{r_f}$ when $t_f = 2$ hrs. (c) $\Delta\vec{r}$	6	144	68
Q3	Calculate the magnitude and direction of the following (a) $\vec{r_i}$, (b) $\vec{r_f}$ (c) $\Delta\vec{r}$	12	107	50
Q4	Use the dot product of vectors to find; $(a)\vec{r_f} . \vec{r_i}$ (b) $r_f r_i$, (c) the angle between the two position vectors $\vec{r_f}$ and $\vec{r_i}$	6	52	25
Q5	Define the following terms mathematically: (a) Displacement vector. (b) Average velocity. (c) Instantaneous velocity (d) Average acceleration. (e) Instantaneous acceleration.	10	114	54
Q6	Calculate the following: (a) Displacement vector. (b) Average velocity. (c) Instantaneous velocity at t = 2 hrs (d) Average acceleration. (e) Instantaneous acceleration.	10	47	22

Table 2 below shows misconceptions that were observed during the study. Five crucial misconceptions have been observed and recorded.

Table 3 is a breakdown of student achievement in tests. The table provides information about three types of student categories and their average marks for each category. These categories are HA (high achievers), MA (medium achievers), and LA (low achievers). Besides the average class mark, the table also contains the highest and lowest marks in the class.

Table 2. Misconceptions observed

Question	Main Challenges/Mistakes
1	Students should have answered this question properly and got it right. The question required students to give names of all physical quantities, their SI units, and dimensional symbols present in the situational case.
2	Students write $\Delta\vec{r} = \overrightarrow{r_f} + \vec{r_i}$ instead of $\Delta\vec{r} = \overrightarrow{r_f} - \vec{r_i}$. They apply the associative law of addition. Students need to write simple algebra involving unit vectors.
3	Some students think that $\overrightarrow{r_f} . \vec{r_i}$, is the same as $r_f r_i$, Others computed the magnitude of $\Delta\vec{r}$ as $\lvert\Delta r\rvert = \sqrt{r_i^2 + r_f^2}$
4	They need to define terms mathematically; they use the instantaneous velocity formula for average velocity and average acceleration.
5	Difficulty in knowing exactly which formula to use was a challenge in attempting to solve Basic mathematics.

Table 3. Detailed student marks per category

Student Category	HA average Marks (%)	MA average Marks (%)	LA average Marks (%)	Class
Average Marks	73	48	23	47
Highest	96	62	36	96
Lowest	64	36	0	0
Number of students	57	98	57	212

The statistical analysis data in Table 4 consists of two indices: the discrimination index and the difficulty index. Test questions are commonly assessed with these indices in educational measurement. The two indices are interpreted: Among the six questions (Q1-Q6), discrimination index values range from 0.1 to 0.7. A question with a higher discrimination index value discriminates more effectively between high and low achievers. The questions have a range of difficulty index values from 0.2 to 0.7. When the difficulty index is 0.5, about half of the test takers correctly answered the question. At the same time, values below 0.5 indicate questions of relatively high difficulty, and values above 0.5 indicate questions of relatively low difficulty [6].

Table 4. Statistical analysis of the individual question

	Q1	Q2	Q3	Q4	Q5	Q6
Discrimination index	0.1	0.6	0.7	0.4	0.6	0.4
Difficulty index	0.7	0.7	0.5	0.2	0.5	0.2

4. Discussion

The discussion section of this research paper contains the analysis of achievement data collected from 212 students test marks on vectors. Students were categorized into three performance levels: High Achievement (HA), Moderate Achievement (MA), and Low Achievement (LA), and their average marks, highest and lowest scores, and total number were examined. Students in the HA category

achieved an average mark of 73%, suggesting relatively good performance. The MA category, on the other hand, achieved an average score of 48%. Students in the LA category achieved an average score of 23 %. The highest scores in each category were 96% in HA, 62% in MA, and 36% in LA respectively; a score of 96% was the highest in the class. However, the lowest scores were recorded in HA, MA, and LA, at 64%, 36%, and 0%, respectively. A score of 0 was earned in the category LA, the lowest in the class. There were 57 students in HA, 98 in MA, and 57 in LA categories.

At least, 92 of the 212 students passed the test; to define the discrimination index and the difficulty index in assessment and measurement context is important for a critical analysis of Table 4. Each index and its implications are analyzed based on the values given in Table 4. An index of discrimination ranges from -1 to 1, with positive values indicating good discrimination and negative values reflecting poor discrimination [6]. Test questions with a high discriminatory index, such as Q3, have a relatively high difficulty index. For Q1, a discrimination index of 0.1 indicates that the question is limited in its discriminative ability. The discrimination between high and low achievers is most effective in Q3 due to its discrimination index of 0.7. The rest of the questions (Q2, Q4, Q5, and Q6) have an average discriminative index.

Thompson et al. [7] state that poorly constructed test questions are associated with low discrimination indices. Test questions must be evaluated during development to ensure test reliability and validity. Based on the difficulty index, which ranges from 0 to 1, it is possible to determine the correct number of people who answered the question. As can be seen in Table 4, difficulty index values range between 0.2 and 0.7. The values indicate how difficult the test questions are. With a difficulty index of 0.2, Q4, for instance, shows a relatively low percentage of people answered this question correctly, indicating a higher difficulty level. It is more helpful to discriminate between these groups if the discrimination index is higher. In support of this analysis, reference is made to a study by [8], which shows that questions with extremely high or low difficulty indices can lead to biased test results and reduce the overall reliability of the test.

5. Recommendations

Based on the results, the following recommendations can be made: there is a need to provide additional support and resources to low-performing students (category LA). To achieve the later, targeted intervention strategies should be introduced. Additional tutoring, learning materials, and personal support could be offered to improve their understanding of vectors.

Reviewing and revising test questions using discrimination and difficulty indices is essential. Questions with low discrimination indices should be modified or replaced to effectively differentiate between high and low-performing students.

To maintain the overall reliability of the test, questions with extremely high or low difficulty indices should be carefully reviewed. If these recommendations are implemented, students will be able to understand better and represent vectors in first-year physics classes, leading to better academic performance.

6. Conclusions

Test questions were ranked according to how well they discriminated between high and low performers. Students who performed well on questions Q2, Q3, and Q5 had higher discrimination indices than those who performed poorly. However, questions Q1, Q4, and Q6 had lower discrimination indices; they could have been more effective in distinguishing between high and low-performing students. A difficulty index was used to determine the difficulty of each test question. From the difficulty index, questions Q4 and Q6 were relatively more complex. Questions Q1 and Q3 were reasonably straightforward, while Q2 and Q5 were of medium difficulty.

References

[1] Shaffer P S and McDermott L C A 2005 *American Journal of Physics.* **73** 921
[2] Pablo B and Genaro Z 2021 *Phys. Rev*. **10** 010121
[3] Wutchana U, Bunrangsri K, and Emarat N 2015 *International Journal of Physics and Chemistry Education* **7** 18
[4] Hisham N B, Kolthoom A A, Hasan M A, and Sami H A 2020 *Health Professions Education* **6** 428
[5] Friatma A and A Anhar 2019 *J. Phys.: Conf. Ser.* **1387** 012063
[6] Taib F and Yusoff M S B 2014 *Journal of Taibah University Medical Sciences*. **9** 110
[7] Thompson R, Adams J and Williams K 2018 *Journal of Educational Measurement*, **45** 267
[8] Smith P and Johnson L 2017 *Educational Assessment* **30** 145

DIVISION

F

APPLIED PHYSICS

Optimization of the Electrical performances of a Silicon Solar cell using a Non-uniform Doping Distribution

A.D. Kapim Kenfack[1], N.M. Thantsha[1] and M. Msimanga[1,2]

[1]PV Nanocomposites R&D Platform, Department of Physics, Tshwane University of Technology, Private Bag X 680, Pretoria 0001, South Africa

[2]NRF-iThemba LABS, P Bag WITS 2050, Johannesburg, South Africa

E-mail: **kapimkenfackad@tut.ac.za**

Abstract. Despite several techniques of optimization of the yield of a solar cell, we remark that the use of a non-uniform doping distribution has not yet been explored. This work studies the effect of a non-uniform doping profile on the output electrical power of a silicon solar cell. The uniform doping profile commonly used in conventional solar cells is replaced by a Gaussian doping distribution. The new doping profile leads to a nonlinear continuity equation that is solved using the Tri Diagonal Matrix Algorithm (TDMA). The simulations made under standard test conditions on the output electrical parameters reveal that the Gaussian doping profile increases the leakage photocurrent at the edge of the solar cell and reduces the photo-generation process of the minority carriers in the solar cell. Additionally, by varying the doping concentration at the top surface and back contact of the solar cell to vary the parameters of the Gaussian profile, the solar cell achieve an electrical power and efficiency of about $50\ mW/cm^2$and 50% respectively.

1. Introduction

Since decade, recombination effect is known as the main performance limiting factor of a solar cell by around 26 % in the laboratory for uniform doping around 10^{16} cm^{-3}[1]. Other factors include temperature [2], shading [3], dust [4]. Research in solar cells has paid much attention to improving the performances using thickness [5] and magnetic field [6] to name a few. From the literature, we observe that the simulations were carried out on the solar cells manufactured with a uniform doping level [1-7]. In the same direction, we decide to explore the optimization of the silicon solar cell by replacing the uniform doping currently used to manufacture the solar cells with a non-uniform doping profile according to the thickness of the solar cell.

The electrical field, diffusion coefficient and lifetime are expressed as function of the non-uniform doping profile, leading to the resolution of the nonlinear continuity equation using Tri Diagonal Matrix Algorithm (TDMA) method. The expressions of the photocurrent, photovoltage lead to establishment of the expression for the electric power. Moreover, the effect of the non-uniform doping on the short circuit photocurrent, open circuit voltage and optimum electrical power are presented.

2. Modeling of the monocrystalline solar cell

Figure 1 shows a schematic representation of the monocrystalline silicon solar cell used in the study. The quantity ***H*** designates the depth of the solar cell, E_z defines the internal electrical field due to the gradient of doping. During our investigations, the current produced by the emitter is far less than that produced by the base due to the fact that the diffusion coefficient (mobility) of holds in the emitter is neglected in front of electron in the base. Therefore, our analysis is developed in the base region [6].

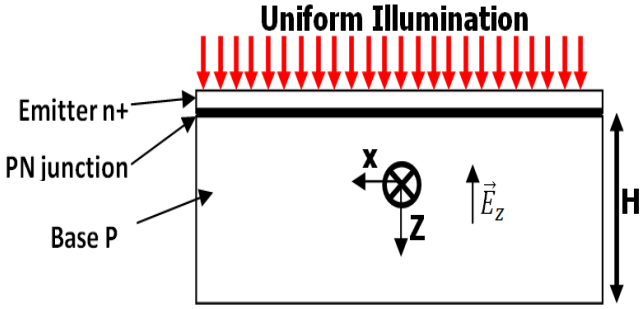


Figure 1: Illustration of the solar cell subjected to a uniform illumination.

The generation rate of the electrons only depends on the depth of the solar cell [5]. The illumination is considered uniform on the solar cell. The defects density of the material created by the doping profile is not considered throughout the simulation

2.1. Excess of minority carriers

The continuity equation of excess of minority carriers $n(z)$, generated in the base region of the solar cell is expressed as follows [5].

$$\frac{1}{q}\frac{\partial J}{\partial z} - \frac{n(z)}{\tau(z)} = -N\sum_{i=1}^{3} a_i e^{-b_i z} \tag{1}$$

with a_i, b_i, N the coefficients for an irradiation under an atmospheric mass of 1.5 (AM 1.5), q defines the elementary charge, $n(z)$ represents the carriers density, the quantity J represents the total density current delivered by the solar cell defined by $J = q \cdot D(z) \cdot \partial n/\partial z + q \cdot E_z \cdot n \cdot D(z)/V_t$ represents the sum of diffusion and drift current. $D(z)$ and $\tau(z)$ designates the diffusion coefficient and lifetime of the carriers depending on the doping distribution [6]. The Gaussian doping profile in the base region is express as follows

$$\mathrm{N_b(z)} = \mathrm{N_{b2}} \cdot \exp\left[-\left(\frac{\mathrm{z-H}}{\varepsilon}\right)^2\right] \tag{2}$$

Wherein $\varepsilon = H/\sqrt{\log(N_{b2}/N_{b1})}$ in which N_{b1} and N_{b2} represent the doping level at the depth $z = 0$ and at $z = H$ respectively. Knowing that the drift current is not more negligible the electrical field is expressed as follows

$$\mathrm{E_z} = \frac{\mathrm{V_t}}{\mathrm{N_b(z)}}\frac{\partial \mathrm{N_b(z)}}{\partial \mathrm{z}} \tag{3}$$

By taking into account the mentioned coefficients $(E_z,\ \tau, D, N_b)$, Eq (1) can be rewritten as follows:

$$\frac{\partial^2 \mathrm{n}}{\partial^2 \mathrm{z}} + \left[\frac{\partial \mathrm{D}}{\mathrm{D}\,\partial \mathrm{z}} + \frac{\mathrm{E_z}}{\mathrm{V_t}}\right]\frac{\partial \mathrm{n}}{\partial \mathrm{z}} + \left[\frac{\partial \mathrm{E_z}}{\mathrm{V_t}\,\partial \mathrm{z}} + \frac{\mathrm{E_z}}{\mathrm{V_t}}\frac{\partial \mathrm{D}}{\partial \mathrm{z}} - \frac{1}{\tau \mathrm{D}}\right]\mathrm{n} = -\frac{\mathrm{N}\sum_{\mathrm{i=1}}^{3} \mathrm{a_i e^{-b_i z}}}{\mathrm{D}} \tag{4}$$

This nonlinear equation (4) can be solved using the TDMA algorithm using the following boundaries conditions:

- PN junction of the solar cell

$$\mathrm{D(z)} \cdot \frac{\partial \mathrm{n(z)}}{\partial \mathrm{z}} + \frac{\mathrm{D(z)}}{\mathrm{V_t}} \cdot \mathrm{n(z)} \cdot \mathrm{E(z)} = \mathrm{S_f} \cdot \mathrm{n(z)} \quad ; \mathrm{z} = 0 \tag{5a}$$

- back side of the solar cell

$$D(z) \cdot \frac{\partial n(z)}{\partial z} = -S_b \cdot n(z) \qquad ; \; z = H \tag{5b}$$

$\mathrm{S_f}$ models the flux of carriers crossing the PN junction. This flux of carriers defines the operating point of the solar cell, S_b represents the recombination velocity of the carriers effectively recombined at the back surface of the solar cell [5].

2.2. Electrical parameters

2.2.1. *Photocurrent.* The photocurrent density crossing the external load depends on the recombination velocity, and the doping level constant N_{b1} and N_{b2} . The photocurrent is expressed as [1]

$$I_{ph} = \left(q \cdot D(z) \cdot \frac{\partial n(z)}{\partial z} + q \cdot D(z) \cdot n(z) \cdot \frac{E(z)}{V_t} \right)_{z=0} \tag{6}$$

2.2.2. *Photovoltage.* Since the excess minority carrier density is known from Boltzmann's law, photovoltage at the outer charge terminals of the solar cell is expressed as follows [5].

$$V_{ph} = \eta \cdot V_t \cdot log\left(1 + \frac{N_b(z)}{n_i^2} \cdot n(z, N_{b1}, N_{b2}) \right)_{z=0} \tag{7}$$

Wherein n_i, η are respectively the intrinsic concentration and the ideal factor. The quantity V_t designates the thermal voltage.

2.2.3. *Electrical power.* The electrical power collected at the terminals of the solar cell, results from the product between the photocurrent and the photovoltage, expressed as follows [1]

$$P_{ph}(S_f, N_{b1}, N_{b2}) = V_{ph} \times I_{ph} \tag{8}$$

As the intrinsic recombination velocity $(\mathrm{S_f})$ is used to determine the operating point of the solar cell, then the maximum electric power can be determined by

$$\mathrm{P_{max}(N_{b1}, N_{b2}) = Max\left(P_{ph}[S_{fmax}, N_{b1}, N_{b2}]\right)} \tag{9}$$

To determine the optimal value of $\mathrm{N_{b1op}}$ and $\mathrm{N_{b2op}}$ for which expression (9) is optimum, this following relation should be satisfy

$$\begin{cases} \frac{\partial P_{max}}{\partial N_{b1}} = 0 \\ \frac{\partial P_{max}}{\partial N_{b2}} = 0 \end{cases} \quad (10)$$

Matlab software offers the possibility to check the contribution of the non-uniform doping profile on the electrical parameter through the resolution of expression (4) [5].

3. Methodology

The contribution of non-uniform doping level on a solar cell is evaluated through its output electrical parameters. The simulations of the electrical parameters (short circuit photocurrent, open circuit photovoltage and electric power) presented in this study were made in the standard test condition, where the irradiation was normalized at 0.1 W/cm^2 and 25°C as the solar cell temperature. The curves were obtained with the system parameters: $n_i = 10^{10}\ cm^{-3}$, $S_b = 10^2 \mathrm{cm}^2$, $H = 0.03\ cm$, and $\eta = 1.2$. We have considered the mononocrystalline silicon solar cell as shown in Figure 1, justified of the ideal factor value η and the also the mathematical expression of the lifetime and diffusion coefficient. The two grandeurs N_{b1} and N_{b2} are considered as the parameters in our system throughout our simulations. The results of the simulation presented in this section were made in MATLAB platform.

4. Results and discussion

4.1. Impact of non-uniform doping on the short circuit photocurrent

This subsection presents the simulations of the short circuit photocurrent obtained by exploiting the analytical expression (6) when $S_f \to +\infty$ and presented in Fig 2. The plotted short circuit photocurrent versus N_{b2} for various values of N_{b1} is shown in Fig. 2 (a) on one hand, and versus N_{b1} for various values of N_{b2} is shown in Fig. 2 (b) on the others hand.

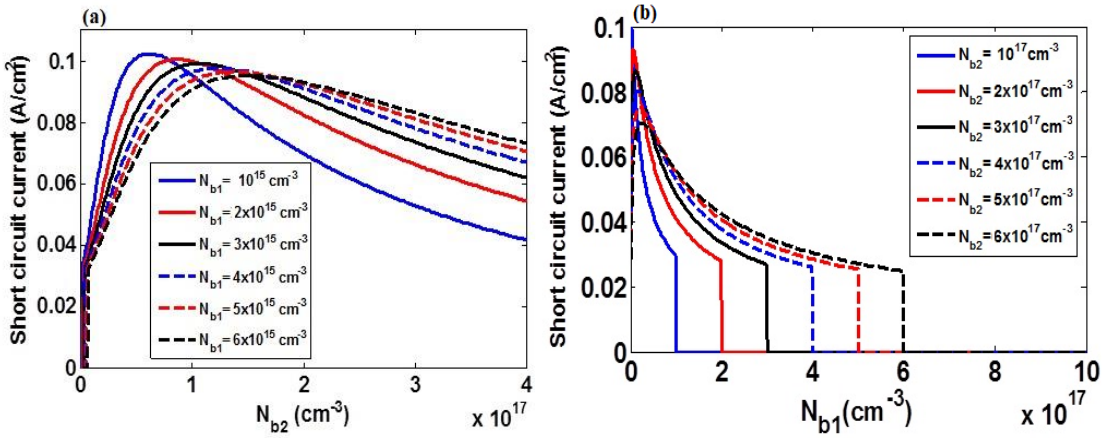


Figure 2: Variation of the short circuit photocurrent versus **a)** N_{b2} for various value of N_{b1}**b)** N_{b1} for various value o f N_{b2}.

From the literature, the short circuit photocurrent decreases when the uniform doping increases [7]. In contrast, with a Gaussian doping profile, the behavior has changed completely as shown in Fig.2. In fact, we remark that near (Fig 2(a)) and far (Fig 2(b)) the vertical axe, the current is nil. In this situation, $N_{b2} < N_{b1}$. Indeed, the drift current in this configuration contributes to completely inhibit the diffusion current. The photo-generated carriers are pushed toward the rear face of the solar cell thanks to the electrostatics force generated by the doping profile. So, no carrier can cross the PN junction and hence no current is produced. Somewhere, when $N_{b2} \geq N_{b1}$, the short circuit photocurrent increases up to the maximum and decreases after. This behavior can be explained by the fact that, drift current has changed the direction

and contributes to the increase in the concentration of the photo-generated minority carriers at the PN junction of the solar cell. The increase in the concentration of minority carriers induces a strong carrier gradient at the PN junction and facilitates the carrier extraction out of the solar cell. In addition, the increase in the population of minority carriers at the PN junction promotes the increase in frictional forces between the carriers reducing the energy in the system. Thus, the photo-generation that takes place in the solar cell is reduced in favor of the recombination process and therefore reduces the photocurrent as shown Fig 2. This high current is also due to the fact the defects density of the material created by the doping profile is not considered throughout the simulations [8].

4.2. Contribution of non-uniform doping on the open circuit photovoltage

The open circuit photovoltage is plotted by exploiting the expression (7) when $S_f \to 0$. The curves are plotted against N_{b2} for various values of N_{b1} and depicted on Fig. 3 (a) on one hand, and versus N_{b1} for various values of N_{b2} on the others hand and depicted on Fig. 3 (b). Cheikh et al in 2015 have shown the open circuit photovoltage increases when the uniform doping increases [7].

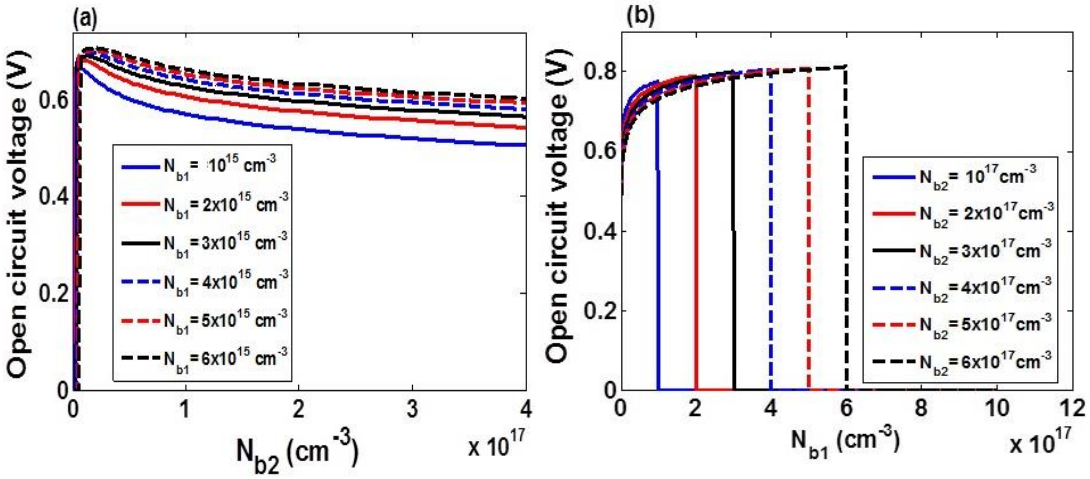


Figure 3: Variation of the open circuit photovoltage versus **a)** N_{b2} for various value of N_{b1} **)** N_{b1} for various value of N_{b2}.

However, in the presence of the non-uniform doping level, we remark that near (Fig. 3(a)) and far (Fig. 3(b)) the vertical axe, $N_{b2} < N_{b1}$ and the open circuit photovoltage is nil. In this situation, all the carriers generated are blocked in the solar cell due to the electrostatic field produced by the doping profile. The direction of the electrostatics force is opposite to the direction of the carriers. Thus, no minority carriers diffuse into the solar cell. Therefore, the solar cell is insensitive to the incident photon and does not operate. In addition, the absence of movement of the free minority carrier in the solar cell leads us to zero photovoltage. In contrast, for $N_{b2} \geq N_{b1}$, the photovoltage increases and decreases afterwards. In this situation, the drift current and diffusion current have the same direction. Thus, the carrier can be easily photo-generated and contribute to the production of the photovoltage. However, the decrease in photovoltage is due to the frictional forces induced by the high population of carriers at the PN junction. Thus this frictional force induces the birth of the recombination process and therefore the carrier becomes less photo-generated hence, the decrease in photovoltage.

4.3. Impact of non-uniform doping on the optimum maximum electric power

The simulations of the optimum maximum electric power delivered from a solar cell are obtained by using expression (9). Fig 4 was obtained for a specific set of combination (N_{b2}, N_{b1}) such that the solar cell deliver a maximum power. The curves were plotted against N_{b2} as depicted in Fig. 4 (a), and versus N_{b1} as depicted in Fig. 4 (b). Under the uniform doping, it appears from the literature that the electric

power and the efficiency delivered by a silicon solar cell in standard test condition was 26 $mW/cm^2$and 26 % respectively [1] for a doping level $N_b = 10^{16}\ cm^{-3}$. In the case of the non uniform doping level considered in this study, we remark that the optimum electric power and efficiency delivered by a silicon solar cell are $50 mW/cm^2$ and 50 % respectively as shown Fig.4. At the optimum power point, the characteristic doping $N_{b1op} = 4 \times 10^{15} cm^{-3}$ and $N_{b2op} = 1.14 \times 10^{17} cm^{-3}$ satisfy expression (10). This high electric power obtained compared to the uniform doping is due to the non consideration of the defect created by the doping profile during the simulations that cause the distortion of the crystal structure and create imperfections in the device [8].

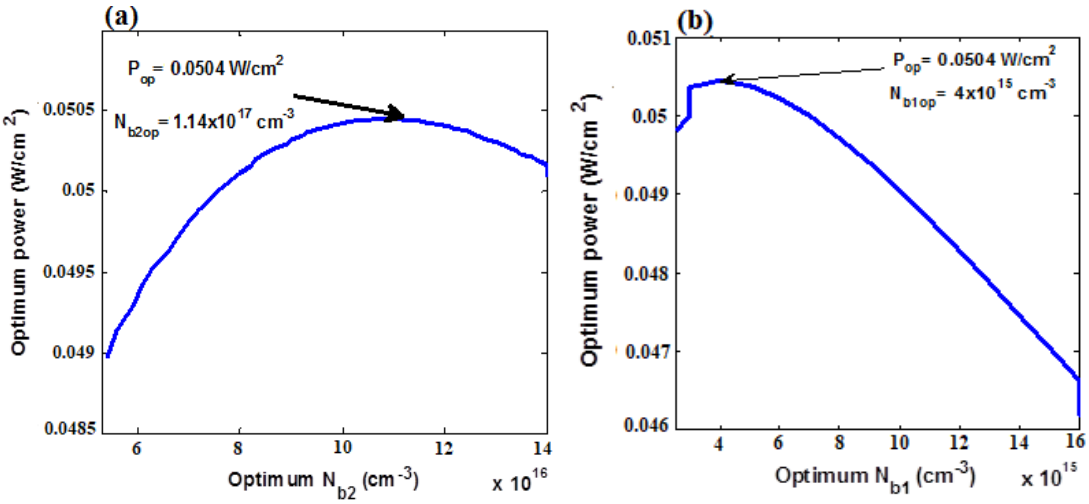


Figure 4: Optimum electric power versus **a)** N_{b2} for various value of N_{b1} **b)** N_{b1} for various value of N_{b2}.

5. Conclusion

The aim of this paper was to evaluate the impact of the non-uniform doping profile on the output electrical parameters of a monocrystalline silicon solar cell. The influence of the non-uniform doping profile was mainly checked on the photocurrent, photovoltage and electric power. The doping profile chosen within our investigations follows the shape of a Gaussian function. We reveal that to theoretically improve the performance of a monocrystalline silicon solar cell up to 50.04 $mWcm^{-2}$ using a Gaussian doping profile, the optimum doping at the PN junction and rear surface should be around $4 \times 10^{15} cm^{-3}$ and $1.14 \times 10^{17} cm^{-3}$ respectively.

Acknowledgements

The authors gratefully acknowledge financial and material support from the Tshwane University of Technology PV Nanocomposites R&D Platform and iThemba LABS of the National Research Foundation of South Africa

References

[1] Green A.M, Yoshihiro H, Hishikawa Y, Warta W, Ewan D.D, Dean HL, Jochen, H.E, Anita W.Y.H 2017 *Progress in Photovoltaics: Research and Applications* **25** 668.
[2] Dinçer F, Meral M.E 2010 *Smart Grid Renew. Energy* **1** 47.
[3] Seyedmahmoudian M, Mekhilef S, Rahmani R, Yusof R and Renani E.T 2013. *Energies* **6** 128.
[4] Sulaiman S.A, Hussain H.H, Leh N.S.H.N and Razali M.S.I 2011 *Int. J. Mech. Mechatron. Eng.* **5** 2028
[5] Pelap F.B, Konga, T.E, Kapim K.A.D 2021 *J.Ren. En.* **24** 25.
[6] Dieng A, Thiam N, Thiam A, Maiga A.S, Sissoko G 2011 *Res. J. Appl. Sci. Eng.Technol.* **3** 602.
[7] Cheikh M.L.O, Seibou B, Moujtaba M.A.O, Faye K, Wade M, Sissoko G 2015 *Int. J. Eng. Trends Technol.* **19** 44.

[8] Ali K, TARIQ Z, Khan H.M, Javed Y, Sultan M, Anmol M 2019 J. Ovonic Res.**15** 215.

Effect of polymer coating on the calcium ferrite ($CaFe_2O_4$) nanoparticles for biomedical applications

WB Mdlalose[1] and SR Mokhosi[2]

[1]Discipline of Physics, [2]Discipline of Biochemistry University of KwaZulu-Natal, Westville campus, P/Bag X54001, Durban 4000, South Africa

E-mail: **MdlaloseW@ukzn.ac.za**

Abstract. Calcium ferrite nanoparticles (NPs) with the chemical formula $CaFe_2O_4$ have been synthesized by glycol-thermal process. Surfaces of the synthesized NPs were then functionalized with biocompatible polymers, chitosan (CH) and polyvinyl alcohol (PVA) to obtain CH-$CaFe_2O_4$ and PVA-$CaFe_2O_4$ NPs, respectively. The X-ray diffraction (XRD) results for all NPs revealed the formation of a cubic phase without any impurity phases. The average crystalline sizes (using Scherrer's equation) decreased from 6.43 nm to 5.61 nm and 6.36 nm after CH and PVA coatings, respectively. Fourier transmission infrared (FTIR) confirmed spinel structures of the NPs and proper polymer coating on the surfaces of $CaFe_2O_4$ NPs. High-resolution transmission electron microscope (HRTEM) and scanning electron microscopy (SEM) images show nearly spherical particles with nano-sizes below 6.9 nm. The energy dispersion spectra (EDS) results confirmed the absence of the impurities and atom contents in the compounds. Magnetization data revealed the superparamagnetic nature of all NPs. The magnetization of the $CaFe_2O_4$ NPs was shielded, while their coercivity increased after including the polymers on the surface. Nanoparticle tracking analysis (NTA) revealed more stability in physiological pH (6.889) for the CH-$CaFe_2O_4$ NPs with the highest zeta potential of -22.5 mV. However, the PVA-coated NPs are also indicated to be more stable compared to the naked $CaFe_2O_4$ NPs. Hence, this suggests that both polymers are suitable surface stabilizers of $CaFe_2O_4$ NPs. Stable NPs are crucial for future biomedical applications such as target drug delivery systems (DDSs) and chemotherapy treatment for various cancer diseases when suspended in liquid. Hence, CH and PVA-coated $CaFe_2O_4$ NPs may be used for the mentioned applications.

1. Introduction

Magnetic nanoparticles (MNPs) have recently been explored in nanomedicine, including magnetic resonance imaging, hyperthermia therapy and targeted drug delivery [1]–[5]. Ferrites are a class of iron-oxide NPs with unique physico-chemical and magnetic attributes [6]. Considerations in the control of the particle size, shape, stoichiometry and surface chemistry are essential in biomedical applications

[7]. They affect the physical and chemical properties of NPs, influencing their overall functionalities in biological settings[8]. The magnetic field parameter has been noted to play a vital role in MNP cytotoxicity, the most significant of which are magnetic field amplitude, frequency, and duration of action [9], [10]. Additionally, large MNPs present a cytotoxic effect in alternating low-frequency magnetic fields compared to the smaller ones [11]. Some of the strategies to improve NP toxicity profile, biodegradability and chemical stability in a physiological environment employ the surface modification of the NPs [8], [12]. Functionalization with organic and biodegradable polymers such as chitosan (CH), poly-ethylene glycol (PEG) and polyvinyl alcohol (PVA) is routine practice for NPs in biomedical applications [13], [14]. $CaFe_2O_4$ NPs are one of the many NPs types that have gained much interest due to their fascinating catalytic, optical, and magnetic properties [15]. Over $MnFe_2O_4$, $NiFe_2O_4$, $ZnFe_2O_4$, $CoFe_2O_4$ and $CuFe_2O_4$ NPs, $CaFe_2O_4$ NPs present an added advantage of being eco-friendly due to the presence of Ca^{2+} instead of heavy metals [15], [16]. For applications in a various application including cancer therapy, MNPs should ideally exhibit biocompatibility, monodispersity, chemical stability, and superparamagnetism [17], [18]. Additionally, NP sizes should be controllable with desirable ranges below 200 nanometers [18]– [20].

In this report, we look at the synthesis of ferrite NPs using the glycol-thermal reaction method. Several synthesis methods are employed in ferrite NPs, such as the sol-gel, hydrothermal, high-energy ball milling, and glycol-thermal reaction methods [21]– [23] However, the selected glycol-thermal route is known to produce small NPs that exhibit superparamagnetism [20]. A comparative investigation of coating the NPs using biopolymers, *viz.* CH and PVA was evaluated to assess their feasibility for biomedical applications.

2. Experiment details

Reagents: Ethylene glycol [$(CH_2OH)_2$, 99 %], ethanol (C_2H_5OH, 99 %), calcium chloride ($CaCI_2.6H_2O$, 99 %), iron (III) chloride ($FeCl_3.6H_2O$, 98 %) hexahydrate salts, chitosan and polyvinyl alcohol all were purchased from Sigma-Aldrich (St. Louis, United States of America).

1.1 Synthesis of $CaFe_2O_4$ NPs

The $CaFe_2O_4$ NPs were produced using the glycol-thermal reaction method discussed elsewhere [24]. Using various stoichiometric formulations of raw materials, $CaCl_2.6H_2O$ and $FeCl_3.6H_2O$ hexahydrates were dissolved in 500 ml of deionized water, producing a homogenous mixture. Precipitation of the metal chlorides was carried out by gradually adding 5 M NaOH solution until a pH of 9 was reached. The residue was then washed several times with deionized water. The clean precipitate was dispersed in 300 ml of ethylene glycol, and then the mixture was placed in a 600 ml stainless steel pressure vessel (Watlow series model PARR 4843 reactor). In that instance, the reactor could run for 6 hours at a soak temperature of 200 ^{0}C with a stirring speed of 300 rpm under the pressure of about 80 psi. The final product acquired was rinsed five times with 200 ml of ethanol over Whatman paper. After that, it was placed under a 200 W infrared light and left to dry overnight. The dried sample was then homogenized using an agate mortar and pestle. The synthesized $CaFe_2O_4$ NPs powdered sample was subsequently coated with biocompatible polymers chitosan (CH) and polyvinyl alcohol (PVA) and named CH-$CaFe_2O_4$ and PVA-$CaFe_2O_4$ NPs.

2.2 Surface functionalization of MNPs

2.2.1 Coating with chitosan (CH)

Coating with chitosan was carried out as per the coating method by Khalkhali *et al*. [25] without any further modifications. Approximately 0.2 grams of the $CaFe_2O_4$ NPs were weighed and placed in a beaker. After that, 0.5 grams of chitosan was dissolved in 100 ml of acetic acid to prepare the 0.5 % chitosan solution. After that, 10 M NaOH solution was added dropwise to the 0.5 % chitosan solution until a pH of 4.8 was reached. This was followed by adding $CaFe_2O_4$ NPs in the chitosan solution, and

the resulting mixture was sonicated using a Scientech Ultrasonic cleaner at 60 ℃ for 1 hour. The mixture was transferred into a beaker and stirred mechanically using an IKA RW 20 Digital Dual-Range Mixer System for 18 hours at room temperature. The black homogeneous mixture attained was separated by centrifugation at 300 rpm for 30 minutes. The sample was centrifuged up to 5 times, and between every centrifuge interval, it was washed with deionized water. The magnet was used to assist in separating CH-coated $CaFe_2O_4$ NPs from the homogenous mixture. Then, the product was dispersed in a petri dish and left to dry overnight in the oven at 30 ℃. Dried samples were then homogenized using an agate mortar and pestle at room temperature for further characterization studies.

2.2.2 Coating with polyvinyl alcohol (PVA)

Coating with polyvinyl alcohol was carried out per the coating method described by Ehi-Eromosele *et al*. [26] with modifications. One gram of dried $CaFe_2O_4$ NPs and 3 grams of PVA were dissolved in 96 grams of deionized water. Complete dissolution of PVA was achieved under vigorous stirring at 80 ℃, and then the speed was slowly reduced. The final solution had the ratio of NPs: PVA = 1:3. The solution remained stirred for 20 hours at room temperature to achieve coating—three wt. % PVA-coated $CaFe_2O_4$ NPs were then separated with a permanent magnet and washed three times with deionized water. Finally, pure PVA-coated $CaFe_2O_4$ NPs were separated and dried at 30 ℃ after removal of all the residual PVA with washing**.** The dried samples were then homogenized using an agate mortar and pestle at room temperature for further characterization studies.

2.3 Characterizations

X-ray powder diffraction patterns were recorded on a D8 Brucker advance X-ray diffractometer using monochromatic Co-Kα (1.7903 Å) radiation at room temperature in the range of 10° to 80° in the 2θ scale, with a scanning speed of 0.02° per second and a step time of 3 seconds. A Perkin Elmer Spectrum 10 FTIR spectrometer was used for the FTIR analyses in which a Universal Attenuated Total Reflectance (ATR) component was bound to it. The analyses were performed at RT, and the data needed was obtained using the Spectrum ® Software. Transmission electron microscopy (TEM) was carried out under a JEM-1010 Transmission Electron Microscope operated at an accelerated voltage of 100 kV. The Mega View III Soft Imaging Systems (SIS) side-mounted 3-megapixel digital camera was used to document the micrographs. Selection and visualization of samples were performed using SIS iTEM software. The surface morphology of t was investigated using Scanning Electron Microscopy (SEM). Samples were coated with gold using a Q150R Rotary-Pumped Sputter Coater and viewed under a Zeiss Ultra Plus FE-SEM (Field Emission Scanning Electron Microscope) at a magnification of 3500 X. Colloidal stability. Zeta-potential measurements were carried out using a Malvern Nano Sight NS500 in distilled water at 25 ℃. Data analyses were performed using the Nano Sight NTA 3.2 Software. Magnetization measurements were performed at room temperature using the Lake Shore Model 735 Vibrating Sample Magnetometer (VSM) in applied fields up to 14 kOe.

3. Results and Discussion

3.1 Structural and phase analysis

Figure 1 shows the XRD patterns of $CaFe_2O_4$, CH-$CaFe_2O_4$ and PVA-$CaFe_2O_4$ NPs. All spectra exhibit single spinel structures corresponding to diffraction peaks of the orthorhombic form according to (JCPDS no. 22-1086). No impurities were observed in the spectra. The broad XRD peaks indicate the presence of fine particles. The crystalline sizes (D) listed in Table 1 were calculated using the Debye Scherrer formula [24], $D = \frac{K\lambda}{\beta \cos\theta}$, where K is the Scherer constant ($K = 0.89$), λ is the X-ray

wavelength, β is the peak width at half maximum, and θ is the Braggs diffraction angle. It's worth noticing that the crystallite size of the as-prepared $CaFe_2O_4$ NPs was 6.43 nm, significantly reduced to 5.61 nm for CH-$CaFe_2O_4$, while slightly declined to 6.36 nm for PVA-$CaFe_2O_4$ NPs. The reduction of the crystallite size after chitosan coating may be associated with a stable shell over the NPs to prevent agglomeration of the magnetic core, while the increase with PVA coating suggests the formation of clusters due to PVA polymer. This behaviour has been earlier reported by Nadeem *et.al* [27]. However, the presence of a non-magnetic Ca^{2+} prevented the formation of these clusters as the interparticle interaction in their composition causes a lesser magnetic moment. Thus, there is no significant increase in the structural and magnetic behaviour in uncoated $CaFe_2O_4$ and CH-$CaFe_2O_4$ NPs. Hence, the PVA coating suggests the presence of atoms with magnetic ions.

The values of the lattice parameter (a) also listed in Table 1, were calculated from XRD data using Bragg's law and the equation [28] $d = \frac{a}{\sqrt{h^2+k^2+l^2}}$, where, (h, k, l) are the Miller indices of the indexing plane of the crystal planes and d is line spacing. The lattice constant for uncoated $CaFe_2O_4$ NPs was obtained to be 8.16 Å, a slight change to 8.11 Å and 8.15 Å was noted for CH-CaF_2O_4 and PVA $CaFe_2O_4$ NPs, respectively. Hence, no effect on the spinel ferrites resulted from the coatings. Therefore, both coatings might have caused other structural properties changes, including the lattice constant values in both polymers coated $CaFe_2O_4$ NPs. Nevertheless, in this instance, both coatings had a negligible effect on the cationic distribution and the spinel lattice parameter constants at the tetrahedral and octahedral sites, respectively.

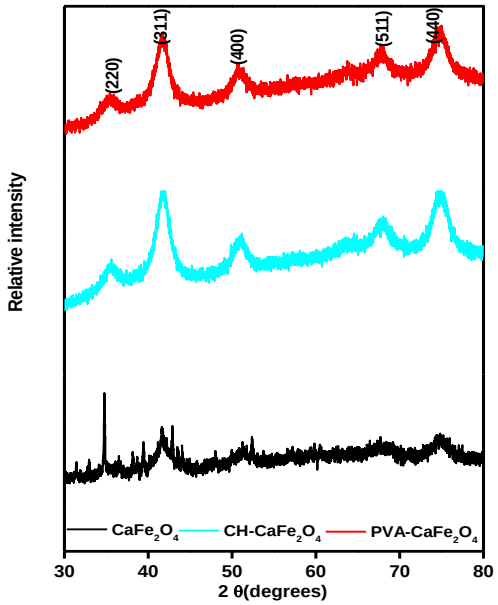


Figure 1: XRD patterns for $CaFe_2O_4$, CH-$CaFe_2O_4$ and PVA- $CaFe_2O_4$ nanoparticles.

Table 1: Average particle measurements of MNPs from XRD data analysis

Samples	D_{XRD} (nm)	Lattice parameters (Å)
$CaFe_2O_4$	6.43	8.16
CH-$CaFe_2O_4$	5.61	8.11
PVA-$CaFe_2O_4$	6.36	8.15

FTIR spectroscopy is the most common vibrational technique used to identify several functional groups introduced by the attachment of the coating polymers on the surface of nanoparticles. FTIR spectrum of the PVA alone in Figure 2 revealed two additional absorption peaks at 2364 cm^{-1} and 2166 cm^{-1},

which were also observed on the PVA-$CaFe_2O_4$ NPs spectrum, confirming the coating. Coating using chitosan broadens the absorption peaks compared to uncoated $CaFe_2O_4$ and PVA-$CaFe_2O_4$ NPs. Other peaks were observed at 1094 cm^{-1} and 1660 cm^{-1}, respectively, indicating the vibration of the CH_2 groups and the bending vibration of N–H of the primary amine groups (NH_2). Also, it confirmed the functionalization of chitosan on the surface of calcium ferrites. Several weaker peaks between 560 cm^{-1} and 910 cm^{-1} correspond to the metal-oxygen bonds (Fe–O or Ca–O) bending or stretching bonds. When a molecule absorbs infrared radiation, its chemical bonds vibrate and confirm the spinel structure, which remains unaffected by the coating. For uncoated $CaFe_2O_4$ NPs, a broad peak at 3401 cm^{-1} was attributed to O–H bonds due to the water and silanol groups engaged in the hydrogen bonds. The peaks around 3296 cm^{-1} were attributed to the H-O-H asymmetrical stretching vibration bonds of the absorbed or free water molecules.

In contrast, the peak at 910 cm^{-1} is attributed to Fe–OH bending bonds. The observed sharp peaks at 377 cm^{-1} and 560 cm^{-1} correspond to the metal-oxygen bending and stretching bonds; when a molecule absorbs infrared radiation, its chemical bonds vibrate. These bonds can stretch and bend due to the metallic particles (Fe^{3+} and Ca^{2+} ions) occupying the tetrahedral and octahedral lattice positions, respectively. Furthermore, it confirmed the spinel orthorhombic structure of $CaFe_2O_4$ NPs.

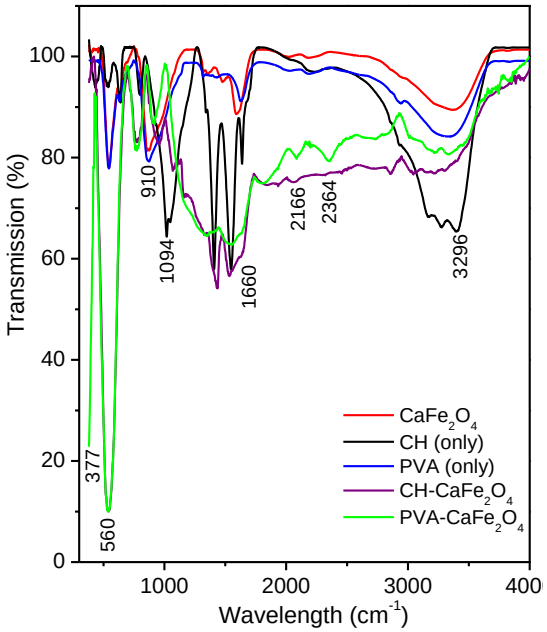


Figure 2: FTIR spectra for $CaFe_2O_4$, CH-$CaFe_2O_4$ and PVA-$CaFe_2O_4$ nanoparticles**.**

3.2 HRTEM, SEM and EDS analysis

The microstructure of the nanoparticles was further studied with the aid of high-resolution transmission electron microscopy (HRTEM), scanning electron microscopy (SEM), and energy dispersion spectra (EDS) at room temperature. HRTEM images in Figure 3 confirm the crystalline nature of the uncoated NPs with spherical shapes. It may be noted that the image of uncoated NPs revealed larger particles, suggesting the agglomeration, as seen by larger clusters and amorphous surface structure in the image below. After polymer-coating, NPs became more spherical with a slight decrease in size (from 6.31 nm to 6.00 nm) for the CH-coated NPs. A compaction of size with chitosan coating has been previously reported [29], [30]. On the other hand, PVA resulted in increased sizes (6.85 nm) as per previous studies by Ramnandan *et al*. [31]. Furthermore, the NPs were monodispersed with clear surface edges, which correlates to the strong peak intensities from XRD.

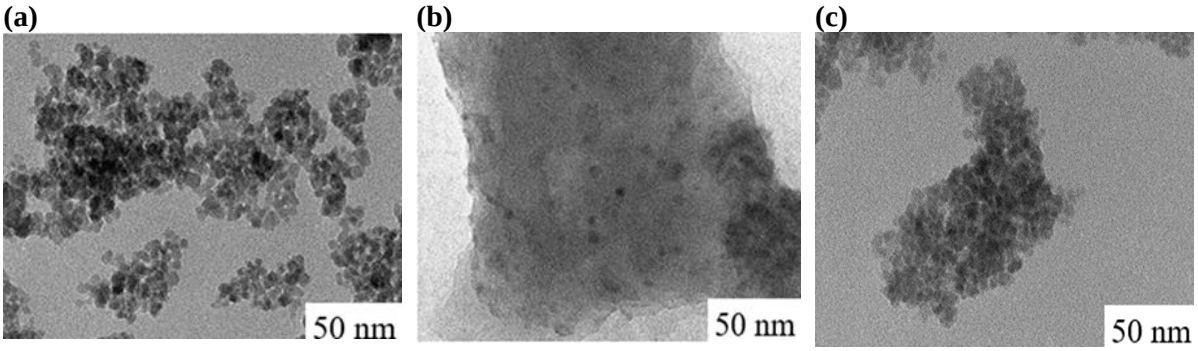


Figure 3: HRTEM micrographs; (a) $CaFe_2O_4$, (b) CH-$CaFe_2O_4$ and (c) PVA-$CaFe_2O_4$ NPs.

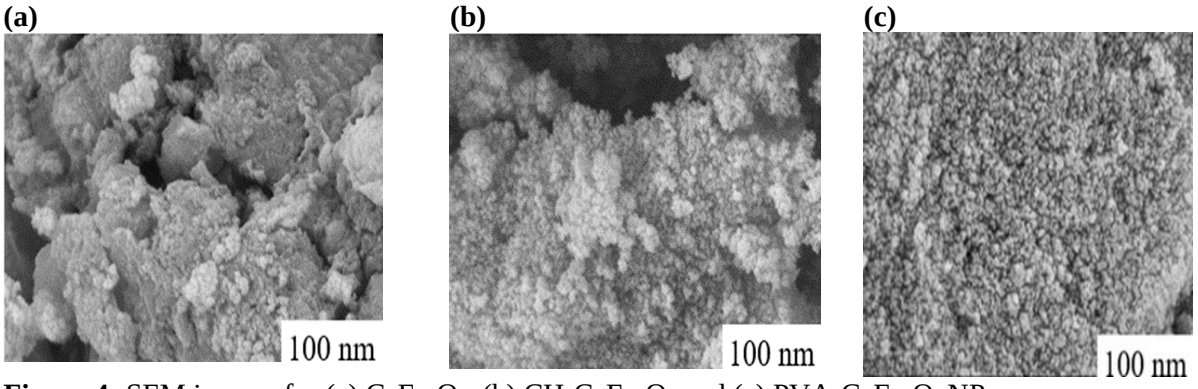


Figure 4: SEM images for (a) $CaFe_2O_4$, (b) CH-$CaFe_2O_4$ and (c) PVA-$CaFe_2O_4$ NPs.

The particles are observed to be well separated upon coating with the polymers; however, they are more pronounced for CH-coated NPs. Hence, chitosan may be preferred since the NPs are early isolated, enhancing the magnetic NPs' stability. The HRTEM results correlated well with the SEM results, where spherical-shaped NPs were obtained seen in Figure 4. Both HRTEM and SEM results suggest good nanocrystalline growth and led to a decrease in particle sizes, also observed from XRD results. Furthermore, hydrophilic CH and PVA coating forms a non-ionic surface layer, reducing the interparticle interaction, thus minimising agglomeration of uncoated $CaFe_2O_4$ NPs and enhancing their dispersibility. EDS data presented in Table 2 reveal no impurities in all samples; however, it confirms the synthesis of pure $CaFe_2O_4$ NPs whereby Ca, Fe and O were the only chemical elements with varied % weight abundance. Hence, the reported quantitative chemical composition correlates with the desired compound ($CaFe_2O_4$).

Table 2: Elemental composition (%); $CaFe_2O_4$ NPs, CH-$CaFe_2O_4$ and PVA-$CaFe_2O_4$ NPs.

Weight (%)	$CaFe_2O_4$	CH-$CaFe_2O_4$	PVA-$CaFe_2O_4$
Calcium (Ca)	15.21	2.23	1.03
Iron (Fe)	9.24	53.52	88.48
Oxygen (O)	75.54	44.24	10.48

3.3 Vibrating Sample Magnetometer (VSM) analysis

The M-H curves measured at room temperature using the VSM for the as-prepared and polymer-coated $CaFe_2O_4$ NPs are represented in Figure 5. The *S*-shape of the curves exhibits a superparamagnetic nature in all samples. This property of NPs suggests that the NPs can only be magnetized in the presence of the external field, which is desired for medical applications [32], [33]. The magnetic parameters, such as the saturation magnetization (M_s) and coercivity (H_c), are listed in Table 3. These parameters were

recorded at the applied magnetic field up to 14 kOe. The value of the M_s of $CaFe_2O_4$ NPs was significantly reduced from 38.520 emu/g to 11.272 emu/g after coating with chitosan, while slightly changed to 37.726 emu/g for PVA-coated NPs. The reduction of saturation magnetization of the NPs after coating could be attributed to non-magnetic substances of the polymers on the surface of NPs. However, the magnetization of the coated NPs was still enough to respond in the presence of an external magnetic field. In addition, there were significant small coercivity values for all NPs, confirming the superparamagnetic characteristic of the nanoparticles. Furthermore, the polymer interaction with the magnetic core of $CaFe_2O_4$ NPs led to the formation of the magnetic dead layers at the surface as they formed weaker bonds, especially with chitosan coating. Hence, chitosan and PVA result in a shielding effect on the magnetic NPs.

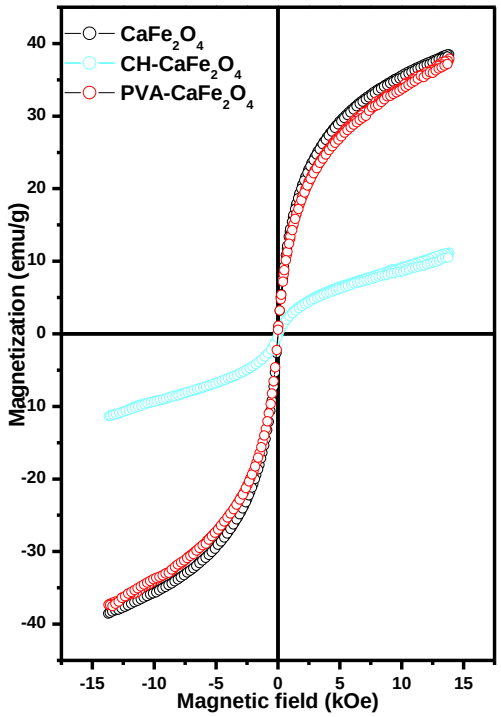


Figure 5: The hysteresis loops $CaFe_2O_4$, CH-$CaFe_2O_4$, PVA-$CaFe_2O_4$ nanoparticles.

Table 3: Saturation magnetization and coercivity measurements of nanoparticles using VSM.

Nanoparticles	Saturation magnetization (M_s) (emu/g)	Coercivity (H_c) (kO_e)
$CaFe_2O_4$	38.520	5.493
CH-$CaFe_2O_4$	11.272	8.190
PVA-$CaFe_2O_4$	37.726	7.271

Magnetic NPs with higher magnetization are essential in biomedical applications as they show higher sensitivity and efficiency for heat generation when combined with the external magnetic field for potential magnetic hyperthermia in cancer treatment [34]. The heating capacity of hyperthermia nanomaterials is obtained by evaluating the specific loss power, which is related to the saturation magnetization and coercivity of the NPs. Moreover, the heating capability of magnetic NPs depends upon various components, such as particle size, shape, and morphology, like outward parameters like the associated AC magnetic field. Furthermore, the coercivity (H_c) value reflects the intensity of the

applied magnetic field necessary to demagnetize that material after an external field has saturated the magnetization of the NPs. This study's H_c value for $CaFe_2O_4$ (5.493 kOe) increased as chitosan and PVA coatings to 8.190 kOe and 7.271 kOe, respectively. This increase is reported by Socoliuc *et al*. to confirm the formation of thick layers over the magnetic core, which shifted their magnetic response [35] .

3.4 Colloidal stability and hydrodynamic size

The colloidal stability of the as prepared and coated NPs was determined using zeta potential measurements. The zeta potential charge distribution and the hydrodynamic sizes of the NPs were measured and are summarized in Table 4. The hydrodynamic sizes for $CaFe_2O_4$ increased with coatings from 92.0 nm to 169.4 nm and 150.6 nm reported for CH-$CaFe_2O_4$ and PVA-$CaFe_2O_4$ NPs, respectively. However, these hydrodynamic sizes below 200 nm are still well-suitable for biomedical applications [36]. The increase in size might be due to polymers forming thick surface layers, and the observed results correlate with that coating led to an increase in particle size. A near-neutral zeta potential value of -5.7 mV was observed for PVA-$CaFe_2O_4$ NPs, compared to -20.2 mV and -22.5 mV, recorded for the as-prepared $CaFe_2O_4$ and CH-$CaFe_2O_4$ NPs. These results suggest that the non-ionic PVA did not improve the chemical stability of the NPs. However, it formed weaker bonds with the magnetic core. These results were in contracts with those reported by Ramnandan *et al*., [31].

Furthermore, chitosan coating improved the colloidal stability of the as-prepared NPs, recording a higher zeta potential value [31]. Also, the solubility in aqueous media or water increases due to the hydrophilic nature of the chitosan functional groups, which form strong bonds with the naked magnetic core. Colloidal stability in aqueous media (water and PBS) is essentially for magnetic NPs, especially for biomedical applications such as magnetic fluid hyperthermia and magnetically targeted drug delivery systems. Hence, CH-coating enhanced the colloidal stability of the uncoated $CaFe_2O_4$ MNPs in physiological pH and showed ideal potential for biomedical applications.

Table 4: Hydrodynamic size and zeta potential of NPs using NTA.

Samples	Hydrodynamic size (nm)	Zeta potential (mV)
$CaFe_2O_4$	92.0 ± 3.5	-20.2 ± 4.5
CH-$CaFe_2O_4$	169.4 ± 33.6	-22.5 ± 4.2
PVA-$CaFe_2O_4$	150.6 ± 14.2	-5.7 ± 0.8

4. Conclusion

This study synthesized a single orthorhombic spinel phase nanocrystalline of $CaFe_2O_4$ NPs via the glycol-thermal reaction method. After that, their surface was functionalized with chitosan and polyvinyl alcohol polymers. From the XRD pattern, a spinel orthorhombic structure remained unaffected after chitosan and PVA coatings, leading to broad and strong peaks observed with a reduction in crystalline sizes. Moreover, chitosan and PVA-coated $CaFe_2O_4$ NPs recorded a decrease in the calculated structural and magnetic properties compared to the uncoated sample. The superparamagnetic properties remained intact after coatings. Improvement of colloidal stability after chitosan coating highlighted the greater potential of CH-$CaFe_2O_4$ NPs for biomedical applications as chitosan forms more stable bonds with $CaFe_2O_4$ NPs. Hence, uncoated $CaFe_2O_4$ NPs showed size-dependent structural and magnetic properties with or without coating. In future, surface coating highlighted favourable potentials for biomedical applications. These results also imply that surface coating of $CaFe_2O_4$ NPs might enhance their dispersion in aqueous media or water and heating capability under physiological conditions to further their potential in biomedical applications.

Acknowledgements

The National Research Foundation (NRF), Pretoria, South Africa, for the funding through the Thuthuka Grant. The Non-Viral Gene and Drug Delivery Research Lab, Biochemistry (UKZN). The data presented in this manuscript was collected and analysed by an Honours student, Sibusiso Mtshali, no longer affiliated with any institution.

References

[1] Y. Wang *et al.*, "Engineering ferrite nanoparticles with enhanced magnetic response for advanced biomedical applications," *Mater Today Adv*, vol. 8, 2020, doi: 10.1016/j.mtadv.2020.100119.

[2] M. Kalubowilage, K. Janik, and S. H. Bossmann, "Magnetic nanomaterials for magnetically-aided drug delivery and hyperthermia," *Applied Sciences (Switzerland)*, vol. 9, no. 14. MDPI AG, Jul. 01, 2019. doi: 10.3390/app9142927.

[3] H. Du, O. U. Akakuru, C. Yao, F. Yang, and A. Wu, "Transition metal ion-doped ferrites nanoparticles for bioimaging and cancer therapy," *Transl Oncol*, vol. 15, no. 1, p. 101264, 2022, doi: 10.1016/j.tranon.2021.101264.

[4] Z. Hedayatnasab, A. Dabbagh, F. Abnisa, and W. M. A. Wan Daud, "Polycaprolactone-coated superparamagnetic iron oxide nanoparticles for in vitro magnetic hyperthermia therapy of cancer," *Eur Polym J*, vol. 133, no. May, p. 109789, 2020, doi: 10.1016/j.eurpolymj.2020.109789.

[5] M. S. A. Darwish, H. Kim, H. Lee, C. Ryu, J. Y. Lee, and J. Yoon, "Engineering core-shell structures of magnetic ferrite nanoparticles for high hyperthermia performance," *Nanomaterials*, vol. 10, no. 5, May 2020, doi: 10.3390/nano10050991.

[6] K. K. Kefeni, T. A. M. Msagati, and B. B. Mamba, "Ferrite nanoparticles: Synthesis, characterisation and applications in electronic device," *Materials Science and Engineering B: Solid-State Materials for Advanced Technology*. 2017. doi: 10.1016/j.mseb.2016.11.002.

[7] S. R. Mokhosi, W. Mdlalose, A. Nhlapo, and M. Singh, "Advances in the synthesis and application of magnetic ferrite nanoparticles for cancer therapy," *Pharmaceutics*, vol. 14, no. 5, 2022, doi: 10.3390/pharmaceutics14050937.

[8] R. G. D. Andrade, S. R. S. Veloso, and E. M. S. Castanheira, "Shape anisotropic iron oxide-based magnetic nanoparticles: Synthesis and biomedical applications," *International Journal of Molecular Sciences*, vol. 21, no. 7. MDPI AG, Apr. 01, 2020. doi: 10.3390/ijms21072455.

[9] M. N. Zharkov *et al.*, "Enhanced cytotoxicity caused by AC magnetic field for polymer microcapsules containing packed magnetic nanoparticles," *Colloids Surf B Biointerfaces*, vol. 199, Mar. 2021, doi: 10.1016/j.colsurfb.2020.111548.

[10] M. E. Maffei, "Magnetic fields and cancer: Epidemiology, cellular biology, and theranostics," *Int J Mol Sci*, vol. 23, no. 3, Feb. 2022, doi: 10.3390/ijms23031339.

[11] Y. I. Golovin *et al.*, "Non-heating alternating magnetic field nanomechanical stimulation of biomolecule structures via magnetic nanoparticles as the basis for future low-toxic biomedical applications," *Nanomaterials*, vol. 11, no. 9. MDPI, Sep. 01, 2021. doi: 10.3390/nano11092255.

[12] H. T. Phan and A. J. Haes, "What Does Nanoparticle Stability Mean? HHS Public Access," *J Phys Chem C Nanomater Interfaces*, vol. 123, no. 27, pp. 16495–16507, 2019, doi: 10.1021/acs.jpcc.

[13] N. Zhu *et al.*, "Surface modification of magnetic iron oxide nanoparticles," *Nanomaterials*, vol. 8, no. 10, pp. 1–27, 2018, doi: 10.3390/nano8100810.

[14] S. Shitole, "Nanoparticle - Novel drug delivery system: A review," *Int J Res Appl Sci Eng Technol*, vol. 10, no. 1, pp. 517–530, 2022, doi: 10.22214/ijraset.2022.39849.

[15] N. Yadav and M. Ahmaruzzaman, "Recent advancements in $CaFe_2O_4$-based composite: Properties, synthesis, and multiple applications," *Energy and Environment*. SAGE Publications Inc., 2023. doi: 10.1177/0958305X231155491.

[16] T. Dippong, E. A. Levei, and O. Cadar, "Recent advances in synthesis and applications of MFe_2O_4 (M = Co, Cu, Mn, Ni, Zn) nanoparticles," *Nanomaterials*, vol. 11, no. 6. MDPI AG, Jun. 01, 2021. doi: 10.3390/nano11061560.

[17] B. Huzum *et al.*, "Biocompatibility assessment of biomaterials used in orthopedic devices: An overview (Review)," *Exp Ther Med*, vol. 22, no. 5, Sep. 2021, doi: 10.3892/etm.2021.10750.

[18] K. Zipare, J. Dhumal, S. Bandgar, V. Mathe, and G. Shahane, "Superparamagnetic Manganese Ferrite Nanoparticles: Synthesis and Magnetic Properties," *Journal of Nanoscience and Nanoengineering*, vol. 1, no. 3, pp. 178–182, 2015, [Online]. Available: http://www.aiscience.org/journal/jnn

[19] K. Zipare, J. Dhumal, S. Bandgar, V. Mathe, and G. Shahane, "Superparamagnetic manganese ferrite nanoparticles: Synthesis and magnetic properties," *Journal of Nanoscience and Nanoengineering*, vol. 1, no. 3, pp. 178–182, 2015, [Online]. Available: http://www.aiscience.org/journal/jnn

[20] H. Hermawan, D. Ramdan, and J. R. P. Djuansjah, "World's largest Science, Technology & Medicine Open Access book publisher Metals for Biomedical Applications," no. 2016, pp. 267–322, 2018.

[21] A. G. Niculescu, C. Chircov, and A. M. Grumezescu, "Magnetite nanoparticles: Synthesis methods – A comparative review," *Methods*, vol. 199, pp. 16–27, Mar. 2022, doi: 10.1016/j.ymeth.2021.04.018.

[22] S. Fayazzadeh, M. Khodaei, M. Arani, S. R. Mahdavi, T. Nizamov, and A. Majouga, "Magnetic properties and magnetic hyperthermia of cobalt ferrite nanoparticles synthesized by hydrothermal method," *J Supercond Nov Magn*, vol. 33, no. 7, pp. 2227–2233, Jul. 2020, doi: 10.1007/s10948-020-05490-6.

[23] N. H. Sulaiman, M. J. Ghazali, B. Y. Majlis, J. Yunas, and M. Razali, "Superparamagnetic calcium ferrite nanoparticles synthesized using a simple solgel method for targeted drug delivery," *Biomed Mater Eng*, vol. 26, pp. S103–S110, 2015, doi: 10.3233/BME-151295.

[24] W. B. Dlamini, J. Z. Msomi, and T. Moyo, "XRD, Mössbauer and magnetic properties of $Mg_xCo_{1-x}Fe_2O_4$ nanoferrites," *J Magn Magn Mater*, vol. 373, pp. 78–82, 2015, doi: 10.1016/j.jmmm.2014.01.066.

[25] M. Khalkhali, K. Rostamizadeh, S. Sadighian, F. Khoeini, M. Naghibi, and M. Hamidi, "The impact of polymer coatings on magnetite nanoparticles performance as MRI contrast agents: A comparative study," *DARU, Journal of Pharmaceutical Sciences*, vol. 23, no. 1, pp. 1–12, 2015, doi: 10.1186/s40199-015-0124-7.

[26] C. O. Ehi-Eromosele, B. I. Ita, and E. E. J. Iweala, "The effect of polyethylene glycol (PEG) coating on the magneto-structural properties and colloidal stability of $Co_{0.8}Mg_{0.2}Fe_2O_4$ nanoparticles for potential biomedical applications," *Dig J Nanomater Biostruct*, vol. 11, no. 1, pp. 7–14, 2016.

[27] M. Nadeem *et al.*, "Magnetic properties of polyvinyl alcohol and doxorubicine loaded iron oxide nanoparticles for anticancer drug delivery applications," *PLoS One*, vol. 11, no. 6, Jun. 2016, doi: 10.1371/journal.pone.0158084.

[28] J. Z. Msomi, W. B. Dlamini, T. Moyo, and P. Ezekiel, "Investigation of phase formation of $(Zn, Mg)_{0.5}Co_{0.5}Fe_2O_4$ nanoferrites," *J Magn Magn Mater*, vol. 373, pp. 68–73, 2015, doi: 10.1016/j.jmmm.2014.01.044.

[29] M. G. Pineda *et al.*, "Chitosan-coated magnetic nanoparticles prepared in one-step by precipitation in a high-aqueous phase content reverse microemulsion," *Molecules*, vol. 19, no. 7, pp. 9273–9287, 2014, doi: 10.3390/molecules19079273.

[30] S. Zeinali, S. Nasirimoghaddam, and S. Sabbaghi, "Investigation of the synthesis of chitosan coated iron oxide nanoparticles under different experimental conditions," *Int. J. Nanosci. Nanotechnol.*, vol. 12, no. 3, pp.183–190, Sept. 2016.

[31] D. Ramnandan, S. Mokhosi, A. Daniels, and M. Singh, "Chitosan, polyethylene glycol and polyvinyl alcohol modified $MgFe_2O_4$ ferrite magnetic nanoparticles in doxorubicin delivery: A comparative study in vitro," *Molecules*, vol. 26, no. 13, Jul. 2021, doi: 10.3390/molecules26133893.

[32] E. M. Materón *et al.*, "Magnetic nanoparticles in biomedical applications: A review," *Applied Surface Science Advances*, vol. 6, Dec. 2021, doi: 10.1016/j.apsadv.2021.100163.

[33] H. M. Williams, "The application of magnetic nanoparticles in the treatment and monitoring of cancer and infectious diseases," *Bioscience Horizons: The International Journal of Student Research*, vol. 10, Aug. 2017, doi: 10.1093/biohorizons/hzx009.

[34] S. R. Mokhosi, W. Mdlalose, A. Nhlapo, and M. Singh, "Advances in the synthesis and application of magnetic ferrite nanoparticles for cancer therapy," *Pharmaceutics*, vol. 14, no. 5. MDPI, May 01, 2022. doi: 10.3390/pharmaceutics14050937.

[35] V. Socoliuc *et al.*, "Magnetic nanoparticle systems for nanomedicine—A materials science perspective," *Magnetochemistry*, vol. 6, no. 1. MDPI AG, pp. 1–36, Mar. 01, 2020. doi: 10.3390/magnetochemistry6010002.

[36] S. M. Mngadi, S. R. Mokhosi, and M. Singh, "Surface-coating of $Mg_{0.5}Co_{0.5}Fe_2O_4$ nanoferrites and their in vitro cytotoxicity," *Inorg Chem Commun*, vol. 108, no. June, p. 107525, 2019, doi: 10.1016/j.inoche.2019.107525.

Effect of concentrator geometry on wind velocity augmentation

Chipo Shonhiwa[1]; Golden Makaka[1]; Patrick Mukumba[1]; Ngwarai Shambira[1]

[1]University of Fort Hare, Physics Department, 1 King Williams Town Road, Private Bag X1314, Alice 5700, South Africa.

E-mail: **cshonhiwa@ufh.ac.za**

Abstract. South Africa is experiencing acute energy shortages leading to prolonged load shedding. Renewable energy can be a solution to this energy problem. From the renewable energy resources, wind has a great potential of increasing the energy mix and help in abatement of climate change. It is one of the fastest growing energy technologies that can substitute fossil fuels in electricity production. Most wind turbines that are available on the market need wind speeds above 5 m/s to produce meaningful electricity. Regrettably some areas don't have such high wind speeds. Several researches have been done in coming up with techniques for improving wind turbine power output in low wind speed areas. Concentrator Augmented Wind Turbines (CAWTs) have been proved to increase the effectiveness of the wind turbines in low wind speed areas by increasing the wind speed upstream of the turbine. However, the effect of concentrator geometry on the velocity augmentation capability of the concentrator is not clearly understood to enhance the designing of competitive CAWTs. This study investigated the effect of concentrator geometry on velocity augmentation. Computational fluid dynamics was used for the investigations. A model was analysed in ANSYS Fluent® software The performance of 45 concentrators was investigated in terms of velocity augmentation ratio ($v_{r,s}$). The concentrators had the same outlet radius and their geometry was varied by changing the concentrator length (L) and the inlet diameter (D_o). It was observed that the changes $v_{r,s}$ was affected by the change in both the L to D_o ratio (L_r) and the difference between inlet and outlet radii to D_o ratio (R_r). As L_r was increased from $L_r = 0.1$, the $v_{r,s}$ increased to pick at $L_r = 0.4$ and then decreased with further increase in L_r. Also, as R_r was increased from 0.025, the $v_{r,s}$ increased to peak at $R_r = 0.1$ and then decreased as R_r was increased. It was concluded that too short and too long concentrators are not ideal for velocity augmentation. It was recommended that very short concentrators such as those with $L_r = 0.1$ should not be used since the concentrator outlet velocity would be almost the same as the inlet velocity.

1. Introduction

Wind has a very low energy density in nature implying that a bigger rotor area is required to harness kinetic energy for generation of meaningful electrical energy [1] According to Betz, the theoretical maximum power coefficient of a conventional wind turbine is limited to about 59% but in practice this figure is reduced to about 40% due to aerodynamic losses [2]. This results in higher costs of producing energy from wind than conventional sources of energy such as fossil fuels and hydro. To lower the cost of wind energy, researchers have come up with several initiatives to boost wind turbine power output.

Several mechanisms, which can be used individually or in combination have been proposed to increase the turbine power output per unit rotor area. These include increasing the mass flow passing through the turbine and boosting the specific energy content of the wind [3]. The concepts that have become popular in research have been found to have a number of advantages over conventional wind turbines such as: an augmented power output per unit area swept by the rotor, meaningful power generation in areas of low wind speeds and reduced costs of electrical energy generation which results from increased rotor rotation per minute hence lowering cost of gearbox, drive-train and generator [4]

The concept of ducted wind turbines, which includes the use of diffusers and concentrators has been on the research agenda for decades, with the first in-depth study around mid-20th century, but to date there has been no successful commercial designs [5]. The rationale behind continued studies is that an insignificant rise in wind speed due to inclusion of ducts on conventional turbine systems can cause a huge increase in power output, since wind turbine power output is proportional to the cube of the wind speed [6]. Much research effort on ducted turbines has focused on diffuser augmented wind turbines (DAWTs) [3,7,8]. A number of studies have been done on use of concentrators on vertical axis wind turbines and proved that a power augmentation factor of about 3.7 can be achieved [9]. There are limited studies involving use concentrators on horizontal axis wind turbines.

The concentrator augmented wind turbines (CAWTs) concept involves the use of a funnel-shaped duct to capture wind from a larger area and deliver it to the rotor through a smaller area thus increasing the mass flow rate. A detailed review of work that has been done an effort to increase wind speed using CAWTs is given by Shonhiwa and Makaka [10].

Currently not much is known about the CAWT concept using horizontal axis wind turbines (HAWTs). More work still needs to be done to understand the mechanisms involved in increasing the CAWTs' power output in order to evaluate the concept and come up with a cost-effective system. There is need to understand the flow mechanism inside the concentrator and the dependence on each other of different concentrator geometry parameters and their influence on power output. Thus, this research sought to investigate the effect of a simple concentrator geometry on the wind velocity augmentation in order to determine the optimum design of a concentrator.

2. Methodology

To reduce cost and time on experimental work, (computational fluid dynamics (CFD) modelling was used to investigate effect of concentrator geometry parameter on concentrator performance and give an insight on the air flow fields. To carry out these investigations, a CFD model was analysed in ANSYS Fluent version 2023 R1 because of the availability of a student version which could be used at no cost. Other relevant software such as Phoenics are relatively expensive. This was done to establish dimensionless variables (optimal concentrator ratios), understand characteristics of the flows present in the concentrator and to determine concentrator efficiency. This aided the designing of the CAWT system.

2.1. The concentrator geometry

The geometry and the mesh were created in ANSYS Workbench 2023 R1. For the computed data to be compared with corresponding experimental data, the model interprets exactly the geometry of the concentrator. A cylindrical virtual wind tunnel is 15 concentrator outlet diameters long and 10 concentrator diameters wide to avoid blockage effects. Due to symmetry of the computational domain and to save on computational time and cost, the flow domain was divided into 90^{o} slices using symmetry. A mesh independence test was carried by changing the number of elements of the mesh to check the validity of the quality of the mesh on the solution. A structured tetrahedral mesh was applied.

The incompressible Reynolds-Averaged Navier-Stokes (RANS) equation given by (1) were solved to get the solution.

$$\nabla\vec{v} = 0; \partial\vec{v}/\partial t + \nabla\vec{v} = 1/\rho_a \nabla p + \nabla\kappa + \vec{\phi} \quad (1)$$

where $\vec{v}$ denotes the average velocity vector, $\vec{\phi}$ represents a body force, κ is the specific Reynolds stress tensor, ρ_a is the air density and p is the static pressure.

The boundary conditions were the velocity inlet, the pressure outlet and symmetric conditions at the symmetrical plane of the domain. The centre of the inlet was set as the reference pressure location. The simple pressure-velocity coupling scheme was used. The standard pressure correction method and second order upwind scheme for turbulent kinetic energy and specific dissipation rate were chosen while the other factors were left at default values. The concentrator and virtual wind tunnel walls were set to be no-slip boundaries. Standard solution initialisation method was used with initial values being computed from the inlet.

2.2. Determination of concentrator ratios

To investigate the effect of concentrator geometry on the performance of concentrator, 45 concentrators were tested. The concentrator geometry parameters are shown in Figure 1. L , D_i, Do, R_d = R-r, θ_c are concentrator length, concentrator length diameter, concentrator outlet diameter, difference between inlet and outlet radii, and concentrator tilt angle respectively.

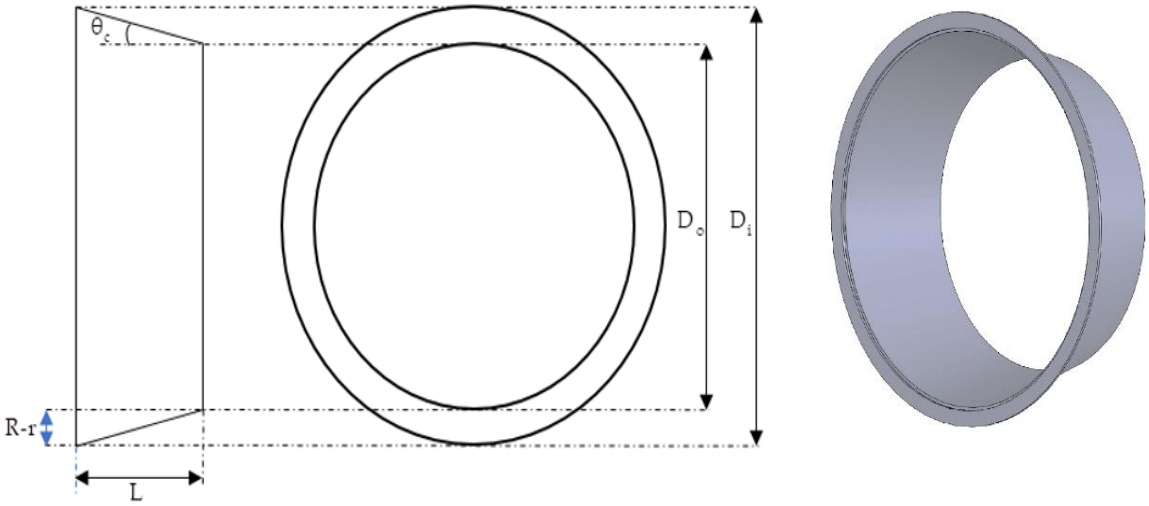


Figure 1: Concentrator geometry parameters.

The performance of the concentrator, which was investigated in terms of velocity augmentation ratio was parameterised by L_r and R_r ratios. Here L_r is a dimensionless length factor which shows the ratio of L to D_o while R_r is the ratio of R_d to D_o. The change in R_d lead to a change in the concentrator inlet area (A_i) to outlet area (A_o) ratio denoted by (A_r).

For making the concentrator six equally spaced L_r were used in the range $L_r = \text{linspace}(0.1,0.6,6)$. Thirteen equally spaced R_r were used in the range $R_r = \text{linspace}(0.025,0.325,13)$. The change in velocity along the concentrator was analysed in Fluent. For each value of L_r , a curve of simulation velocity ratio ($v_{r,s}$) against R_r was plotted on the same graph for all values of L_r using the MATLAB 'hold on' command. This graph was used to determine an optimum L_r by taking the curve with highest values of $v_{r,s}$ at most points. A graph of $v_{r,s}$ against R_r was plotted for the optimum L_r . The optimum R_r was obtained at the maximum turning point of the curve.

3. Results

Figure 2 shows the variation of $v_{r,s}$ with R_r. The velocity augmentation ratio changed with the change in both L_r and R_r. As L_r was increased from $L_r = 0.1$, the $v_{r,s}$ increased to a peak at $L_r = 0.4$ and then started to decrease. With reference to Figure 2, the velocity augmentation capability of short concentrators is low as shown by $v_{r,s}$ having varied from 0.1096 to 1.1 for $L_r = 0.1$ and was in the range $1.128 \le v_{r,s} \le 1.184$ for $L_r = 0.2$. It was also clear that the $v_{r,s}$ curves for $L_r = 0.5$ and 0.6 were lower than the curves for $L_r = 0.3$ and 0.4; thus, the velocity augmentation capability of long concentrators was lower than that of medium length concentrators. For the observed L_r range, the $v_{r,s}$ was low for low R_r; it increased to a peak at optimum R_r and then decreased as R_r was increased. Table 1 shows the $v_{r,s}$ values at the beginning of the L_r range, at optimum R_r and at the end of the L_r range.

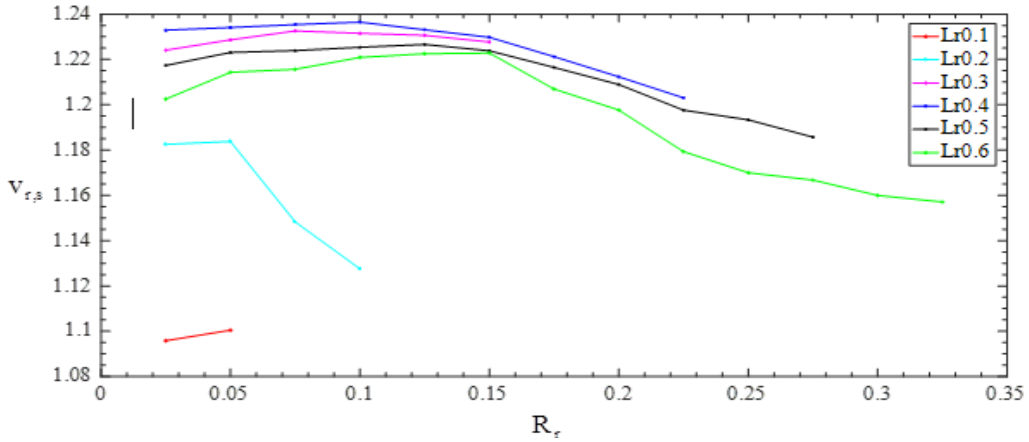


Figure 2: Concentrator velocity augmentation for different L_r and R_r

Table 1: The $v_{r,s}$ values at the beginning of range, at optimum and the end of range

L_r	$v_{r,s}$ at $L_r = 0.1$	Optimum R_r	Optimum $v_{r,s}$	$v_{r,s}$ at end of L_r range
0.1	1.096	0.05	1.1	1.1
0.2	1.182	0.05	1.184	1.128
0.3	1.229	0.075	1.235	1.228
0.4	1.233	0.1	1.236	1.203
0.5	1.217	0.125	1.227	1.186
0.6	1.202	0.15	1.223	1.157

Figure 3 and 4 show the velocity distribution along the concentrator for different L_r. All the concentrators have the same outlet radius. For short concentrators (L_r = 0.1 and 0.2), the highest velocity was attained behind the concentrator outlet for as shown in Figure 3. The value of the highest velocity increased with increasing L_r. As the length of the concentrator was increased, for and above, maximum velocity was attained at the concentrator outlet as shown in Figure (4). However, the value of the highest velocity started to decrease after for all.

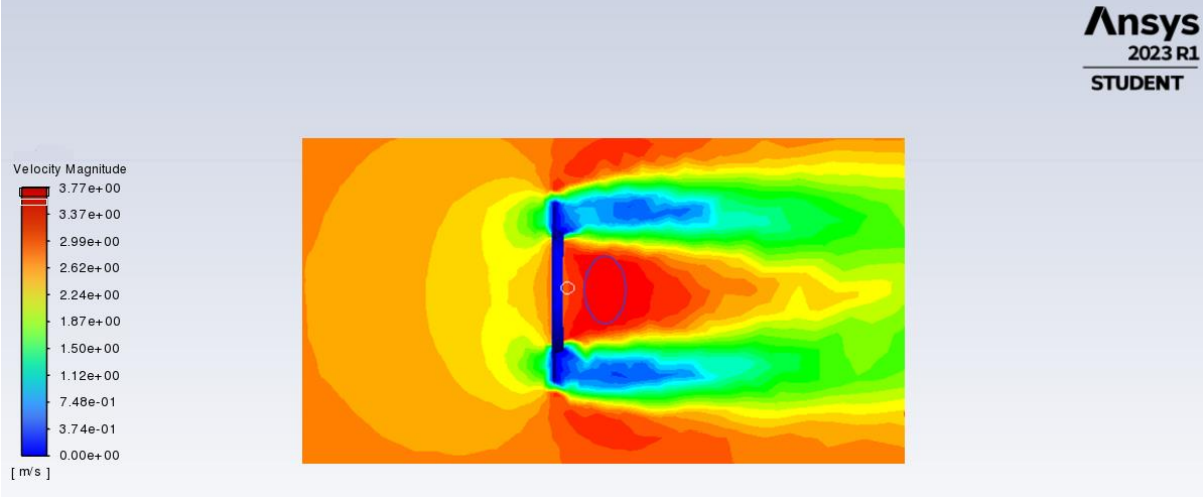


Figure 3: Concentrator outlet velocity $L_r = 0.1$

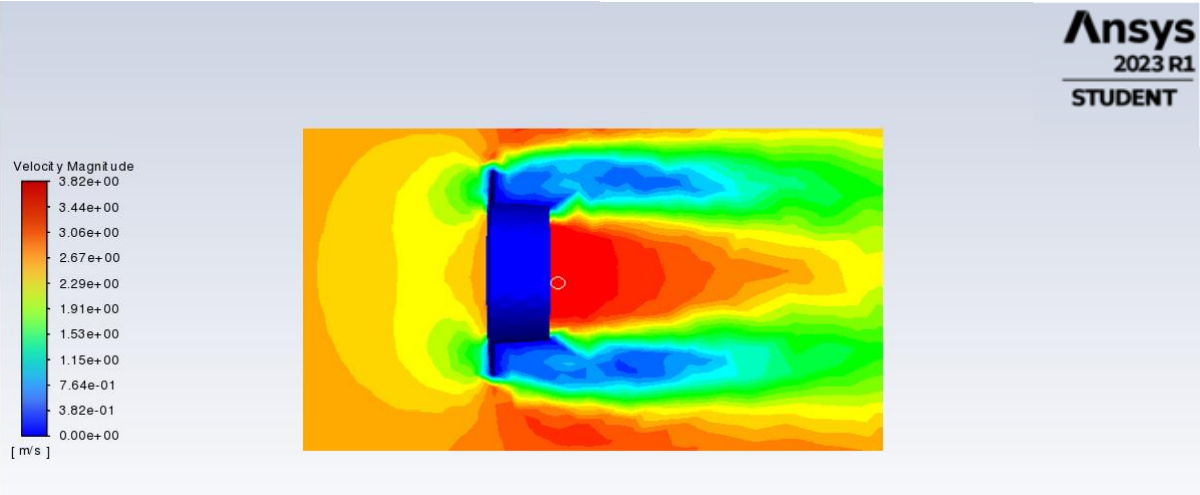


Figure 4: Concentrator outlet velocity $L_r = 0.4$

Figure 5 shows the variation of velocity augmentation ratio with R_r for $L_r = 0.4$. It is shown that $v_{r,s}$ increased from a minimum of 1.233 at $R_r = 0.0125$ to a maximum of 1.234 at $R_r = 0.1$ and decreased thereafter. It is also clear that when R was increased by increasing R_r, the percentage rate of velocity augmentation increase $\left|dv_{r,s}/dr\right|_{inc}$ was less than the percentage rate of velocity augmentation decrease $\left|dv_{r,s}/dr\right|_{dec}$ that is $\left|dv_{r,s}/dr\right|_{inc} < \left|dv_{r,s}/dr\right|_{dec}$. For example, an increase from $R_r = 0.025$ to 0.05 resulted in $\left|dv_{r,s}/dr\right|_{inc} = 4.614\%$ while an increase from $R_r = 0.1$ to 0.125 resulted in $\left|dv_{r,s}/dr\right|_{dec} = 13.447\%$.

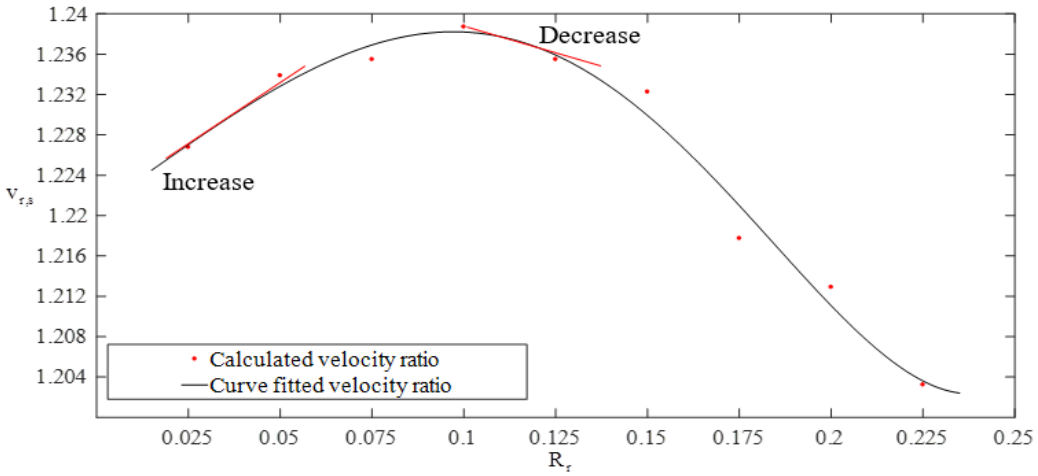


Figure 5: Variation of velocity augmentation with R_r for $L_r = 0.4$

4. Discussion, conclusions and recommendations

The angle θ_c influenced the angle of incidence of wind on the wall surface, which determined the amplification of wind velocity [9]. For short concentrators (low L_r), the friction on the walls occurred on a small range such that the energy losses due to friction were insignificant. However, the angle of incidence was large resulting in reduced component of wind velocity parallel to the concentrator wall. Thus, as the wind entered the concentrator more of it bombarded on the concentrator wall and lost kinetic energy resulting in deceleration which caused a decrease in velocity augmentation. Since wind had moved a relatively small distance, it continued to accelerate after the outlet and attained the maximum speed behind the concentrator outlet.

As L was increased by increasing L_r the incidence angle decreased resulting in an increased component of wind velocity parallel to the wall. Thus, energy loss due to wind bombarding on the concentrator wall was insignificant. However, the frictional losses along the walls occurred on an increased range because the air moved along a longer distance. Therefore, too short or too long concentrators are not good for wind velocity amplification. Thus, length optimisation is necessary to minimise energy losses.

Optimum $v_{r,s}$ of 1.236 was attained at $L_r = 0.4$ and $R_r = 0.1$. The researchers recommended that the parameters be used in the designing of concentrators. They ruled out the use of very short concentrators such as those with $L_r = 0.1$since the concentrator outlet velocity would be almost the same as the inlet velocity.

Acknowledgements

The authors express their gratitude to Govan Mbeki Research & Development Centre (GMRDC) for funding this research.

References

[1] Cloete, S. The fundamental limitations of renewable energy. http://www.theenergycollective.com/schalk-cloete/257351/fundamental-limitations-renewable-energy (Accessed Jun 21, **2014**).

[2] Lilley, G.M.; Rainbird, W.J. A Preliminary Report on the Design and Performance of a Ducted Windmill; Cranfield, **1956**.

[3] Dick, E. Momentum Analysis of Wind Energy Concentrator Systems. Energy Convers. Manag., **1984**, 24, 19–25.

[4] Phillips, D.G. An Investigation on Diffuser Augmented Wind Turbine Design, The University of Auckland, PhD Theses, **2003**.

[5] Shives, M.; Crawford, C. Developing an Empirical Model for Ducted Tidal Turbine Performance Using Numerical Simulation Results In: Proceedings of the Proceedings of the Institution of Mechanical

Engineers, Part A: Journal of Power and Energy; **2011**; Vol. 226, pp. 112–125.
[6] van Bussel, G.J.W. The Science of Making More Torque from Wind: Diffuser Experiments and Theory Revisited. J. Phys. Conf. Ser., **2007**, 75, 012010.
[7] Politis, G.K.; Koras, A.D. A Perfomance Prediction Method for Ducted Medium Loaded Horizontal Axis Windturbines. Wind Eng., **1995**, 19, 272–288.
[8] de Vries, 0. Fluid Dynamic Aspects of Wind Energy Conversion; AGARD: London, **1979**.
[9] Shikha, S.; Bhatti, T.S.; Kothari, D.P. Air Concentrating Nozzles: A Promising Option for Wind Turbines. Int. J. Energy Technol. Policy, **2005**, 3, 394–412.
[10] Shonhiwa, C.; Makaka, G. Concentrator Augmented Wind Turbines: A Review. Renew. Sust. Energ. Rev., **2016**, 59, 1415–1418.

Fast neutron transmission spectroscopy for the non-destructive analysis of concrete

Naledi Segale[1*], Tanya Hutton[1], Andy Buffler[1], Sizwe Mhlongo[1]

[1]Metrological and Applied Sciences University Research Unit (MeASURe), Department of Physics, University of Cape Town, South Africa, 7700.

[*]E-mail: **SGLNXX001@myuct.ac.za**

Abstract. The non-destructive measurements of the elemental composition of existing concrete in nuclear facilities are needed to ensure compliance with international safety standards. In this work, an experimental and simulated verification of the use of fast neutron transmission spectroscopy for the characterization of sand, a critical component of any concrete, is demonstrated. Transmitted neutron energy spectra were measured for a collimated beam of neutrons produced by an ^{241}Am-^{9}Be radioisotopic source incident on samples of sand, and its constituents SiO_2 and $CaCO_3$, and used to determine the effective removal cross section for each. An equivalent setup was simulated in FLUKA to validate the use of Monte Carlo derived removal cross sections. Many of the expected features associated with the different elements were evident in both measured and simulated energy dependent removal cross sections. Minor variations in intensity and a reduced energy resolution were observed in the measured removal cross sections relative to the simulated equivalent and methods to improve agreement are currently underway.

1. Introduction

In nuclear power plants (NPPs), concrete is a preferred choice of construction material, in part, due to its excellent ability to shield against neutron and gamma ray radiation. During its service, concrete is exposed to extreme conditions that become relevant when considering the service life extension of ageing NPPs, particularly with respect to a changing water content [1]. The non-destructive evaluation of existing concrete in terms of radiation shielding and structural properties is thus important to ensure compliance with international safety standards [2].

Fast neutrons can be used for the characterization of bulk samples as they are highly penetrating and are sensitive to light mass elements such as hydrogen, carbon, and oxygen. By measuring the unique radiation signatures produced from neutron interactions, both in energy and intensity, the elemental composition of a sample can be deconvolved [3, 4]. In fast neutron transmission spectroscopy (FNTS), a beam of neutrons is incident on sample of interest, and the neutrons that do not interact within the sample pass through and can be detected. The effective removal cross section (Σ_R) quantifies the probability of a neutron being removed from the fast neutron beam as it goes through a sample and is dependent on the neutron energy and target nucleus [5]. This cross section is related to the total interaction cross section (Fig. 1) of each nuclide, which has distinct features that aid in element identification. The attenuation of a neutron beam through a sample can be described by Eq. 1 where $I(t)$

is the intensity of the attenuated neutron beam for sample thickness *t*, and I_0 is the incident neutron intensity.

$$I(\mathrm{t}) = I_0 \exp(-\Sigma_{\mathrm{R}} t) \quad (1)$$

The removal cross section for a composite sample can be expressed as a linear combination of the individual removal cross sections for each constituent as seen in Eq. 2, where m_k is the mass ratio of component *k*, with corresponding density ρ_k [6].

$$\Sigma_R/\rho = \sum_k m_k \, \Sigma_{R,k}/\rho_k \quad (2)$$

By measuring the energy dependent Σ_R of an unknown sample and comparing that to a data library of Σ_R of the constituents it is possible to determine the mass ratios of the sample constituents through a fitting procedure [7].

A study has been previously undertaken at the University of Cape Town to demonstrate the use of FNTS for the characterization of concrete in which the energy dependent effective removal cross section for sand, cement and water were measured and used to deconvolve the composition of well-defined concrete samples. Nevertheless, this technique can only be applied to recently installed concrete as some of the constituents such as sand are region specific. As such, further experimental verification is still required to generalize the technique [7, 8]. The overall aim of this study is to build a repository of elemental response functions, thus determine the elemental composition of aging concrete used in nuclear facilities. In this work we demonstrate the measurement and simulation processes required to determine the removal cross section and validate the use of these cross sections in the analysis of sand, an integral part of any concrete, with respect to its constituents SiO_2 and $CaCO_3$.

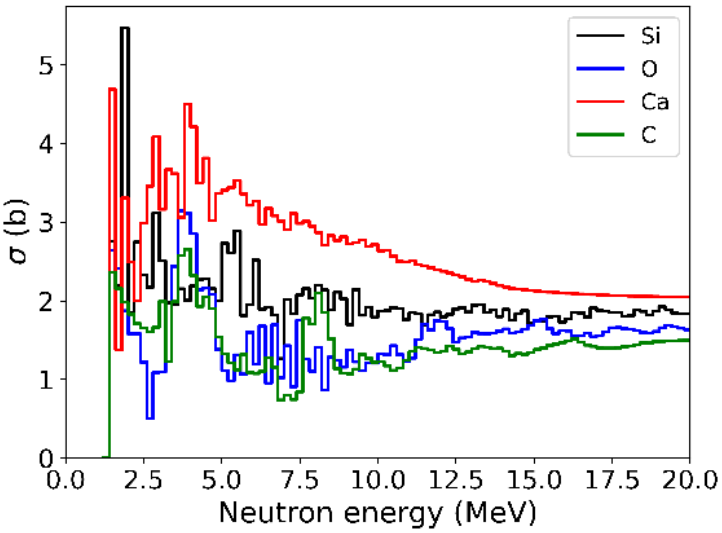


Figure 1: Total microscopic neutron cross sections as a function of energy for the most abundant nuclide of the four elements of interest, ^{28}Si, ^{16}O, ^{40}Ca and ^{12}C, taken from the ENDF/B-VIII.0 nuclear data library [9].

2. Experimental measurements

Measurements were taken at the n-lab, a fast neutron facility in the Department of Physics at the University of Cape Town using a pencil beam of broad-spectrum neutrons produced by a 220 GBq ^{241}Am-^{9}Be (AmBe) radioisotopic source. Figure 2(a) shows a typical experimental setup for FNTS

measurements. The samples consisted of sand, SiO_2 (granular) and $CaCO_3$ (powder) prepared in identical Perspex containers with dimensions $\emptyset$ 6.5 cm $\times$ 10.0 cm as shown in Figure 2(b), with measured mass and density given in Table 1.

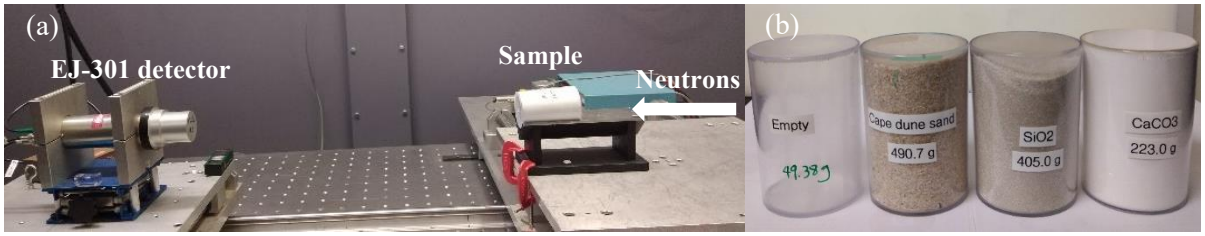


Figure 2: (a) A collimated beam of ^{241}Am-^{9}Be neutrons is incident on a cylindrical sample with dimensions $\emptyset$ 6.5 cm $\times$ 10.0 cm. Transmitted neutrons are measured by an EJ-301 detector. (b) Samples of sand, SiO_2, $CaCO_3$ and an empty container (used for the no sample measurement).

Table 1: Measured mass and calculated density for sand, SiO_2 and $CaCO_3$

Sample	Mass (g)	ρ (g/cm^3)
Sand	490.1 ± 0.1	1.6 ± 0.2
SiO_2	404.9 ± 0.1	1.4 ± 0.1
$CaCO_3$	223.0 ± 0.1	0.8 ± 0.1

Transmitted neutrons were measured with a $2" \times 2"$ EJ-301 organic liquid scintillator optically coupled to an ETL 9214 12-stage photomultiplier tube and base supplied by Scionix and operated at a negative bias of 1.1 kV. The anode and dynode signals were digitized using a CAEN DT5730 [10] and the signals were acquired and processed by the QtDAQ software [11]. The light output parameter L was defined as the pulse height of the dynode after amplification and shaping and was calibrated in units of MeV electron equivalent (MeV_{ee}) through measurements of gamma ray sources such as ^{137}Cs, ^{22}Na, and ^{60}Co. Pulse shape discrimination (PSD) was implemented in QtDAQ using the fast anode pulse to produce a pulse shape parameter S, which was used to separate neutron and gamma ray events. Figure 3 shows a typical L-S distribution measured for an ^{241}Am-^{9}Be source using an EJ-301 detector [12].

After the neutron-only events were isolated the energy spectrum was determined through spectrum unfolding. Spectrum unfolding is a fitting procedure used to deconvolve the energy spectrum from the measured pulse height spectrum using monoenergetic detector response functions. For this work, neutron pulse height spectra were unfolded with GRAVEL, an iterative algorithm based on least squares minimization [13]. The removal cross section was calculated for each energy bin in the spectrum according to Eq. 1, where $I(t)$ was defined as the fluence rate in that bin measured for each sample and I_0 was defined using measurements with an empty container.

3. Monte Carlo Simulations

Simulations were undertaken with FLUKA [14], a Monte Carlo radiation transport code that can simulate interactions and transport of hadrons in any material. An equivalent setup to the one in the n-lab was constructed in FLUKA, with a beam of fast neutrons (1.0 - 11.0 MeV), incident on cylindrical samples of sand, SiO_2 and $CaCO_3$ ($\emptyset$ 6.5 cm $\times$ 10.0 cm), using the same densities as the physical samples (Table 1). Transmitted neutrons were scored in a cylindrical detector region of radius 6.5 cm at 43.0 cm from the sample. The type of scoring used was USRTRACK which scores the average energy dependent neutron fluence in a specified region. For each simulation, 2×10^5 primaries were used over

25 cycles to ensure that statistical uncertainty associated with each energy bin was less than 10 %. The attenuated neutron beam $I(t)$ was defined as the scored fluence distribution for each sample and for I_0 the sample was replaced with vacuum, which were used to determine the energy dependent removal cross section according to Eq. 1.

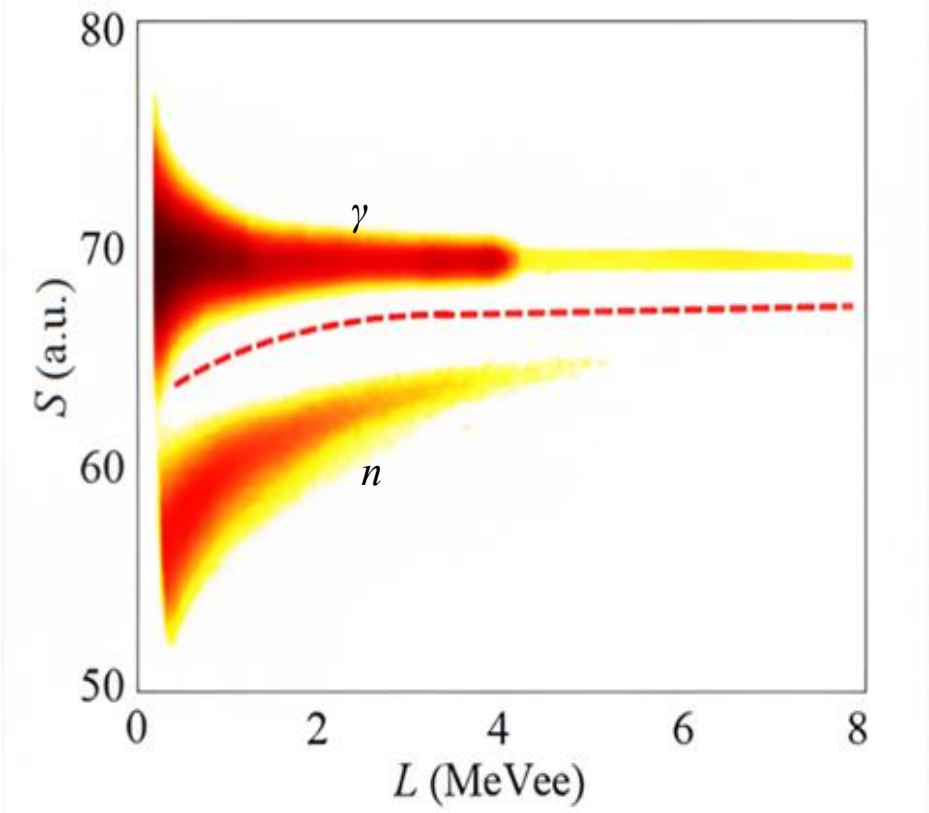


Figure 3: Event density as a function of pulse shape parameter (*S)* and light output parameter (*L*) for an ^{241}Am-^{9}Be source measured with the EJ-301 detector. The dashed line separates neutron (*n*) and gamma ray (γ) induced events.

4. Results and analysis

4.1 Experimental

Calibrated neutron light output *L* spectra for the empty container (no sample) and all three samples are shown in Fig. 4(a). For each sample, the neutron energy spectrum was unfolded with GRAVEL within the *L* range of 0.3 – 6.0 MeV$_{ee}$, and energy range of 1.2 – 11.0 MeV, with solution χ^2 values of 2.0 or less. Unfolding with GRAVEL requires an initial estimate of the spectrum, and here we have used the ISO-recommended AmBe spectrum. Unfolded neutron energy spectra for the no sample and with sample cases are shown in Fig. 3(b) and the no sample case is well matched to the ISO-recommended spectrum. The features observed in the no sample are consistent in the AmBe ISO spectrum in terms of position and intensity. Subtle changes in intensity are expected in the no sample spectrum as our AmBe neutron source does not have the same physical construction as the source used to produce the ISO recommended

reference spectrum. Resulting neutron energy spectra may vary from one ^{241}Am-^{9}Be source to another because the production of neutrons from the $^{9}Be + \alpha \rightarrow {}^{12}C^{*} + n$ reaction relies on the energy of alpha particles, which can be influenced by shielding and other external variables[15, 16].

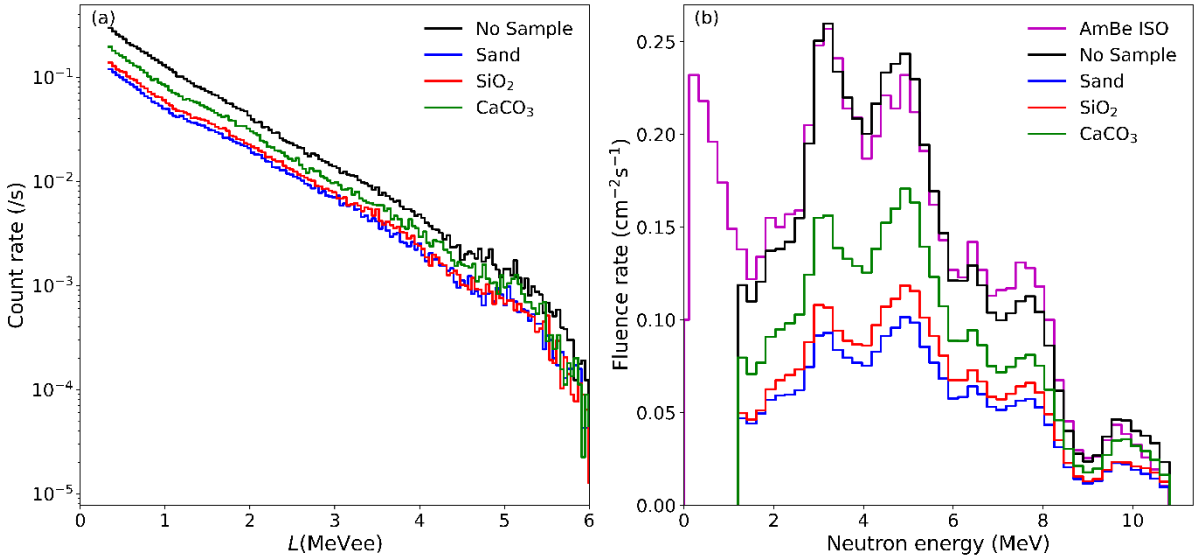


Figure 4: (a) Light output spectra for AmBe neutrons incident on an empty container (no sample measurement), sand, SiO_2 and $CaCO_3$ measured with an EJ-301 detector. (b) Neutron energy spectra unfolded with GRAVEL for the no sample, sand, SiO_2 and $CaCO_3$ measurements.

The removal cross section Σ_R for each sample was determined using Eq.1 where the measured fluence rate in each energy bin for the empty container (no sample) was taken as I_0, and with each of the samples taken as the attenuated neutron beam $I(t)$. Thereafter, Σ_R was normalised with respect to the measured density to get energy dependent Σ_R/ρ for all three samples as seen in Fig. 5(a). The energy dependent removal cross sections demonstrate some of the unique features expected from each element (Fig. 1) such as the peak around 3 MeV evident in all samples is attributed to the presence of oxygen and the enhancement around 8 MeV in $CaCO_3$ is attributed to the carbon content. Other features include the drop just after 10 MeV in $CaCO_3$ characteristic of carbon and the enhancement around 10 MeV in SiO_2 due to silicon. A peak around 3 MeV due to calcium is expected, however due to oxygen presence, the two peaks can't be distinguished from each other at this stage.

The removal cross sections measured for SiO_2 and $CaCO_3$ were used to reconstruct the removal cross section for sand according to Eq. 2 using mass ratios determined from previous XRF measurements, where the sand was found to contain approximately 80 % SiO_2 and 20 % $CaCO_3$ by mass [8]. A comparison of measured and reconstructed sand is shown in Fig. 5(b) and there is a good agreement between the two results. There are, however, some minor differences, particularly in the low energy (< 3 MeV) and high energy range (> 9 MeV), which are attributed to the unfolding process. In these regions the measured removal cross section is sensitive to minor variations that may arise during the measurement process such as small changes in the detector gain, which are carried over into the unfolding process. Due to this we consider the 3-9 MeV region to be most reliable.

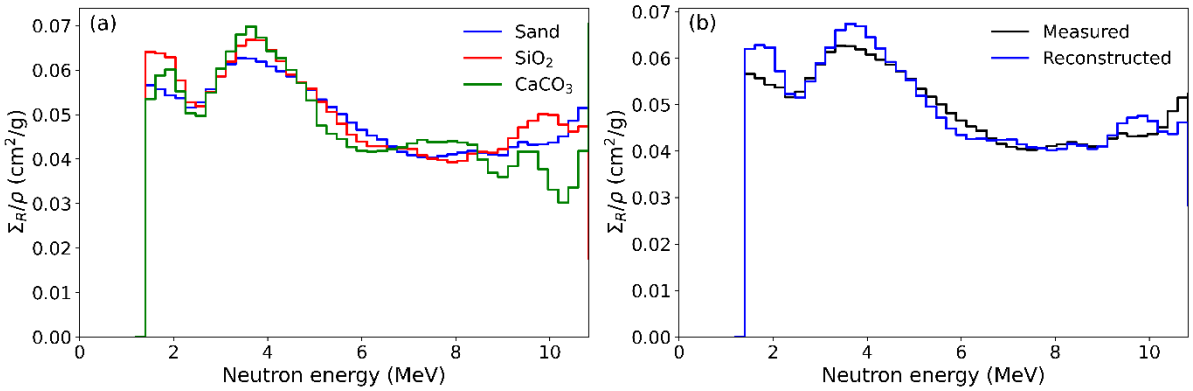


Figure 5: (a) Energy dependent Σ_R/ρ for sand, SiO_2 and $CaCO_3$. (b) Measured and reconstructed energy dependent Σ_R/ρ for sand.

4.2 Simulations

The simulated Σ_R/ρ for sand, SiO_2 and $CaCO_3$ were determined from the energy dependent fluence, as with the measured data, and are shown in Fig. 6 (a). The energy dependent Σ_R/ρ for sand is almost identical to that simulated for SiO_2 as expected from the mass ratios found via XRF. The simulated and reconstructed simulated energy dependent removal cross sections for sand are shown in Fig. 6 (b). Overall, there is excellent agreement between the simulated and reconstructed simulated Σ_R/ρ for sand.

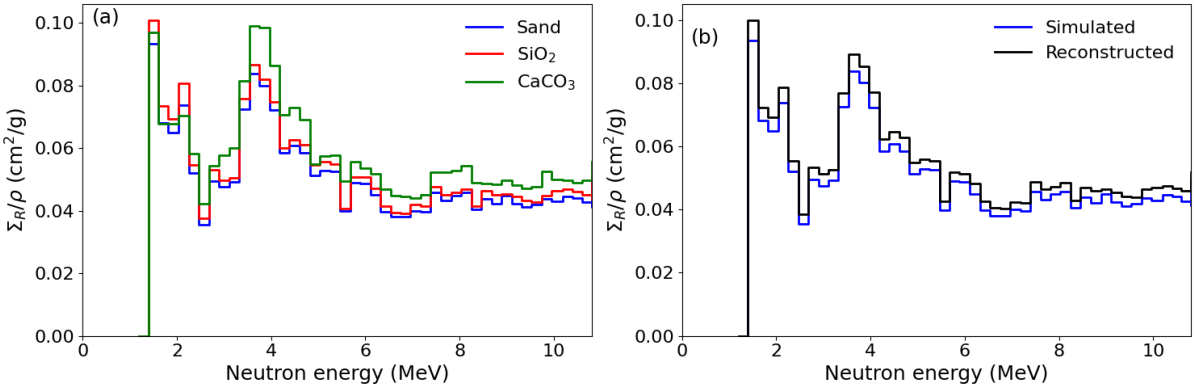


Figure 6: (a) Simulated energy dependent Σ_R/ρ for sand, SiO_2 and $CaCO_3$. (b) Simulated and reconstructed simulated energy dependent Σ_R/ρ for sand.

4.4 Comparison

A comparison between the simulated and measured Σ_R/ρ for all samples are shown in Fig. 7, which demonstrate a consistent scaling and include key features such as the enhancement between 3-4 MeV due to oxygen. However, the features are much broader in the measured data and are attributed to reduced energy resolution relative to the simulations. In the simulations the scoring region does not account for the detection process and any neutrons that traverse the boundary are scored. When taking measurements with an organic scintillator and using spectrum unfolding, there are several processes that limit the energy resolution of the neutron spectrum, and hence removal cross sections.

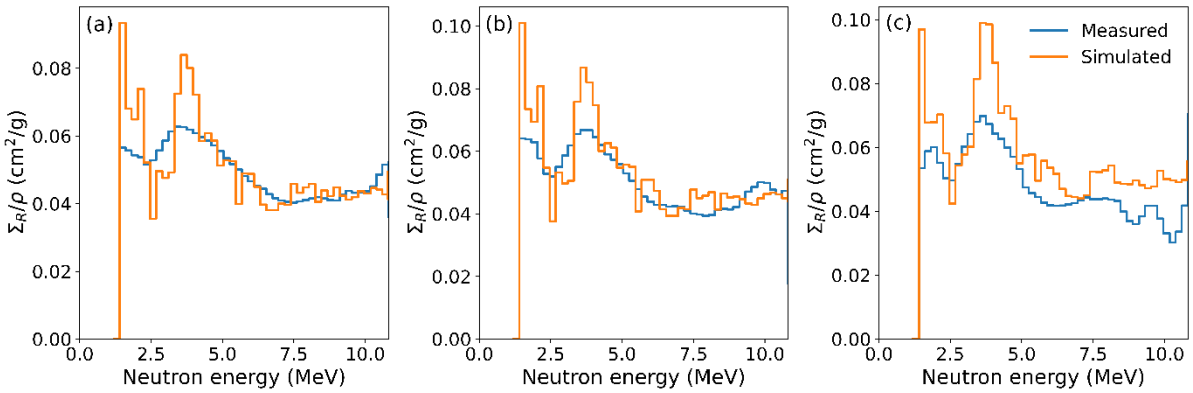


Figure 7: Comparison between simulated and measured Σ_R/ρ for (a) sand, (b) SiO_2, and (c) $CaCO_3$.

5. Conclusion

The use of fast neutron transmission spectroscopy to deconvolve the elemental composition of concrete is underway at the n-lab, and in this work measurements and simulations for sand and its constituents were made. Energy dependent removal cross sections were determined via measurement and FLUKA simulations. Features like the enhancement between 3-4 MeV due to oxygen in the measured Σ_R/ρ for all samples are consistent with expectations. From XRF results, measured Σ_R/ρ for sand was reconstructed and overall, there good agreement between the measured and reconstructed Σ_R/ρ was observed. For all samples Σ_R/ρ was also calculated from FLUKA simulations and compared to the measured results. Overall, most of the expected features appear in the same positions in both measured and simulated data. However, there are some minor differences in intensity and a broadening effect. These broadened features in the measured data are attributed to the measurement process, and investigations to improve the agreement are underway. By building a repository of element dependent removal cross sections, the elemental composition of an unknown sample can be deconvolved via a secondary unfolding process. The next phase of this work will involve using both transmitted neutrons and gamma rays for the characterization of unknown samples as the ratio of the neutron to gamma ray attenuation varies uniquely with the atomic mass of the sample constituents.

Acknowledgements

We would like to extend our thanks to the IAEA Marie Sklodowska-Curie Fellowship Programme.

References

[1] Naus D J, 2009. The management of aging in nuclear power plant concrete structures. *JOM*, **61**(7), p. 37.
[2] Trampus P, 2019. Role and importance of NDE in nuclear power plant life extension. *Procedia Structural Integrity*, **16**, pp. 161–163.
[3] Micklich B J *et al*, 1994. System design considerations for fast neutron interrogation systems. *Substance Detection Systems*, pp. 1–3.
[4] Hutton T, Buffler, A and Alexander M, 2022. Elemental analysis of concrete via fast neutron transmission and scattering spectrometry. *EPJ Web of Conferences*, **261**, p. 2.
[5] Fink C L *et al*, 1995. Evaluation of few-view reconstruction parameters for illicit substance detection using fast-neutron transmission spectroscopy. *1995 IEEE Nuclear Science Symposium and Medical Imaging Conference Record*, p. 748.
[6] El Abd A *et al*, 2017. A simple method for determining the effective removal cross section for fast neutrons, *Journal of Radiation and Nuclear Applications*, **2**(2), p. 53.
[7] Buffler A *et al*, 2020. Neutron transmission studies for concrete used in the nuclear industry. *International Journal of Modern Physics: Conference Series*, **50**, p. 2060015.
[8] Segale N et al, 2022. Measurement of fast neutron removal cross sections for the elemental analysis of concrete. *SAIP2022 Proceedings*, pp 601.
[9] Brown D A et al, 2018. ENDF/B-VIII.0: The 8 th major release of the Nuclear Reaction Data Library with

Cielo-Project Cross Sections, new standards and thermal scattering data. Nuclear Data Sheets, **148**, pp. 13–15.
[10] CAEN SpA, 2021. User manual UM3148, DT5730/DT5730S – 8 Channel 14 bit 500/200 MS/s Digitizer
[11] Comrie A, 2015. QtDAQ git repository, [online] Available: https://bitbucket.org/veggiesaurus/qtdaq
[12] Buffler A *et al*, 2015. Neutron spectrometry with EJ299-33 plastic scintillator for E_n=10-100 MeV, *IEEE Transactions on Nuclear Science*, **62**(3), pp. 1423–24.
[13] Nobs C R *et al*, 2019. Neutron spectrum unfolding for the development of a novel neutron detector for Fusion. *Fusion Engineering and Design*, **146**, p. 2659.
[14] Battistoni G *et al*, 2015. Overview of the fluka code. *Annals of Nuclear Energy*, **82**, pp. 10.
[15] International Organization for Standardization, 2021. Reference neutron radiations, part 1: characteristics and methods of production. ISO 8529-1
[16] Kostal, M. *et al*, 2020. Validation of heavy water cross section using Ambe Neutron source. *EPJ Web of Conferences*, **239**, p. 18008.

Comparative Analysis of Numerical Methods for Assessing Wind Potential in Fort Beaufort, South Africa, using Two-Parameter Weibull Distribution Model.

N Shambira, G Makaka, P Mukumba and C Shonhiwa.

Department of Physics, Faculty of Science and Agriculture, University of Fort Hare, P. Bag X1314, Alice, 5700, South Africa.

E-mail: **201927103@ufh.ac.za**

Abstract. The utilisation of wind energy in South Africa promises to address the country's energy crisis and extend electricity access to remote, off-grid regions. Wind power is a clean and renewable resource that can significantly reduce carbon emissions and help combat the impacts of climate change, including heatwaves and cyclones, which are increasingly affecting South Africa and its neighbouring nations. South Africa's potential for wind energy generation is particularly promising in coastal areas, notably the Eastern Cape Province. Accurate wind distribution models are vital for assessing an area's wind characteristics and potential. This study delves into this by examining eight numerical methods for estimating Weibull parameters to establish a suitable model. The study utilises five-and-a-half years of hourly wind data collected from the Fort Beaufort weather station at a height of 10m. The results indicate that the Open Wind algorithm outperforms other algorithms, providing Weibull shape (k) and scale (c) values of 1.67905 and 3.35800, respectively. The predicted overall average wind speed and wind power density for Fort Beaufort stands at 2.999 m/s and 38.45 W/m², suggesting that small-scale wind power projects for purposes like lighting, battery charging, and water pumping are ideal. To enhance wind turbine feasibility, augmentation systems such as concentrators, diffusers, and Invelox can be employed to reduce the cut-in wind speeds required for operation, considering that most turbines on the market typically need wind speeds above 5 m/s to start in this region. The study also identifies the prevailing wind direction as predominantly from the south-east (SE) in the Fort Beaufort area.

1. Introduction

Electricity generation from wind energy presents a viable solution to South Africa's ongoing national energy crisis, offering potential benefits beyond simply addressing power shortages. By harnessing wind energy, extending electricity supply to remote regions that currently lack access to the national grid becomes feasible [1]. This approach capitalises on the advantages of wind energy, a clean, readily accessible, and perpetually renewable resource, thereby contributing to reducing carbon dioxide emissions and mitigating climate change effects [2]. In the 1990s, South Africa employed approximately 30,000 windmills primarily for water supply and agricultural needs [3]. However, the adoption of wind energy for electricity generation remained limited due to coal's prevailing availability and affordability

[4]. Regions such as Eastern Cape, Western Cape, Northern Cape, and KwaZulu-Natal exhibit remarkable wind energy potential, with coastal areas experiencing an annual mean wind speed of 6 m/s at a height of 10 meters above ground level [5]. Recent years have witnessed significant investment in South Africa's wind energy sector, amounting to R209.7 billion in 2020 alone [6]. The nation is now home to thirty-three wind farms, of which 22 are fully operational and commercially active, generating tangible economic and environmental benefits. This growth has created 2723 jobs, facilitated by commercialising 22 wind-independent power producers and a corresponding reduction of 6.4 million tonnes in carbon dioxide emissions [7].

In wind energy research, various probability distribution methods (Weibull, Gamma, Rayleigh, and Exponential) are employed. However, the two-parameter Weibull distribution is the most popular and widely used due to its flexibility, simplicity, and adaptability to a wide range of data [2,8]. Its effectiveness relies on the accurate estimation of scale and shape parameters [9,10]. Numerous studies have employed various numerical methods to address this fundamental aspect. Hussain et al. (2023) investigated wind power density potential in coastal areas of Pakistan, revealing weak performance of the energy trend and graphical methods [8]. Patidar et al. (2022) focused on assessing wind characteristics and potential in offshore locations within Gujarat, India, employing the Weibull density function and recognising the Multivariate MLM method as remarkably accurate for wind potential evaluations [11]. Yaniktepe et al. (2023) investigated wind characteristics and energy potential in Osmaniye, Turkey, utilising data from an 80m meteorological mast [12]. Their analysis of the two-parameter Weibull distribution involved a comparative evaluation of eight numerical methods, with the modified maximum likelihood method (MMLM) identified as the most effective and the graphical method (GM) as the least effective. Bulut and Bingöl (2023) conducted a comparative study on numerical methods and a metaheuristic algorithm for Weibull distribution parameter estimation in the Isparta region, Turkey [13].

The manta ray foraging optimisation method (MRFO) outperformed the energy pattern factor method (EPM) and maximum likelihood method (MLM). Aziz et al. (2023) compared fourteen methods for wind potential estimation, highlighting the Energy Pattern Factor method as the most suitable for all wind speeds, while the Weighted Least Squares Method (WVM) and the Moroccan Method (MoroM) were deemed unsuitable [10]. Patidar et al. (2023) explored accurate wind speed distribution selection for enhanced wind energy utilisation in India [9]. Comparing five numerical methods and three metaheuristic optimisation algorithms, the WAsP method demonstrated the best curve fit. In contrast, social spider optimisation (SSO) and particle swarm optimisation (PSO) outperformed the genetic algorithm (GA). GOK exhibited the highest wind power density. Rehman et al. (2021) analysed wind speed patterns within South Africa, identifying Port Elizabeth as having the highest average wind speed [14]. Due to their favourable characteristics, Cape Town, East London, and Port Elizabeth were promising locations for wind power deployment.

Wind potential estimations are uncertain due to temporal and spatial wind variations. Thus, site-specific wind resource assessments are crucial for accurately evaluating wind potential and mitigating investment risks in wind energy projects [2]. The reviewed literature emphasises the need to employ various methods for accurate predictions of wind characteristics and potential, considering that the best method for one location might not be suitable for another. The novelty of this study lies in utilizing eight different numerical methods, with seven being employed for the first time in the Fort Beaufort area, to estimate Weibull scale and shape parameters. This provides a more precise Weibull distribution that fits the observed wind data for this location. Additionally, a goodness-of-fit test with six performance indicators will be used for the first time to evaluate the best method that accurately estimates Weibull scale and shape parameters.

2. Methodology

2.1. Site description and wind speed data

The study utilised five-and-a-half years of hourly wind speed data (January 2015 to July 2020) obtained from Fort Beaufort weather stations through South African Weather Services. The wind data was collected at a standard anemometer height of 10m above the ground (AGL). The geographical coordinates of these weather stations are given in Table 1.

Table 1. Geographical coordinates of Fort Beaufort weather station.

Weather stations	Latitude	Longitude	Height (m)
Fort Beaufort	-32.7880	26.6290	455

2.2. Fitting probability distributions to observed wind data

The two-parameter Weibull distribution has been widely applied to many wind research studies to describe wind speed data because it is simple, adaptable and precise when compared to other distribution functions [15,16]. Weibull Cumulative distribution function (W) is shown in equation (1) below:

$$W(v) = 1 - e^{\left[-\left(\frac{v}{c}\right)^k\right]} \tag{1}$$

Differentiating $W(v)$ with respect to v gives a Weibull probability distribution function (w) (bi-parameter):

$$w(v) = \frac{dW(v)}{dv} = \left(\frac{k}{c}\right)\left(\frac{v}{c}\right)^{k-1} \bullet e^{\left[-\left(\frac{v}{c}\right)^k\right]} \tag{2}$$

In the equation, with no units, k represents the Weibull shape, $c(\geq 1)$ represents the Weibull scale parameter, v represents wind speed, and $\frac{dW(v)}{dv}$ is the derivative of the Weibull Cumulative Distribution Function with respect to v. If c increases, the wind speed will also follow the same increasing trend, and k values indicate wind stability. Equation (2) becomes a cumulative Rayleigh distribution when $k = 2$. Additionally, calm conditions must be excluded since $c(\geq 1)(\mathrm{m/s})$ [17].

2.3. Algorithms to calculate scale (c) and shape (k) for Weibull distribution

There are various methods to estimate the parameters in the Weibull distribution. In this study, for calculating the parameters of the Weibull distribution the following eight different numerical methods were utilised: mean standard deviation method (Msdm), method of multi-objective moment (MofMoM), probability-weighted moments based on power density method (PwmbpdM), WAsP method (WM), method of Mabchour (Momab), OpenWind method (Owm), energy pattern factor method (Epfm), and novel energy pattern factor method (Nepfm) as shown in [2,18–21].

2.4. Goodness of fit test.

To evaluate the suitability and effectiveness of the eight numerical estimation algorithms in investigating the accuracy of the theoretical frequency distribution compared to the observed frequency distribution of wind speed, we employ the following statistical metrics: Mean absolute Bias error (MaBE), Root mean square error (RMsE), Wind power density error (WPDE), Kolmogorov–Smirnov test (KS), Anderson darling test (AD), Chi-squared test (χ^2) as shown in [2,18,22]. Low values indicate a good fitting between the frequency distributions and the observed actual wind data.

2.5. Wind Power Density (WPD)

WPD is expressed in two forms using parameters Weibull scale and shape or average wind speed derived from collected data as given by equation (3) and equation (4):

$$WPD_{wbl} = \frac{1}{2}\rho c^3 \Gamma\left(\frac{3}{k}+1\right) \quad (3)$$

$$WPD_{obs} = \frac{1}{2}\rho v^3 \quad (4)$$

where WPD_{wbl} is the Weibull wind power density (W/m^2), WPD_{obs} is the wind power density (W/m^2) of actual wind speed data, v is wind speed (m/s) and ρ is the air density (= 1.225kg/m^3). For this current study to determine the wind resource availability using wind power densities, the Fazelpour et al. [23,24] classification is used, as presented in Table 2.

Table 2. Classification of wind resources.

Class	Fair	Fairly Good	Good	Very good
P /(W/m^2)	$P < 100$	$100 \leq P < 300$	$300 \leq P < 700$	$P \geq 700$

3. Wind energy analysis

3.1. Assessment of wind speed properties

Table 3 illustrates shows that the overall average wind speed is 2.999 m/s with a standard deviation (Std) of 1.77 m/s. A positive skewness of 1.37 indicates that the wind speed data is skewed to the right. It also implies that the measured wind speed values are above the average wind speed (2.999 m/s), which reflects better wind performance at the site. A positive kurtosis of 1.75, below 3, was obtained, depicting wind speed data with few extreme values

.

Table 3. Statistical wind data of Fort Beaufort.

Statistic	Range	Mean	Variance	Std.	Skewness	Kurtosis
Value	13.9	2.999	3.14	1.77	1.37	1.75

3.2. Fitting probability distributions to observed wind data

Figure 1 compares the frequency distribution function of wind speed and the actual observed wind data at a height of 10m AGL. The eight algorithms exhibited varying degrees of fitting to the data, with the Openwind and WAsP algorithms showing the best fit in that order. Conversely, the mean standard deviation algorithm performed the least effectively overall, a result consistent with our error analysis in Table 4.

3.3. Daily averages of wind speed and wind power density

Figure 2 displays the daily average wind speed values at 10m AGL. In the afternoon, the average wind speed profile is dome-shaped; between 21:00 and 06:00, the mean wind speed has the lowest values. It can be concluded from this analysis that Fort Beaufort has wind speeds above 3m/s between 12:00 and

3.3. Daily averages of wind speed and wind power density

Figure 2 displays the daily average wind speed values at 10m AGL. In the afternoon, the average wind speed profile is dome-shaped; between 21:00 and 06:00, the mean wind speed has the lowest values. It can be concluded from this analysis that Fort Beaufort has wind speeds above 3m/s between 12:00 and
20:30.

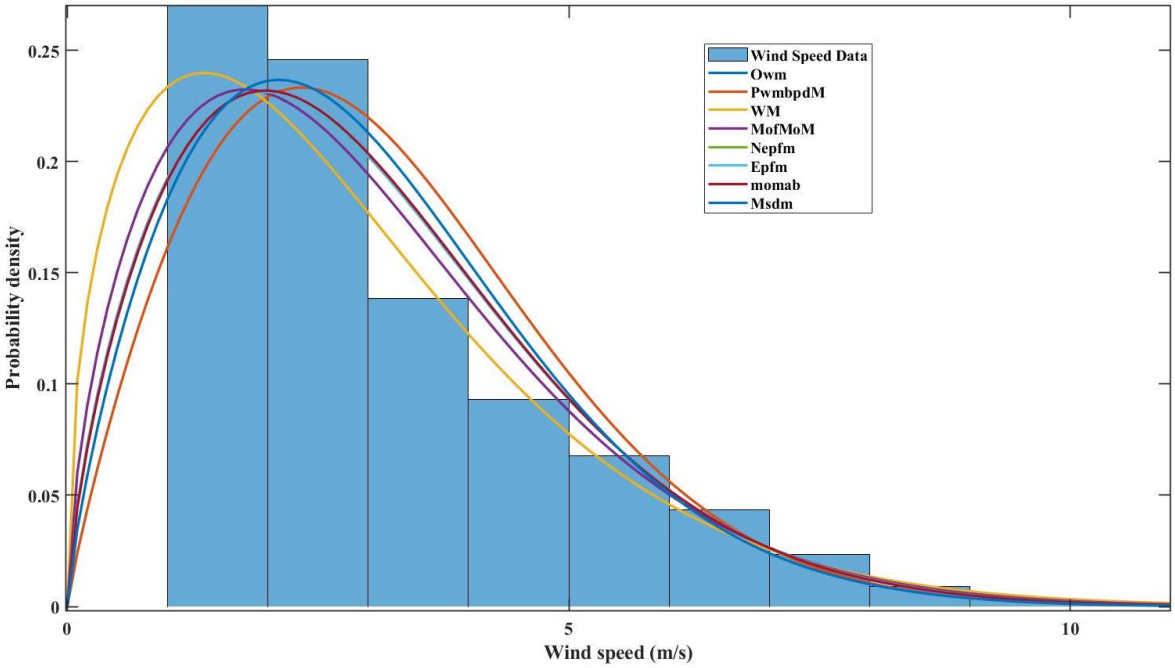


Figure 1. Probability distribution using eight different numerical algorithms at 10m height.

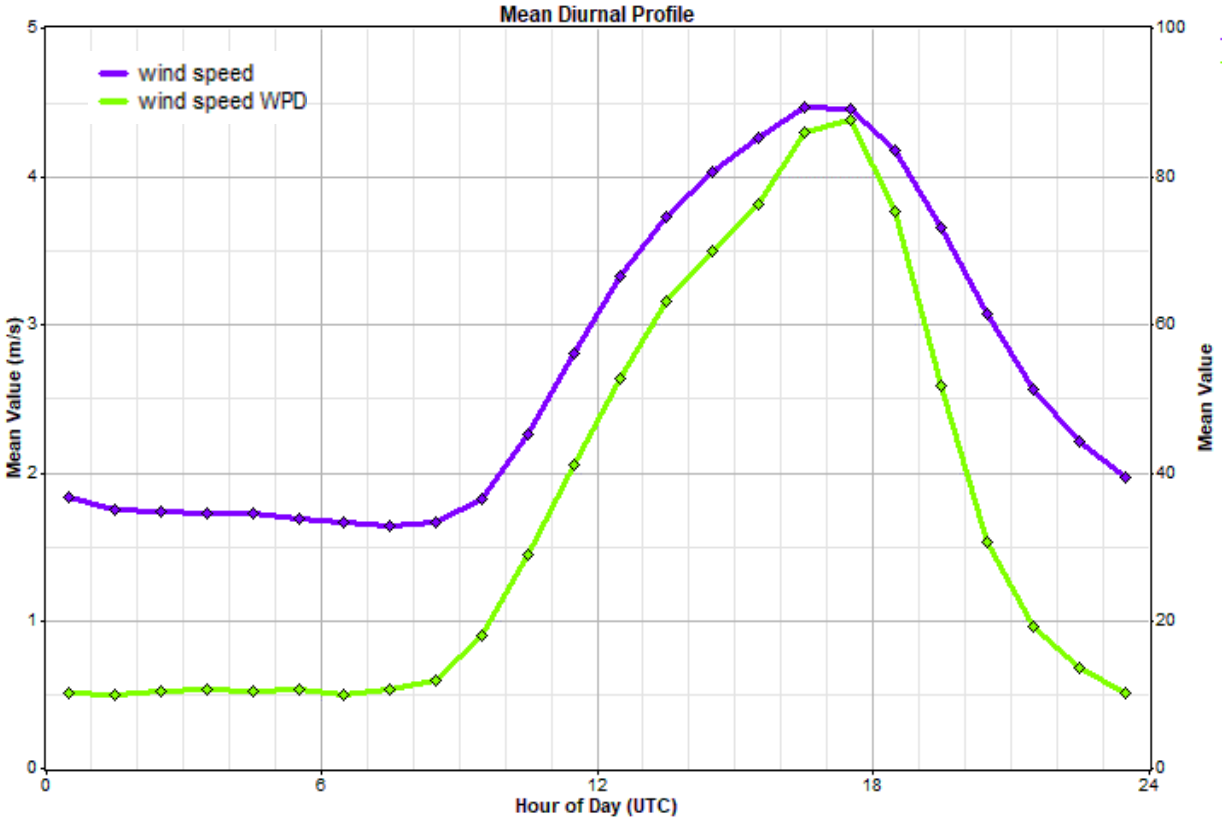


Figure 2. Daily mean wind speed and wind power density

3.4. Wind direction

Wind Rose presents data on wind speed and direction occurrences [25]. This information is crucial for site selection, as it helps identify the optimal locations for installing wind turbines to maximise wind power utilisation [26]. The dominating wind mostly comes from the South-East direction (SE) of the Fort Beaufort area, as shown in Figure 3.

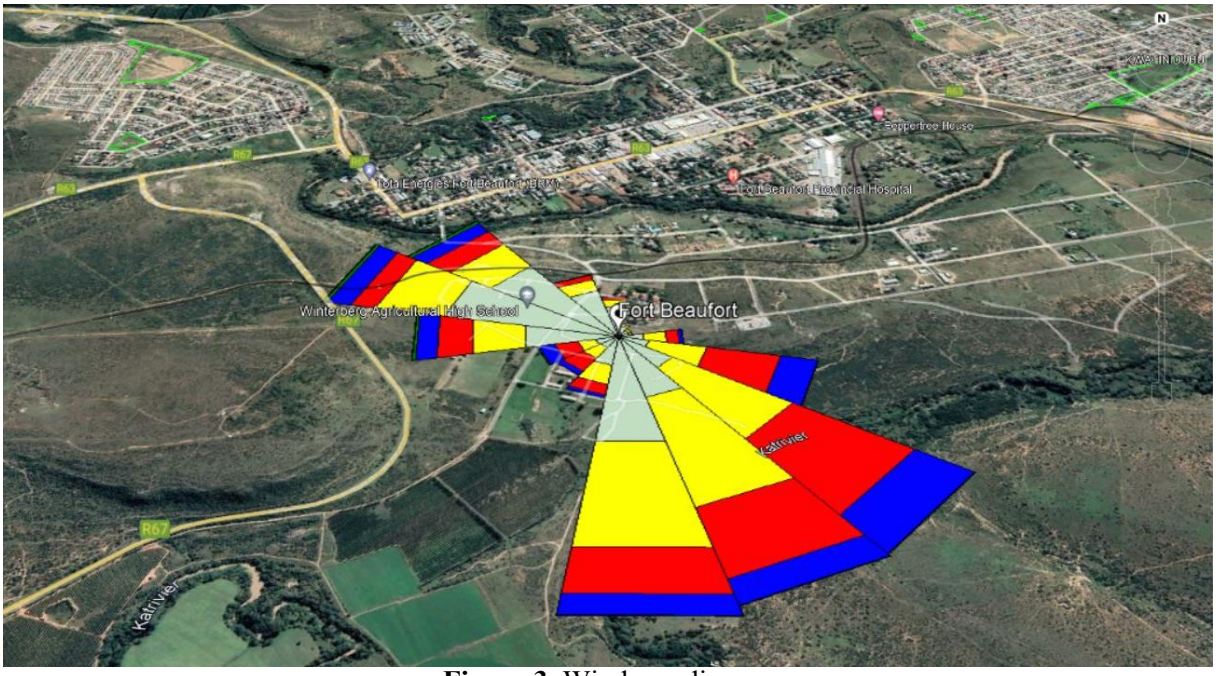


Figure 3. Windrose diagram.

Table 4. Goodness of fit test.

Algorithm	KS	No	AD	No	χ^2	No	RMsE	No	WPDE	No	MaBE	No	Ave
Owm	0.13002	1	1014.6	1	8745.3	1	0.07361	1	9.948E-14	1	0.4392	1	1
PwmbpdM	0.13416	2	1049.6	2	9919.9	3	0.07755	2	6.519E-06	2	0.4944	3	2
WM	0.14451	6	1058	4	9102.2	2	0.07941	3	2.774E-05	3	0.4692	2	3
Nepfm	0.14422	4	1056.7	3	9931	4	0.07942	4	1.117E-02	5	0.4957	4	4
MofMOM	0.14281	3	1164.3	6	9933.3	5	0.08220	7	8.603E-04	4	0.4979	5	5
Epfm	0.14446	5	1058.3	5	9983.4	6	0.07947	5	5.294E-02	6	0.4981	6	6
momab	0.16372	7	1293	7	12836	8	0.07973	6	3.017E-01	7	0.5224	7	7
Msdm	0.2078	8	1825.4	8	10746	7	0.08354	8	2.443E+00	8	0.5695	8	8

3.5. Calculation of wind power density and classification

Owm is the best algorithm, with a value of 38.452176 W/m^2, closely matching the actual wind power density (WPD_{obs}) for observed wind speed data, which is also 38.452176 W/m^2, with a negligible error margin of 9.95E-14%. In contrast, the mean standard deviation method (Msdm) performed poorly, yielding an error margin of 2.44%. It's worth noting that this method was exclusively employed in a previous study [27] conducted at the same area to estimate wind power density. This underscores the importance of employing multiple algorithms to determine the most accurate result.

Table 5. Wind power density of Fort Beaufort.

		Weibull parameters		Wind Power Density	
	Algorithm	Shape (k)	Scale (c)	WPdwbl	WPdobs
1	Owm	1.679048803	3.358002058	38.452176	38.452176
2	PwmbpdM	1.867853000	3.522606000	38.452182	38.452176
3	WM	1.444533000	3.079657000	38.452148	38.452176
4	MofMoM	1.599217000	3.273809000	38.451315	38.452176
5	Nepfm	1.679380000	3.358006000	38.441002	38.452176
6	Epfm	1.680868000	3.358266000	38.399233	38.452176
7	momab	1.689512843	3.359496866	38.150480	38.452176
8	Msdm	1.771723996	3.369759722	36.009513	38.452176

4. Conclusion.

The analysis has revealed that the Open Wind algorithm is the most effective method for calculating the Weibull scale and shape parameters for the two-parameter Weibull distribution. Consequently, it is advisable to utilise the two-parameter Weibull distribution to model the observed wind data in the Fort Beaufort area, with the condition that the k and c parameters of the Weibull distribution are estimated using the Open Wind algorithm. The overall power density estimated for the Fort Beaufort area, at a height of 10 meters, is 38.452176 W/m², falling within the 'fair' category according to Fazelpour's classification. Therefore, it is recommended to consider small-scale wind power generation projects in this area, which could serve lighting, battery charging, or water pumping purposes. To enhance the feasibility of wind turbines in this region, augmentation systems such as concentrators, diffusers, and invelox should be considered. Additionally, evolutionary metaheuristic algorithms are recommended to enhance the accurate determination of the Weibull scale and shape parameters.

Acknowledgements

I would like to express my gratitude for the invaluable guidance, wisdom, and unwavering support I received from my supervisors, Prof G Makaka and Dr P Mukumba. Additionally, I would like to extend my heartfelt appreciation to the Govan Mbeki Research and Development Centre and the Department of Physics at the University of Fort Hare for their financial support. Lastly, we also want to give special thanks to the South African Weather Service for providing us with the weather data utilised in this study.

References

[1] Ravanbach B, Kuhnel M, Hanke B, Von Maydell K, Van Dyk EE, Vumbugwa M, et al. Development of a Smart Monitoring and Evaluation Framework for Hybrid Renewable Mini-grids. 2020 15th Int Conf Ecol Veh Renew Energies, EVER 2020 2020. https://doi.org/10.1109/EVER48776.2020.9243124.

[2] Shambira N, Makaka G, Mukumba P, Lesala M, Roro K, Julies J, et al. Wind Resource Assessment in the Upper Blinkwater area in the Province of Eastern Cape, South Africa. Int J Eng Res Technol 2020;9:387–402.

[3] Akinbami OM, Oke SR, Bodunrin MO. The state of renewable energy development in South Africa: An overview. Alexandria Eng J 2021;60:5077–93. https://doi.org/10.1016/j.aej.2021.03.065.

[4] Merem EC, Twumasi Y, Wesley J, Olagbegi D, Crisler M, Romorno C, et al. The Evaluation of Wind Energy Potentials in South Africa. Energy and Power 2022;12:9–25. https://doi.org/10.5923/j.ep.20221201.02.

[5] Alavi O, Mohammadi K, Mostafaeipour A. Evaluating the suitability of wind speed probability distribution models: A case of study of east and southeast parts of Iran. Energy Convers Manag 2016;119:101–8. https://doi.org/10.1016/j.enconman.2016.04.039.

[6] Macingwane ZN. Wave energy a new energy mix to produce green hydrogen: a potential future maritime shipping fuel : a study on the Port of Ngqura, Southern Africa's "Green Status Port." World Marit Univ 2021.

[7] McKenna R, Pfenninger S, Heinrichs H, Schmidt J, Staffell I, Bauer C, et al. High-resolution large-scale onshore wind energy assessments: A review of potential definitions, methodologies and future research needs. Renew Energy 2022;182:659–84. https://doi.org/10.1016/j.renene.2021.10.027.

[8] Hussain I, Haider A, Ullah Z, Russo M, Casolino GM, Azeem B. Comparative Analysis of Eight

Numerical Methods Using Weibull Distribution to Estimate Wind Power Density for Coastal Areas in Pakistan. Energies 2023;16:1515.
[9] Patidar H, Shende V, Baredar P, Soni A. Comparative evaluation of optimal Weibull parameters for wind power predictions using numerical and metaheuristic optimization methods for different Indian terrains. Environ Sci Pollut Res 2023;30:30874–91. https://doi.org/10.1007/s11356-022-24395-6.
[10] Aziz A, Tsuanyo D, Nsouandele J, Mamate I, Mouangue R, Abiama PE. Influence of Weibull parameters on the estimation of wind energy potential. Sustain Energy Res 2023;10:1–20. https://doi.org/10.1186/s40807-023-00075-y.
[11] Patidar H, Shende V, Baredar P, Soni A. Comparative study of offshore wind energy potential assessment using different Weibull parameters estimation methods. Environ Sci Pollut Res 2022;29:46341–56. https://doi.org/10.1007/s11356-022-19109-x.
[12] Yaniktepe B, Kara O, Aladag I, Ozturk C. Comparison of eight methods of Weibull distribution for determining the best - fit distribution parameters with wind data measured from the met - mast. Environ Sci Pollut Res 2023;30:9576–90. https://doi.org/10.1007/s11356-022-22777-4.
[13] Bulut A, Bingöl O. Analysis and Comparison of Weibull Parameters for Wind Energy Potential Using Different Estimation Methods: A Case Study of Isparta Province in Turkey. Electr Power Components Syst 2023;0:1–17. https://doi.org/10.1080/15325008.2023.2210574.
[14] Rehman S, Natarajan N, Mohandes MA, Meyer JP, Alam MM, Alhems LM. Wind and wind power characteristics of the eastern and southern coastal and northern inland regions, South Africa. Environ Sci Pollut Res 2021. https://doi.org/10.1007/s11356-021-14276-9.
[15] Adem Çakmakçı B, Hüner E. Evaluation of wind energy potential: a case study. Energy Sources, Part A Recover Util Environ Eff 2022;44:834–52. https://doi.org/10.1080/15567036.2020.1811810.
[16] Ali B, Abbas G, Memon A, Mirsaeidi S, Koondhar MA, Chandio S, et al. A comparative study to analyze wind potential of different wind corridors. Energy Reports 2023;9:1157–70. https://doi.org/10.1016/j.egyr.2022.12.048.
[17] Lopez-Rodriguez AF, Serrano-Rico JC, Florez-Serrano EG. Statistical methodologies for wind resource analysis , case : Catatumbo region - Norte de Santander , Colombia. 18th Int. Conf. Renew. Energies Power Qual., Granada , Spain: RE&PQJ; 2020.
[18] Safari MAM, Masseran N, Majid MHA. Wind energy potential assessment using Weibull distribution with various numerical estimation methods: a case study in Mersing and Port Dickson, Malaysia. Theor Appl Climatol 2022;148:1085–110. https://doi.org/10.1007/s00704-022-03990-0.
[19] Usta I, Arik I, Yenilmez I, Kantar YM. A new estimation approach based on moments for estimating Weibull parameters in wind power applications. Energy Convers Manag 2018;164:570–8. https://doi.org/10.1016/j.enconman.2018.03.033.
[20] Akdag SA, Güler Ö. A novel energy pattern factor method for wind speed distribution parameter estimation. Energy Convers Manag 2015;106:1124–33.
[21] Derome D, Razali H, Fazlizan A, Jedi A. Application of a New Method "AGM" to Estimate Weibull Parameters for Low Wind Speed. Int J Renew Energy Res 2022;12:1362–71. https://doi.org/10.20508/ijrer.v12i3.13164.g8520.
[22] Altmimi AI, Al-Swaiedi SJ, Abdullah OI. Estimating Weibull Parameters Using Mabchour's Method (MMab) for Wind Power at RAWA City, Iraq. Appl Syst Innov 2022;5:1–12. https://doi.org/10.3390/asi5010014.
[23] Fazelpour F, Markarian E, Soltani N. Wind energy potential and economic assessment of four locations in Sistan and Balouchestan province in Iran. Renew Energy 2017;109:646–67.
[24] Fazelpour F, Soltani N, Soltani S, Rosen MA. Assessment of wind energy potential and economics in the north-western Iranian cities of Tabriz and Ardabil. Renew Sustain Energy Rev 2015;45:87–99. https://doi.org/10.1016/j.rser.2015.01.045.
[25] Tahir ZUR, Kanwal A, Afzal S, Ali S, Hayat N, Abdullah M, et al. Wind Energy Potential and Economic Assessment of Southeast of Pakistan. Int J Green Energy 2021;18:1–16. https://doi.org/10.1080/15435075.2020.1814298.
[26] Hulio ZH, Wei J, Khan GM. Statistical analysis of wind resource and energy potential assessment of Sanghar site, Sindh Pakistan. Int J Energy Sect Manag 2022;16:389–413. https://doi.org/10.1108/IJESM-04-2020-0002.
[27] Shonhiwa C, Makaka G, Munjeri K. Estimation of Wind Power Potential of Six Sites in Eastern Cape Province of South Africa. Phys Sci Int J 2015;6:209–18. https://doi.org/10.9734/psij/2015/16699.

Assessment of CO_2 emission in the rural set-up of the Vuwani area in Limpopo, South Africa

L Takalani[1], T S Mulaudzi[1], N E Maluta[1]
124 Department of physics, university of Venda, Limpopo in South Africa, Thohoyandou 0950, South Africa

E-mail: lufunotakalani@gmail.com/lufuno.takalani@univen.ac.za

Abstract

The world is currently battling climate change through various mitigation and adaptation measures. Some of these measures include reducing greenhouse gas emissions, transitioning to renewable energy sources, implementing sustainable agriculture and forestry practices, and enhancing climate resilience in vulnerable communities. The Vuwani area plays a significant role in carbon emissions, where agriculture and land use change are major drivers of emissions. This area often has high levels of deforestation, burning of fossil fuels, and use of traditional biomass for cooking and heating, which contribute to carbon emissions. The assessment of CO_2 emissions in the Vuwani area is a growing concern due to its potential impact on climate change. This study focuses on the rural area of Vuwani in Limpopo, South Africa, where CO_2 emissions are high due to various anthropogenic activities. However, there was a gap in knowledge regarding the magnitude and sources of CO_2 emissions in this area. The availability of data on CO_2 emissions from the Vuwani eddy covariance flux tower filled that gap and shows the area as a source of carbon which can assist the country in formulating effective policies to mitigate climate change. The eddy covariance technique is used as a possible solution to estimate CO_2 emissions from the Vuwani area. The method measured the exchange of CO_2 between the surface and the atmosphere and provided accurate and continuous measurements of daily means of Net Ecosystem Exchange ranging from $-5\ gCm^{-2}/day$ and $+54\ gCm^{-2}/day$ that included meteorological parameters radiation(daily mean) of minimum *250 w.m^{-2}* and maximum *504 w.m^{-2}*. The daily CO_2 flux ranged from$\pm$ 7 $\mu mol/m^2s^1$, and close to *100%* energy closure. The findings of this study assist in formulating policies and strategies to reduce CO_2 emissions in rural areas and mitigate climate change.

1. Introduction

The Vuwani town ecosystem's contribution to the global carbon budget and the variables affecting their influence on the temporal and geographical variance of terrestrial carbon uptake and emission are still very much in question. Despite the fact that numerous eddy Covariance experiments have been carried out in Africa over the years as part of numerous research projects, including the Sahelian Energy Balance Experiment (SEBEX) (Wallace JS et al. 1991), the Hydrological and Atmospheric Pilot Experiment-Sahel in Niger (HAPEX-Sahel) (Verhoef A et al. 1997, Friborg T et al. 1997, Hanan NP et al. 1998), the African Monsoon Multidisciplinary Analyses (AMMA) project (Redelsperger J-L et al. 2006), and the CARBOAFRICA project (Bombelli A et al.2009). The Eddy Covariance method

provides measurements of gas emission and consumption and allows measurements of fluxes of sensible heat, latent heat, and momentum, integrated over an area.

2. Materials and methods

The Vuwani Eddy Covariance flux tower, established in 2016 at the University of Venda Vuwani Science Resource Center as part of a South Africa-Germany collaborative research project, plays a crucial role in comprehending carbon dioxide's movement within the ecosystem. By recording fluxes, the tower aids in unravelling CO_2 exchanges involving plant uptake, soil release, and human activities like breathing and combustion. The site's significance lies in understanding whether it acts as a carbon sink or source, depending on carbon binding and release ratios. This peri-urban ecosystem, representing a mesic savanna, is equipped with a 15-meter measurement mast and various sensors capturing meteorological variables, such as rainfall and wind speed. Buried sensors gauge soil temperature and water content at different depths (Figure 1).

The eddy flux can be calculated as:

$$F = \bar{\rho}_a \overline{W'S'} \tag{1}$$

The eddy flux (F) is approximately equal to the mean air density ($\bar{\rho}_a$) multiplied by the mean covariance between deviations in instantaneous vertical wind speed (W') and mixing ratio (S'). Carbon dioxide flux (F_C) that will be obtained is presented as the mean covariance ($\bar{\rho}_a$) between deviations in instantaneous vertical wind speed (W') and density of CO_2 in the air (ρ_c'). (Grünwald T, Bernhofer C et al.2007, Mauder M, Foken T et al.2011)

$$F_C = \rho_a \overline{W'\rho_c'} \tag{2}$$

$$H = \rho_a c_p \ \overline{W'T_a'} \tag{3}$$

$$\lambda E = \lambda \overline{W'\rho_v'} \tag{4}$$

where ρ_a is the density of dry air (kg m^{-3}) at a given air temperature, c_p is the specific heat capacity of dry air at constant pressure (J kg^{-1} K^{-1}), λ is the latent heat of vaporization (J kg^{-1}), ρ_c is the molar density of CO_2 gas (mol m^{-3}) and ρ_v is the molar density of water vapour (mol m^{-3}). T_a is the air temperature derived from the sonic anemometer (K) and W is the vertical wind velocity component (m s^{-1}), and S' mixing ratios. Over bars denote time averages and primes indicate fluctuations about the averages.

3. Results

3.1. Daily Net Ecosystem exchange (NEE)

The net amount of carbon dioxide (CO_2) that is exchanged between an ecosystem and the atmosphere over the course of a single day is analyzed with the use of Equation 1. It represents the balance between the CO_2 that is taken up by plants through photosynthesis and the CO_2 that is released back into the atmosphere through respiration and other processes.

A CO2 flux Windrose visually illustrates the distribution and intensity of carbon dioxide emissions and uptake within the Vuwani area. It focuses on depicting the direction and magnitude of CO2 fluxes. Each segment's length indicates the strength or frequency of CO2 emissions or uptake from that direction. This tool identify key emission sources and uptake sinks areas contributing to the understanding of carbon balance dynamics. CO2 flux Windrose provide insight into seasonal variations and temporal changes in carbon dynamics, supporting environmental assessments and climate change mitigation strategies. In essence, they serve as a valuable graphical representation for carbon movement with measured speeds.

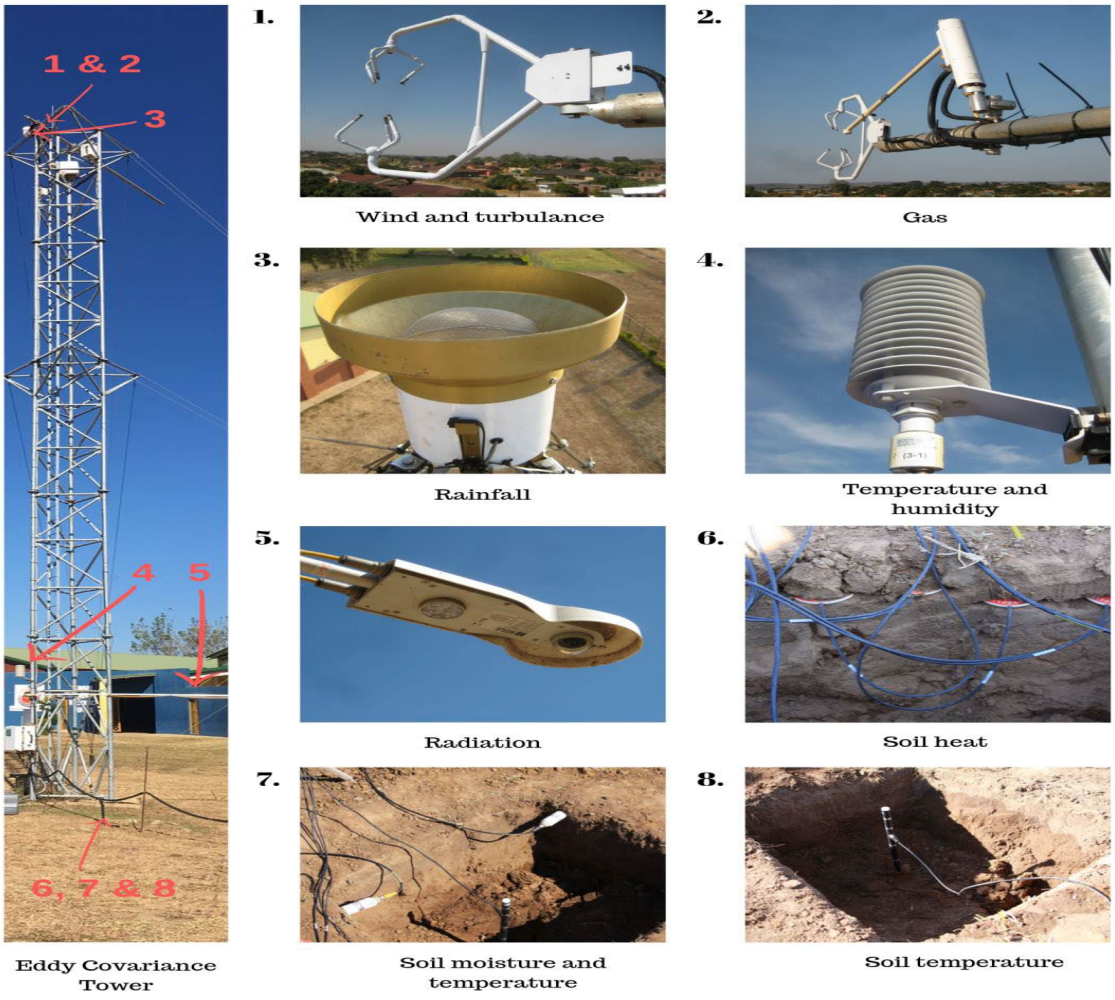


Figure 1. The Vuwani Eddy Covariance flux tower with its meteorological parameters at Limpopo province South Africa.

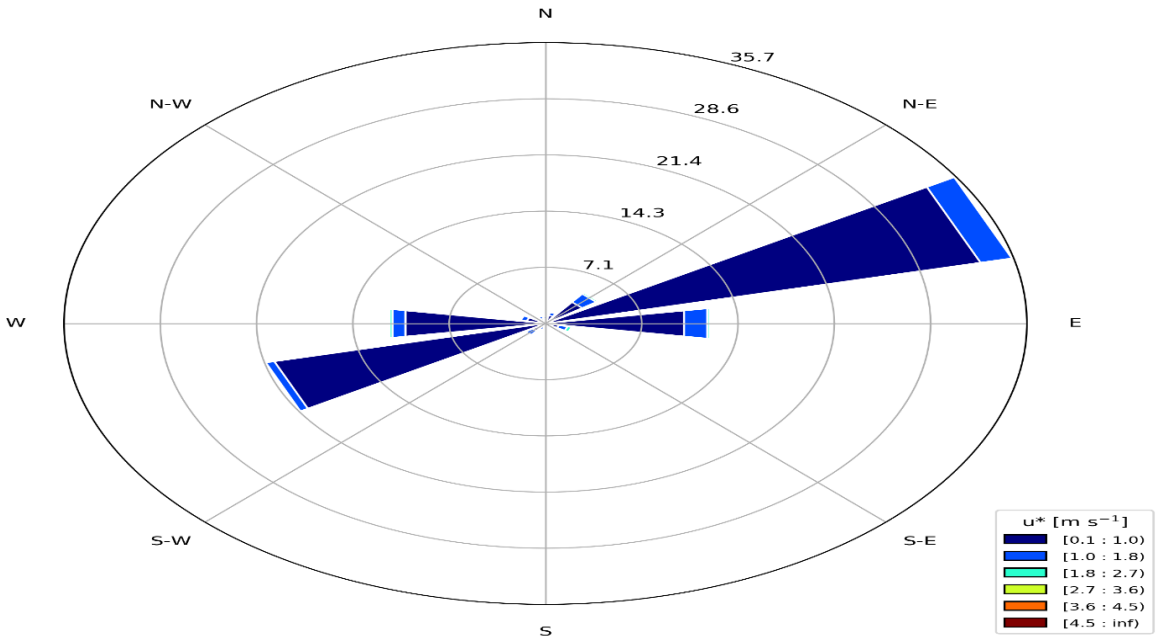


Figure 2. Windrose measurements from the Vuwani eddy covariance flux tower

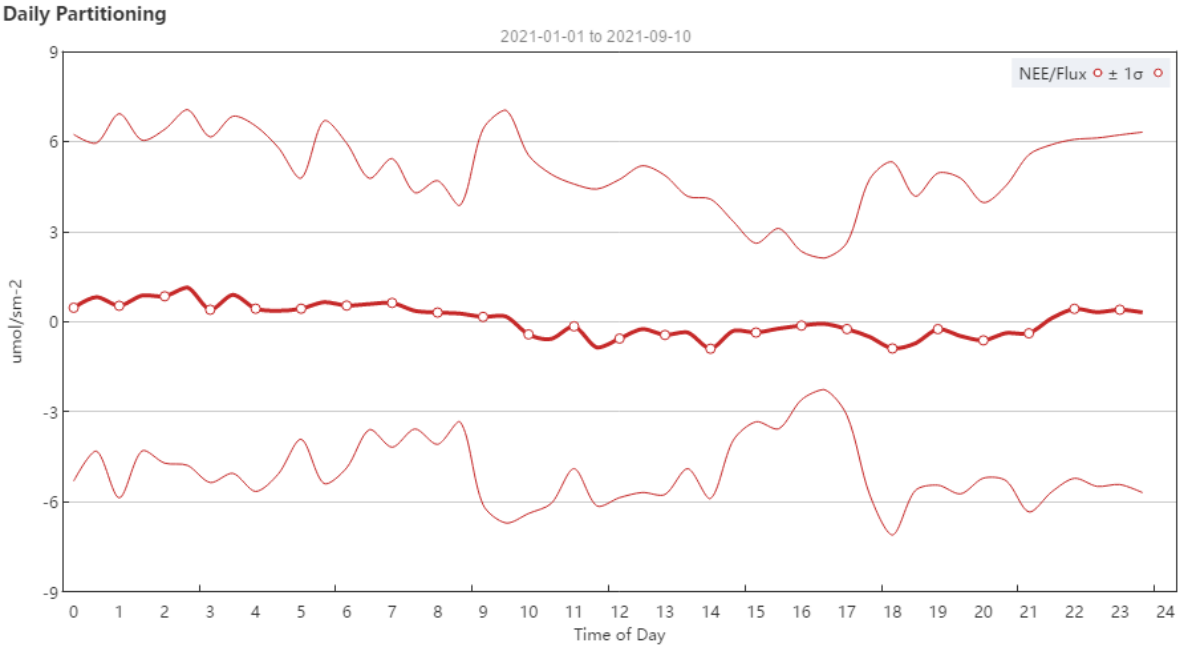


Figure 3. The Net ecosystem exchange (NEE), daily partition flux movement between 01 January 2021 to 10 September 2021 at Vuwani**.**

Figure 3 illustrates the variations in daily NEE sums throughout the seasons at the study site. Data analysis shows between March and April, there were exchanges of NEE, signifying the shift from the rainy to the dry season. This transition reached its peak at +6 $\mu mol.m^{-2}s^{-1}$ and -3 $\mu mol.m^{-2}s^{-1}$, respectively. Subsequently, from May to August, there was a predominant net uptake, reaching its zenith at -6 $\mu mol.m^{-2}s^{-1}$ by the end of August. This pattern aligned with the period of transition between the rainy and dry seasons. In contrast, for the Vuwani location, both uptake and flux were observed within the rainy season, with the highest uptakes occurring in April, August, and September. The occurrence of the highest daily NEE uptake during the rainy season throughout the study period underscores the significant roles that precipitation, vegetation, and soil moisture play in influencing the rates of CO_2 uptake and fluxes within the studied ecosystem.

3.2. Daily NEE and radiation pattern analysis

Figure 4 displays the relationship between the measured NEE and the incoming long-wave radiation, which spans from 300 w/m^2 to 520 w/m^2. In this range, the carbon dioxide exchange showcases a downward trend, reaching as low as -30 $\mu mol.m^{-2}s^{-1}$, while also demonstrating an upward uptake trend of 40 $\mu mol.m^{-2}s^{-1}$. Notably, there is a noticeable positive correlation observed between NEE and radiation across the entirety of the study duration. Throughout the investigation period, the average NEE stands at +6 $\mu mol.m^{-2}s^{-1}$.

3.3. Energy balance closure

The First Law of Thermodynamics, known as the conservation of energy, is relevant to energy balance. It asserts that energy can neither be created nor destroyed, only converted, or transferred between different forms. When applied to an energy balance, it signifies that the overall energy entering a system must match the energy exiting the system, accounting for various forms like potential, kinetic, and thermal energy.

Energy closure balance will be obtained by use of energy budget represented by equation 5, data is stored at the eddy covariance flux tower every 30 minutes.

$$\mathrm{Rn} - \mathrm{G} = H + \lambda \mathrm{E} \tag{5}$$

Where H is sensible heat flux, (λE) is Latent heat flux, Rn is net radiation, and G is ground heat flux. While $\mathrm{Rn} - \mathrm{G}$ is available energy and combination (H + LE) represents the total turbulent heat flux.

$$\text{Closure Residual} = (\mathrm{Rn} + \mathrm{G}) - (\mathrm{H} + \lambda \mathrm{E}) \tag{6}$$

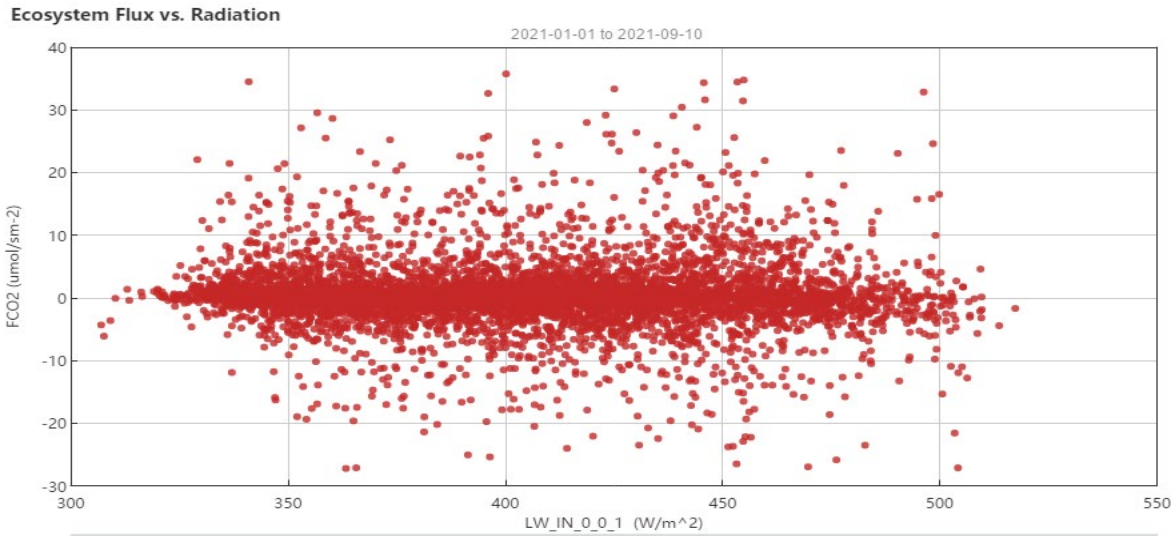


Figure 4. The scatter plot of ecosystem flux and radiation of long wave in the ecosystem of study.

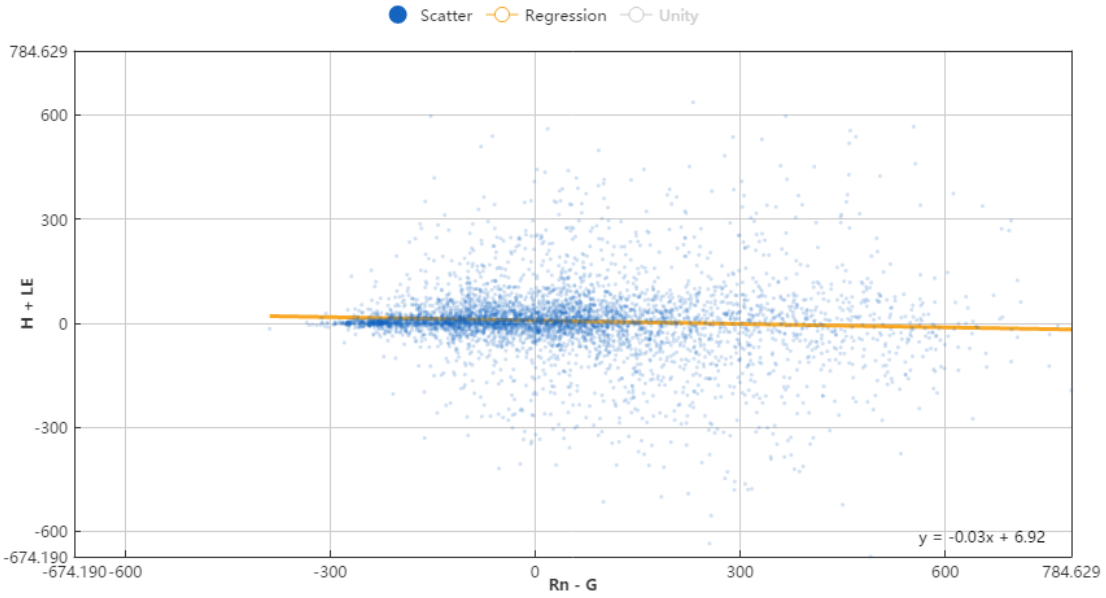


Figure 5. The energy closure based on net radiation and ground heat flux (Rn – G) and sensible heat flux and latent heat flux (H + LE).

The recorded energy fluxes, which encompass sensible heat flux, latent heat flux, net radiation, and soil heat flux throughout daylight hours, depict the motion of energy. There exists a prominent peak in the time series between 06:00 AM and 06:00 PM, followed by a decline in energy fluxes from 19:00 PM to 5:00 AM. The regression analysis between (Rn − G) and (H + LE) yielded slopes and intercepts of 0.03 and 6.92 W m^{-2}, as depicted in Figure 5.

The moderate slope indicated a reasonable energy balance closure for Vuwani. Our findings align with the typical observation that the sum of turbulence fluxes (H + LE) is usually 10 to 30% lower than the available energy (Rn – G) due to unaccounted large-scale transport and measurement errors in eddy covariance tower measurements.

The regression line approaching zero suggests potential total energy closure at zero. This implies that changes in meteorological conditions during dry and rainy seasons might have influenced the energy balance closure over the study period. The scattered nature of the observed points during the study could be attributed to the sensitivity of the EC measurement device to raindrops and dust. Unlike our focus on carbon dioxide fluxes, no adjustments were made in this study to compensate for energy balance closure imbalances, in line with Foken et al. 2012.

4. Conclusion

The study examined the Daily Net Ecosystem Exchange (NEE) of carbon dioxide, revealing seasonal variations and correlations with radiation. Figure 3 depicted NEE shifts between rainy and dry seasons, while Vuwani showed uptake during rainy months suggesting the influence of factors such as precipitation, vegetation, and soil moisture on CO_2 uptake and fluxes. The correlation between NEE and incoming radiation further underscored the intricate relationship between CO_2 exchange and radiation levels. Figure 4 highlighted a positive NEE-radiation correlation, with CO_2 exchange ranging from -30 to +40 $\mu mol.m^{-2}s^{-1}$. Energy balance closure was explored in Figure 5, indicating a moderate closure with slopes of 0.03 and 6.92 W m^{-2}. The regression line's approach to zero suggested potential energy closure, potentially influenced by meteorological changes. The study acknowledged device sensitivity and upheld Foken et al. 2012's approach to addressing energy balance disparities. Collectively, these findings provide valuable insights into the interplay between CO_2 fluxes, energy balance, and environmental variables in the area. This knowledge holds significant implications for climate change mitigation and environmental management strategies, contributing to informed decision-making for sustainable land use and carbon management efforts.

Acknowledgment

I extend my gratitude to Oksana Rybchak for data processing, Jens Jüdt meticulous tower maintenance, and the partnership with EMSAfrica and Univen significantly contributed to the success of this research project. Their expertise, support, and financial backing were essential in obtaining reliable results and insights.

References

[1] Bombelli, A., Henry, M., Castaldi, S., Adu-Bredu, S., Arneth, A., de Grandcourt, A., ... & Valentini, R. (2009). An outlook on the Sub-Saharan Africa carbon balance. Biogeosciences, 6(10), 2193-2205.

[2] Foken, T., Aubinet, M., Finnigan, J. J., Leclerc, M. Y., Mauder, M., & Paw U, K. T. (2012). Results of a panel discussion about the energy balance closure correction for trace gases. Bulletin of the American Meteorological Society, 93(4), ES73-ES76.

[3] Grünwald, T., Bernhofer, C., Goldberg, V., Wohlfahrt, G., & Kolle, O. (2007). Eddy covariance measurements of carbon dioxide, latent and sensible heat fluxes above a meadow on a forest–meadow transect in the Ore Mountains, Germany. Meteorologische Zeitschrift, 16(3), 275-288.

[4] Mauder, M., Foken, T., Clement, R., Elbers, J., Eugster, W., Grünwald, T., ... & Oncley, S. (2011). The energy balance experiment EBEX-2000. Part II: Intercomparison of eddy-covariance sensors and post field data processing methods. Boundary-Layer Meteorology, 139(1), 1-28.

Configuration of power chips for the TileCom for phase-II upgrades at CERN

MGD Gololo[1, 2]; CTM Tsima[2]; B Mellado[1, 3]

[1] CERN, Switzerland

[2] University of Johannesburg, South Africa

[3] University of Witwatersrand, South Africa

mpho@uj.ac.za216077192@student.uj.ac.za;bruce.mellado.garcia@cern.ch

Abstract. The large Hadron Collider (LHC) has four main experiments along the ring. The Toroidal LHC Apparatus (ATLAS) experiment along with the other experiments enable scientists to collect and analyse data. The LHC needs to be upgraded to a High Luminosity LHC to account for the increase in luminosity that will come with the upgrade and permit scientists to explore phenomena beyond the standard model. South Africa will contribute to the phase-II upgrade of the Tile Calorimeter, by producing the GbE switch and TileCom PCBs. These boards will connect the upgraded electronic chain to the network to monitor and allow for data transmissions at high speeds respectively. A TPS65086100 component needs to be configured before being populated onto the TileCom board for power purposes. This paper will illustrate the methodology for configuration, the tabulated output voltages for the voltage regulators, and the power-up sequence for the PMIC TPS65086100 chip.

1. Introduction

The Large Hadron Collider (LHC) is one of CERN's key achievements. It was constructed by CERN between 1998 and 2008 and it is the largest and most powerful particle collider in the world [1]. There are four places around the accelerator ring, which correspond to the locations of the four particle detectors ATLAS, CMS, ALICE, and LHCb that are used to cause the LHC's internal beams to clash and explain phenomena [1]. The Large Hadron Collider (LHC) will require a significant upgrade to maintain and increase its discovery potential. This will result in a five-fold increase in instantaneous luminosity (rate of collisions) over the original design value and a ten-fold increase in integrated luminosity (total collisions produced) [2]. This upgrade must be carefully planned out because the LHC is currently a very complicated and exquisitely optimized machine. This second-generation LHC (HL-LHC) was approved by the CERN Council in 2016 and is anticipated to begin operations after 2025. The High Luminosity LHC (HL-LHC) will enable advanced physics measurements but brings forth challenges to the detector and to the data and trigger systems respectively and will prompt the Tile Calorimeter's physics performance to require upgrades.

The goal of the technical design document of HL-LHC is to serve as the foundation for the comprehensive engineering design of the HL-LHC by describing the technologies and components that will be employed to realize the project [2].

2. ATLAS Tile Calorimeter Phase-II Upgrade

The ATLAS (A Toroidal LHC Apparatus) is one of the four experiments/detectors that are placed around the accelerator ring which are used to induce the internal beams of the LHC to collide [3]. The ATLAS experiment uses the compact muon solenoid for standard model research, which includes the Higgs Boson as well as search for additional dimensions and particles that might be involved in the formation of dark matter [4]. A large amount of data is generated by the interactions in the ATLAS detectors. ATLAS uses a sophisticated response system to instruct the detector which events to capture and which to ignore to process the data. After then, the recorded collision incidents are analysed using sophisticated data-acquisition and computation technologies.

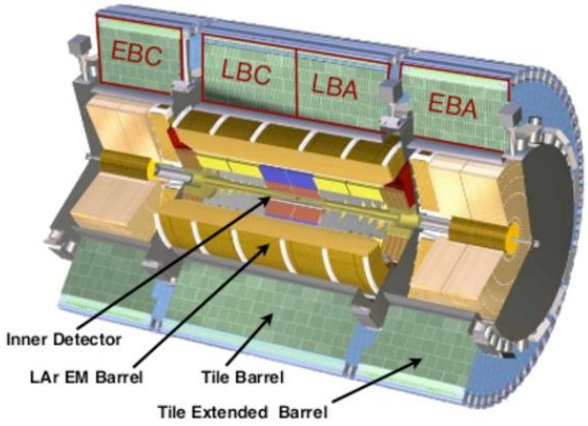


Figure 1: ATLAS Tile Calorimeter and Liquid Argon Calorimeter.

Figure 1 illustrates the ATLAS calorimetry system which consists of the Liquid Argon Calorimeter and the Tile Hadronic Calorimeter. The Liquid Argon Calorimeter measures the energy of electrons, photons, and hadrons around the ATLAS Inner Detector [5]. It is made up of many metal layers that trap incoming particles and change them into a collection of brand-new, lower-energy particles. These particles ionise the liquid argon between the layers, producing a measurable electric current [5]. By considering all the recorded data, physicists may determine the energy of the initial particle that struck the detector.

Hadronic particles do not completely deposit all their energy in the Liquid Argon Calorimeter, thus the Tile Calorimeter, which surrounds the Liquid Argon Calorimeter, monitors the energy of hadronic particles. Layers of steel and plastic tiles make up its structure. Particles create a shower of new ones as they strike the layers of steel. In turn, plastic scintillators generate photons, which are then transformed into an electric current whose strength is inversely proportional to the energy of the original particle.

TileCal is the ATLAS experiment's heaviest component. A considerable upgrade of practically all ATLAS detection systems is required since a luminosity increase of approximately an order of magnitude over the original design will result in a comparable rise in particle fluxes in the detector [6]. As a result of the increased trigger rates and detector occupancy brought on by the high luminosity of the HL-LHC for Phase-II, the detector as well as the trigger and data collecting systems must overcome significant obstacles [7].

Front-end electronics and photomultipliers (PMT) for 48PMT channels are found in several drawers which are mechanical structures and are housed in each tile module. [2] Each drawer has feeders mainly cooling lines, data links and low voltage power [2]. The HL-LHC phase-II upgrade will increase reliability of cooling circuits and easy access to the electronics behaviour during detector openings. The on-detector electronics of the TileCal are found in a mini-drawer and are connected to the mainboard which supplies LV and controls an amplifier card, digitalizes the signal and pushes data to the daughterboard which is responsible for comms at high speed with the PPr (back-end electronics pre-processor) [6]. The off-detector electronics aim to reduce the latency. The PMT digital signals are

transferred to the PPr at every event, data is processed at the PPr and sent to the trigger system. Trigger systems monitors and decides which data is kept and which is disposed.

The current ATLAS Tile Calorimeter cannot accommodate the increase luminosity due to the increased number of collisions. This results in frequent maintenance and an increase in maintenance services. Current architecture of the Tile Calorimeter is not capable of handling the amount of data that will be acquired for Phase-II upgrades. Some disadvantages of the present system include noise, poor quality factor, loss in latency and energy (overheat electronics). As part of the Phase-II upgrade of the Tile Calorimeter, two modules of the off-detector electronics for the TileCal electronic chain are currently in production. These two modules are called the TileCom and the GbE Switch. These two modules will enable the electronic chain to connect to the network, monitor certain data and enable the transmission of data within the chain. The TileCoM PCB currently in production will use a power chip (TPS65086100) to power the TileCoM. This paper presents the configuration of the TPS65086100 device which will be populated onto the TileCoM PCB to provide power.

3. Experimental setup

The TPS650861 device family is a single-chip Power-Management Integrated Circuit (PMIC) that may be set up for optimal output voltages and sequencing. The TPS650861 includes three controllers to offer adjustable power up to 30A with big external FETs for high power applications, which can scale down in size and cost for smaller systems [8]. The TPS650861 can supply system power for a range of applications when combined with three 3 A converters, three general purpose LDOs, a termination LDO for DDR, and three load switches [8]. The BOOSTXL-TPS650861 Booster Pack plug-in module for the TPS65086100 contains two banks of non-volatile one-time programmable (OTP) memory that can be programmed using an MSP430F5529 Launch Pad development kit or directly on the chip during manufacture [9].

A modular, programmable, and adaptable power solution is offered by the TPS650861 PMIC, which can power a variety of CPUs, DDR3/DDR4 memory, and other peripherals [8]. Three step-down controllers (Buck 1, 2, and 6), three step-down converters (Buck 3, 4, and 5), a source or sink LDO (VTT LDO), three low-voltage VIN LDOs (LDOA1-LDOA3), and three load switches (SWA1, SWB1, and SWB2) are all included into the PMIC [8]. The configuration of each rail for the default output value, power-up sequence, fault management, and Power Good mapping into a GPO pin are all easily configurable with on-chip one-time programmable (OTP) memory [8]. The adjustable current limit (controlled by an external resistor at the ILIMx pin) of the buck controllers' integrated gate drivers for external power stages enables the best possible selection of external passive components depending on the required system load [8].

The power input, inductor, and output voltage feedback input pin requirements for buck converters with integrated power stages are minimal [8]. All of these qualities, when combined with high-frequency switching, enable the use of inductors in compact form factors, lowering the cost and size of the whole system. The default V_{OUT} can be utilised with all controllers and converters, or their voltage can be dynamically altered at any moment. So that the device starts up with the default voltage, the rails may be default-programmed for any available VOUT via OTP programming locally or at the factory. Alternatively, the rails can be configured by I2C while the device is in operation to another functioning V_{OUT} while the rail is enabled or disabled [8]. There are different software and hardware required to configure the TPS65086100 power chip before it can be populated onto the TileCom PCB. The rest of the paper will give the methodology, expected and measured results.

The MSP430F5529 LP Evaluation board and the BOOSTXL-TPS650861 Booster Pack previously mentioned are shown in figure two respectively. These two boards that will be used for the testing stages for the power chips before the chips will be populated onto the TileCoM board.

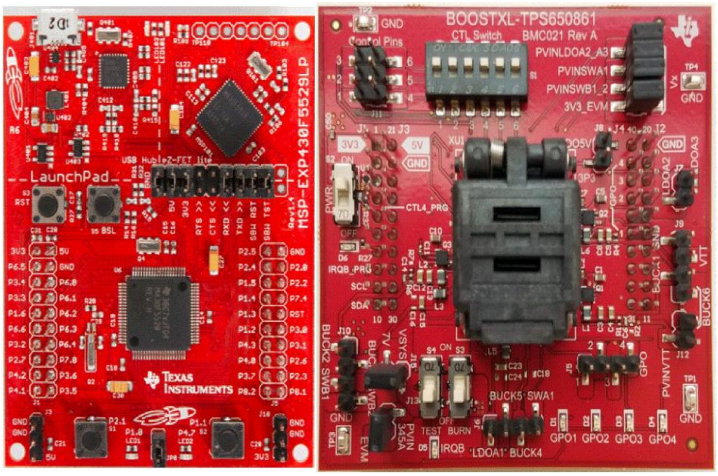


Figure 2:MSP430F5529LP Development board and BOOSTXL-TPS650861 BoosterPack plug-in module.

4. Methodology

The TileCom and GbE Switch board have been designed and manufactured. The PMIC device needs to be configured before it can be populated onto the TileCom to provide power for the FPGA. The block diagram below shows the methodology followed to configure the TPS65086100 PMIC:

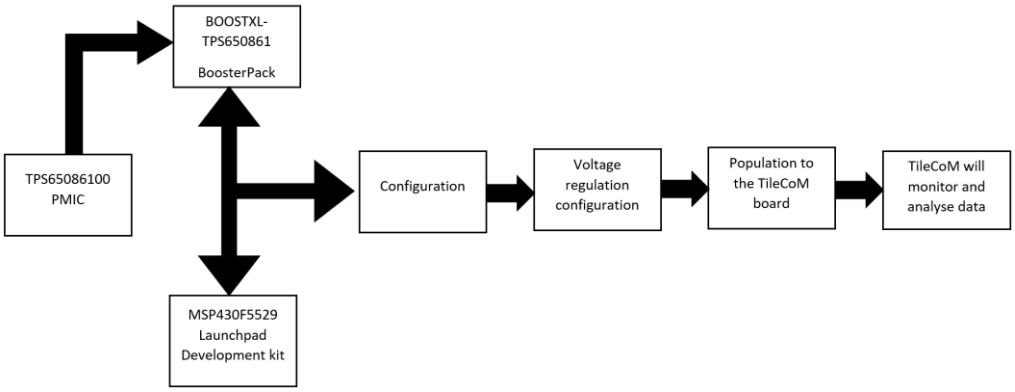


Figure 3: Block diagram showing the methodology followed for the PMIC configuration.

4.1. Programming Steps

Use the firmware example to update the MSP430F5529LP development board firmware using the USB2ANY text file. The TPS65086100 OTP generator tool (TPS6508641 settings) is used to match the desired settings when the overview, sequencing and additional tabs in the OTP generator are modified. Under the "script" tab copy the B2-B96 cells onto a TPS650861-Script-1p0.js file which will be extracted in the GUI. The booster pack plug-in is plugged into the Launchpad, the GUI is started and use the select device option to go drag the TPS650861 device file previously downloaded. Turn ON S2 on the BOOSTXL-TPS650861 plug-in module. By accessing the SHUTDNSRC register (0x05) from the Register Map or Register Controls tabs in the IPG-UI EVM GUI, you may verify the connection between the PMIC device and the IPG-UI EVM GUI. It ought to come back as 0x04, denoting that under voltage lockout was the reason of the most recent shutdown [10]. On the devices controls programming tab ensure all the five steps are done making sure that the correct OTP program (0x141) is chosen. Ensure that the measured output voltages of the regulators are favorable before programming the OTP onto the PMIC.

5. Results and discussion

The figure below shows the experimental setup for the TPS65086100 PMIC configuration using the evaluation board and the BoosterPack:

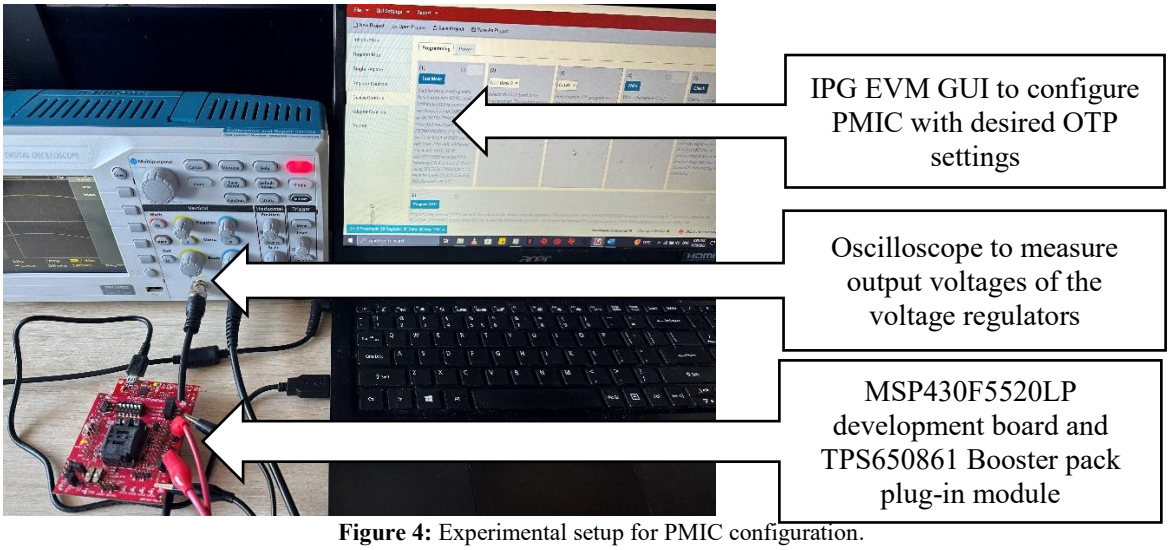


Figure 4: Experimental setup for PMIC configuration.

Table 1 below shows the output voltages for the voltage regulators after the configuration of the power chip using an OTP program:

Table 1: Voltage Regulators Output voltage.

Rail	Type	Input Voltage (V)		Output Voltage Range (V)		Measured Voltage (V)
		MIN	MAX	MIN	MAX	MAX
BUCK1	Step-down controller	4.5	21	0.41	3.575	3.2
BUCK2	Step-down controller	4.5	21	0.41	3.575	3.1
BUCK3	Step-down converter	3	5.5	0.41	3.575	2.8
BUCK4	Step-down converter	3	5.5	0.41	3.575	2.8
BUCK5	Step-down converter	3	5.5	0.41	3.575	2.9
BUCK6	Step-down controller	4.5	21	0.41	3.575	3.0
LDOA1	LDO	4.5	5.5	1.35	3.3	2.7
LDOA2	LDO	1.62	1.98	0.7	1.5	0.9
LDOA3	LDO	1.62	1.98	0.7	1.5	1.0
SWA1	Load Switch	0.5	3.3			
SWB1/SWB2	Load Switch	0.5	3.3			
VTT	Sink and source LDO	1.1	1.8			

For LDOA1, LDOA2 and LDOA3 outputs, ground is bypassed with a 4.7uF ceramic capacitor but when not in use you leave floating. These load switches can monitor their output voltage and report good and fault conditions. Can configure GPO to "Power good" mode, when the unmasked power good signals are high it will go high as well by monitoring the unmasked power good signals it is given in the sequencing tab of the excel OTP tool.

The power sequence of a particular rail is dependent on the "power good" of the previous rail resulting in a clean staggered power up sequence. Ensure that the power rails are not enabled at the same time to decrease the inrush current of the supply. The rails are also enabled by CTLx pins and I2C software command. Power down sequences are enabled by using falling delays. When analysing the power down sequence for Buck1-Buck4, the bucks will have different falling edge delays, for example Buck 1 and Buck 2 falling edge delay is 4ms for the power sequence and for Buck 3 and Buck 4 it is 2ms and 0ms respectively. The falling edge delay decreases as the power down sequence number increases. The buck controllers have high frequency, react fast and a capable of driving two external N-MOSFETS. The buck converters provide excellent transient and AC load regulation and reduces switching noise of external filter components.

6. Conclusion

The off-detector is the scope of the research, where two boards mainly the TileCom and GbE Switch PCB boards will be designed and built by SA-CERN for the LHC Phase-II upgrade. Phase-II upgrade enables the system to work efficiently and fast compared to the system that's currently operating. The output voltages of the buck converters, buck controllers and LDOs are desirable for the TPS65086100 PMIC device. Once populated on the TileCom board the PCB will be powered accordingly.

The TileCom board will enable the new system to transmit, monitor and store data at larger quantities with no loss of energy and latency. The GbE Switch will enable the electronic chain to connect to the internet. There are other PCBs under the off detector which will work hand in hand with the TileCom and the GbE switch PCBs. The collaboration with different countries for the ATLAS TileCal Phase-II upgrade will bring forth increased knowledge,

References

[1] CERN, "Large Hadron Collider," *CERN*, vol. 1, pp. 1–32, 2012, [Online]. Available: https://en.wikipedia.org/wiki/Large_Hadron_Collider.

[2] T. D. R. V, O. Brüning, M. Lamont, L. Rossi, and L. Tavian, "HL-LHC Technical Design Report," vol. 4, 2017.

[3] Encyclopedia, "CERN," *Wikipedia*. Wikipedia, pp. 1–50, [Online]. Available: https://en.wikipedia.org/wiki/CERN#:~:text=The European Organization for Nuclear,physics laboratory in the world.

[4] C. J. Rhodes, "Large Hadron Collider (LHC)," *Sci. Prog.*, vol. 96, no. 1, pp. 95–109, 2013, doi: 10.3184/003685013X13623370524107.

[5] ATLAS; Joao Pequenao, "Calorimeter," *ATLAS experiment*, 2008. https://atlas.cern/Discover/Detector/Calorimeter (accessed Feb. 17, 2023).

[6] B. Collaboration, "Technical Design Report," *Cern/Lhcc*, vol. 11, no. July, pp. 1–26, 2007.

[7] CERN, "ATLAS," *CERN*, 2022. https://home.cern/science/experiments/atlas (accessed Feb. 17, 2023).

[8] C. Each *et al.*, "TPS650861 Programmable Multirail PMU for Multicore Processors , FPGAs , and Systems," 2018.

[9] U. Guide, "TPS65086100 Non-Volatile Memory Programming Guide," no. August, pp. 1–12, 2018.

[10] U. Guide, "BOOSTXL-TPS650861 EVM User ' s Guide," no. August, pp. 1–8, 2018.

O_2 adsorption on $PtSb_2$ (100) surface

S.S.Mangwejane, P.P.Mkhonto and P.E.Ngoepe

Materials Modelling Centre, University of Limpopo, Private Bag x 1106, Sovenga, 0727, Republic of South Africa

E-mail: **seshupo.mangwejane@ul.ac.za**

Abstract. In order to understand the oxidation of the $PtSb_2$ minerals by either weathering or exposure to air, we considered the adsorption of oxygen molecules on mineral surfaces. The adsorption properties will give more insight into the interaction of the oxygen with the atoms on the surface. Generally, minerals are exposed to atmospheric air during crushing or storage. It had been reported previously that oxidation depresses sulphide minerals, resulting in poor grade-recovery performance and the brittle nature of $PtSb_2$ may result in difficulty of its recovery from such ores. The surface energy of (100) has been found to be 0.920 Jm^{-2} as compared to (110) 1.194 Jm^{-2} and (111) 0.939 Jm^{-2}. We performed oxidation on the $PtSb_2$ (100) surface where O_2 was adsorbed in a peroxo, superoxide and Pt-O-O-Sb bridging on the surface Pt and Sb atoms.

1. Introduction

$PtSb_2$ is used in magneto-resistance products and display superconductivity and in medical and aerospace industries and jewellery [1]. In order to understand the oxidation of the $PtSb_2$ minerals either by weathering or exposure to air, we considered the adsorption of oxygen molecules on mineral surfaces. The adsorption properties will give more insight into the interaction of the oxygen with the atoms on the surface. Generally, minerals are exposed to atmospheric air during crushing or storage. It had been reported previously that oxidation depresses sulphide minerals [2], resulting in poor grade-recovery performance and the brittle nature of $PtSb_2$ may result in difficulty of its recovery from such ores. Furthermore, it had also been highlighted that the use of oxygen as a bubbling gas, was effective since it increased the pulp potential [3].

The ground state of isolated O_2 molecule was experimentally found to be a spin triplet state, with the equilibrium bond length of 1.21 Å [4]. This corresponded to the current calculated equilibrium oxygen molecule bond length of 1.211 Å. We investigated the preferential oxidation site using DFT on the $PtSb_2$ and described the bonding character of O_2 on the surfaces. On this adsorption, we positioned the O_2 on the surface in the vertical orientation. This is in line with previous studies of oxygen-iron/nickel interaction where the superoxide isomers of the iron and nickel favour the vertical orientation [5].

The adsorption energies of the oxidations on the three surfaces will also be discussed. Firstly, we started by adsorption of the O_2 molecule on the surfaces and examined their bonding geometric behaviour (bond length and angles). Then the Mulliken population charges were analyzed and discussed to fully describe the chemistry of the adsorptions. Other paragraphs are indented.

2. Methodology

We have employed the DFT within CASTEP code of Materials Studio, using the plane-wave with PBE exchange-correlation functional. The plane-wave basis set with a cut-off energy of 500 eV was set, which demonstrated convergence to within 0.2 meV/atom. We employed the ultrasoft pseudopotentials [7, 8] and the electron configurations considered for mineral surface were: Pt: $[Xe]4f^{14}5d^{9}6s^{1}$ and Sb: $[Kr]4d^{10}5s^{2}5p^{3}$ and for adsorbates we focused on the atoms that binds to the surface, which were: S:$[Ne]3s^{2}3p^{4}$, O: $[He]2s^{2}2p^{4}$. A k-point grid of 6x6x6 for the bulk and 4x4x1 for surfaces were generated using the Monkhorst-Pack [9] scheme. The size of the k-point grid used for mineral represents a k-point sampling spacing no greater than 0.05 $Å^{-1}$. The convergence tolerances for force, ionic displacement and energy were 0.05 $eVÅ^{-1}$, 0.001 Å and 0.01 meV/atom, respectively. Slab depth for each surface model will be listed for each model. The vacuum height (as discussed in the following section) for all surface models in this study was set at 20.0 Å in order to avoid the interaction of the adsorbates with the upper repeating slab model. Throughout this work model, total energies (in eV) will be quoted to three(3) decimal places to allow for future in-depth model comparison and optimisation by other researchers, while the surface and adsorption energies will be quoted to three decimal places.

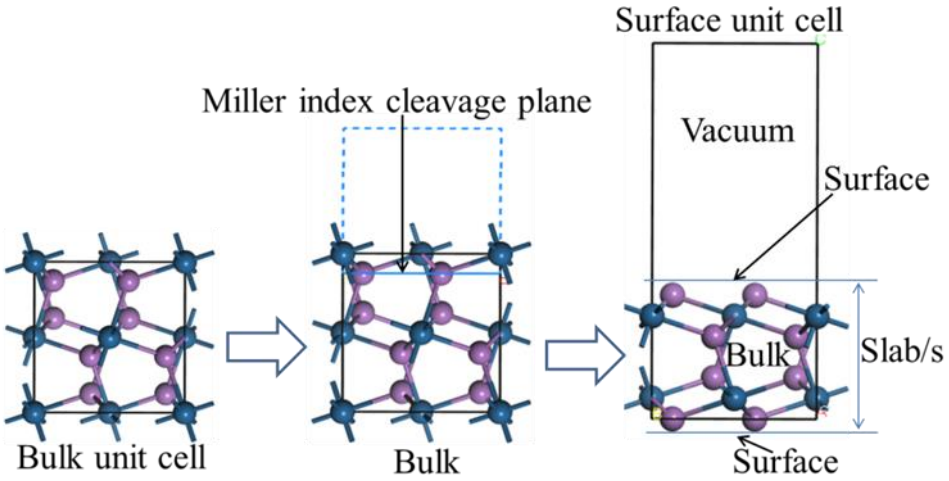


Figure 1: Bulk and cleavage for surfaces.

3. Results and Discussion

We observed that the peroxide adsorption on Pt resulted in migration of oxygen molecule to the Sb atoms forming a Sb-mononuclear bonding (superoxide) and gave a Sb–O–O bond angle of 113.80º. The peroxide adsorption on Sb showed an over stretched O–O bond to 1.541Å, which was due to electron transfer from the Sb atoms to the LUMO p-orbital of the oxygen thus weakening the O–O bond. This is evident from atomic charges in Table 3, where the oxygen had more negative charge indicating electron gain. The two Sb–O bond lengths were noted to have one slightly larger than the other and the O–Sb–O was found to be 43.95º. Furthermore, the Sb–O–O–Pt bridging was observed to have poor adsorption on the Pt and thus results in Sb–O–O bonding. The O–O bond was also stretched to O–O = 1.417 Å and the Sb–O–O bond angle was found as 102.88º.

.

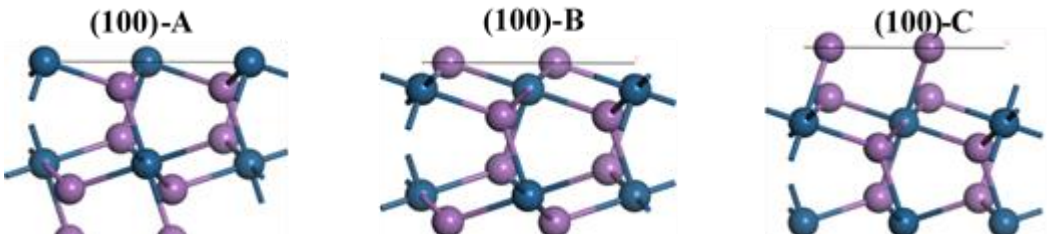


Figure 2: The different surface terminations for $PtSb_2$ cleaved from low miller index (100) plane.

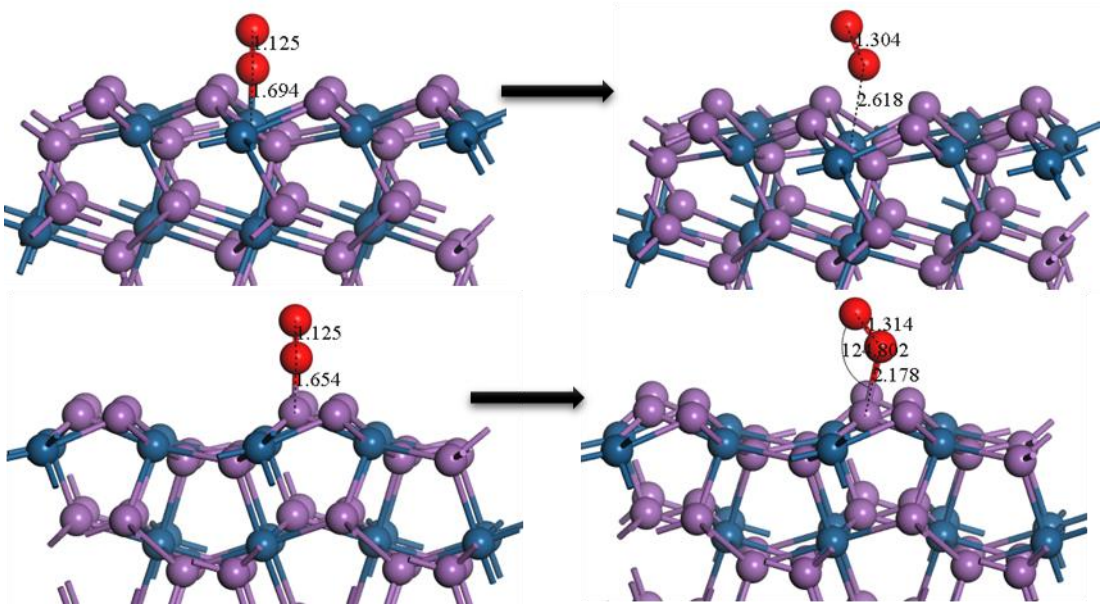


Figure 3: The side view of superoxide O_2 molecule adsorption on Pt-top and Sb-top sites on the $PtSb_2$ (100) surface.

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

Table 1: Surface energies for different terminations

Miler Index Plane	Surface Termination	Stoichiometry	Surface energy ($J.m^{-2}$)
			Unrelaxed
(100)	(100)-A	$PtSb_2$	1.569
	(100)-B	*$PtSb_2$*	*0.920*
	(100)-C	$PtSb_2$	1.569

The model for the adsorbed adsorbate plus surface, surface, adsorbate and number of adsorbate are referred to as [S + A], [S], [A], and *n*, respectively. The superscript numbers will represent any charged states. The adsorption energies ($E_{ads.}$) of an uncharged adsorbate (O_2) to a surface was calculated from the difference in total energy between the neutral adsorbate surface ($[S + A]^0$) and the sum of the neutral adsorbate ($[A]^0$) and neutral surface ($[S]^0$) models, as expressed in:

$$E_{ads.} = [[\mathrm{S} + \mathrm{A}]^0 - ([\mathrm{S}]^0 + n[\mathrm{A}]^0)]/n$$

This equation is applicable for the neutral adsorbate as the atom and electron counts for $[S + A]^0$ and $([A]^0 + [S]^0)$ are the same. However, if [A] carries a formal negative charge, then the [S + A] must also carry a negative charge in order to balance the electrons. This presents a problem, as the periodic-boundary-condition box sizes required to demonstrate effective energy convergence would be prohibitively large as also reported by Waterson et al. [7]. Neugebauer and Scheffler [10], which utilises the work function (ϕ), to correct the extra electron in the electronic band structure, reported a solution to the problem. This was obtained by replacing the total energy of the $[S + A]^-$ system by $[S + A]^0$. We have calculated this correction for the [S + A] system and full details of our results are shown in the sections discussed as reported by Waterson et al [7]. As such the adsorption energies for the charged OH^- and xanthates adsorbate $[A]^-$to a neutral surface $[S]^0$ is computed from:

$$E_{ads.} = [([\mathrm{S} + \mathrm{A}]^0 - \phi_{[S+A]^0}) - ([\mathrm{S}]^0 + n[\mathrm{A}]^-)]/n$$

Within these calculations, negative adsorption energy corresponds to an exothermic adsorption process and a positive adsorption energy corresponding to an endothermic process. The adsorption energies are given in kJ/mol, a convention by 1.0 (eV) = 96.485 (kJ/mol)/eV.

Table 2: Adsorption energies.

Adsorption	Atom site	$[S+A]^0$ (eV)	$[A]^0$ (eV)	$[S]^0$ (eV)	Total $E_{ads.}$ (kJ.mol^{-1})
Superoxide	Pt1	–49897.441	–868.856	–49028.608	+2.21
	Sb2	–49897.335	–868.856	–49028.608	+12.49
	Sb3	–49897.637	–868.856	–49028.608	–16.63
Peroxide	Pt2	–49897.717	–868.856	–49028.608	–24.38
	Sb2	–49897.910	–868.856	–49028.608	–43.03
	Sb3	–49898.192	–868.856	–49028.608	–70.16
Bridge	Sb3–O–O–Pt2	–49897.992	–868.856	–49028.608	–50.89
	Sb2–O–O–Sb3	–49898.830	–868.856	–49028.608	–131.79

Table 3: The calculated atomic population (Mulliken) charges of superoxide and peroxide O_2 molecule adsorption on $PtSb_2$ (100) surface**.**

Adsorption	Atom	Adsorptions state	Mulliken population charges (|e$^-$|)				
			s	p	d	Total	Charge
	Pt	Before adsorption	1.00	0.85	9.08	10.97	–0.97|e$^-$|
		After adsorption	0.97	0.92	9.02	10.91	–0.91|e$^-$|

$Pt–O_2$ Superoxide	O1	Before adsorption	1.88	4.13	0.00	6.00	0.00|e⁻|
		After adsorption	1.88	4.33	0.00	6.21	–0.21|e⁻|
	O2	Before adsorption	1.88	4.13	0.00	6.00	0.00|e⁻|
		After adsorption	1.90	4.27	0.00	6.17	–0.17|e⁻|
$Sb–O_2$ Superoxide	Sb	Before adsorption	1.84	2.95	0.00	4.79	+0.20|e⁻|
		After adsorption	1.81	2.70	0.00	4.51	+0.49|e⁻|
	O1	Before adsorption	1.88	4.13	0.00	6.00	0.00|e⁻|
		After adsorption	1.87	4.39	0.00	6.27	–0.27|e⁻|
	O2	Before adsorption	1.88	4.13	0.00	6.00	0.00|e⁻|
		After adsorption	1.90	4.30	0.00	6.20	–0.20|e⁻|
$Pt–O_2$ Peroxo	Sb	Before adsorption	1.84	2.95	0.00	4.79	+0.20|e⁻|
		After adsorption	1.81	2.72	0.00	4.53	+0.47|e⁻|
	O1	Before adsorption	1.88	4.13	0.00	6.00	0.00|e⁻|
		After adsorption	1.90	4.43	0.00	6.33	–0.33|e⁻|
	O2	Before adsorption	1.88	4.13	0.00	6.00	0.00|e⁻|
		After adsorption	1.89	4.38	0.00	6.27	–0.27|e⁻|
$Sb–O_2$ Peroxo	Sb	Before adsorption	1.84	2.95	0.00	4.79	+0.20|e⁻|
		After adsorption	1.74	2.41	0.00	4.15	+0.85|e⁻|
	O1	Before adsorption	1.88	4.13	0.00	6.00	0.00|e⁻|
		After adsorption	1.94	4.52	0.00	6.47	–0.47|e⁻|
	O2	Before adsorption	1.88	4.13	0.00	6.00	0.00|e⁻|
		After adsorption	1.95	4.50	0.00	6.44	–0.44|e⁻|
$Pt–O_2–Sb$ Bridge	Sb	Before adsorption	1.84	2.95	0.00	4.79	+0.20|e⁻|
		After adsorption	1.80	2.64	0.00	4.43	+0.57|e⁻|
	O1	Before adsorption	1.88	4.13	0.00	6.00	0.00|e⁻|
		After adsorption	1.91	4.74	0.00	6.37	–0.37|e⁻|
	O2	Before adsorption	1.88	4.13	0.00	6.00	0.00|e⁻|
		After adsorption	1.90	4.40	0.00	6.31	–0.31|e⁻|

The atomic charges computed for (100) surface were analysed and clearly indicated that the Sb atoms adopted more positive charge (lose charges), while the O_2 molecule adopted more negative charge (gain charges) as shown in Table 3. Note that all Sb-peroxide adsorptions resulted in Sb-superoxide bonding mode, except for the $Sb–O_2$ peroxide which did not change the bonding mode. We found that for all superoxide modes, the Sb-bonded oxygen was more negative than the terminal or end-bonded oxygen atom. This charge distribution behaviour for superoxide has been previously reported [11, 12]. The charge on O_2 (i.e., charge sum of O1 and O2), for lowest ground state were as: for $Pt–O_2$–superoxide ($-0.38|e|$), for $Sb–O_2$–superoxide ($-0.47|e|$), for $Pt–O_2$–peroxide ($-0.60|e|$), $Sb–O_2$–peroxide ($-0.91|e|$) and $Pt–O_2$–Sb-bridge ($-0.68|e|$). It was clear that the $Sb–O_2$ (peroxide) had more negative charge than all the superoxide modes, suggesting that there was more charge gain on the O_2 molecule. This was due to both oxygen atoms bonded to the Sb atom. For the superoxide, we observed that the O_2 on $Pt–O_2–Sb$ bridge, gained more charge than all other superoxide modes. The negative charge on the oxygen atoms was also confirmed by experiment for superoxide or end-bonded O_2 [13].

Table 3 clearly indicated that there was a greater loss of electrons in the 5p-orbital of Sb atoms with more gain of electrons in the 2p-orbital of O_2 molecule. This suggested that there was a strong hybridisation of the 5p-orbital of Sb and 2p-orbital of O_2 molecule. These effects may suggest that the Sb atoms take the form $[Kr]5s^25p^2$ electron configurations. However, there was some small charge loss on the 5s-orbital of Sb atoms and these may have been distributed to the 2s-orbital of O_2 molecule.

4. Conclusions

The oxidation showed that the oxygen molecules preferred interacting with the Sb atoms than the Pt atoms for all surfaces. We observed different bonding modes, where bridging and dissociation into atomic bonding occurred. For the (100) surface we found that the Pt-O_2 peroxide initial adsorption site gave the strongest adsorption. In the case of the (110) surface the Sb2-O-O-Sb3 bridging yielded the most exothermic adsorption. The (111) surface showed the Sb2-O-O-Sb2 bridging as the strongest exothermic adsorption, which dissociated and resulted in atomic bonding. The atomic charges indicated that the oxygen molecules gain charges from the Pt and Sb atoms. These findings suggested that the 2p-orbital spin-down un-occupied orbital (LUMO) of O_2 was fully occupied. We also observed that the Sb/Pt-bonded oxygen molecules were more negative than the terminal or end-bonded oxygen atom for superoxide bonding modes. Furthermore, based on the charge sum of the O_2, i.e. charge addition of O1 and O2, we found that in all cases the O_2 interacting with Sb gained more charges, thus showing that the Sb atoms were the preferred sites for oxygen adsorption.

Acknowledgements

This work was supported and performed at the Materials Modelling Centre (MMC), University of Limpopo. Computing resources were provided by the Centre for High Performance Computing (CHPC). We acknowledge the National Research Foundation (NRF) for financial support.

References

[1] Zhang K, Wang X, ,Zheng P.Y.,Huang K.K.,Hu C.Q.,Feng S.H ,Wen M and Zheng W.T.,(2017), The investigation of magneto-transport properties of $PtSb_2$ single crystal synthesized by sb-flux method, Journal of Alloys and Compounds **694** 935-938.

[2] P. Clarke, D. Fornasiero, J. Ralston and R. St. C. Smart (1995), A Study of the removal of oxidation products from sulphide mineral surfaces. Minerals Engineering, Vol. 8, No. 11, pp. 1347-1357.

[3] Ferihan G. (2002). Effect of pH on pulp potential and sulphide mineral flotation. Turkish Journal Engineering Enviromental Science, **26**, 309–318.

[4] Weast R. (1985). CRC Handbook of Chemistry and Physics. FL, USA: CRC, Boca Raton.

[5] Gutsev G.L, Rao B.K. and Jena P (2000), Experimental and theoretical study of the photoelectron spectra of MnO^{-1}_{x}(x=1–3) clusters, Journal of Chemical Physics, Vol 113, Number 4.

[6] Kleinman L. and Bylander D. M. (1982). Efficacious form for model pseudopotentials. Physical Review Letter, **48**, 1425–1428.

[7] Vanderbilt D. (1990). Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. Physical Review B., **41**, 7892–7895.

[8] Monkhorst H.J and Pack J.D., Special points for Brillonin-zone integrations, Physical Review B, **13,** 12, 5188-5192.

[9] Waterson C. N, Tasker P. A., Farinato R., Nagaraj D. R., Shackleton N. and Morrison C. A., (2016). A Computational and Experimental Study on the Binding of Dithio Ligands to Sperrylite, Pentlandite, and Platinum. Journal Physical Chemistry C, **120(39)**, 22476–22488.

[10] Neugebauer J. and Scheffler M., (1992), Adsorbate-substrate and adsorbate-adsorbate interactions of Na and K adlayers on Al (111), Physical Review B, **46**, 16067–16080.

[11] Mkhonto P. P., Chauke H. R. and Ngoepe P. E., (2015) Ab-initio Studies of O_2 Adsorption on (110) Nickel-Rich Pentlandite ($Fe_4Ni_5S_8$) Mineral Surface, Minerals, **5**, 665–678.

[12] Uzunova, E. L.; Mikosch, H.; Nikolov, G. St. (2008). Electronic structure of oxide,peroxide, and superoxide clusters of the 3d elements: A comparative functional study. Journal of Chemical Physics., **128**, 094307-1–094307-8.

[13] Lever A. B. P., Ozin G. A. and Gray H. B. (1980). Electron transfer in metal–dioxygen adducts. Inorganic Chemistry, **19**, 1823–1824.

Setting up an environment for extracting and analyzing data from the DCS ATLAS experiment for the behavior of High Voltage channels

Sanele S Gumede[1], Lungisani S Phakathi[1], Betty Kibirige[1], Juraj Smiesko[2] and Filipe Martins[3]

[1]Department of Physics and Engineering, University of Zululand, KwaDlangezwa Campus, KwaZulu Natal, South Africa.
[2]Institute of Particle and Nuclear Physics, Charles University.
[3]Laboratório de Instrumentação e Física Experimental de Partículas.

E-mail: sanele.scelo.gumede@cern.ch

Abstract. A hadronic calorimeter called Tile Calorimeter (TileCal) can be discovered in the center of the ATLAS detector. A sample calorimeter called TileCal employs steel plates as the absorber and plastic scintillating tiles as the active medium. Because about 30 percent of the total energy jets produced in a proton-proton collision are deposited in TileCal, it is crucial in the accurate reconstruction of the kinematics of the physics events. The main objective of this project is to develop the plugin that will be used to retrieve data from the Detector Control System (DCS) Data Viewer (DDV) server and analyze it for anomalies in the HV channels. The plugin will then be incorporated into the Tile-in-One platform, which unifies all TileCal work onto a single platform. The results shows that the plugin is able to extract the data from the DDV serve and visualize it.

1. ATLAS Tile Calorimeter

ATLAS [1] stands for A Toroidal LHC ApparatuS, it is one of the most massive detector at the Large Hadron Collider (LHC). The ATLAS detector's function is to investigate wide range of physics such as study of Higgs and boson, and other Standard Model phenomena to the searches of extra dimensions and particles that could make up dark matter. Figure 1 shows the ATLAS detector, the detector is 44 m long barrel with a diameter of 25 m, where the sub-detectors and magnets are arranged in cylindrical layers around the beam pipe. The sub-detectors are arranged in this order: Inner Detector (ID), Solenoid Magnet, Electromagnetic Calorimeter, Hadronic Calorimeter, Toroid Magnet and Muon Spectrometer.

In this poject we look inside of the ATLAS detector, where we look at the high voltage system of the Tile calorimeter (TileCal) of the ATLAS experiment. This system is appropriate to deliver a reliable power distribution into particles physics detectors using a huge quantity of PhotoMultiplier Tubes (PMTs). The high voltage system has been evolved to deliver every of the approximately ten thousand PMTs of the ATLAS TileCal with a voltge starting from 500 to 900 V, the relation among the gain G of PMTs and the implemented high voltage is given via way of means of G is proportional to HV^{β}, where β is particular to every PMT [2].

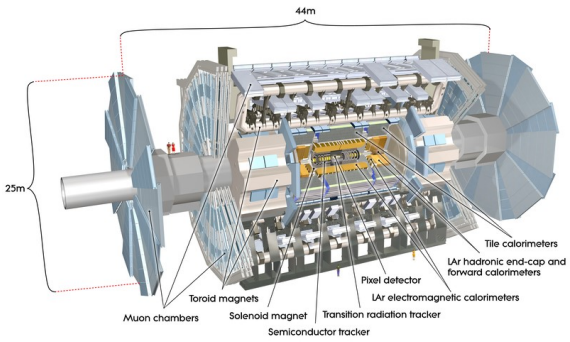


Figure 1. ATLAS detector shown in grey [4]

Note that any variation of high voltage would induce a gain variation deteriorating the measurement of the energy deposited by the particles in the calorimeter cells. The aim of this project is to build a web based tool (plugin) that will monitor and control high voltage data offline for better analysis. This web tool will be integrated to the platform called Tile-in-One (TiO) that is described in section 2 further. TiO plugin will be developed in order to visualize and convey complex information easily of the data that will be extracted from the Detector Control System Data (DCS) using DCS Data Viewer (DDV) server. Once the data is being visualized then the analysis take place for the behavior of HV channels in order to acquire a better understanding of the detector and detecting anomalies.

2. Tile-in-One Platform

A number of tools were developed in order to monitor and maintain TileCal activities, these were developed by different group of collaboraters. Since the tools were developed by different groups, it make use of distinctive technologies and data sources require different ways of data recovery. This results in time consuming and costly work since collaborators have to browse among a number of tools in order to perform a given task [3].

The Tile-in-One (TiO) web platform, is a platform that aim integrate all TileCal web tools into one platform, which will share the computing infrastructure and access to same data and services inside or outside the TileCal Collaboration.

2.1. TiO Architecture

The TiO platform is configured with flexibility and ease of maintenance [4] as shown in Figure 2. TiO design is based on one main server that is in charge of the secure connection to the platform, authentication of the users and also taking care of the user requests and responses. Plugin are are small host web applications that are behind the main server, they receive the user's requests, process the request and returns the response to the user.

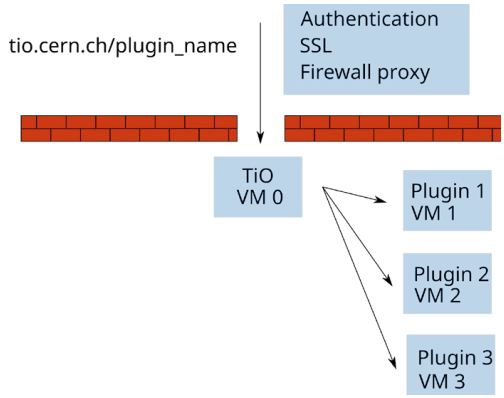


Figure 2. Architecture of TiO platform

The important key elements of TiO design to note, the plugins run in separate Virtual Machines (VMs); they are based on a template provided by the platform developers and there is a group or person responsible for the maintenance of the plugin etc.

As stated that there is a template provided for the plugin developed. This is due to the fact that the majority of the plugin developers they are less skilled in web development. The platform so far provide two templates: Simple Static Plugin Template and Simple Dynamic Plugin. Here is the screenshot of one of the template provided in Figure 3.

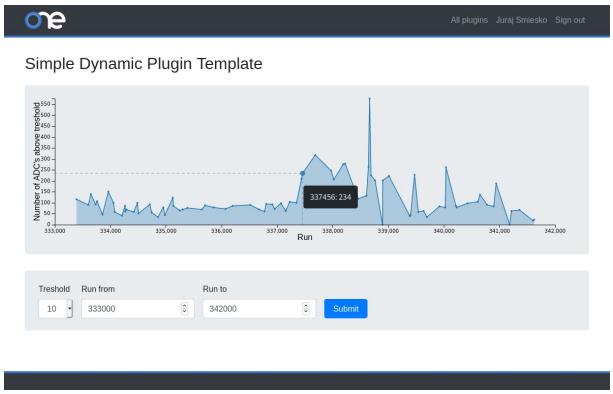


Figure 3. Screenshot of the Simple Dynamic Plugin Template

3. Preliminary Results and Discussion

Preliminary results show the successful of querying the Tile DCS high voltage data collected from Oracle database using DDV tool. HTTP requests python script is created to the data directly from the DDV server for any high voltage channel of the module to be visualised and analysed using plotly library. Figure 4 shows the web interface set of HV plugin on the user side. The plot shown in Figure 4 is of the different drawer and channel number, it is important to note that, each drawer number (64 drawers) consist of 47 channels. Therefore, there are at least 10 000 channels will be visualised and analysed in order to observer the behaviour of the HV channel.

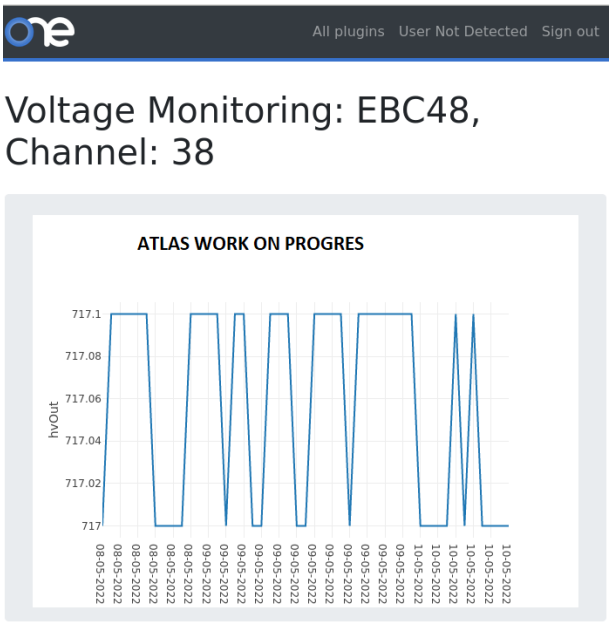


Figure 4. The plot visualised by the user at web interface of the high voltage plugin

4. Conclusion

The web interface (environment) was successfully created for high voltage using necessary templates provided at CERN for web developers. The extraction of data from the DCS DDV server was accomplished and the data was visualised as seen in Figure 4. The displayed data is for 24hrs. Note that the plugin is still on development for proper proposed title.

References

[1] ATLAS Collaboration, JINST 3, The ATLAS Experiment at the CERN Large Hadron Collider, S08003 (2008)
[2] The High Voltage distribution system of the ATLAS Tile Calorimeter and its performance during data taking; arXiv:1804.05036v3 [physics.ins-det] 16 Aug 2018
[3] Juraj Smiesko, "An integrated system for data quality and conditions assessment for the ATLAS Tile Calorimeter," EPJ Web of Conferences 214, pp 2-3, 01030 (2019)
[4] Yuri Smirnov and Juraj Smiesko, "Tile-in-One: an integrated system for data quality and condition assessment for the ATLAS Tile Calorimeter," EPJ Web of Conferences 245, page 2-5, 01010 (2020)
[5] A. Sivolella, F. Ferreira, C. Maidantchik1, C. Solans, A. Solodkov, B. Burghgrave and Y. Smirnov, Tile-in-ONE: A web platform which integrates Tile Calorimeter data quality and calibration assessment.

Comparative analysis of performance optimally designed on-grid and off-grid hybrid power systems for a Limpopo, South Africa Community Development Centre energy system

S Muronga, D Tinarwo, TS Mulaudzi

Department of Physics University of Venda: University Rd, Thohoyandou,0950

E-mail:**david.tinarwo@univen.ac.za**

Abstract. The current global energy crises threaten achieving sustainable development goals (SDGs). The recent closure of numerous small enterprises in South Africa proves that the country's economic progress is being hampered by the country's increasing number of power outages (load shedding). Several residential buildings and businesses have converted their roof space to photovoltaics renewable energy producers, part of an avalanche in implementing various renewable energy technologies. A Photovoltaic (PV)/battery hybrid power system operated at Masia Development Centre was modelled with and without being connected to the grid using the Hybrid Optimization Model for Electric Renewable (HOMER)software. The comparative performance analysis of the on-grid and off-grid configurations of the hybrid power systems was carried out using HOMER software. Based on the results, it can be concluded that the PV/Grid/Converter system configuration is more reliable, sustainable, and economically feasible with the lowest net present cost (NPC) and levelized cost of energy (LCOE) of R992,237.00 and R1.36/kWh, respectively, in comparison to other comparable configurations. The investigation has shown that while reducing reliance on the national grid, using a battery bank may not considerably lower the net present cost of a hybrid power system. It was established that, as compared to standalone PV/Batt/Converter and PV/Grid/Batt/Converter hybrid power systems, the adoption of a hybrid PV/Grid/Converter system without battery storage may achieve much lower NPC and LCOE.

1. Introduction

The current global energy crisis is threatening sustainable development in any developing nation. Thus, governments in developing nations must achieve high energy security to sustain their development trajectory. In South Africa, the continuous increase in power outages (load shedding) hinders the country's economic growth, as evidenced by the recently witnessed closure of several small businesses across the country [1, 2]. Due to the increase in energy demand and load-shedding-related power outages, the South African government needs to find alternative energy resources. South Africa must use available renewable energy potential to cover the energy shortfall and sustain its economic growth rate. Regardless of the fluctuating nature of renewable energy systems (RES), especially solar photovoltaic (PV) and wind, South Africa is in the best position to harness the available renewable energy, as evidenced by an avalanche in implementing different renewable energy technologies, with several residential buildings and firms turning their roof space to photovoltaics renewable energy generators [3, 4]. According to Statista global data and business intelligence platform and the South Africa Solar Photovoltaic (PV) Market Report from the global data, South Africa leads the continent

with a cumulative PV installed capacity of 7021 MW with 4412 MW as industry and residential rooftops as of June 2023, [5]

Solar energy, which is clean and limitless, can offer a sustainable and long-term solution to the current global energy crises [6, 7]. PV generators can immediately convert solar energy into helpful power [6]; therefore, as many studies/researchers have proposed, solar energy is frequently included in hybrid systems. A hybrid system that is either grid-tied or off-grid can combine fuel cells, diesel generators, wind turbines and solar energy [8, 3]. Such hybrid systems can substantially support the utility grid by injecting power to or taking part of the load from the grid, mitigating greenhouse gas emissions from conventional power generation. Concerning sustainability, efficiency, and economic feasibility, solar photovoltaic (PV) is one of the most promising alternative energy sources, which is being touted as a leading solution to long-term electrification and development problems worldwide [9, 10].

However, local municipalities have different by-laws for connecting distributed renewable energy generators to the local grid, making it challenging to implement these PV systems [11, 12]. Additionally, PV systems configurations require an understanding of optimal configuration and proper scheduling to match the load demand for the facility or building. Thus, off-grid, grid-tied, and grid-interactive hybrid PV energy systems design, scheduling, and performance analysis are gaining increasing attention from academia and industry to enhance their integration as an alternative energy source. The overall performance analysis is the right approach to evaluate the solar photovoltaic power potential. It benefits from assessing upcoming solar photovoltaic plants' design, maintenance, and economic performance [13, 14, 12].

Few authors have conducted comparative study of on-grid and off-grid hybrid power systems. The earlier studies of different solar energy systems show that the grid-tied and the hybrid energy systems perform better in terms of energy reliability and investment since more than one energy source contributes to the system's energy [15]. Tomer and Tiwari (2017) conducted a techno-economic feasibility study of a grid-connected PV hybrid system. The results proved that the proposed hybrid approach is economically viable to meet the household’s demand [16].

This study aims to provide a comprehensive understanding of the operation of grid-connected photovoltaic and hybrid power systems design using a recently installed hybrid power system at Masia Development Centre, Limpopo, as a case study. The focus of the study is not to claim or confirm the performance of an installed PV system but to show the benefits of off-grid and grid-interactive hybrid renewable energy systems (HRES) in the provision of reliable power with no carbon emission or other contaminants in a residential house or on a small-scale power supply.

2. Methodology

This study used HOMER software to model and optimise the grid-tied and off-grid hybrid power system architecture for electricity supply in the Masia Community Development Centre in Limpopo, South Africa. The facility is currently not getting electricity from the national grid.

2.1. Description of study area

The Masia Community Development Centre, located in Masia Tshikwarani village in Collins Chabane local municipality in Limpopo Province, South Africa, with geographical position system (GPS) coordinates 23°11'03.5"S 30°18'37.6"E, was chosen as the research site for this investigation. Currently, a 20-kW solar system serves as the only source of power for selected electrical needs in one of the buildings at the community development centre. The facility has all the utility grid system infrastructure but is not operating. The facility is already wired for grid supply with a separate wiring for the PV system. However, the simulation is done assuming the two power systems use the same wiring infrastructure.

2.2. System architecture

The hybrid components (PV, battery bank, inverter) of HOMER models were used with and without connecting to the grid. This study has guided the system configuration design and components sizing by economical reliability factors. The components and solar resources data were fed into the HOMER software and simulated to obtain the optimal configuration with the lowest levelized cost of energy (LCOE) net present cost (NPC) and high reliability. The national grid and PV modules are the energy generation components, hence the utmost substantial solid requirements when selecting an optimally designed hybrid power system. The architecture of the proposed hybrid system is shown in Figure 1.

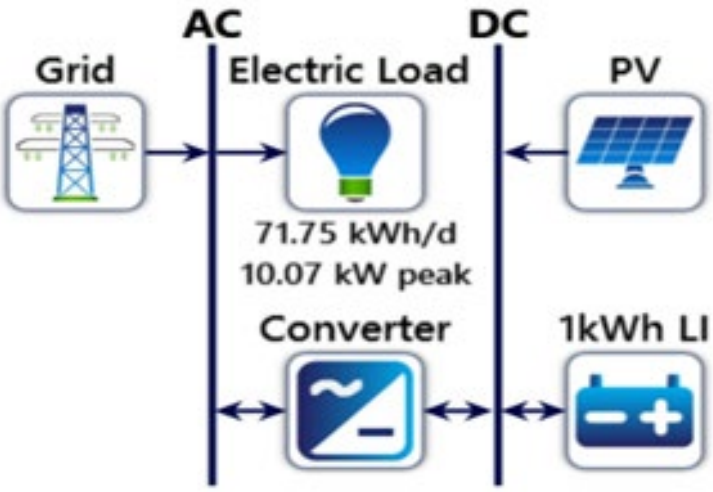


Figure 1. Proposed hybrid system architecture.

2.3 Renewable source assessment

The monthly solar irradiation data of the designated site were acquired from the National Renewable Energy Laboratory (NREL) [8] using the geographical coordinates of the location, Masia Community Development Centre, into HOMER software. The research site is one of the regions receiving the highest solar irradiation in the country, demonstrating the excellent potential for producing electric energy from the installed PV system. The site received maximum sun irradiance $6.34\text{Wh}/(\text{m}^2\text{day})$ in January and the lowest solar irradiance was obtained in June at $395\ \text{Wh}/(\text{m}^2\text{day})$ th the scaled annual average solar irradiaion of $5.26\ \text{Wh}/(\text{m}^2\text{day})$.The research area's clearness index and sun irradiation are shown in Figure 2.

The NASA surface meteorology was used to get the crucial temperature information. This information is essential for the HOMER software process and aids in comparing the effectiveness of various microgrid power systems. The temperature information calculates the solar PV derating factor and evaluates the system's performance under challenging weather or climate conditions. The area's scaled annual average temperature is 21.10 °C. The lowest daily temperature is 15.58 °C in July, and the maximum is 24.52 °C in January, as shown in Figure 3.

2.4 Load assessment

In these studies, the assessment of energy demand is one of the essential parameters to consider in the performance analysis and was accomplished through an inspection survey and an analysis of energy flows for energy conservation in a facility [15]. The used loads or appliances were selected based on

necessity, and all air conditioning units were omitted. In this study, the appliances used, their power rating, and operating hours data of used loads in this facility were collected and entered in HOMER software for the daily, seasonal, and yearly load profile generation, as shown in Figure 4.

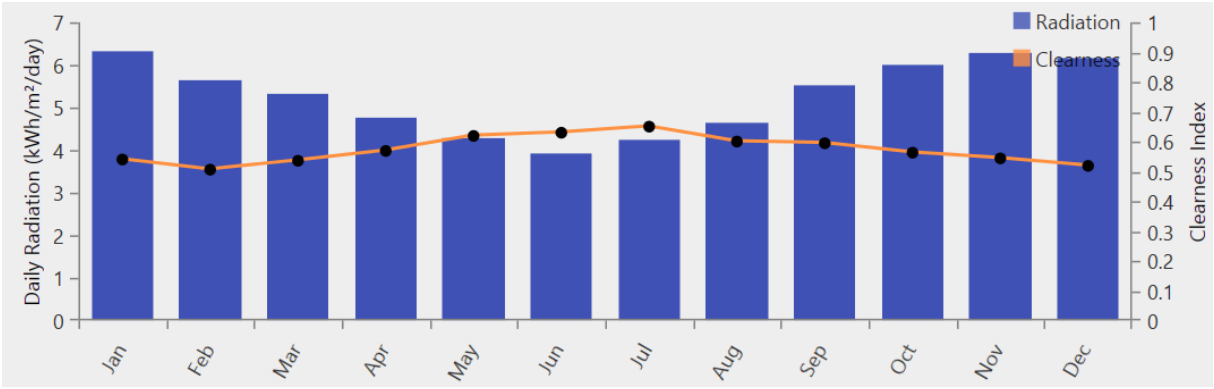


Figure 2. Irradiation and clearness index.

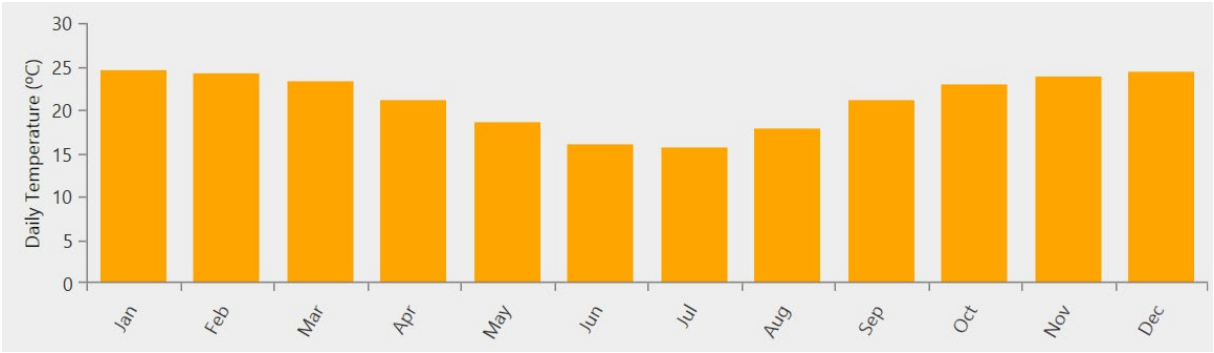


Figure 3. Monthly average temperature.

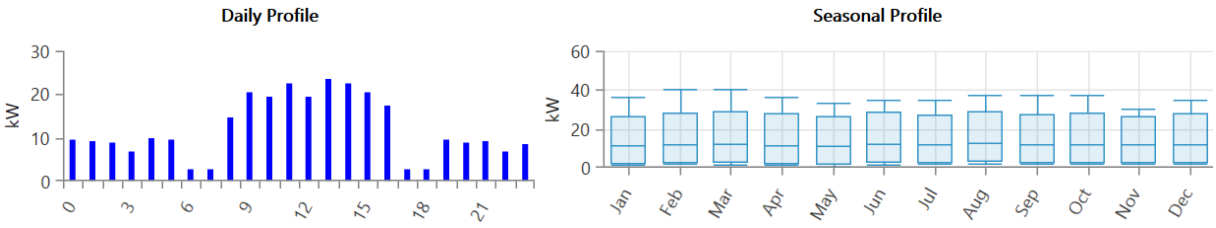


Figure 4. Daily and seasonal load profile.

2.5 System components

For the objectives of simulating and analysing an off-grid and grid-tied hybrid system, this article considers and uses the PV, national grid, battery, and system converter as essential system components. HOMER software was used to model and simulate the performance of the off-grid and grid-tied hybrid

systems. Solar PV and batteries are connected on the DC side, while the grid and the loads are connected to the AC side. AC to DC or DC to AC conversions are made using the 3-phase inverters system. In the grid-tied configuration, when the solar energy is insufficient, the facility will use the utility grid to supply the required amount of electricity, and when PV generates more than needed, that is, after fully meeting the load, the PV array's extra energy recharges the energy storage devices first, [3] and then feed-in to the power grid. The facility's electrical needs should be able to be fulfilled by the architecture of the ideal microgrid system. However, in the off-grid configuration, when more energy is produced from the PV generator, some inverter shutdown or reduction of power output mechanisms are employed where the inverter adjusts its operating parameters to match the energy demand or system conditions, for example, power curtailment or just load dumping.

Below is a description of the system components model utilised in this investigation.

(a) PV generator model: The power generated by the solar photovoltaic (PV) panels is given by [3, 17]

$$P_{PV} = P_{R-PV}\left[\frac{G}{G_{ref}}\right]\left[1 + K_T\left(T_c - T_{ref}\right)\right], \quad (1)$$

where: P_{PV} is the power generated by the solar PV) panels, P_{R-PV} is the rated power at reference condition, G is the solar radiation in W/m^2, G_{ref} is the solar radiation at a reference condition. It is given by $G_{ref} = 1000\,W/m^2$, T_{ref} is the total cell temperature t referenc coditions ($T_{ref} = 25℃$), K_T is temperature coefficient of the PVpanels, T_c the cell temperature.

(b) Battery model: The battery capacity, C(t) at a point in time t, is calculated by HOMER with the equation [18]:$C(t) = C(t-1) - \eta_{batt}\left(\frac{P_B(t)}{V_{Bus}}\right)\Delta t$ (2)

where: $C(t-1)$ is battery capacity at the previous increment, $P_B(t)$ battery input ($E_i(t)$) /output power ($E_g(t)$) given by: $P_B(t) = E_g(t) - E_i(t)$, η_{batt} is the battery efficiency , V_{Bus} is the bus voltage , $\Delta t = t - t_0$ is the time interval at any time *t*.

(c) Grid: The utility (ESKOM) grid is the back-up source of electrical energy in this study. Integrating renewable energy sources (RES) solar photovoltaic (PV) and battery storage into the grid is considered principal to the system. The grid systems supply power when the solar system fails to provide adequate/enough ability to meet the facility's energy demand. The grid absorbs excess electricity when it is generated.

2.6 Financial metrics

The net present cost (NPC) and levelised cost of energy (LCOE). HOMER calculates net present cost (NPC) by the formula [18, 19, 20]:

$$\mathrm{NPC} = \frac{\mathrm{C_{ann.tot}}}{\mathrm{CRF(i\times proj)}} \quad (3)$$

where: $\mathrm{C_{ann.tot}}$ is the total annualized cost of the system, CRF is the capital recovery factor, $\boldsymbol{i}$ is the interest rate in %, and proj is the project lifetime.

The capital recovery factor (CRF) is determined by [18, 19, 20] :

$$\mathrm{CRF(i,N)} = \frac{\mathrm{i(1+i)^N}}{\mathrm{(1+i)^N - 1}} \quad (4)$$

where: i is the real interest rate, and N is the period/number of years.

HOMER uses the following equation to calculate the Levelized cost of energy (LCEO [18, 19]:

$$\mathrm{LCOE} = \frac{\mathrm{C_{ann.tot}}}{\mathrm{E_{prim}+E_{def}+E_{grid.sales}}} \quad (5)$$

where: $C_{ann.tot}$ is the total annualized cost, E_{prim} is the total amount of primary load that the system serves per year, E_{def} is the total amount of deferrable load that the system does per year, and $E_{grid.sales}$ is the energy sold to the Grid per year.

2.7 System reliability

There are usually situations where the proposed energy system would not meet consumer's energy demand requirements. This may be due to insufficiency in energy resources. Such circumstances will lead to a loss of power, resulting in the consumer's load not being served at such a moment. This study uses a loss of energy probability (LOEP)—Equation (6) and availability (Equation (7)) indices to evaluate the reliability of the proposed systems. [21]

$$\mathrm{LOEP} = \frac{\sum_{j=0}^{8760} \mathrm{Unmet_j}}{\sum_{j=0}^{8760} \mathrm{Demand_j}}, \quad (6)$$

$$\mathrm{Availability} = 1 - \mathrm{LOEP} = 1 - \frac{\sum_{j=0}^{8760} \mathrm{Unmet_j}}{\sum_{j=0}^{8760} \mathrm{Demand_j}}. \quad (7)$$

3. Results and discussion

In this section, the simulation results from HOMER software are presented and discussed

3.1. Optimization results and optimal sizing of the proposed hybrid power system

NPC and LCOE were used to determine the hybrid power system's ideal system component size. HOMER ranks the system configurations in their order starting with the one with the lowest NPC and LCOE**.** The decision of the wining system will involve in addition to the two criteria the designer or owners considered benefits best and winning system. The optimal hybrid power system architecture with component sizes and results of the financial analysis are shown in Table 1. The optimum rated capacity of 20 kW for solar PV array power production provides PV penetration of 138% with operating hours of 4,338 hrs/yr and a levelized cost of 0.184 R/kWh. The power converter capacity is 20 kW with 4,388 hrs/yr operating hours, and energy loss of 750kWh/yr. The grid has a sale capacity of approximately 999 999 kWh with energy purchased of 10,540 kWh/yr (from the grid) and energy sold of 19,655 kWh/yr (to the grid). Although the PV/Grid/Converter configuration of the hybrid system shows the most favourable on economic parameters, like lowest net present cost and levelized cost of energy, the reliability criteria is at stake given the current with an NPC of R992 237.00 and LCOE of R1.36, and a minimum operation cost (OC) (R/yr) of R52 183.00, the PV/Grid/Converter (Cases 1) design is considered the best/wining option. The two cases are ranked as best: case 1 and case 2. This optimal hybrid power system architecture, both off-grid and grid-connected, is considered the most affordable and trustworthy.

3.2 Electric energy production

The national grid and solar PV are the two most efficient hybrid power system's monthly average electrical energy output. Based on a monthly analysis, the graph demonstrates how solar PV and national grid systems produce different amounts of electricity depending on the time of year. Solar PV has the most electricity, followed by the national grid. These ideal hybrid power system components continuously provide enough electricity to meet or load the facility's necessary power demand. Case 4

has a more challenging time meeting the load power requirement at night. It is found that the ideal grid-connected hybrid micro-grid power system produces 36 053kWh annually.

Table 1. Optimisation results of the proposed hybrid power system component sizes.

Case	Architec-ture	PV (kW)	Batt (kWh)	Grid (kW)	Conv (kW)	NPC (R)	LCOE (R)	OC (R)	Energy purchased (kWh)	Energy sold (kWh)
1	PV/Grid /Conv	20	0	999999	20	99 237.00	1.36	52 183.00	10540	19655
2	PV/Batt/ Grid/Co nv	20	16	999999	20	1.28M	1.76	65 320.00	10540	19622
3	PV/Batt/ Grid/Co nv	20	34	999999	20	1.60M	2.19	76 293.00	10540	19579
4	PV/Batt/ Conv	20	98	0	20	32.8M	83.73	2.01M	0	0

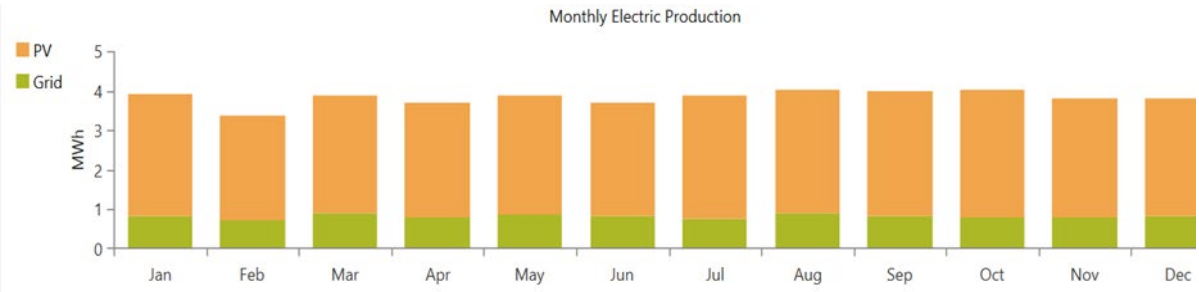


Figure 5. Monthly electric production.

Figure 6 displays the daily energy sales and purchases from the grid throughout the year. According to the findings, more energy was bought from the grid between 18:00 and 6:00 a.m. and less throughout the day. The facility will be supplied with electricity from the grid at night. With a peak around noon, more energy is sold back to the grid during the day. The findings indicate that more power was sold back to the grid than was consumed, which will assist with future energy credits.

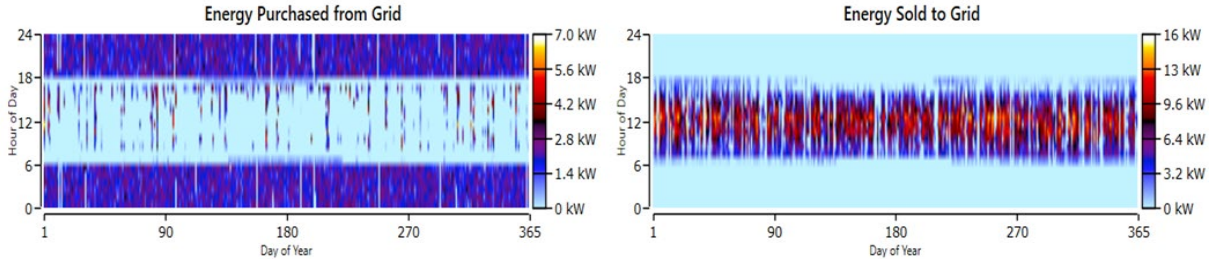


Figure 6. Energy purchased from the grid and energy sold back to the grid.

5. Conclusion and future studies

The simulations of the power systems configurations showed that a PV/Grid/Converter hybrid system without batteries may considerably lessen reliance on the national grid (20 kW PV array, 999 999 kW

grid). Based on the results, it can be concluded that the PV/Grid/Converter system configuration is the most economically feasible with the lowest net present cost (NPC) and levelized cost of energy (LCOE), of R992,237.00 and R1.36, respectively, in comparison to other comparable configurations. However, due to the reliability requirement of the electricity supply to the loads and the current load-shedding experienced with the grid, a design with storage would be best. The high PV penetration of 138% shows the oversized generation to the connected load. It was demonstrated that even while using a battery bank lessens the hybrid power system's reliance on the national grid, the net per cent cost may not be significantly reduced. It was discovered that, when compared to standalone PV/Batt/Converter and PV/Grid/Batt/Converter hybrid power systems, the adoption of a hybrid PV/Grid/Converter system without battery storage might yield much lower NPC and LCOE during a 25-year projection period. To sum up, the hybrid PV/Batt/Converter system may be helpful in remote locations and the present power crisis experience in the country.

In future work, a study of the entire centre loads and generators needs to be simulated, and results to inform the benefits and operation of the typical example of microgrid set up at the site. A hydrogen fuel cell is already installed, so its performance analysis will be ideal.

References

[1] V. Mlambo, "Living in the Dark: Load Shedding and South Africa's Quest for Inclusive Development," *IAHRW International Journal of Social Sciences Review,* vol. 11, no. 2, pp. 153-160, 2023.

[2] G. Heffner, L. Maurer, A. Sarkar and X. Wang, "Minding the gap: World Bank's assistance to power shortage mitigation in the developing world.," *Energy,* vol. 35, no. 4, pp. 1584-1591, 2020.

[3] B. Adebanji, O. Atoki, T. Fasina, O. Adetan and A. Abe, "Comparative study of off-grid and grid-connected hybrid power system: issues, future prospects, and policy framework," *Indonesian Journal of Electrical Engineering and Computer Science,* vol. 22(2), pp. 752-759, 2021.

[4] S. Asamoah, S. Gyamfi, F. Uba and G. Mensah, "Comparative assessment of a stand-alone and a grid-connected hybrid system for a community water supply system: A case study of Nankese community in the eastern region of Ghana," *Scientific African,* vol. 19, p. e01331, 2022.

[5] GlobalData, "South Africa Solar Photovoltaic (PV) Market Analysis by Size, Installed Capacity, Power Generation, Regulations, Key Players and Forecast to 2035," GlobalData, London, United Kingdom, 2023.

[6] H. Al-Najjar, C. Pfeifer, R. Al-Afif and H. El-Khozondar, "performance Evaluation of a Hybrid Grid-Connected photovoltaic Biogas-Generator Power System," *Energies 2022,,* vol. 15, 2022.

[7] K. Ismaila, A. Sahin, B. Yilbas and A. Al-Sharafi, " Thermo-economic optimization of a hybrid photovoltaic and thermoelectric power generator using overall performance index," *J. Therm. Anal. Calorim. ,* vol. 144, pp. 1815-1825, 2021.

[8] H. Al-Najjar, H. El-Khozondar, C. Pfeifer and R. Al Afif, "Hybrid grid-tie electrification analysis of bio-shared renewable energy systems for domestic application.," *Sustain. Cities Soc.,* vol. 77, 2022.

[9] A. Chand, K. Prasad, K. Mamun, K. Sharma and K. Chand, "Adoption of grid-tie solar system at residential scale," *Clean Technologies,* vol. 1(1), pp. 224-231, 2019.

[10] S. Sen, S. Ganguly, A. Das, J. Sen and S. Dey, "Renewable energy scenario in India: opportunities and challenges," *J Afr. Earth Sci.,* vol. 122, p. 25–31, 2016.

[11] B. Scharfetter and M. Van Dijk, "Legislation governing the implementation of small-scale hydropower projects for rural electrification in South Africa.," *Journal of Energy in Southern Africa,* vol. 28, no. 2, pp. 14-28, 2017.

[12] J. Guo, V. Fast, P. Teri and K. Calvert, "Integrating land-use and renewable energy planning decisions: A technical mapping guide for local government.," *ISPRS International Journal of Geo-Information,* vol. 9, no. 5, p. 324, 2020.

[13] M. Dahmoun, B. Bekkouche, K. Sudhakar, M. Guezgouz, A. Chenafi and A. Chaouch, "Performance evaluation and analysis of grid-tied large scale PV plant in Algeria," *Energy for Sustainable Development,* vol. 61, pp. 181-195, 2021.

[14] M. Adaramola and E. Vågnes, "Preliminary assessment of a small-scale rooftop PV-grid tied in Norwegian climatic conditions," *Energy Conversion and Management,,* vol. 90, pp. 458-465, 2015.

[15] J. Barzola-Monteses and M. Espinoza-Andaluz, "Performance analysis of hybrid solar/H2/battery renewable energy system for residential electrification.," *Energy Procedia,* vol. 158, pp. 9-14, 2019.

[16] V. Tomer and G. Tiwari, "Techno-economic evaluationof grid-connected PV system for households with feed in tariff and time of day tariffs regulation in New Delhi-a sustainable approach," *Renewable Sustainable Energy,* vol. 70, pp. 822-835, 2017.

[17] K. Okedu, R. Uhunmwangho and N. Bassey, "A Comparative Study of on and off Grid Tied Integrated Diesel Solar PV Generation System," *INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGIES,* vol. 1, 2015.

[18] T. Lambert, P. Gilman and P. Lilienthal, "Micropower system modeling with HOMER," *Integration of alternative sources of energy,* vol. 1(1), pp. 379-385, 2006.

[19] R. Dash, L. Behera, B. Mohanty and P. K. Hota, "Cost and sensitivity analysis of a microgrid using HOMER-Pro software in both grid-connected and standalone mode," *ICRIEECE,* pp. . 3444-3449, 2018.

[20] S. Asamoah, S. Gyamfi, F. Uba and G. Mensah, "Comparative assessment of a stand-alone and a grid-connected hybrid system for a community water supply system: A case study of Nankese community in the eastern region of Ghana," *Scientific African,* vol. 19, 2022.

[21] O. Babatunde, J. Munda and Y. Hamam, "Selection of a hybrid renewable energy systems for a low-income household.," *Sustainability, ,* vol. 11, no. 16, p. 4282., 2019.

Low-budget atmospheric monitoring system

Marco Mariola[1], Mazibuko Kwanele[1], and Venkataraman Sivakumar[1,2]

[1]University of KwaZulu-natal, School of Chemistry and Physics, Westville Campus, Private Bag X54001, Durban, 4000, South Africa
[2]National Institute for Theoretical Sciences, UKZN, Durban, South Africa

E-mail: mariolam@ukzn.ac.za

Abstract. Thanks to the use of weather balloons, it is possible, through the instruments on board, to carry out measurements of pressure, humidity, and temperature in a capillary way, from sea level, up to an altitude determined by technological limits, or as a function of the specific experiment. The cost of an experiment carried out with a weather balloon depends on the types of equipment on board and the weight of the payload. In addition to scientific instruments, a geolocation device and a transmitter capable of transmitting measurements in real time are essential. In this work, we want to explore a low-cost modular system in which the instruments can be added, removed, or modified according to the needs of the experimenters. This device can represent a valid aid for research or teaching. In the proposed system, all the device works properly, but the optimization at the software level still need to be improved to run the Global Positioning System and the GSM simultaneously. In this project, the behavior of the humidity sensor was tested.

1. Introduction

One of the first valuable uses of the balloon for science was proven in early 1911-12 by physicist V.F. Hess, which proved the existence of cosmic rays [1]. Over time, weather balloons have found use in the academic field and in civilian sectors such as aviation support. Like any other aerospace system, consideration must be given to the environment in which the system operates. The design of this first payload prototype was made around the International Telecommunication Union (ITU)[2], which give the order of magnitude of pressures and temperatures for the different layers of the atmosphere. The knowledge of the temperatures is essential to understand which electronic component or bus system needs to ensure the system's survival. In first approximation, the temperature from an initial altitude z_B to an altitude z, changes according to the Equation 1.

$$T(z) = T_B + \alpha(z - z_B). \tag{1}$$

The temperature T_B is the temperature at the initial altitude z_B, α is a coefficient, and z is the altitude of the ballon. The Equation 1 is only valid for altitude lower than 11km from the sea level [3]. Since the temperature at the altitude z depends on the initial temperature T_B, the ITU recommendations are used to calculate temperature and pressure for mid-latitudes.

From the International Telecommunication Union (ITU) recommendation, it is possible to estimate temperature and pressure as a function of the altitude, as represented in Figure 1. A mid-latitude was chosen because of the lower temperature than South Africa, and it is possible

to guarantee the survival of the system. The estimations of the temperature and pressures considered also in function of the seasons.

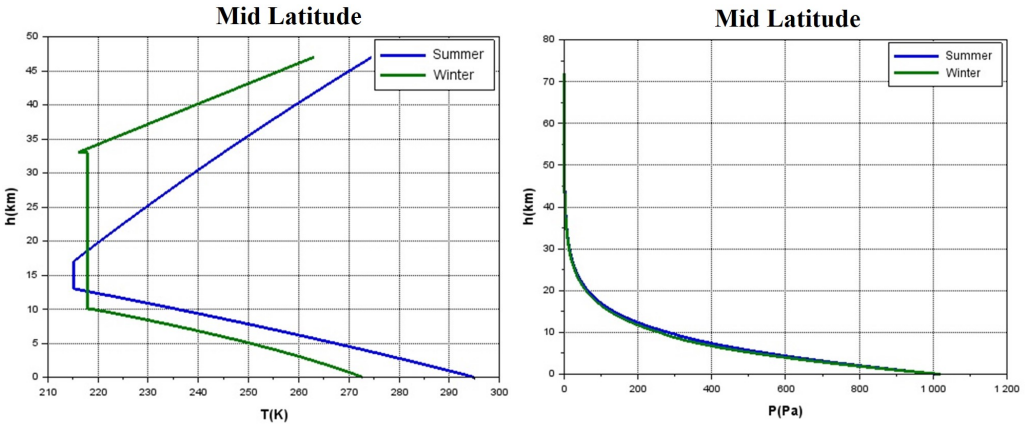


Figure 1. Pressure and temperature as function of the altitude h, for mid latitude between 22° and 45° degrees. The estimation was performed for winter and summer.

2. General scheme of the sensors system

The main goal of the design consists of a low-cost system and allows modifying the configuration according to the researcher needs. The structure of the system is represented in Figure 2. The "Board 1" represents the core of the system. The board includes a Micro Controller Unit (MCU), which controls the experiments and performs calculations. This board contains the USB to serial converter to program the MCU and check the system before launching the balloon. The MCU is an Atmega328P [4] and can be programmed using the Arduino IDE, configured for Arduino Uno[5].

Board 2 includes a Global Positioning System, a power system, and a transmission device. The Board 3 consists of the experiments. The Arduino Environment was chosen for the following reasons:

- Well known in science environment and easy to use.
- Rich of modules available from third parties.
- Microcontroller available in Dual In-line Package (DIP).

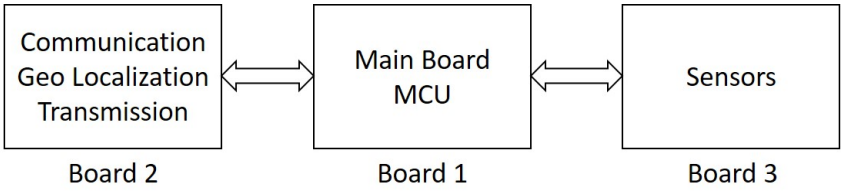


Figure 2. Board 1, is the Micro Controller Unit and is the core of the system.Board 1 is used for data elaboration and controls. Board 2 includes the data transmission, the Geo localisation and the power control.

The packaging of the micro-controller is important if the Printed Circuit Board is produced in house, and can be easily made trough a toner-transfer method or lithography. On the other hand, this family of microcontroller has limited performance in terms of calculation and Random Access Memory (RAM) space. The lack of calculation performances are easily mitigate by coding optimisation but the RAM can be only mitigate by adding an external Serial Peripheral Interface SPI-RAM, or by changing the MCU to an Atmega2560.

3. Board 1

As specified in the previous section, Board 1 allocates the microcontroller unit. The electronic Scheme of the Board is represented in Figure 3. The MCU, Atmega328, is typically configured at 16MHz. A serial to USB converter has been used to program the board and data communication between the computer and the balloon before launch. The Integrated Circuit (IC) used for the serial communication is the CH340G Ic. The vantage of the mentioned Ic is the cost, and it is simple to place in a Printed Circuit Board (PCB).

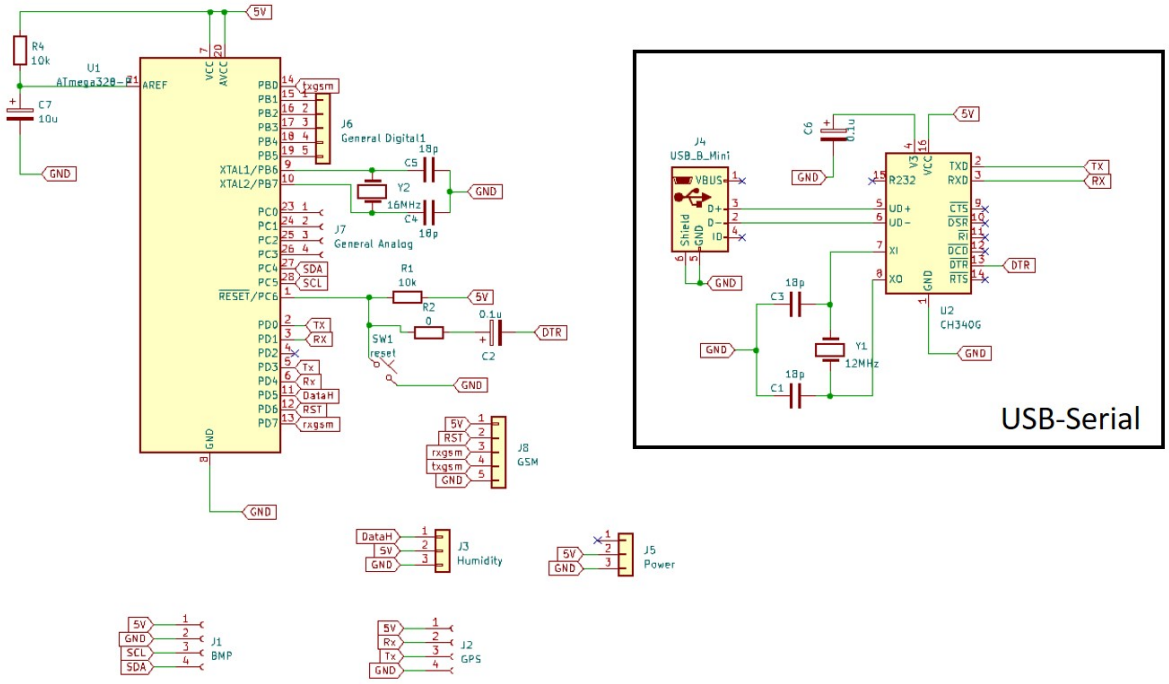


Figure 3. Electronic scheme of the Board 1.

4. Board 2

One of the most critical parts which characterize the weight and the period of the mission is the power subsystem. The second essential part is the device's transmission and geo-localization components. The Board 2 includes these mentioned parts. The schematic of the Board 2 is represented in Figure 4. The two voltage regulators work as two independent power sources. One power source, specifically the voltage regulator 1, is used to power the microcontroller of Board 1 and any other service that runs at 5V. The second power source is dedicated to the transmission system, which includes a Global System for Mobile communication (GSM) Sim800L [6]. At the existing design, all the modules can operate to a voltage between 3.3 and 5V, except the GSM module, which requires a maximum of 4.4V. The system can ideally work

only with 4.4V, but if different instruments need to be placed in Board 3, maybe 5V is necessary. Removing the voltage regular 1 is possible by simply connecting the nodes A1-A2 and B1-B2. The GPS module, GY-GPS6MV2, communicates to the MCU through serial communication as the GSM module. The SoftwareSerial library of Arduino IDE can program the transmission for the GSM and the GPS module.

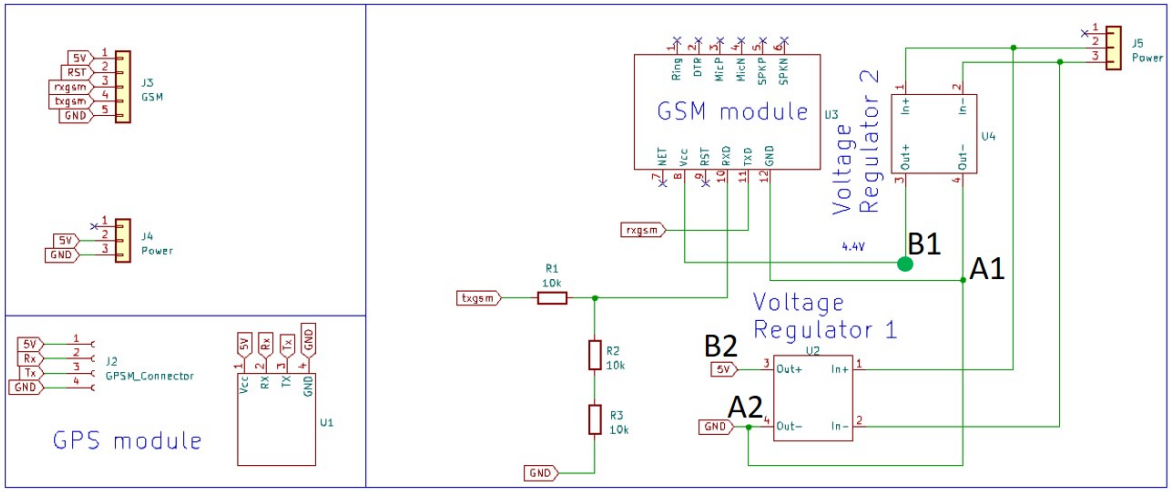


Figure 4. Schematic of the board number 2.

5. Board 3

Board 3 is just a connector board where the atmospheric instruments are placed. The most important feature to note is the pin connected to the MCU. Board 3 is depicted in Figure 5. One humidity sensor, one temperature, and a pressure sensor are connected to the board 3. Whatever sensor or instrument can communicate via the SPI[7] protocol, such as the temperature sensor made by the MAX31865[8] and the probe PT1000. Other SPI systems can be placed on Board 3 (previous modification) by controlling the different chip selectors (CS) using the available digital/analog pins on Board 1. The pressure sensor BMP180, communicates trough the I2C protocols.

6. Experimental testing of the humidity sensor

The pressure sensor BMP180, is a sensor that comes from the SIEMENS, and is also able to measure the temperature. The sensor is calibrated, and for now, only a qualitative analysis is performed on the humidity sensor, for which we do not have many details. Since relative humidity depends on the temperature [9], the qualitative analysis is performed by considering the inverse proportionality between temperature and relative humidity. The measure was conducted at the sea level and is represented in Figure 6. From the plot, the linear relation between the two quantities can be observed and the appropriate calibration will be done using an appropriate reference sensor. The temperature calibration of the sensor PT1000 is not performed yet, and further investigation will be done to characterize the MAX31865 and the thermocouple PT1000.

7. Summary

The presented system represents a modular balloon system that can be reconfigured to the researcher's needs and is easy to program. The proposed system can only be used for low

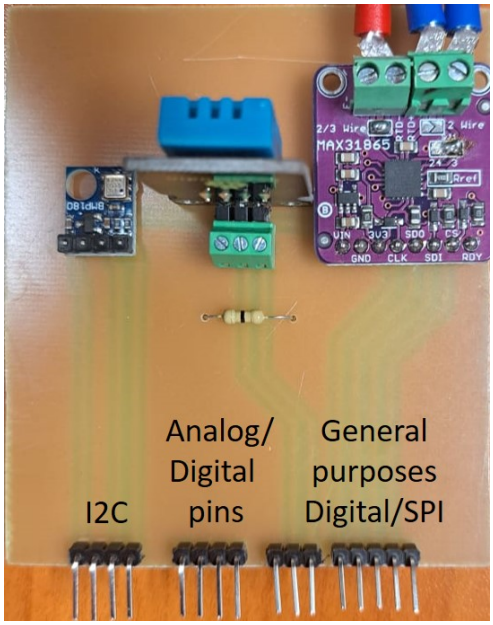


Figure 5. Atmospheric sensors board.

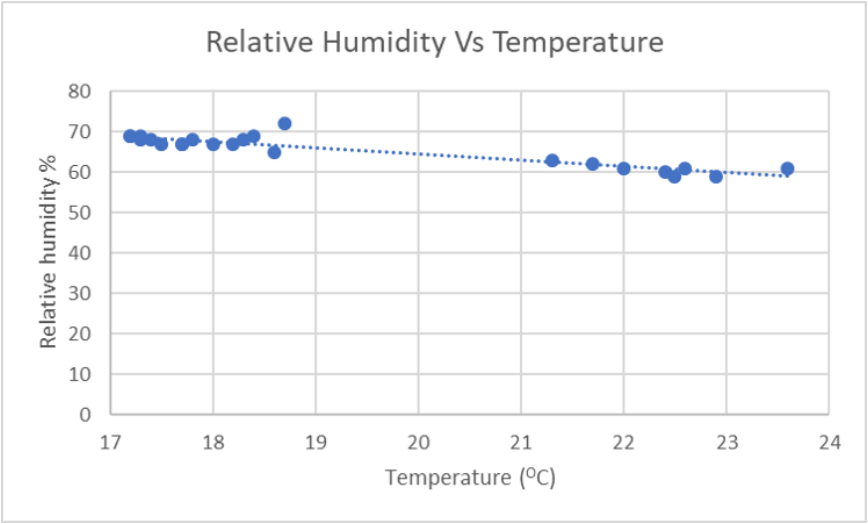


Figure 6. Measured temperature and humidity. The temperatures were collected by the sensor BMP180, while the humidity from the sensor DHT11.

altitudes or as a ground station at this design stage. The reasons are related to the commercial sensor that cannot operate, except the temperature probe, at temperatures lower than zero Celsius. The system needs to be reconfigured with other sensors to be used at high altitudes, and the GSM needs to be substituted with a Radio transmitter. These changes can be operated easily since the structure of the design only needs replacements with other modules.

Acknowlegements
This work is based on research supported by the South African Research Chair Initiative of the Department of Science and Innovation. This work is also supported by NASP.

References

[1] Benton W (ed) 1973 *Balloon* vol 1 (Chicago: William Benton)
[2] International Telecommunication Union 2017 *Recommendation ITU-R P.835-6*
[3] Barboni R 1998 *Costruzioni Aeronautiche* (Rome: Esagrafica)
[4] Atmel 2015 *Atmega328P*
[5] Arduino *Arduino Uno Rev3* URL `https:store.arduino.ccproductsarduino-uno-rev3`
[6] SIMCom 2013 Sim800l hardware design v1.0 URL `https://www.robotics.org.za/SIM800L-V2`
[7] Technology M *SPI Overview and Use of the PICmicro Serial Peripheral Interface* URL `https://ww1.microchip.com/downloads/en/devicedoc/spi.pdf`
[8] Max31865 URL `https://www.analog.com/media/en/technical-documentation/data-sheets/MAX31865.pdf`
[9] Service N W *Relative Humidity* URL `https://w1.weather.gov/glossary/index.php?word=relative+humidity`

Doping of Co and Ni to ZnO (101) surface using DFT Method: For gas sensing application.

K Shandukani[a], S.R Dima[a,b], M.R Letsoalo[a] ,T.S Ranwaha[a], N.E. Maluta[a,c] and R.R Maphanga[b,c]

[a] Department of Physics, University of Venda, P/Bag X5050, Thohoyandou, 0950, South Africa
[b] Next Generation Enterprises and Institutions, Council for Scientific and Industrial Research, P.O. Box 395, Pretoria 0001, South Africa
[c] National Institute for Theoretical and Computational Sciences (NITheCS), Gauteng 2000, South Africa

E-mail: **SDima@csir.co.za**

Abstract. Gas detection involves identifying harmful gases in the surroundings or atmosphere to manage air quality and safety. This is particularly crucial in sectors like food security. The working principle is centered on converting the effects of gas adsorption on active material surfaces into detectable signals, such as altered electrical, optical, thermal, mechanical, magnetic (magnetization and spin), and piezoelectric properties. To overcome the sensitivity challenges of sensing materials, strategic approaches have been developed, including noble metal doping and heterostructures. Doping, a technique involving the introduction of certain elements, has proven more effective in enhancing ZnO's gas sensing capability compared to other methods. It achieves this by inducing changes in the material's microstructure. In this work, Co and Ni were used to dope ZnO(101) surfaces, aiming to understand how the efficiency of doping impacts electronic and structural properties, as well as the gas sensing mechanism. The outcomes demonstrate that all doped ZnO(101) surfaces exhibit enhanced conductivity due to narrowed band gaps and the presence of impurity levels within these gaps. This heightened conductivity is vital for materials used in gas sensing. These research findings hold significance for comprehending the gas sensing mechanism of modified ZnO and offer insights into its potential as a sensing material.

1. Introduction

Having a comprehensive understanding of the gases present in a specific area within our daily lives holds significant importance. Gas detection involves a process that aids in identifying and overseeing various gases in our vicinity, utilizing a device called a gas sensor. These sensors can be categorized into groups like paramagnetic, solid electrolyte, electrochemical, thermal conductive, and metal oxide semiconductor sensors. They operate based on diverse characteristics of the materials employed, including electrical, optical, and gas chromatographic properties [1]. Among these sensors, the semiconductor metal oxide gas sensor stands out due to its exceptional sensitivity, precision, selectivity, response time, stability, durability, portability, and cost-effectiveness, with ratings ranging from excellent, good to fair [1]. The mechanism of these semiconductor metal oxide (SMO) gas sensors encompasses two key functions: the receptor function and the transducer function. The receptor function identifies the target gas, leading to electrical alterations on the metal oxide's surface. Subsequently, the transducer function converts these surface changes into variations in electrical resistance within the sensor. These SMOs are based on general properties of two groups, where the first group are the

materials that exhibit surface conductance effects operating at lower temperature between 400 ° C and 600 ° C and second group are the materials that shows bulk conductance effects operating at temperature greater than 700° C). The first group is made up of materials such as Sn_2O and ZnO (zinc oxide), whereas the second group is made up of materials such as TiO_2, CeO_2, and Nb_2O_5 [2].

Zinc oxide has gained recognition as an exceptional SMO due to its remarkable physical, optical, chemical, and electrical attributes. To effectively utilize ZnO, it is necessary to cleave the bulk structure to expose its most prominent surfaces, since ZnO follows surface conductance effects. The (101) surface of ZnO has been identified as the most exposed and stable surface in the material's bulk structure [3]. However, challenges exist in utilizing zinc oxide-based gas sensors, including high power consumption and inadequate selectivity towards target gases. Addressing these challenges, Patil *et al* [4]. conducted research involving the doping of ZnO with aluminium (Al), resulting in improved performance of ZnO-based gas sensors and rapid detection of nitrogen dioxide (NO_2). There are few numbers of studies on doping ZnO surfaces with different elements such as Ag and Al [4]. Recently, it has been reported that the doping ZnO nanoparticles with Cobalt (Co) leads to distinct changes in their emission characteristics. The addition of Co triggers a red shift in UV near band emissions and introduces an additional green emission peak, with its intensity increasing alongside the Co doping level. This alteration in emission properties supports the use of Co as a dopant in enhancing the hole carrier concentration within the material, a crucial factor in establishing ferromagnetism in ZnO-based systems [5]. There are no literature reports on theoretical studies of Co doped ZnO (101) surface. Nickel (Ni) falls under the same group IV as Co, exhibiting same electronic configuration, hence, doping ZnO (101) surface with Ni is investigated in along with Co doping. This study focuses on the utilization of ZnO semiconductor for gas sensor applications, incorporating a doping mechanism to enhance sensor performance. The investigation also delves into the manipulation of band gap and the doping process. This is achieved by analyzing the structural and electronic characteristics of ZnO (101) surface doped with Co and Ni.

2. Computational method

Density functional theory calculations were conducted using the Cambridge Serial Total Energy Package (CASTEP) code [6], which was implemented in the Material Studio software developed by BIOVIA Inc. For the exchange-correlation functional, the Perdew–Burke–Ernzerhof (PBE) model was employed, following the generalized gradient approximation (GGA) [7,8]. Plane-wave cutoff energy of 500 eV and a k-points grid of 3×4×4 was utilized using the Monkhorst–Pack approach. To generate surface models, these were derived from the optimized bulk structure of the ZnO crystal with a space group F-43m and lattice parameters a = b = c = 4.625 Å. Specifically, ZnO (101) surfaces were constructed based on the optimized ZnO bulk structure, followed by introducing doping, and subsequently, evaluating structural and electronic properties. Throughout the structural optimization process, medium quality convergence criteria were set, with tolerances of 2.0106 eV/atom, 2.0105 eV/atom, 0.05 eV/atom, and 2.0103 eV/atom for the self-consistent field (SCF), energy, maximum force, and maximum displacement. The ZnO (101) surface was made consisting of a 32-atom (2×1×1) supercell. To improve the surface structures of both doped and undoped ZnO (101), surface atoms were relaxed to remove surface atomic tension. A periodic slab was used to represent the surface, with each slab being separated from its neighbours by a vacuum. The optimal vacuum space, which avoids interaction with their periodic image, was found to be 10 Å. However, it was set to 15 Å to provide sufficient space for the adsorbate. The ZnO (101) surface was mono doped by replacing one Zn3C atom with either (Co or Ni) as shown in Figure 1

3. Results and Discussion

3.1. Atomic structures and structural properties

For the cubic ZnO structure, various properties have been analysed, such as the F-43m space group, the bulk lattice parameters have been found to be a=b=c = 4.63, which corresponds to the values reported in the literature [9]. After cleaving the bulk to the ZnO (101) surface, the most exposed and stable surface

in the ZnO crystalline structure [10], the atoms were reoriented, and new surfaces with broken atomic bonds were introduced. This caused surface reconstruction, which changed the atomic positions and bonding close to the surface, affecting the lattice parameters. Additionally, because this substance is a polar crystal with surface defects, the distribution of non-unformal charges may shift, changing the lattice parameters.

A mixture of three- and four-fold coordinated oxygen ions is found in both the top and bottom layers of the Zn (101) surface, which is a step surface. Additionally, the coordination of Zn ions varies between three-fold and fourfold. The lengths of the Zn–O bonds in ZnO (101), Co–O bonds in Co-doped ZnO (101) and Ni–O bonds in Ni-doped ZnO (101) are listed in table 1. Clearly, the Zn3c–O bonds (i.e., A and C bonds marked in Fig. 1 (a)) are longer than the corresponding Sn–O ones in the surface plane. Along the direction perpendicular to the surface, the Co–O (A and B) (1.791 Å), Ni–O (A) (1.782 Å) and Ni–O (B) (1.784 Å) bonds are much shorter than the Zn–O (A) (1.875 Å) and Zn–O (B) (1.876 Å) ones, respectively, indicating that Co and Ni atoms had moved inward and outward from the surface when Zn3c atom is substituted, respectively. This alteration in the crystal structure is due to the distinct sizes and electronic configurations of the substituting Co and Ni atoms. These variations cause a shift in the interatomic distances, that shows the way atoms may respond to maintain stability in the crystal lattice. As such the changes can be an outline of how different atomic properties influence the arrangement of atoms in a crystalline structure

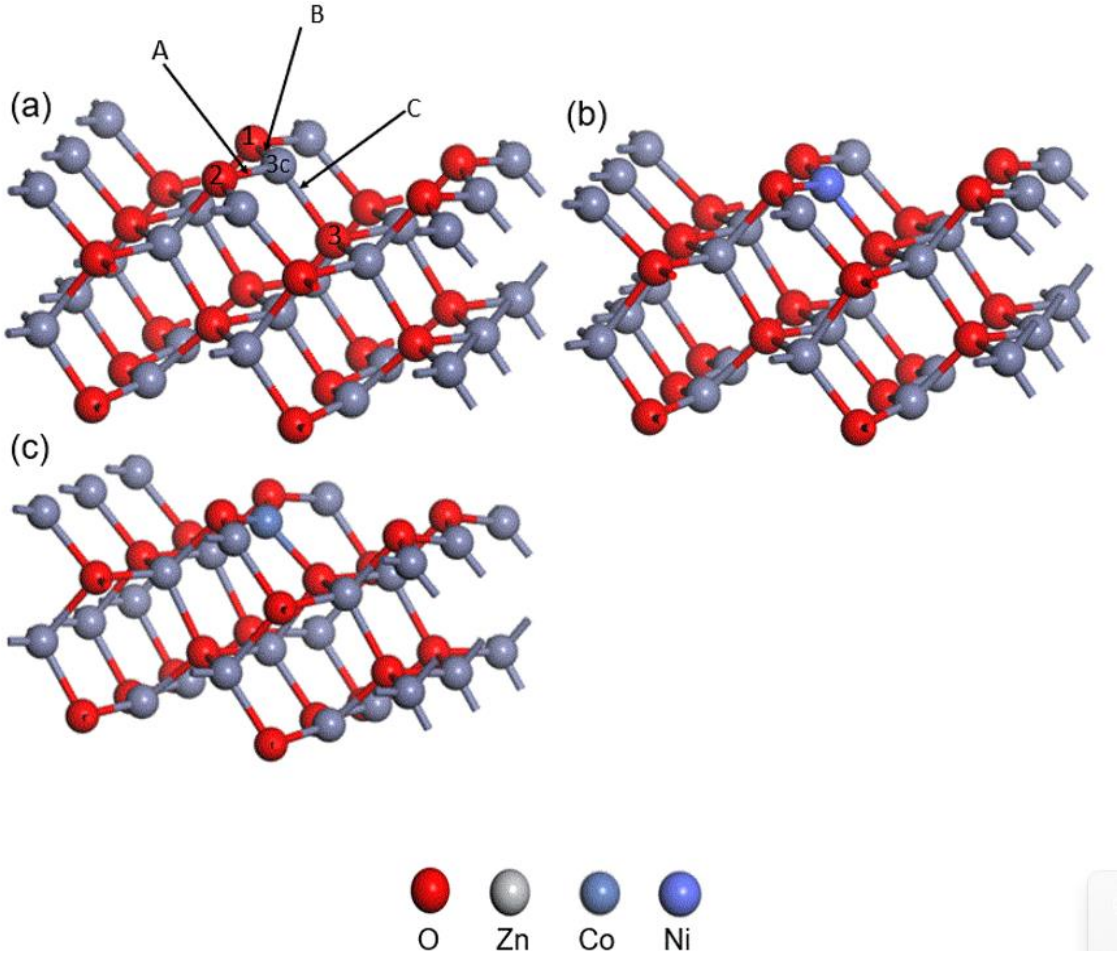


Figure 1. The atomic surface structure for (a) ZnO (101), (b) Ni-ZnO (101) and (c) Co-ZnO (101).

Table 1. Comparison of Zn–O bond lengths in ZnO (101), Ni–O and Co–O bond lengths in both Ni and Co doped ZnO (101) surface.

Bond length (Å)	A	B	C
Zn-O	1.875	1.876	1.970
Co-O	1.791	1.791	1.825
Ni-O	1.782	1.784	1.861

3.2. Atomic structures and structural properties

Fundamentally, comprehending the classification of materials into the three groups of metals, semiconductors, and insulators depends on a system's electronic properties. Figures 3, 4 and 5 show the density of states of the bulk, ZnO (101), doped Co-ZnO (101) and doped Ni-ZnO (101) structures, respectively. The calculated band gap was found to be 1.65 eV, which is lower than the reported experimental literature value of direct band gap of approximately 3.37 eV at room temperature [11], the difference in the value of band gap is due the fact that GGA exchange correlation functional underestimates the band gap of the materials. Although that's the case this can still be overcome using a Hubbard U approximation term that is adopted to accurately describe the electronic structures [12].

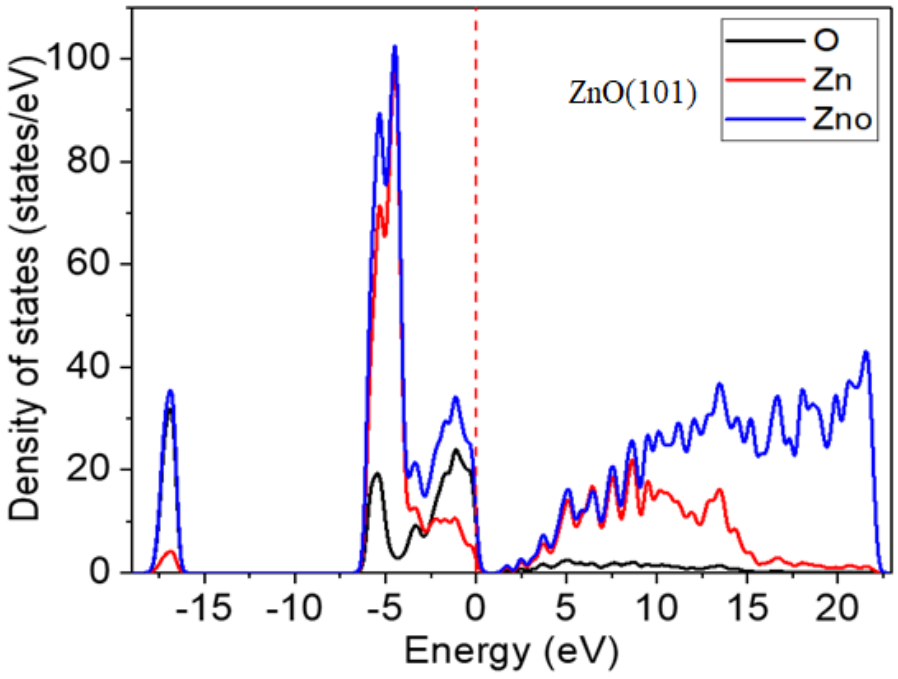


Figure 2. Density of states of undoped ZnO (101) surface.

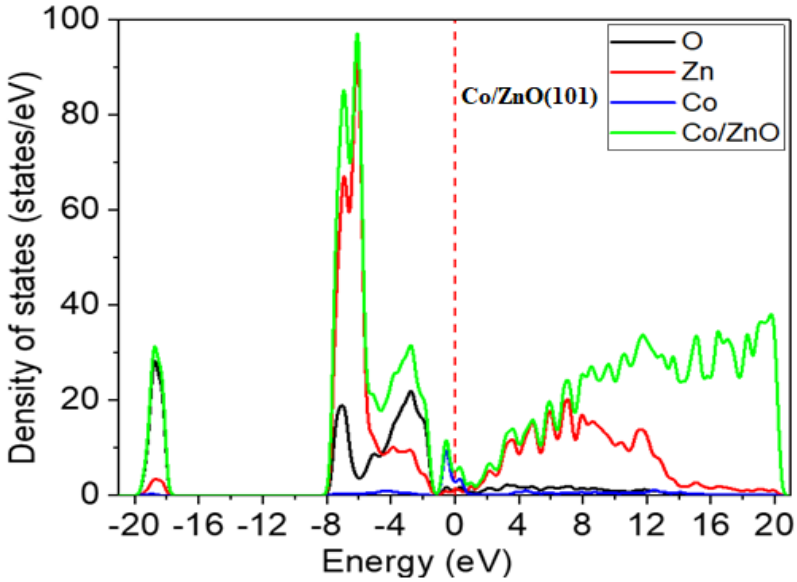


Figure 3. Density of states of Co doped ZnO (101) surface.

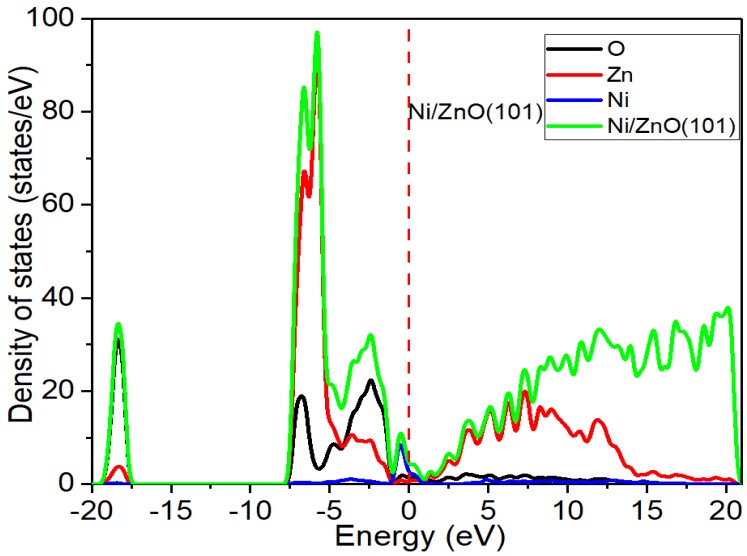


Figure 4. Density of states of Ni doped ZnO (101) surface.

The calculated band gap value for pure ZnO (101) surface was found to be lower than that of the bulk structure. The valence band of pure ZnO (101) surface, as seen from the density of states (DOS) in Figure 2, had two peaks between 0 and 6.5 eV as well as between 15.9eV and 17.0 eV. Doping ZnO (101) with Co and Ni caused modifications in the electronic structure illustrated in Figures 3 and 4, as evidenced by the decrease in the energy band gap from 1.364 eV to 0.074 eV and 0.0 71 eV, respectively. This shows that doping with Co and Ni increases the electrical conductivity of ZnO (101), which is desirable for gas sensing materials. Further investigation into the electronic characteristics of Co and Ni doped ZnO (101) surfaces, specifically pertaining to the Zn, O, Co and Ni atoms can be derived through the characterization of Mulliken population analysis. Table 2 displays the partial atomic charges of Zn, O, Co and Ni atoms. The obtained calculations indicate that the partial charges for O range from around -0.82 to -0.84eV, while for Zn they range from +0.87 to 0.86 eV, and Co and Ni are +0.65 eV and +0.71 eV, respectively. It is noteworthy that all of these charges are below +1. These results suggest that the O atoms in the pure and Co and Ni doped ZnO (101) surfaces have a slightly negative charge, indicating

that they have gained electrons. On the other hand, the zinc atoms show a mixture of positive and negative charges, suggesting a redistribution of electrons between the zinc and oxygen atoms. The Co and Ni atoms, on the other hand, exhibit positive charges, indicating that they have lost electrons. These findings are consistent with the notion that Co and Ni atoms act as acceptors in the ZnO lattice, while the oxygen atoms act as donors. Overall, the low partial charges observed in this study suggest that the Co and Ni doping has not significantly altered the overall charge distribution on the ZnO (101) surfaces.

4. Conclusion

This study was performed using density functional theory calculations to investigate structural and electronic properties of Co and Ni doped ZnO (101) surfaces to understand doping mechanisms and band gap engineering. The partially increasing side of the valence band after doping indicates that the material has a wider valence band, allowing for a greater number of available states for electrons to occupy. This suggests that the addition of Co and Ni element on ZnO (101) surface have enhanced its conductivity and increases its electronic mobility, which is beneficial for gas sensing. Furthermore, low partial charges observed in this study suggest that the Co and Ni doping has not significantly altered the overall charge distribution on the ZnO (101) surfaces.

Acknowledgement

We would like to thank the Centre for High Performance Computing (CHPC) for using their computing facilities. We would also like to acknowledge the Department of Science and Innovation (DSI) Foundational Digital Capability Research and Council for Scientific and Industrial Research for financial support.

References

[1] Korotcenkov, G. (2007). *Materials Science and Engineering B*, vol *139*(1), p 1.
[2] Moseley, P. T. (1992). *Sensors and Actuators B: Chemical*, vol 6(1-3), p 149.
[3] More reference
[4] Dima, R. S., Tshwane D. M., Shingange K., Modiba R., Maluta N. E., and Maphanga R. R. (2022). Processes, vol 10(10), p 2027.
[5] Patil, V. L., Dalavi, D. S., Dhavale, S. B., Tarwal, N. L., Vanalakar, S. A., Kalekar, A. S., and Patil, P. S. (2022). *Sensors and Actuators A: Physical*, vol 340, p 113546.
[6] Gandhi, V., Ganesan, R., Abdulrahman Syedahamed, H. H., and Thaiyan, M. (2014). *The Journal of Physical Chemistry C*, vol 118(18*)*, p 9715.
[7] Segall, M. D., Lindan, P. J., Probert, M. A., Pickard, C. J., Hasnip, P. J., Clark, S. J., and Payne, M. C. (2002). *Journal of Physics: Condensed Matter*, vol 14(11), p 2717.
[8] Perdew, J. P., Burke, K., and Ernzerhof, M. (1996). *Physical Review Letters*, vol 77(18), p 3865.
[9] Onkhorst, H. J., and Pack, J. D. (1976). *Physical Review B*, vol 13(12), p 5188.
[10] Juhnevica, I., Masonkina, M., Mezinskis, G., and Gabrene, A. (2015). *Materials Science and Applied Chemistry*, vol 31, p 33.
[11] Jafarova, V. N., and Orudzhev, G. S. (2021). *Solid State Communications*, vol 325, p 114166.
[12] Jaramillo-Fierro, X., Capa, L. F., Medina, F., and González, S. (2021). *Molecules*, vol 26(13), p 3780.

First-principle study of sodium de-intercalation from di-chalcogenides $NaNbSe_2$

S Patel[1], T.S Ranwaha[1], R.S Dima[1,3], N.E Maluta[1,2] and R.R. Maphanga[2,3]

[1]Department of physics, University of Venda, P/Bag X 5050, Thohoyandou, 0950
[2]The National Institute for Theoretical and Computational Sciences (NITheCS), Gauteng 2000, South Africa
[3]Next Generation Enteprises and Institutions, Council for science and Industrial Research, P.O Box 395, Pretoria, 0001

E-mail: **Sdima@csir.co.za**

Abstract. The importance of energy generation and storage cannot be overstated, given the rising energy demands and the depletion of fossil fuels. Because of their low cost, abundance of elements, strong reversibility, and moderate energy density, sodium-ion batteries have piqued interest as a possible alternative for large-scale electrochemical energy storage. To provide critical fundamental insights into electrode materials and to facilitate the development of materials for sodium ion batteries, computational techniques have been widely used in connection with experimental investigations. In this work, we investigate the electrochemical performance of the transition-metal chalcogenide $NaNbSe_2$ as a material for sodium ion batteries using systematic first-principles calculations based on density functional theory employing the generalized gradient approximation. Comparison of the lattice parameters of the relaxation structure, reported DFT results, and previously experimental results was compared and found to be consistent. It was also found that the system's behaviour showed a slight change during de-intercalation.

1. Introduction

The energy storage system has become an area of interest for many researchers because of the increase in demand for energy as well as the increased cost of fossil fuels and the damage they cause to the environment. Therefore, to establish a sustainable energy system and to overcome energy and environmental crises, renewable energy technologies have been seen as an alternative to solve the problem. Electrochemical energy storage system (ESS) technologies are renewable and environmentally responsible solutions and have shown great promise due to attractive features such as high round trip efficiency, long cycle life, low maintenance, and flexible power and energy characteristics to meet different functions of the grid [1].

Furthermore, rechargeable batteries are a key enabler for clean energy storage systems, and one of the most dominating rechargeable batteries today is lithium-ion batteries (LIBs). LIBs are widely used because of their high-power density and are used as energy storage systems (ESSs) for several portable devices such as cellphones, drones, laptops, etc. [2]. However, the scarcity, uneven distribution, and high price of lithium resources make them difficult to apply in large-scale energy storage systems [3-6]. Thus, a proper alternative for LIBs, especially with abundant natural resources and low cost, is imminently required to meet the development of large-scale applications [7].

Sodium ion batteries (SIBs) have received widespread attention as an attractive alternative to lithium-ion batteries, particularly for large-scale and grid energy storage applications, as sodium sources are more abundant and cost-effective compared to lithium [8].

A sodium ion battery is an energy storage system based on electrochemical charge/discharge reactions that occur between a positive electrode (cathode) composed of sodium-containing layered materials and a negative electrode (anode). The working principle is similar to that of LIBs, only that the larger size of the Na ion compared to the Li ion induces slower ion transport and electrochemical reaction kinetics [9].

One of the most required components of an SIB is the cathode materials. Over the past decade, a lot of scientific publications have been done on sodium-ion batteries to enhance their performance by introducing new electrode materials, yet it is still difficult to stabilise and enhance the electrochemical performance of cathode material to satisfy the requirements of practical application [9-12]. Amongst various cathode materials, high-capacity material such as transition-metal chalcogenide has been explored as SIB anodes, especially the cathode, to enhance the cell density and electrochemical performance of SIBs [9,10]. The characteristics of bulk TMDs exhibit a wide range of features. These qualities encompass insulating behaviour, as observed in HfS_2, semiconducting behaviour, as seen in MoS_2 and WS_2, semi-metallic behaviour, as found in WTe_2 and $TiSe_2$, and metallic behaviour, as observed in NbS_2 and VSe_2 [9]. Several bulk transition-metal dichalcogenides (TMDs), such as $NbSe_2$ and TaS_2, manifest several low-temperature phenomena, including superconductivity, charge density wave (CDW) formation (which involves a periodic distortion of the crystal lattice), and Mott transitions (characterised by a shift from a metallic to a nonmetallic state). In this work, we investigated the electrochemical performance of the transition metal chalcogenide $NaNbSe_2$ and the de-intercalated systems $Na_{0.5}NbSe_2$ and $NbSe_2$ as potential cathode material for SIBs.

2. Computational Method

The density functional theory (DFT) calculations were performed using generalized gradient approximation (GGA) classified by Perdew-Burke Ernzerhof (PBE) to describe the exchange-correlation functional [13]. All calculation were caried out using Cambridge Serial Total Energy Package (CASTEP) simulation codes on Material Studio of BIOVIA. Geometric optimization was done by moving the atoms of a molecule to get the most stable structure with the lowest possible ground-state energy. The valence states of the orbitals Na (2p), Nb (2s), and Se (2p) were studied. The cut-off energy of 520 eV was utilized for the plane set wave basis and a 4×4×2 Monkhorst-Pack mesh was carried out for the integration over the irreducible Brillouin zone. Convergence was obtained when the highest force on all the ions was less than 0.01 eV/Å, which implied that the atomic locations and lattice parameters were totally relaxed.

3. Results and Discussion

3.1. Structural properties

The crystal structure of sodium niobium selenide ($NaNbSe_2$) is a hexagonal crystal structure with a space group of $P6_3/mmc$. The main building block of $NaNbSe_2$ is based on $NbSe_2$, where all elements are stacked on top of each other. The Nb atoms are coordinated to six selenium (Se) atoms in a trigonal prismatic arrangement, and sodium is inserted between the layers of $NbSe_2$ without interrupting their arrangement. The simulation lattice was set to P1 space group, because of the slight distortion of the octahedron.

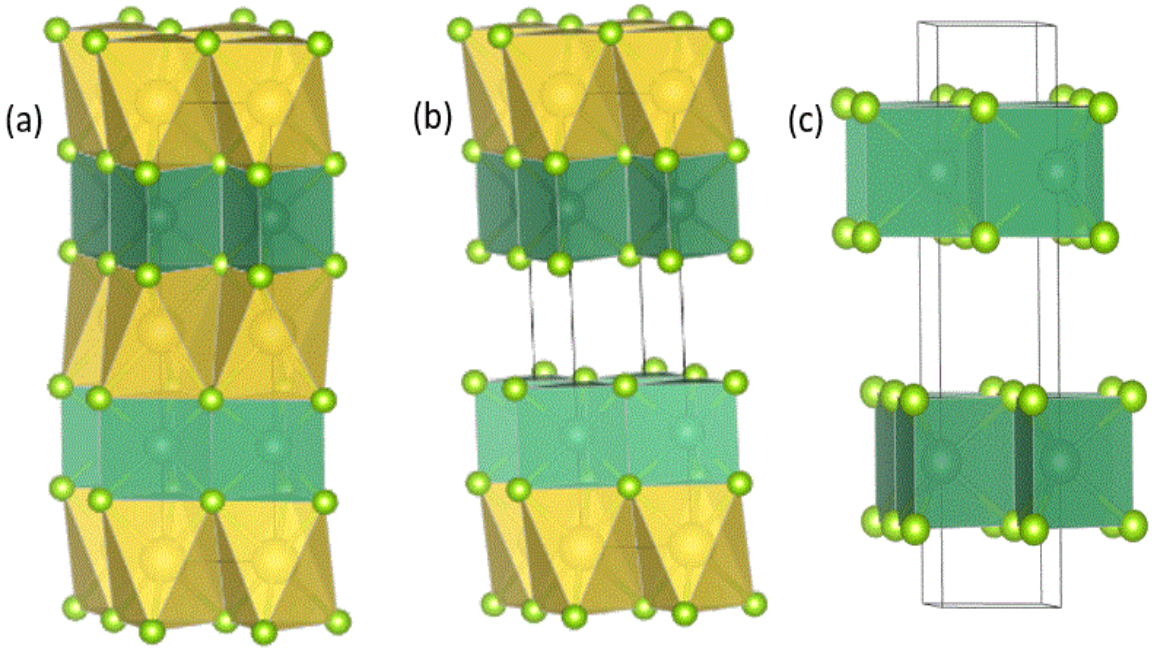


Figure 1. Crystal structure of (a) $NaNbSe_2$, (b) $Na_{0.5}NbSe_2$ and (c) $NbSe_2$.

Table 1 shows the lattice parameter of $NaNbSe_2$, the calculated lattice parameters, volume, and band gap are in good agreement with the literature work and are within the deviation of 1%, which implies that our results are comparable with the previous theoretical study [2] and our calculation methods are reasonable. The lattice parameters of the sodiated and half de-sodiated systems are the same, which shows that there is no re-orientation of atoms during this deintercalation stage. The lattice parameter of sodiated and fully de-sodiated are slightly the same, which indicate a minimal change in the atom orientation at this deintercalation stage. It also shows that the Na diffusion channel is not affected because there is no expansion but slight compression of volume for fully de-intercalated system.

Table 1. Calculated structural properties of $NaNbSe_2$, $Na_{0.5}NbSe_2$ and $NbSe_2$ compared with literature work.

Lattice parameters	Literature [2]	This work		
		$NaNbSe_2$	$Na_{0.5}NbSe_2$	$NbSe_2$
a (Å)	3.570	3.755	3.755	3.491
b (Å)	3.570	3.755	3.755	3.491
c (Å)	14.810	14.810	14.810	14.776
Volume ($Å^3$)	163.880	163.876	163.876	155.713
Band gap (eV)	0.500	0.579	0.511	0.463

3.2. Electronic properties

To gain insights about electrochemical performance, the partial density of state and band structure are plotted along the symmetry path in the first Brillouin zone as shown below in Figures 2,3 and 4.

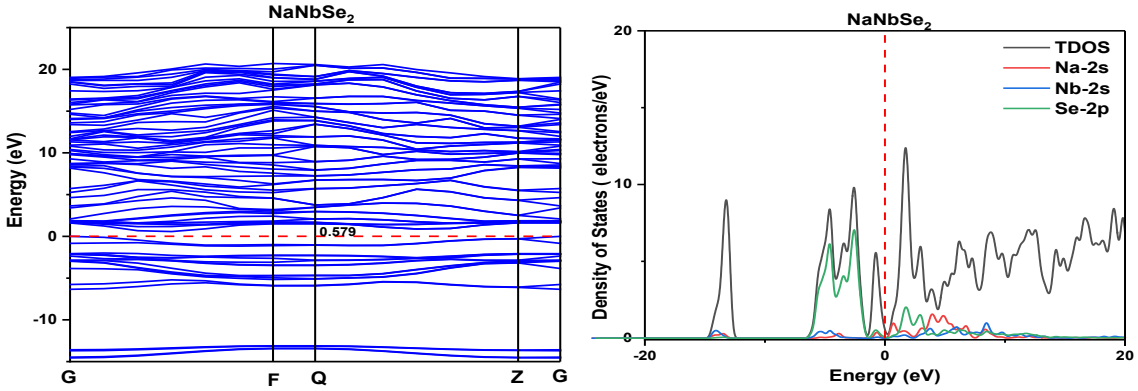


Figure 2: Band structure and PDOS of $NaNbSe_2$.

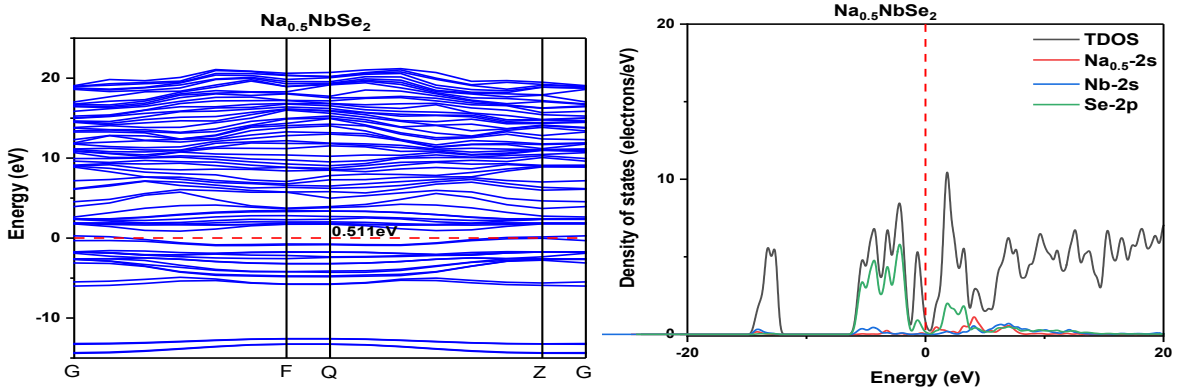


Figure 3. Band structure and PDOS of $Na_{0,5}NbSe_2$.

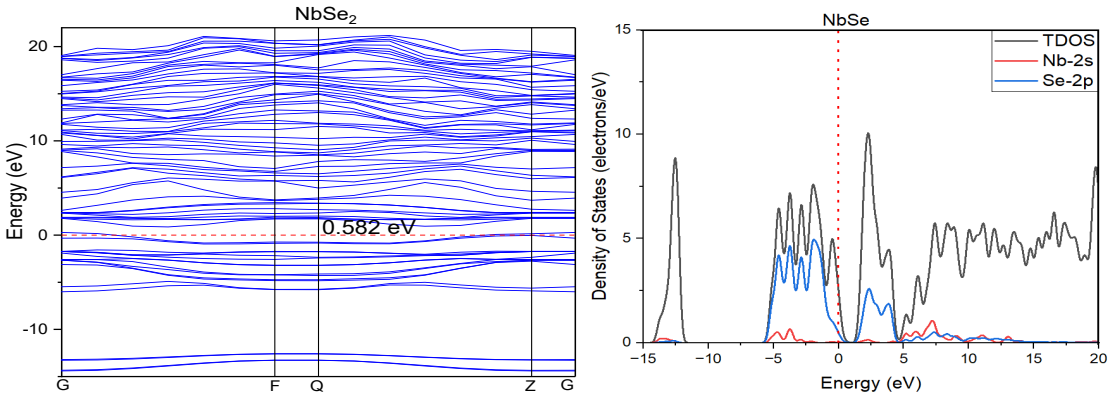


Figure 4. Band structure and PDOS of $NbSe_2$.

The band gap, which refers to the disparity between the highest energy level in the valence band and the lowest energy level in the conduction band, was determined to be 0.579 eV for the fully sodiated system ($NaNbSe_2$), and 0.511 eV for the half-desodiated system ($Na_{0.5}NbSe2$). A decrease in band gap was noticed when half Na+ ions were removed. It is the movement of valence band towards the higher energy level and the movement of conduction band towards the lower energy level that may lead to a reduction in the band gap, the opposite leads to an increase in the band gap. The findings from the DOS indicate that the highest energy level of the valence band does not align with the lowest energy level of the conduction band; this shows that the material has an indirect band gap. The fully detached system ($NbSe_2$) in Figure 4 showed a different trend from the half-desodiated system ($Na_{0.5}NbSe_2$) where the system band gap increased to 0.582 eV. The projected density of states (PDOS) analysis reveals that the valence band is completely occupied, whereas the conduction band exhibits partial occupancy across all systems, hence demonstrating half-metallic characteristics.

3.3. Voltages/Redox Potential and Formation Energy

The energy change of the individual Na atom at every deintercalation stage of the NaNbSe structure is calculated using analytical expression:

$$\Delta E = E\,(Na_{x-y}NbSe_2) + yE\,(Na) - E\,(Na_xNbSe_2) \quad (1)$$

Where: ΔE is the energy change of the process that y (Na) atoms are extracted between the NaNbSe_2 layers, $E\,(Na_{x-y}NbSe_2)$, $E\,(Na)$ and $E\,(Na_xNbSe_2)$ are the energy of $(Na_{x-y}NbSe_2)$, Na metal and Na_xNbSe_2.

Accordingly, the cathode voltage could be calculated by:

$$U = \Delta E / ye \quad (2)$$

Where: U is the cathode voltage, e is the absolute value of the electron charge and y is the number of the extracted Na atom.

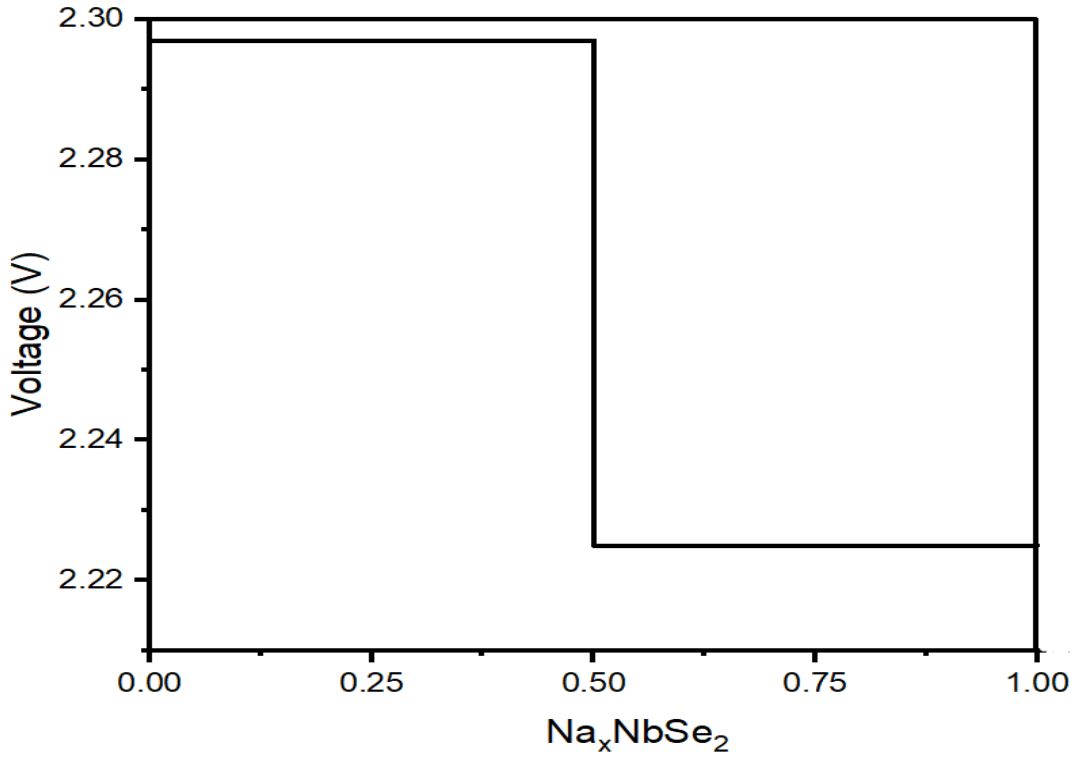


Figure 5. Deintercalation potentials (V) for Na_xNbSe_2 (x = 1, 0, 0.5, 0).

The average voltage/average potential at each deintercalation stage was calculated using equation (2) and plotted in Figure 5. The voltage or potential required to remove the Na ions from Na_xNbSe_2 ranges between 2.225 and 2.374 V, which implies that the voltage window is suitably applied as SIBs cathodes.

4. Conclusion

The structural and electronic properties of $NaNbSe_2$, $Na_{0.5}NbSe_2$ and $NbSe_2$ were examined using first-principle calculations. The lattice parameters and volume are not altered during the 1st deintercalation stage; however, during the 2nd deintercalation stage there was a minimal change in the lattice parameters and volume, which is not enough to cause irreversible distraction to the system. The voltage window of Na_xNbSe_2 was found to be between 2.225 and 2.374 V. A reduction in the band gap was noticed in the 1st deintercalation, which lead to the increase in the metallicity of the system. However, the band gap showed a slight change during the 2nd deintercalation stage, which is the fully desodiated system.

Acknowledgements

The authors would like to acknowledge the Centre for High-Performance Computing (CHPC) for the computing resources. We would also like to acknowledge the Department of Science and Innovation

(DSI) Foundational Digital Capability Research and Council for Scientific and Industrial Research for financial support.

References

[1] Kundu, D., Talaie, E., Duffort, V. and Nazar, L.F., 2015. The Emerging Chemistry of Sodium Ion Batteries for Electrochemical Energy Storage. *Angewandte Chemie International Edition*, 54(11), pp.3431-3448

[2] Palchoudhury, S., Ramasamy, K., Han, J., Chen, P., & Gupta, A. (2023). Transition metal chalcogenides for next-generation energy storage. *Nanoscale Advances*, *5*(10), 2724-2742.

[3] Konarov, A., Kim, H. J., Jo, J. H., Voronina, N., Lee, Y., Bakenov, Z., ... & Myung, S. T. (2020). High-voltage oxygen-redox-based cathode for rechargeable sodium-ion batteries. *Advanced Energy Materials, 10(24),* 2001111.

[4] Liu, K., Tan, S., Moon, J., Jafta, C. J., Li, C., Kobayashi, T., ... & Dai, S. (2020). Insights into the enhanced cycle and rate performances of the F-substituted P2-type oxide cathodes for sodium-ion batteries. *Advanced Energy Materials*, *10*(19), 2000135.

[5] Sommerville, R., Shaw-Stewart, J., Goodship, V., Rowson, N., & Kendrick, E. (2020). A review of physical processes used in the safe recycling of lithium ion batteries. *Sustainable Materials and Technologies*, *25*, e00197.

[6] Pellow, M. A., Ambrose, H., Mulvaney, D., Betita, R., & Shaw, S. (2020). Research gaps in environmental life cycle assessments of lithium ion batteries for grid-scale stationary energy storage systems: End-of-life options and other issues. *Sustainable Materials and Technologies*, *23*, e00120.

[7] Xiao, J., Li, X., Tang, K., Wang, D., Long, M., Gao, H., ... & Wang, G. (2021). Recent progress of emerging cathode materials for sodium ion batteries. *Materials Chemistry Frontiers*, *5*(10), 3735-3764.

[8] Xu, B., Ma, X., Tian, J., Zhao, F., Liu, Y., Wang, B., ... & Xia, Y. (2019). *Ionics*,*25*, 4171-4177.

[9] Dima, R. S., Maleka, P. M., Maluta, N. E., & Maphanga, R. R. (2022). Structural, electronic, mechanical, and thermodynamic properties of na deintercalation from olivine NaMnPO4: first-principles study. *Materials*, *15*(15), 5280.

[10] Moring, J., & Kostiner, E. (1986). The crystal structure of NaMnPO4. *Journal of Solid State Chemistry*, *61*(3), 379-383.

[11] Dima, R. S., Maleka, P. M., Maluta, N. E., & Maphanga, R. R. (2022). Structural, electronic, mechanical, and thermodynamic properties of na deintercalation from olivine NaMnPO4: first-principles study. *Materials*, *15*(15), 5280.

[12] Perdew, J. P., Burke, K., & Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Physical review letters*, *77*(18), 3865.

[13] Zhang, Y. and Yang, W., 1998. Comment on "Generalized gradient approximation madesimple". *Physical Review Letters*, *80*(4), p.890.

Density functional theory study of $Na_xMn_{0.5}Ti_{0.5}O_2$ as a cathode material

T S Ranwaha[1], R S Dima[1,3], N E Maluta[1,2] and R R Maphanga[2,3]

[1]Physics Department, University of Venda, Thohoyandou, 0950, South Africa
[2]The National Institute for Theoretical and Computational Sciences (NITheCS), Gauteng 2000, South Africa
[3]Next Generation Enterprises and Institutions, Council for Scientific and Industrial Research, Pretoria 0001, South Africa

E-mail: **tshifhiwa.ranwaha@univen.ac.za**

Abstract. Due to the accessibility of Na resources and comparable performance to lithium-ion batteries, rechargeable sodium-ion batteries are popular for large-scale electric energy storage and smart grids. Na_xMnO_2 is a promising cathode material for Na-ion rechargeable batteries. However, Na has a lower ionisation potential than Li, resulting in operating voltages and energy densities that are lower than those of LIBs. Na^+ ions are heavier and larger than Li^+ ions, resulting in slow diffusion within a solid electrode during SIB cycling and often larger electrode volume expansion. In this study, the density functional theory was used to investigate the structural and electronic properties of Ti doped $NaMnO_2$. Our findings show that the expansion of volumes is induced by the effect of dopant. The partial density of states underlines the new states near the Fermi level are contributed from the d-orbital of Ti. The magnetism is attributed to the hybridisation of d orbitals of the dopant and the Mn atom with O-p states, namely, p-d exchange hybridisation. The lowest conduction band and highest valence band are mostly contributed from the Mn atom and Ti dopants, which are responsible for the electronic conductivity. The mechanical properties results shows that $Na_xMn_{0.5}Ti_{0.5}O_2$ can be fully deintercalated and still maintain its stability and can bend without deformation.

1. Introduction

The growing interest in the integration of sustainable energy sources into the electricity grid has spurred immense interest in the field of energy storage applications in recent years. The predominant commercial battery system is the lithium-ion battery (LIB), which has become the standard energy source for portable electronic devices. In recent years, driven by the increased desire for 'green' technologies, the use of LIB has expanded from portable electronics to large-scale applications, such as electric vehicles (EVs) [1]. However, there is a concern that the amount of Li resources that are buried

in the earth would not be sufficient to satisfy the increased demands on LIB. Therefore, sodium-ion batteries are the alternative to replace existing lithium-ion technology due to the enormous abundance and low expenditure of sodium batteries [1-4]. Sodium transition metal oxides are suitable candidates for energy storage applications because its wide abundance on Earth crust and low poisoning [3,4]. These make the specific performance of $NaMnO_2$ the cathode material for Na-ion rechargeable batteries. Although SIBs are a promising alternative, they typically have poor electrochemical activity compared to LIBs. This is due to two intrinsic shortcomings associated with Na. First, Na has a lower ionization potential than Li, leading to lower operating voltages and thus lower energy densities in comparison to LIBs [5]. Second, Na +ions are heavier and larger than Li + ions, leading to slow diffusion within a solid electrode during the cycle of SIBs and often larger volume expansion of the electrode, as well as lower energy densities when compared with those of LIBs. These intrinsic deficiencies give rise to the need for unique crystalline structures with an open framework structure or 3D structures with large insertion channels to facilitate the intercalation/deintercalation process of the larger sodium ions [5].

Recently, several fundamental research have focused on understanding the structural and electronic properties with the goal of describing the electrochemical properties of the $NaMnO_2$ compound [3,5,6]. Due to its chemical similarity to Li-ion batteries, strategies to improve the electrochemical performance have been implemented into the manipulation of Na-ion batteries. One of the strategies to advance the physical properties of cathode materials is to dope with transition metals. According to previous literature [6], this approach has been widely applied to standard Li-based cathode materials. In the present study, the monoclinic $NaMnO_2$ compound with Mn atom randomly substituted by Ti are calculated and analysed using spin polarized density functional theory with the aim of understanding structural, electronic and mechanical properties.

2. Methodology

All density functional theory (DFT) calculations were performed using the VASP (Vienna Ab Initio Simulation Package) code. A spin-polarized generalized gradient approximation (GGA) was used to solve the Kohn–Sham equations with the Perdew–Burke–Ernzerhof functional (PBE) exchange–correlation functional, with plane-wave pseudopotential, Geometry optimisation for the ground state energy of $2 \times 1 \times 1$ $NaMnO_2$ was performed. The plane wave cutoff energy of 520 eV and the Monkhorst-Pack scheme was used, with a $2 \times 2 \times 2$ **k**-point mesh. The random substitution doping technique was used to introduce Ti atoms into the $2 \times 1 \times 1$ $NaMnO_2$ supercell, followed by calculating structural properties, mechanical and electronic properties of the bulk structure of $NaMn_{0.5}Ti_{0.5}O_2$, and $2 \times 1 \times 1$ $NaMn_{0.5}Ti_{0.5}O_2$ and $Na_xMn_{0.5}Ti_{0.5}O_2$ supercells ($where\ x = 1, 0.75, 0.5, 0.25, 0$).

3. Results and Discussion

3.1. Structural properties

The calculated structural properties of $NaMnO_2$ are listed in Table 1, The calculated lattice parameters are consistent with experimental values, and are within 0.6% deviation for both experimental [6] and theoretical [7] reported results, which indicates that the calculation methods are reasonable, though the calculated cell volume deviation was found to be 2.59% and 4.09% for experimental and theoretical reported values respectively, this deviation is within 5% which is still accepted for DFT calculations [8].

Table 1. Calculated structural parameters of $NaMnO_2$ compared with literature results.

	Experimental [6]	Literature [7]	This work	Deviation (%)	
				Experimental	Literature
a (Å)	5.63	5.71	5.13	0.50	0.58
b (Å)	2.86	2.88	2.85	0.01	0.02
c (Å)	5.77	5.74	5.67	0.10	0.07

The lattice parameters and volumes are clearly altered by the dopant because of the difference in atomic radius between Mn host and doping element, Ti. The calculated lattice parameters of Ti doped system

increased compared to that of pure system, this results in the expansion of volume in the presence of the dopants. The stretching of lattice parameters and volume expansion is again observed when the supercell is created. The volumetric expansion by dopants can widen Na+-ion diffusion channel and thus improve the electronic conductivity [5,6]. During Na extraction there is a slight contraction of lattice parameters which also results in small contraction or volume change.

Table 2. Lattice parameters of $NaMnO_2$, $NaMn_{0.5}Ti_{0.5}O_2$ and $2 \times 1 \times 1$ $Na_xMn_{0.5}Ti_{0.5}O_2$ supercell during Na extraction after relaxation ($x = 1, 0.75, 0.5, 0.25, 0$).

BULK	a (Å)	b (Å)	c (Å)	volume(Å^3)
$NaMnO_2$	5.158	2.946	6.414	79.735
$NaMn_{0.5}Ti_{0.5}O_2$	5.488	3.046	6.446	90.152
SUPERCELL				
$Na_1Mn_{0.5}Ti_{0.5}O_2$	10.981	3.048	6.443	180.376
$Na_{0.75}Mn_{0.5}Ti_{0.5}O_2$	10.806	2.988	6.325	180.127
$Na_{0.5}Mn_{0.5}Ti_{0.5}O_2$	10.871	2.922	6.358	175.257
$Na_{0.25}Mn_{0.5}Ti_{0.5}O_2$	10.131	3.007	6.603	167.988
$Mn_{0.5}Ti_{0.5}O_2$	10.159	2.944	6.110	152.914

3.2. Electronic properties

The density of states (DOS) calculations were computed using the spin polarised PBE-GGA. The spin-up states are displayed on the positive scale, while the lower spin states are seen on the negative scale of the DOS plot in figure 1.

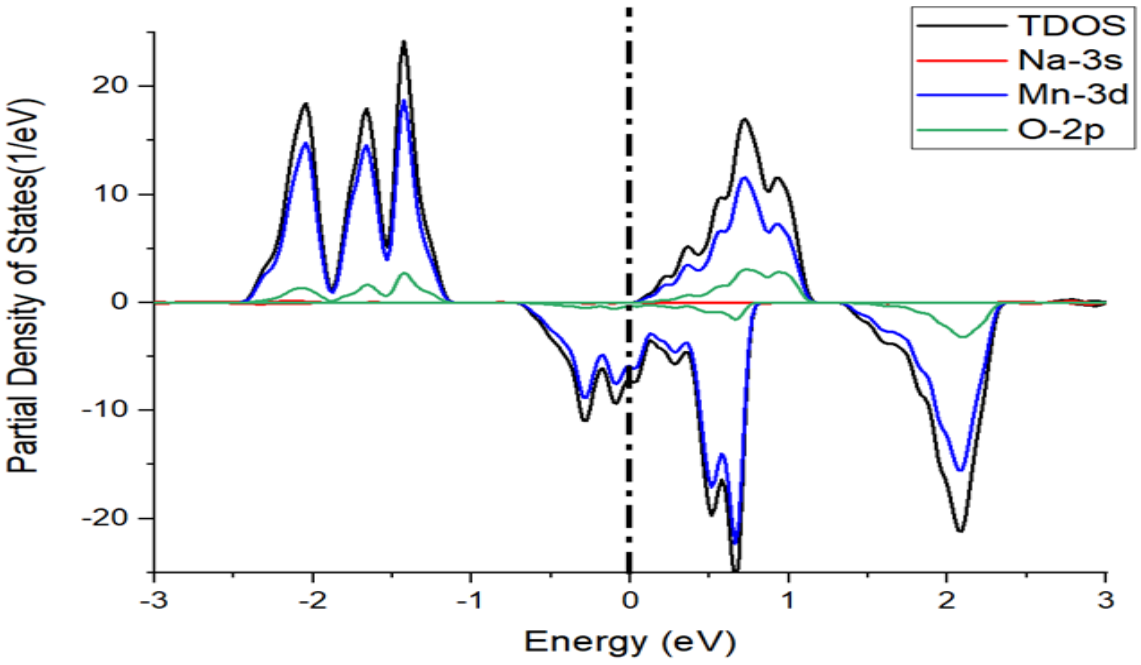


Figure 1. Spin-polarized DOS near Fermi level of $NaMnO_2$.

The total density of states polarised by spin and the partial density of the pure bulk structure is plotted on Figure 1, The total contributions of all atoms in the structure are indicated by the TDOS, whereas the PDOS indicates the contribution of individual atoms.

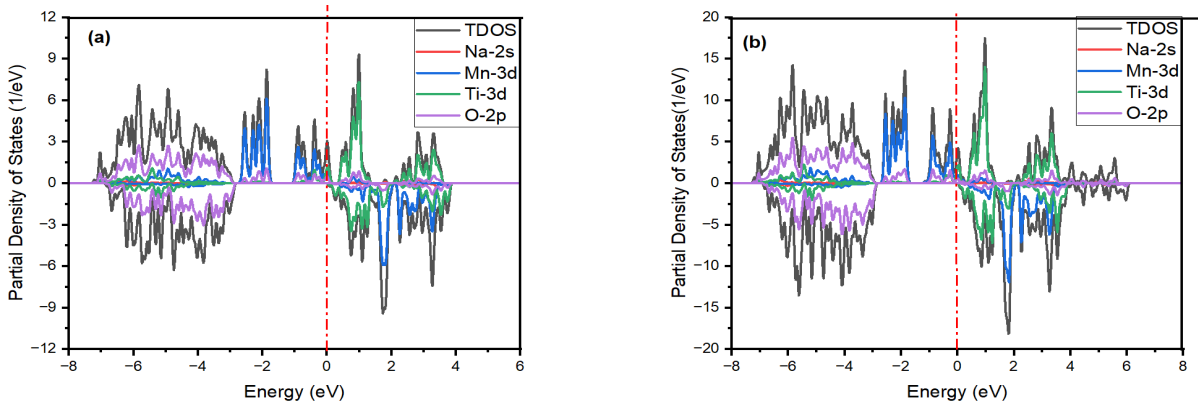


Figure 2. Spin-polarized PDOS of $Na_xMn_{0.5}Ti_{0.5}O_2$ bulk structure and $2 \times 1 \times 1$ $Na_xMn_{0.5}Ti_{0.5}O_2$ supercell.

The total density of states polarised by spin and the partial density of the bulk structure of $Na_xMn_{0.5}Ti_{0.5}O_2$ states and $2 \times 1 \times 1$ Na_x $Mn_{0.5}Ti_{0.5}O_2$ are plotted in Figure 2. For the compounds studied, the spin-polarised density of states is not symmetric, showing magnetism. Doping $NaMnO_2$ with Ti introduces the dopant states at the Fermi level. The partial density of states underlines that these states near the Fermi level are contributed from the d-orbital of Ti. The magnetism is attributed to the hybridisation of d orbitals of the dopant and the Mn atom with O-p states, namely, p-d exchange hybridisation.

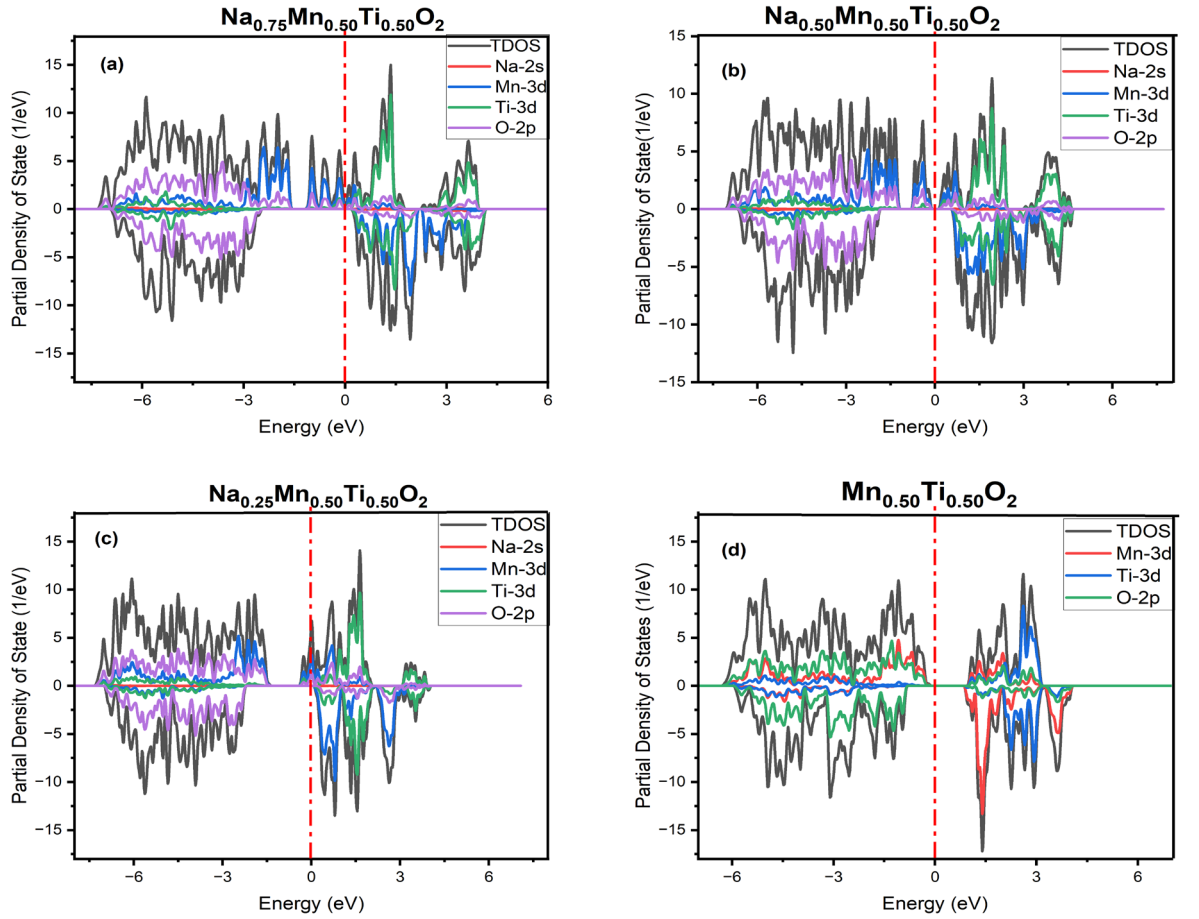


Figure 3. Spin-polarized PDOS of $2 \times 1 \times 1$ $Na_xMn_{0.5}Ti_{0.5}O_2$ supercell (a) $x = 0.75$, (b) $x = 0.5$, (c) $x = 0.25$, (d) $x = 0$.

During Na^+ extraction, the O-2p, Mn-3d, and Ti-3d states slightly cross the Fermi level as illustrated in Figure 3. The up-spin and down-spin states in Figures 3a and 3b are fully occupied, characterising the semiconducting performance, with band gaps of 0.165eV and 0.416eV respectively, while the PDOS depicted in Figure 3c shows the spin-up states that are partially occupied and thus become metallic, whereas the spin-down states are fully filled and present the semiconducting behaviour resulting in the half metallic behaviour for the compound

3.3. Mechanical Properties

$NaMnO_2$ is a hexagonal structure, therefore there are five independent elastic constants, That is, C_{11}, C_{12}, C_{13}, C_{33}, and C_{44}. For the hexagonal crystal, the mechanical stability criterion can be expressed as:

$$C_{11} > 0, C_{44} > 0, C_{11} - C_{22} > 0, \text{and } (C_{11} + C_{12})C_{33} > C_{13}^2.$$

Table 3. Elastic constants (C_{ij}) for $NaMn_{0.5}Ti_{0.5}O_2$ and $2 \times 1 \times 1$ $Na_xMn_{0.5}Ti_{0.5}O_2$ (x = 1, 0.75, 0.5, 0.25, 1).

	C_{11}	C_{12}	C_{13}	C_{22}	C_{23}	C_{33}	C_{44}	C_{55}	C_{66}
Na $Mn_{0.5}Ti_{0.5}O_2$	191.96	80.93	53.20	266.44	59.00	190.85	50.31	40.98	62.95
Na_1 $Mn_{0.5}Ti_{0.5}O_2$	195.40	68.44	50.36	243.84	44.15	183.62	46.16	43.78	62.91
$Na_{0.75}Mn_{0.5}Ti_{0.5}O_2$	160.42	68.25	24.22	301.40	31.43	142.13	33.22	22.70	59.45
$Na_{0.50}Mn_{0.5}Ti_{0.5}O_2$	147.21	59.17	31.01	320.03	18.26	84.84	17.84	18.34	77.74
$Na_{0.25}Mn_{0.5}Ti_{0.5}O_2$	266.99	67.48	21.26	260.63	15.66	54.62	10.73	15.41	76.09
$Mn_{0.5}Ti_{0.5}O_2$	279.77	51.50	6.34	309.47	4.20	11.92	14.58	23.45	115.63

Table 3 demonstrates that the stability condition has been met, indicating that the structures remain mechanically stable throughout and after Na extraction. This phenomenon indicates that $Na_xMn_{0.5}Ti_{0.5}O_2$ (x = 1, 0.75, 0.5, 0.25, 0) can undergo full deintercalation while maintaining stability.

The bulk and Young's moduli determine hardness and stiffness of the material, respectively, while the shear determines the resistance to deformation under shear stress [9]. The bulk modulus also measures resistance to volume change under pressure. Elastic constants, bulk moduli (B) and shear moduli (G) of the hexagonal crystal were calculated using Voigt–Reuss–Hill (VRH) approximation. The expressions of bulk moduli (B_V, B_R, B) and shear moduli ($G_V, G_R G$) are:

$$B_V = \frac{2}{9}\left(C_{11} + C_{12} + 2C_{13} + \frac{C_{33}}{2}\right) \quad (1)$$

$$B_R = \frac{1}{(2S_{11} + S_{33}) + 2(S_{12} + 2S_{13})} \quad (2)$$

$$G_V = \frac{1}{15}(2C_{11} + C_{33} - C_{12} - 2C_{13}) + \frac{1}{5}(2C_{44} + C_{66}) \text{ and} \quad (3)$$

$$G_R = \frac{15}{4(2S_{11} + S_{33}) - (S_{12} + S_{13}) + (S_{44} + S_{66})}, \quad \ldots\ldots\ldots \quad (4)$$

where S_{ij} is the elastic compliance constant.

$$B = \frac{1}{2}(B_V + B_R),\ G = \frac{1}{2}(G_V + G_R). \quad (5)$$

Bulk moduli describe the resistance to volume change and provides the response to hydrostatic pressure. Shear moduli evaluates the resistance to plastic deformation of polycrystalline materials. Young's moduli usually provide a measurement of the stiffness for a material, the larger the value of the shear moduli, the stiffer the material [9,10].

The bulk, shear, and Young's moduli of $Na_xMn_{0.5}Ti_{0.5}O_2$ structures are relatively large as shown in Table 4, implying hardness, great resistance to volume change, deformation, and stiffness, respectively. In addition, the Pugh ductility and brittleness criterion was investigated. The bulk-to-shear modulus (B/G) ratio for polycrystalline phases was calculated with the knowledge that that the shear modulus

represents plastic deformation resistance, and the bulk modulus represents fracture resistance [8]. Brittleness is related with a low B/G value, whereas ductility is related with a high B/G value. The value 1.75 is the critical number that distinguishes ductility from brittleness. Because B_H/G_H is greater than 1.75, the structures of the $Na_1Mn_{0.5}Ti_{0.5}O_2$, $Na_{0.75}Mn_{0.5}Ti_{0.5}O_2$, $Na_{0.50}Mn_{0.5}Ti_{0.5}O_2$ and $Na_{0.25}Mn_{0.5}Ti_{0.5}O_2$ structures are ductile, which implies that these materials can bend without deformation, resulting in fewer to no cracks during battery operation.

Table 4. Bulk, shear and Young's modulus, Poisson's and Pough's ratio for $Na(Mn,Ti)O_2$ and $2 \times 1 \times 1$ $Na_x(Mn,Ti)O_2$ $(x = 1, 0.75, 0.5, 0.25, 0)$.

	B_H	G_H	E_H	Poisson's ratio	Pugh's ratio
$Na\ Mn_{0.5}Ti_{0.5}O_2$	112.71	58.52	149.64	0.2787	1.9260
$Na_1\ Mn_{0.5}Ti_{0.5}O_2$	104.25	58.82	148.50	0.2626	1.7725
$Na_{0.75}Mn_{0.5}Ti_{0.5}O_2$	73.81	41.15	103.92	0.2649	1.7937
$Na_{0.50}Mn_{0.5}Ti_{0.5}O_2$	88.19	48.49	122.87	0.2677	1.8188
$Na_{0.25}Mn_{0.5}Ti_{0.5}O_2$	66.98	37.62	94.96	0.2634	1.7798
$Mn_{0.5}Ti_{0.5}O_2$	46.11	43.00	97.04	0.1443	1.0723

4. Conclusions

We investigated the structural, electronic, and mechanical properties of $Na_xMn_{0.5}Ti_{0.5}O_2$ bulk's structure and $2 \times 1 \times 1$ $Na_xMn_{0.5}Ti_{0.5}O_2$ supercell as a cathode material for SIBs using the first principle method based on the density functional theory. The lattice parameters and volumes are clearly altered by the dopants due to the difference in atomic radius between Mn host and doping element Ti. The expansion of the volumes induced by dopant is observed, and that widens the Na-ion diffusion channel and thus improves the electronic conductivity. Our results show that the spin-up and spin-down states in $Na_xMn_{0.5}Ti_{0.5}O_2$ bulks structure and $2 \times 1 \times 1$ $Na_xMn_{0.5}Ti_{0.5}O_2$ supercell are not symmetric, showing the magnetism. In the case of de-intercalation, it is clearly shown that the compounds are becoming more metallic as the number of Na atoms extracted increases, hence the compound charge carrier density increases, and as a result the electronic conductivity improves. The calculated elastic constants suggested mechanical stability for this compound since the stability criteria were satisfied for all deintercalated systems. All de-intercalated systems were found to be ductile except the fully de-intercalated system, which was found to be brittle according to the Pugh criterion of ductility and brittleness.

Acknowledgements

This work is financially supported through the NRF Thuthuka, Department of Science and Innovation (DSI) Foundational Digital Capability Research and Council for Scientific and Industrial Research for financial support. We would also like to thank the University of Venda for the support to conduct this research and Centre for High performance Computing (CHPC) for using their computing facilities.

References

[1] Tarascon, J.M. and Armand, M., 2001. *nature*, *414*(6861), pp.359-367. Tarascon, J. M., & Armand, M. (2011). *Materials for sustainable energy: a collection of peer-reviewed research and review articles from Nature Publishing Group* (pp. 171-179).

[2] Bucher, N., Hartung, S., Gocheva, I., Cheah, Y. L., Srinivasan, M., & Hoster, H. E. (2013). *Journal of Solid State Electrochemistry*, *17*(7), 1923-1929.

[3] Ma, X., Chen, H., & Ceder, G. (2011). Electrochemical properties of monoclinic $NaMnO_2$. *Journal of The Electrochemical Society*, *158*(12), A1307.

[4] Dima, R. S., Maleka, P. M., Maluta, N. E., & Maphanga, R. R. (2022). First principles study on sodium de-intercalation from $NaMnPO_4$. *Materials Today: Proceedings*, 62, 7-11.

[5] Sukkabot, W. (2020). First-principles calculations of the monoclinic transition-metal doped $NaMnO_2$ cathode material. *Philosophical Magazine*, *100*(7), 917-926.

[6] Zhang, R., Lu, Z., Yang, Y., & Shi, W. (2018). First-principles investigation of the monoclinic $NaMnO_2$ cathode material for rechargeable Na-ion batteries. *Current Applied Physics*, *18*(11), 1431-1435.

[7] Jo, I. H., Ryu, H. S., Gu, D. G., Park, J. S., Ahn, I. S., Ahn, H. J., ... & Kim, K. W. (2014). The effect of electrolyte on the electrochemical properties of Na/α-$NaMnO_2$ batteries. *Materials Research Bulletin*, *58*, 74-77.

[8] Amador, U., Gallardo-Amores, J. M., Heymann, G., Huppertz, H., Moran, E., & y de Dompablo, M. A. (2009). High pressure polymorphs of LiCoPO4 and LiCoAsO4. *Solid state sciences*, *11*(2), 343-348.

[9] Dima, R. S., Maleka, P. M., Maluta, N. E., & Maphanga, R. R. (2022). Structural, electronic, mechanical, and thermodynamic properties of na deintercalation from olivine $NaMnPO_4$: first-principles study. *Materials*, *15*(15), 5280.

[10] Morukuladi, M. T., Lethole, N. L., Masedi, M. C., Ngoepe, N. N., & Ngoepe, P. E. (2022). Structural, Electronic, Elastic and Dynamical Properties of MCO_3 (M: Mn, Co, Ni) Precursor Materials for Li-Ion Batteries: A First-Principles Study. *Journal of The Electrochemical Society*, *169*(2), 020540.

Developing a Nuclear Orientation Thermometer for the UCT Dilution Refrigerator

Yanga Ntolosi[1,2], Steve Peterson[2], Mark Blumenthal[2] and Dominique Gouveia[2]

[1] NMISA, Meiring Naudé Road, Brummeria, Pretoria, South Africa
[2] Department of Physics, University of Cape Town, University Avenue, Cape Town, South Africa

E-mail: YNtolosi@nmisa.org

Abstract. A significant challenge in low temperature thermometry is the accurate measurement of temperatures below 1 K. Nuclear Orientation (NO) is a non-electronic technique to measure ultra-low temperature accurately as opposed to traditional resistive thermometers. The NO method relies on the measurement of the alignment of the nuclear spin in a radioactive nucleus, with the degree of alignment varying with temperature according to Boltzmann distribution. The aim of this study is to develop a NO thermometry system using the recently procured gamma-ray anisotropy thermometer ^{60}CoCo(hcp) source for use in the University of Cape Town Department of Physics dilution refrigerator. The UCT dilution refrigerator is able to achieve these ultra-low temperatures (down to 8 mK) by taking advantage of the properties of both ^{3}He and ^{4}He gas. The ^{60}CoCo(hcp) radiation source, irradiated using the SAFARI-1 research reactor at NECSA, is incorporated into the dilution fridge by thermally mounting it onto the plate in which the mixing chamber is positioned. The data acquisition system, a Sodium Iodide (NaI) scintillation detector, is placed in line with the source allowing it to detect the radiation as accurately as possible. The ratio of the detected radiation at various temperatures provides the measurement of nuclear spin alignment and thus the absolute temperature of the system.The 60CoCo(hcp) thermometer is expected to cover a temperature range of 10 mk to 100 mK with acceptable accuracy. The preliminary measurements are promising, but more work needs to be done in order to develop a fully-functioning NO temperature measurement system.

1. Introduction

Using traditional resistance thermometers to measure low temperature below 1 K can give inaccurate measurements due to self-heating which causes unwanted temperature fluctuations at ultra-low temperatures [1]. In order to overcome this problem, the Nuclear Orientation technique presents an alternative method that allows us to measure temperature below 1 K accurately [1]. NO is based on the ordering of a nuclear spin system in thermodynamic equilibrium, from which the temperature can be derived using the Boltzmann distribution [1]. The self-heating of a gamma-ray anisotropy thermometer, which is caused by the radioactivity of the nucleus, is negligible as compared to a resistance thermometer [1].
NO has a number of favourable properties such as being able to be attached to an experimental set up with good thermal contact because of its metallic property and small dimensions. No wires are attached to it and specific gamma anisotropy thermometers can be used without a

magnetic field [1]. This allows the NO technique to be used in low temperature experiments that need to be performed in a field-free environment.

2. Theory

2.1. Nuclear Orientation

The degree of ordering of the populations in the nuclear spin is governed by the Boltzmann distribution. When the system interacts with its surrounding environment, thermodynamic equilibrium can be achieved [1]. Once this is the case, the spin temperature can be derived from the Boltzmanns factor. The Boltzmann distribution is given as:

$$P = \frac{e^{-Em/kT}}{\sum_m e^{-E_m/kT}} \tag{1}$$

where T is the temperature, E_m is the energy at a substate m and k is the Boltzmann constant [1]. The degree of orientation can be described by using a spin-density matrix [2]. Since the symmetry of the ensemble of nuclear spin is cylindrical, the description will be simplified [1]. The axis of symmetry of a nuclear spin system is the c-axis (axis of quantization) therefore a well-defined crystal is important [1]. The directional distribution of the emitted gamma ray which is normalized can be defined as a function in terms of θ, which is the angle defined between the axis of quantization and the emission of the gamma ray [1]. The directional distribution function is defined as follows,

$$W(\theta) = \sum_{\lambda=0}^{\lambda_{max}} B_\lambda(I,T) U_\lambda A_\lambda Q_\lambda P_\lambda cos(\theta). \tag{2}$$

In the equation $W(\theta)$, B_λ is the orientation parameter which contains the spin-density matrix, A_λ is the angular distribution coefficient, U_λ is the angular momentum de-orientation coefficient, Q_λ is the solid angle correction and the $P\cos(\theta)$ Legendre polynomial [1]. U_λ takes into account the de-orientation that occurs during the intermediate transitions that results as a emitted gamma ray. U_λ can be calculated from the decay scheme of ^{60}CoCo(hcp) which is the radioactive nuclei in use. When the coefficients in equation $W(\theta)$ are known and $B_\lambda(I,T)$ can be determined, $W(\theta)$ can be calculated for a range of temperatures and angle as shown in Figure 1.

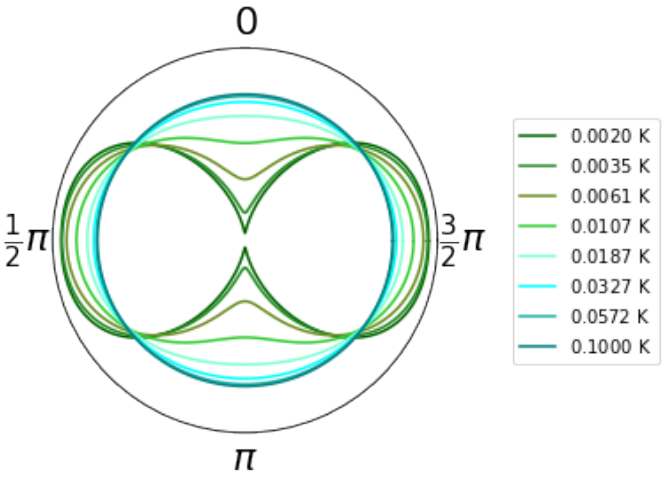


Figure 1. The $W(\theta)$ distribution as a function of θ is plotted for different temperatures [3].

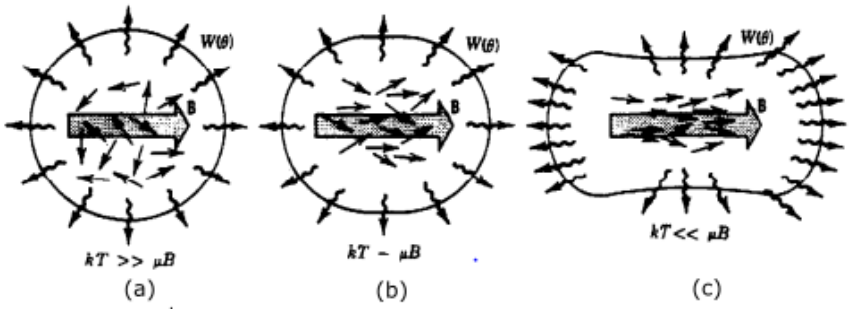


Figure 2. The directional distribution of radiation at different temperatures [4].

In Figure 2 the directional distribution of radiation at different temperatures, kT, is shown. In Figure 2(a), high temperatures (1 K) kT $\gg \mu$B, where B is the overall effective magnetic field generated by the nucleus and μ is the magnetic dipole moment caused by the proton and neutron nuclear spin. In this case, the degree of ordering of the populations is even. The internal magnetic field B has no effect on the nuclear spins due to the thermal motion of the spin system [4]. In Figure 2(b), the radiation distribution at kT $\approx \mu$B (approximately 100 mK) begins to be nonuniform due to the uneven population. Figure 2(c) shows the radiation distribution at very low temperatures kT $\ll \mu$B (approximately 10 mK) where the nuclear spins align with the internal magnetic field B. The temperature range that a gamma-ray anisotropy thermometer can be determined by calculating the partial differential of $W(\theta, T)$ as a function T [1].

2.2. Dilution Process

In order to achieve these ultra-low temperatures, a continuously operating refrigerator is needed. A helium dilution refrigerator is able to achieve these ultra-low temperatures by taking advantage of the properties of ^{3}He and ^{4}He. ^{3}He and ^{4}He has a boiling point of 3.2 K and 4.2 K at atmospheric pressure and in order to achieve lower temperatures, the vapor pressure can be reduced using pumps. The principle of evaporative cooling is applied which uses the latent heat of evaporated liquid [5]. To attain lower temperatures than the boiling points, a mixture of ^{4}He and ^{3}He is considered. A phase separation occurs at 0.8 K with a ^{3}He concentration of 66.9% [6]. When the temperature approaches zero Kelvin, the concentration of ^{3}He will reach a limit of 6.4%. At this point the bottom layer will be ^{4}He rich with a 6.4% of ^{3}He and the top layer will be pure ^{3}He. The dilution process occurs when ^{3}He is removed from the bottom layer which will reduce the concentration of ^{3}He in the bottom layer of ^{4}He. This causes the ^{3}He atom from the top layer to migrate across the phase boundary to replace the amount of ^{3}He that has been removed in the ^{4}He and therefore increase the cooling power [5].The process will repeat until the helium dilution refrigerator reaches base temperature.

3. Method

3.1. ^{60}CoCo Source

^{60}Co is an unstable isotope artificially produced from a stable isotope ^{59}Co. ^{60}Co has a half-life of 5.2714 years and produces two unique gamma ray peaks at 1.17 MeV and 1.3 3MeV. The cobalt source that has been procured is of high-quality with a of purity of 99.9%. This cobalt source was cut into size (approximately $10\times1\times1mm^3$) with the long side being the c-axis which will be accurately determined. The source has been activated at NECSA using the SAFARI-1 nuclear

reactor rabbit which is a PTS (pneumatic transport system) that is used for precisely-timed irradiation.

3.2. Detector and Energy Calibration

Scintillation detectors detect ionizing radiation by converting the incoming gamma radiation into light. The light produced by the scintillating material is converted to a electrical pulse in order to register the presence of any radiation with inorganic solid scintillator being the commonly used detector [7].
Before using the NaI detectors, an energy calibration is needed to convert the voltage pulse heights to energy peaks. In order to accurately measure radiation, the NaI scintillation detector is calibrate using radioactive sources with known energy peaks which are matched with the voltage pulse heights [8]. A plot of the identified energy peaks with respect to the channel is plotted and a straight line is used to fit the data. The straight line function is used to convert the channels to energy in MeV.

3.3. UCT Dilution Refrigerator

The UCT dilution refrigerator is a Leiden Cryogenics Cryogen-free dilution refrigerator system that consists of a dilution cryostat unit, a GHS (gas handling system) and a pulse tube compressor. The dilution cryostat unit consist for a OVC (Outer vacuum chamber), 50 K radiation shield, and a IVC (Inner vacuum chamber). The IVC contains the dilution refrigerator unit which is attached to the 3 K flange. The IVC houses the still plate, the 50 mK cold-plate and the Mixing chamber which is located on the 10 mK cold-plate. The OVC houses the 50 K radiation shield and the IVC is attached to the top external flange.

3.4. Detector and Source placement

The general set up for the nuclear orientation technique consists of the dilution refrigerator, the ^{60}CoCo(hcp) source, detector and data acquisition system. The main uncertainty contributors are the Dilution refrigerator metallic shielding used to insulate and maintain the vacuum, the detector to source distance, the amount of time take measurements and the activity of the source [9].

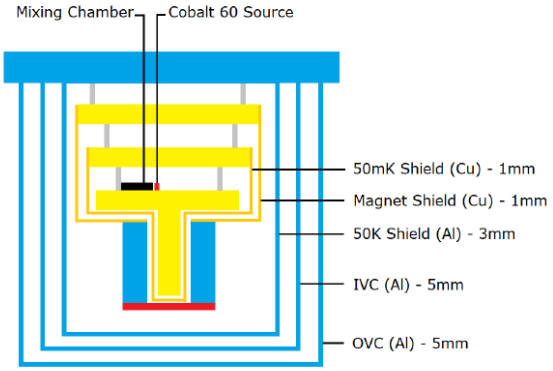


Figure 3. Cross section of the UCT dilution refrigerator with the ^{60}CoCo(hcp) crystal placed next to the mixing chamber [9].

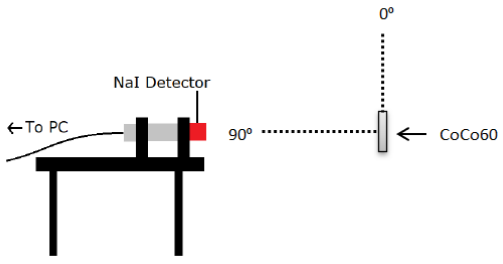


Figure 4. The NaI scintillation detector positioned at a 90° angle from the ^{60}CoCo(hcp) crystal.

The ^{60}CoCo(hcp) source is placed next to the mixing chamber on the 10 mK cold-plate which is the coldest plate on the cryostat unit as shown in Figure 3. The mixing chamber is where the dilution process occurs making it the lowest temperature in the dilution refrigerator. A sample holder was designed to house the ^{60}CoCo(hcp) source due to its small dimension and also helps to mount the source onto the 10mk cold-plate with good thermal contract.
According to Figure 1, the greatest change in $W(\theta, T)$ distribution occurs at 0° and 90° with respect to the c-axis therefore it will be preferable for measurements to be taken at one of the two angles. The design of the sample holder positions the source in an upright vertical position where the 90° will be perpendicular as shown in Figure 4. The detector will be positioned 40.00 ± 0.02 cm away from the ^{60}CoCo(hcp) source with a vertical height of the detector choose to align the center of the Nal detector with the ^{60}CoCo(hcp) source [9].

4. Calculations

In order to measure the temperature of the system, the pre-defined theoretical gamma function W(θ,T) is compared to measured gamma counts to find an experimental W(θ) function. The warm gamma-ray count is measured at a temperature where the anisotropy does not occur, typically above 100 mK for a ^{60}CoCo(hcp) gamma-ray anisotropy thermometer [1]. A temperature of 2 K was chosen to measure the warm count which is the stage where the dilution refrigerator has pumped out all the air molecules to create a vacuum. The cold count will be measured at the desired temperature of interest which is the base temperature of the dilution refrigerature. To account for radiation noice from the surrounding environment, a background count will be subtracted from our cold and warm count. The experimental W(θ) can be rewritten as:

$$W(\theta) = \frac{C_c - B}{C_w - B} \tag{3}$$

where C_c is the cold count, C_w is the warm count and B is the background count [1]. The importance of subtracting the background noise is to account for the gain instability caused by the temperature and magnetic fluctuations in the room. Room temperatures need to be kept constant (>0.5°C) to minimize gain instability and because photomultiplier tubes are sensitive to magnetic fields, shielding material should be used to reduce the external magnetic noise [1]. The second gamma ray peak (1332.5 keV) will be used to reduce the background noise caused by the Compton effect [4].
Once the experimental W(θ,T) has been measured, the temperature can be calculated by taking the difference between the theoretical W(θ,T) and the experimental W(θ) as with:

$$W(\theta, T)' = W(\theta, T) - W(\theta). \tag{4}$$

Once the difference has been calculated, a unique temperature value can be determined by solving for T. Solving T is mathematically challenging because the variable T is found in the exponential terms in the equation. A better alternative is to generate a directional distribution function vesus Temperature lookup graph as shown in Figure 5 at a fixed angle $\pi/2$.

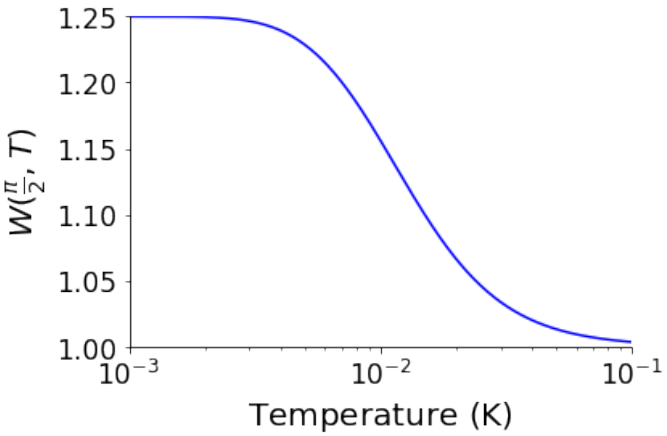


Figure 5. W(θ,T) as a function of temperature at a fixed angle $\pi/2$.

5. Conclusion

In Figure 5, at low temperatures (approximately 10 mK) the W(π/2,T) is distinguishable with a value of 1.25 and at warmer temperatures (approximately 100 mK) the W(π/2,T) trends to 1. The NaI detector was the preferable detector due to resolution making it suitable to resolve W(π/2,T) at warmer temperatures. Two possible angles were identified for measurement and the vertical 90° position was preferable due to a simplified source-to-detector orientation. The final steps in the process involves executing the work outline above. The goal of successfully measuring temperatures below 100 mK using a radioactive ^{60}CoCo(hcp) source at UCT is near at hand.

References

[1] Marshak H 1983 *Journal of Research of the National Bureau of Standards* **88** 175
[2] Scheck F 2007 *Quantum physics* (Springer)
[3] New T 2016 *^{60}CoCo(hcp) nuclear orientation thermometry* (Cape Town: University of Cape Town)
[4] Krane K S 1991 *Introductory nuclear physics* (John Wiley & Sons)
[5] Batey G and Teleberg G 2015 Principles of dilution refrigeration: A brief technology guide
[6] Chaffin K N 1978 *A He3-He4 dilution refrigerator system* Ph.D. thesis Texas Tech University
[7] Knoll G F 2010 *Radiation detection and measurement* (John Wiley & Sons)
[8] Gilmore G 2008 *Practical gamma-ray spectroscopy* (John Wiley & Sons)
[9] Hyslop N 2017 *Practical Considerations when Performing Nuclear Orientation Thermometry for the Helium Dilution Fridge in the Physics Department of the University of Cape Town* (Cape Town: University of Cape Town)

Density Functional Theory Study of Porphyrin Dye Molecule Adsorbed onto TiO_2 (101) Anatase Surface

R.R Randela[1], T.S Ranwaha[1], L.M Mathomu[2], R.R Maphanga[3,4], N.E Maluta[1,3]

[1]Physics Department, University of Venda, Thohoyandou, Limpopo 0950, South Africa
[2]Department of Biochemistry and Microbiology, University of Venda, Thohoyandou, 0950, South Africa
[3]National Institute for Theoretical and Computational Sciences (NITheCS), Gauteng 2000, South Africa
[4]Next Generation Enterprises and Institutions Cluster, Council for Scientific and Industrial Research, Pretoria 0001, South Africa

E-mail: **ronelronella.randela@univen.ac.za**

Abstract. Dye-sensitized solar cells (DSSCs) have attracted considerable attention in recent years as they offer the possibility of low-cost conversion of photovoltaic energy. DSSCs use the dye molecules adsorbed on the TiO_2 semiconductor in nanoarchitecture with the role of absorbing photons from the sun. In this study, Density functional theory (DFT) has been used to study the geometric, electronic, and optical properties of Porphyrin dye molecule and its adsorption behaviour on (101) TiO_2 anatase surface. The generalized gradient approximation (GGA) was used in the scheme of Perdew-Burke Ernzerhof to describe the exchange-correlation function as implemented in the CASTEP package in the material studio of BIOVIA. The results obtained show that Porphyrin dye molecules can improve the efficiency of DSSCs as they show an absorption shift to the near-infrared region, which increases the absorption range on the visible solar spectrum. UV-vis absorbance was found to be between 450, 550 and 620 nm. The porphyrin dye exhibited energy gaps of 2.8 eV. This study demonstrates that the dye molecule synthesized from *B. pilosa* is an efficient sensitizer for DSSCs.

1. Introduction

Electrical energy supply and distribution are key concerns in industrialized and developing countries.The global energy issue is expected to rise as energy consumption rises [1]. Developing countries receive most of their energy from fossil fuels such as coal, which damage the environment and causes climate change by emitting too much carbon dioxide [2]. Climate change and non-renewable resource use are humanity's biggest problems, therefore rapid access to clean and renewable energy is overdue. Renewable energy can replace fossil fuels. It has been widely predicted that fossil fuels will deplete shortly [2]. Alternative promising technology is Dye Sensitized Solar Cells (DSSCs) [3]. Thus, studying DSSCs is motivated. Solar cells have a long-term potential since DSSCs are environmentally friendly, cheap to fabricate, and easy to produce [4,5]. Previously, ruthenium dye was used; however, ruthenium dye is expensive due to their scarcity of ruthenium metal and the chemical processes used in their manufacture.

Researchers are investigating organic metal-free dyes, which are natural dyes and can be easily manufactured [6-8]. Natural dyes assist in converting sunlight into electricity [4,6]. Green plants contain a lot of chlorophyll, which helps absorb photons from the sun.

Since TiO_2 is less expensive and can be synthesised simply, there has been numerous attempts to enhance its transportation of electrons to improve solar energy conversion in DSSCs. Porphyrin dye has

gained an interest as the organic dye sensitizer to be used on a DSSCs, because Porphyrin compounds are naturally occurring [9,10]. Porphyrins' chemical and physical mechanisms are responsible for electron transport, oxygen binding, photochemical reactions, and photosynthesis process. Porphyrin absorption spectra comprise two distinct bands: 380-500 nm (blue region) and 500-650 nm (red region) [8-10]. The working principle of DSSCs is shown in Figure 1, where the sunlight travels through the transparent electrode into the dye layer and is absorbed by the dye, which transmits and injects electrons into the TiO_2 semiconductor. The electrons then travel towards the transparent electrode, where they are captured and used to power a load. After passing via the external circuit, the electrons are reintroduced into the metal electrode flowing into the electrolyte. Then electrolyte returns the electrons to the dye molecules, regenerating an electron flow and producing current.

The structure of Porphyrin($C_20H_{12}MgN_4$) dye molecule is shown in Figure 2. Porphyrin molecule consist of four pyrrole rings and at the center of the porphyrin ring, a magnesium ion helps to maintain the chlorophyll molecule and promote the photosynthetic process by transferring electrons down an electron transport chain. In this study we used computational approach to analyze the adsorption of Porphyrin dye molecule onto TiO_2 (101) nanoparticle surface using density functional theory (DFT)

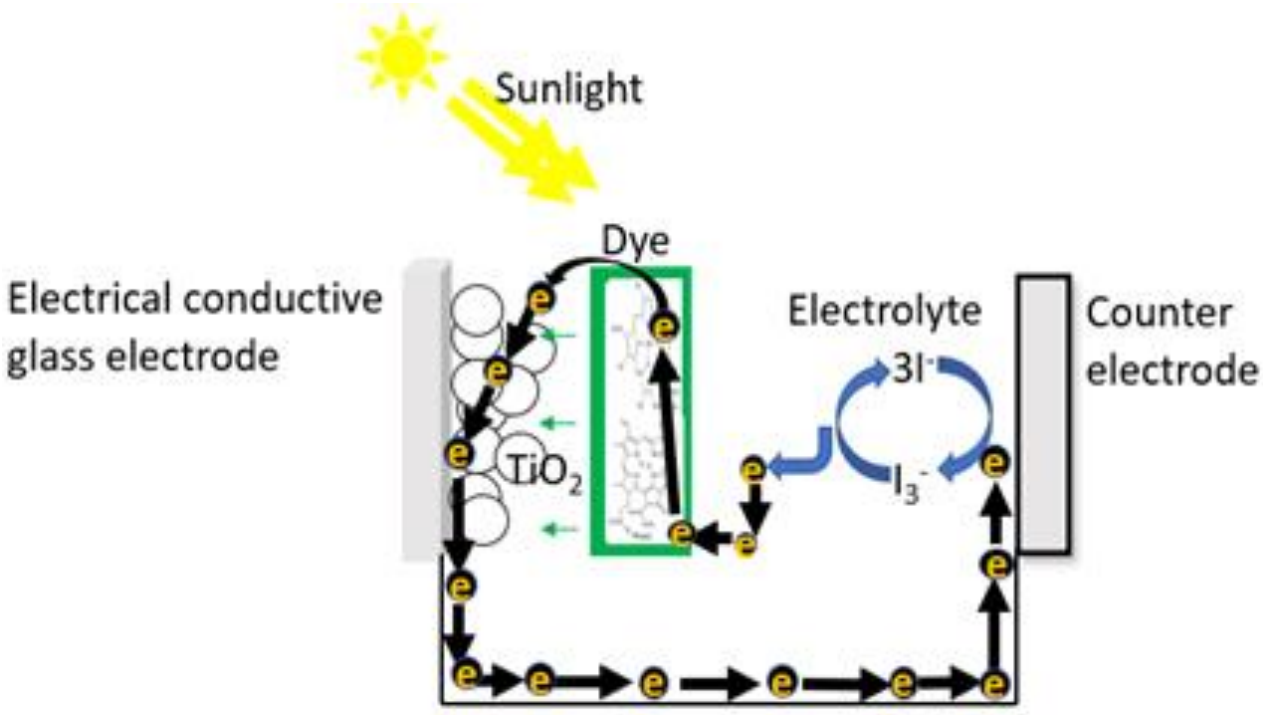


Figure 1. Working principle DSSC.

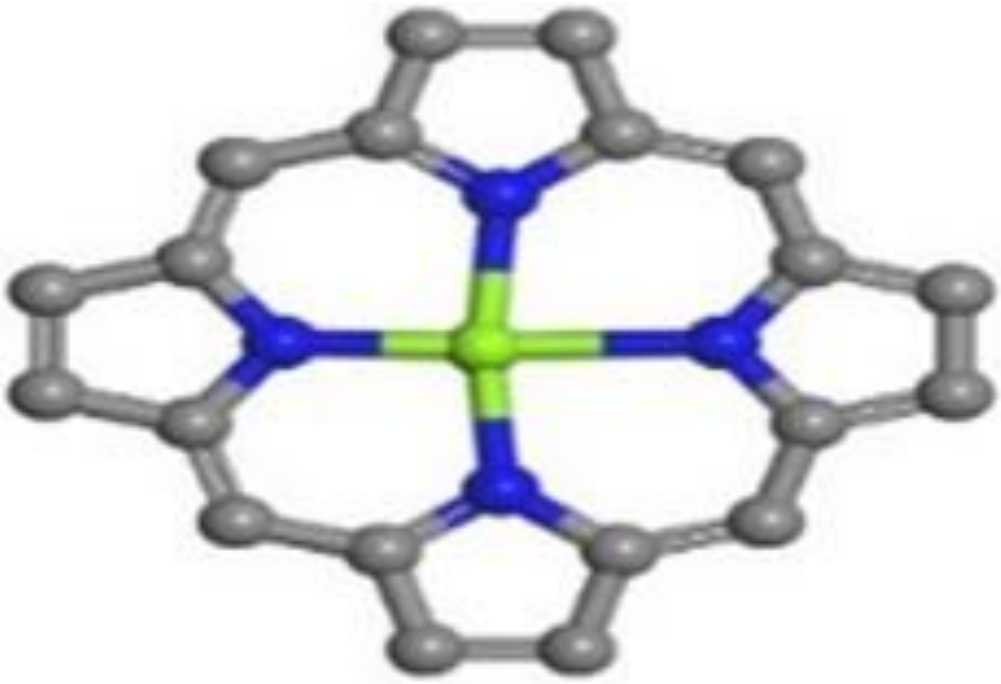


Figure 2. Chemical structure of Porphyrin dye molecule.

2. Computational methodology

The structure of the dye molecule was built using Material Studio on 3D atomistic window, the molecule was cleaned so that the atoms are re-oriented on their lattice position. Geometrical optimizations of the dye molecule were performed using the first principles calculations based on density functional theory (DFT) which uses a plane-wave pseudopotential method. A generalized gradient approximation (GGA) in the scheme of Perdew-Bruke-Ernzerhof (PBE) was used to describe the exchange-correlation functional using the coarse quality and all band/EDFT as electronic minimizer.

The calculations were carried out using Cambridge Serial Total Energy Package (CASTEP) code, which is a package within the framework of the Materials Studio software which is used for periodic system as it utilizes Plain-wave Basis Set, was used to obtain the ground state properties of the dye molecule. Then the calculations of electronic properties of dye molecule were done using Vulnerability Analysis Methodology Program (VAMP). The dye molecules were then imported to a new 3D atomistic window, the electronic levels and energy orbitals of the dye molecules were then calculated to determine the UV-VIS spectrum graph of the dye molecule and on the orbitals, we get to view the HOMO (Highest Occupied Molecular Orbital) and LUMO (Lowest Unoccupied Molecular Orbital) values for the dye molecules. The cutoff energy for plane-wave basis set to 500 eV, 5 x 2 x 2 Monkhorst-pack. All structure were inside the vacuum slab of 10.00 Å vacuum thickness, 17.552 Å and 0.00 Å slab position.

Anatase TiO_2 bulk structures were also optimized using CASTEP code within the framework of the Material Studio package to obtain the ground state properties of the anatase TiO_2 structure, GGA-PBE functional for determination of convergence parameters and geometrical optimization of anatase TiO_2 structure were used, the fixed basis set and ultra-soft pseudopotential was used throughout the study. The ground state structures obtained through geometrical optimization with the convergence parameters were imported into a new 3D atomic window and the surfaces were cleaved from the bulk structure and a vacuum slab of appropriate size was built for the surface structures. The surfaces were optimized using the same convergence parameters obtained for the bulk structures. The dye/TiO_2 complex was optimized to obtain the ground state structures, and the electronic properties were calculated using CASTEP code.

3. Results and discussion

3.1. Absorption Spectrum of dye molecule

Absorption is important for an organic dye to be recognized as the ideal sensitizer for usage in DSSCs. Organic dyes should be able to absorb photons in both the visible and near infrared region.

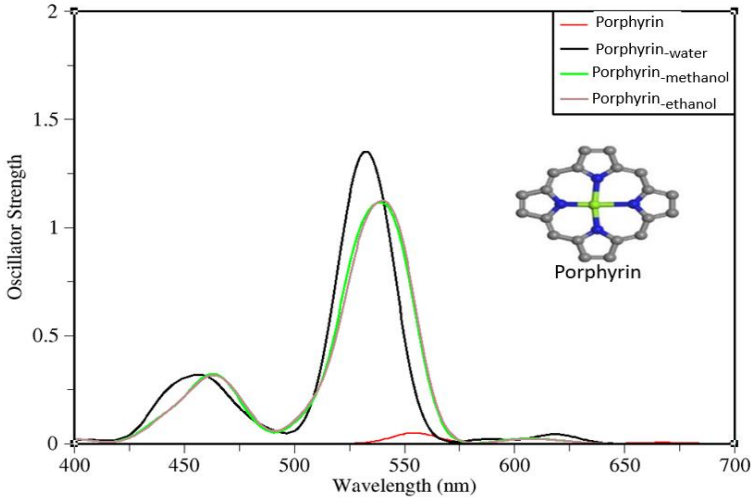


Figure 3. Absorption spectrum of porphyrin dye molecule and porphyrin dye with solvents

Before considering the effect of solvents, the porphyrin molecule exhibits minimal absorption in the visible 550 nm which contributes 6% of its Light Harvesting Efficiency. When the porphyrin molecule was absorbed with solvents (water, methanol, ethanol) a stronger absorption was observed in the visible region at 534,538 and 550 nm, respectively. Figure 3 shows that porphyrin molecule is absorbed with effects of solvents such as water, ethanol, and methanol, it tends to harvest more photons in the visible region of the solar spectrum but absorbs less photons in the nearer infrared region because its absorbance slows down but can still absorb photons in the nearer infrared region [9,11].

3.2. Energy level and isodensity of the Porphyrin dye

The HOMO-LUMO gap refers to the energy disparity between the highest occupied molecular orbital and the lowest unoccupied molecular orbital. The HOMO signifies the electron-donating capacity, while the LUMO symbolizes the electron-accepting propensity, table 1 shows the HOMO, LUMO, and HOMO-LUMO energy gap of Porphyrin dye molecule, equation (1) was used to calculate the energy gap.

$$E_{gap} = HOMO - LUMO \quad (1)$$

Table 1. The HOMO, LUMO and HOMO-LUMO energy gap of porphyrin dye molecule.

Molecules	HOMO (eV)	LUMO (eV)	E_{gap} (eV)
Porphyrin	-7.031	4.26	2.8
Porphyrin-$_{water}$	-8.742	-5.08	3.7
Porphyrin-$_{methanol}$	-8.703	-5.091	3.6
Porphyrin-$_{ethanol}$	-8.703	-5.091	3.6

A smaller energy difference between the sensitizer's HOMO and LUMO increases photon absorption in higher spectral regions of the solar spectrum [12,13]. The porphyrin dye molecule has a significantly lower HOMO-LUMO energy gap of 2.8 eV. High HOMO energy means that the compounds or molecules are very reactive, which makes it easier for electrons to move between them and for chemical bonds to form. A HOMO with lower energy compared to that of semiconductor Valence Band indicate

the stability of the molecule and its less reactive while LUMO higher than semiconductor Conduction Band energy for dye molecules allows electron injection from the semiconductor, resulting in faster dye recombination to ground state complex [12,13].

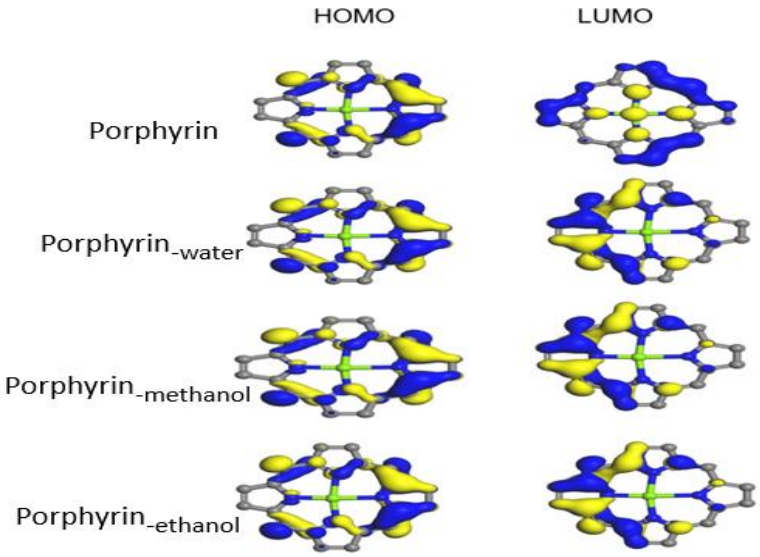


Figure 4. Isodensity of porphyrin dye molecule

Figure 4 shows the location of HOMO and LUMO orbitals on a dye molecule. The frontier orbitals of the molecules are centered on the porphyrin ring. This gives a consequence that the structure of porphyrin ring will be responsible for most of the electronic transition involved such as absorption, emission, and charge transfer. The distribution of electrons between HOMO and LUMO mostly contributes to electronic transitioning of the molecules involving photons that provide high oscillating strength for both HOMO and LUMO transitions [14].

3.3. Optical absorption

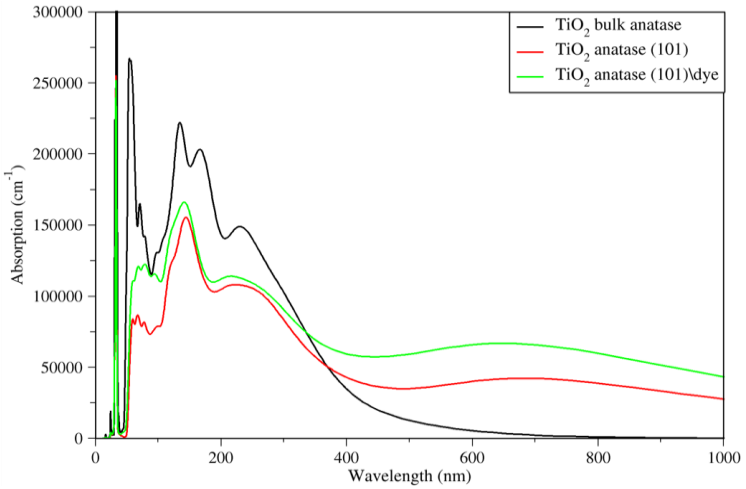


Figure 5. Optical absorption of bulk TiO_2 anatase, supercell (4× 4 × 1) on a (101) surface, and adsorbed TiO_2/dye.

The absorption coefficient defines how far light of a specific wavelength can penetrate before being absorbed. Three separate energy peaks of absorption spectra occur at similar peaks with part of the dielectric function and extinction coefficient. The optical absorption coefficients for the TiO_2 anatase bulk structure and supercell (4 × 4 × 1) of (101) anatase TiO_2 surfaces are compared in Figure 5. The optical band gap differs significantly from the electronic band gap. The electrical band gap is reasonable to assume a single particle, whereas optical absorption requires activation beyond a single particle. The optical absorption of TiO_2 anatase bulk structure and supercell (4 × 4 × 1) of (101) anatase TiO_2 surfaces is shown in Figure 5, revealing that the bulk structure and surfaces can absorb photons in the ultra-violet and near infrared regions, because Figure 5 shows photons with stronger peaks in the ultraviolet area of the solar spectrum and weaker peaks in the visible to near infrared range. However, when the dye is adsorbed onto TiO_2 anatase, substantial absorption in the near infrared range is seen. TiO_2 bulk anatase absorbs only in the ultra-visible and visible bands due to its bulk structure. Adsorption of dye molecules onto TiO_2 can improve the absorption coefficient of TiO_2/dye-based solar cells for renewable energy applications.

3.4. Optical Reflectivity and Refractive index

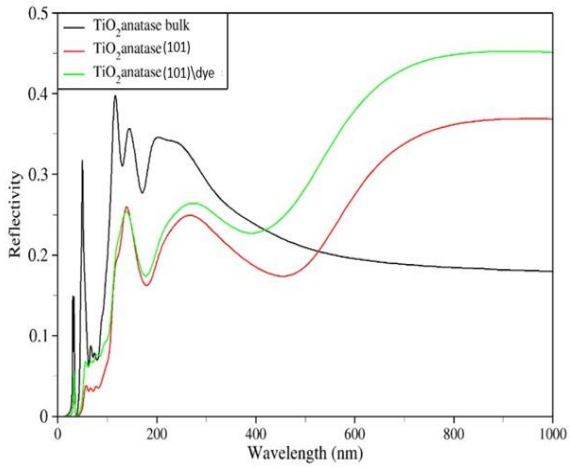


Figure 6a. Optical reflectivity of TiO_2 anatase bulk, supercell (4×4×1) on a (101) surface and adsorbed TiO2/dye.

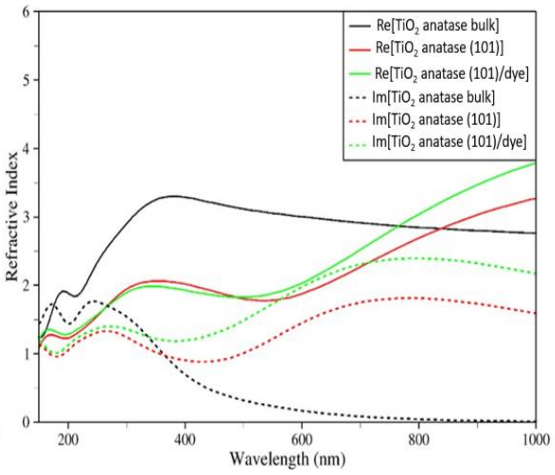


Figure 6b. Optical refractive index of TiO_2 anatase bulk, supercell (4×4×1) on a (101) surface and adsorbed TiO_2/dye.

.

In Figure 6a, a higher reflection of light in the visible region and higher reflection in the near infrared region for supercell and TiO_2/dye complex was observed, whereas the bulk TiO_2 structure displays a decrease from 600 - 1000 nm of solar spectrum. The increased reflection of radiation in the near infrared range is due to atom re-orientation during cleaving of the structure [15]. As expected, the reflectivity for the bulk structure of the TiO_2 anatase curve is substantially higher in the ultra-visible region, our results agree well with the work previously done by Ranwaha *et al.* [15] and Phuthu *et al* [16] they reported higher reflectivity on the ultra-violet region, and lower reflectivity on visible to near infrared region[15,16]. Additionally, in Figure 6b the refractive index graphs for TiO_2 anatase bulk, supercell (4 x 4 x 1) of TiO_2 anatase on a (101) surface and TiO_2/dye complex. For bulk structure, the real refractive index part increases steadily in the visible region, while the imaginary refractive index part drops from 400 to 600 nm. For both supercell and adsorbed dye, the real refractive index displays a distinct absorption of photons or light from the visible to near-infrared region, whereas the imaginary

refractive index exhibits less absorption. If the refractive index is high, light or photon absorption slows and tends to decrease. For both the imaginary refractive index of the bulk structure, supercell, and the adsorbed dye, we observed increased absorbance at 200-380 nm and lower absorbance at 400-1000 nm. However, the dye adsorbed on TiO_2 has a higher refractive index for both the real and imaginary graphs. This demonstrates that our dye molecule adsorbed onto TiO_2 can improve solar spectrum absorption from visible to near infrared.

4. Conclusion

The result shows strong probability of Porphyrin dye molecule to be regarded as the best sensitizer application in DSSCs. The results showed good optical, excitation and spectral properties. The Porphyrin molecule attributed to a broad absorption in the visible region of the spectrum, The UV-vis spectra showed the maximum absorption peak at 550 nm within the edge of the visible region thus favourable for DSSC. The HOMO and LUMO energy levels suggests an energy gap of 2.8 eV. The stimulated results showed that the adsorption of porphyrin dye molecules onto the TiO2 semiconductor can improve the performance of DSSCs.

Acknowledgement

We would like to acknowledge the University of Venda, Centre for High Performance Computing (CHPC), and National Institute for Theoretical and Computational Sciences (NITheCS) for financing and allowing us to conduct this research.

References

[1] Hamann, T. W., Jensen, R. A., Martinson, A. B., Van Ryswyk, H., & Hupp, J. T. (2008). Advancing beyond current generation dye-sensitized solar cells. Energy & Environmental Science, 1(1), 66-78.

[2] Mahlangu, N., & Thopil, G. A. (2018). Life cycle analysis of external costs of a parabolic trough Concentrated Solar Power plant. Journal of cleaner production, 195, 32-43.

[3] Kalyanasundaram, K., & Grätzel, M. (2009). Efficient dye-sensitized solar cells for direct mconversion of sunlight to electricity. Material Matters, 4(4), 88-90.

[4] Grätzel, M. (2003). Dye-sensitized solar cells. Journal of photochemistry and photobiology C: Photochemistry Reviews, 4(2), 145-153.

[5] Li, N., Pan, N., Li, D., & Lin, S. (2013). Natural dye-sensitized solar cells based on highly ordered TiO2 nanotube arrays. International Journal of Photoenergy, 2013.

[6] Syafinar, R., Gomesh, N., Irwanto, M., Fareq, M., & Irwan, Y. M. (2015). Chlorophyll pigments as nature-based dye for dye-sensitized solar cell (DSSC). Energy Procedia, 79, 896-902.

[7] Yang, H., Li, J., Zhou, G., Chiang, S. W., Du, H., Gan, L., & Duan, W. (2015). First principles study of ruthenium (II) sensitizer adsorption on anatase TiO2 (001) surface. RSC advances, 5(74), 60230-60236.

[8] Yella, A., Lee, H. W., Tsao, H. N., Yi, C., Chandiran, A. K., Nazeeruddin, M. K., ... & Grätzel, M. (2011). Porphyrin-sensitized solar cells with cobalt (II/III)–based redox electrolyte exceed 12 percent efficiency. science, 334(6056), 629-634.

[9] Campbell, W. M., Burrell, A. K., Officer, D. L., & Jolley, K. W. (2004). Porphyrins as light harvesters in the dye sensitized TiO2 solar cell. Coordination Chemistry Reviews, 248(13-14), 1363-1379.

[10] Birel, Ö., Nadeem, S., & Duman, H. (2017). Porphyrin-based dye-sensitized solar cells (DSSCs): a review. Journal of fluorescence, 27, 1075-1085.

[11] Bevilaqua, R. C., Zanella, I., & Fagan, S. B. (2010). Chlorophyll a and pheophytin a as gas sensors of CO_2 and O_2 molecules. Chemical Physics Letters, 496(4-6), 310-315.

[12] Rahmatika, Z. (2021, February). Theoretical Modification of Pheophytin Using Cu, Ni, and Zn Atoms as a Sensitizer for Dye Sentized Solar Cell (DSSC). In Journal of Physics: Conference Series (Vol. 1788, No. 1, p. 012006). IOP Publishing.

[13] Karuppasamy, A., Stalindurai, K., Peng, J. D., Ho, K. C., & Ramalingan, C. (2016). Novel metal-free organic dyes possessing fused heterocyclic structural motifs for efficient molecular photovoltaics. Physical Chemistry Chemical Physics, 18(43), 30105-30116.

[14] Suendo, V., & Viridi, S. (2011). Ab initio calculation of UV-Vis absorption spectra of a single molecule chlorophyll a: Comparison study between RHF/CIS, TDDFT, and semi-empirical methods. arXiv

preprint arXiv:1105.3766.
[15] Ranwaha, T. S. (2019). Density functional theory study of adsorption of cronconate dyes on TiO2 Anatase (010) and (100) surfaces.
[16] Phuthu, L., Dima, R. S., Maluta, N. E., Kirui, J. K., & Maphanga, R. R. (2022). DFT study of TiO2 brookite (210) surface doped with silver and molybdenum. Materials Research Express, 9(9), 095901.

Nuclear Orientation Thermometer for the UCT Dilution Refrigerator

Yanga Ntolosi[1,2], Steve Peterson[2], Mark Blumenthal[2] and Dominique Gouveia[2]

[1]NMISA, Meiring Naud´e Road, Brummeria, Pretoria, South Africa
[2]Department of Physics, University of Cape Town, University Avenue, Cape Town, South Africa

E-mail: **YNtolosi@nmisa.org**

Abstract. The development of a Nuclear Orientation (N O) thermometer system for use at the University of Cape Town is essential to address the significant challenge in the accurate measurements at ultra-low temperatures (down to 8 mK). A ^{60}CoCo(hcp) gamma ray thermometry source was irradiated using the NECSA SAFARI-1 research reactor for 6 minutes and a preliminary activity value of 1.3 MBq was achieved. The activity of this source has been validated using an absolute gamma-ray coincidence technique and verified using a well-type ionizing chamber. Preliminary temperature measurements were taken by placing the $_{60}$Co source within the University of Cape Town Department of Physics dilution refrigerator using a vertical sample holder and measuring the anisotropy of the radiation at a 90° angle using a Sodium Iodide (NaI) scintillation detector. These results were promising, but inconclusive, prompting a second set of measurements. Modifications were made to the experimental set up by re-designing the sample holder to hold the source horizontally and take measurements at 0° angle along the *c*-axis. A Lanthanum Bromide (LaBr $_3$) scintillation detector was also used to measure the radiation at temperatures ranging from 10 mK to 100 mK. These experimental improvements provided more accurate and conclusive results.

1. Introduction

Nuclear orientation is a technique that allows us to measure ultra-low temperatures (down to 10 mK) by measuring the degree of ordering of a nuclear spin system gamma-ray anisotropy thermometer. The absolute temperature is derived from the Boltzmann factor by allowing the nuclear spin system to be in thermodynamic equilibrium with its environment. The advantages of using a nuclear orientation thermometer are, (1) it is small (dimensions of 1x1x10 mm^3) which makes it easy to insert into an experimental setup, (2) it has good thermal contact because of its metallic properties and (3) the self-heating of gamma-ray anisotropy thermometer which is caused by the radioactivity of the nuclide is small as compared to resistance thermometers [1].

2. Theory and Calculations

The degree of ordering of the nuclear spin system is governed by the Boltzmann distribution where the temperature can be determined from the Boltzmann's factor [1]. A crystal with a well-defined *c*-axis of symmetry is used. To describe the degree of orientation, an orientation function which consists of a spin-

density matrix is used where the diagonal of the matrix is the relative population of the nuclear substates [2]. A directional distribution function, $W(\theta)$, in terms of the c-axis angle is defined as,

$$\mathrm{W}(\theta)=\sum_{\lambda=0}^{\lambda max} B_\lambda(I,T)U_\lambda A_\lambda Q_\lambda P_\lambda \cos(\theta) \quad (1)$$

Where A_λ is the angular distribution coefficient, U_λ is the angular momentum de-orientation coefficient, Q_λ is the solid angle correction coefficient and $P_\lambda\cos(\theta)$ is a Legendre polynomial. To determine the experimentally calculated $W(\theta)_e$, a cold and warm count ratio measurement is taken. The warm count measurement is taken at a temperature where there is no anisotropy and the cold count measurement is taken at a temperature where there is anisotropy. The experimentally determined $W(\theta)_e$,

$$W(\theta)e = \frac{C_c - B_c}{C_w - B_w} \quad (2)$$

Where C_c is the cold count, C_w is the warm count and B is the background count. The importance of subtracting the background noise is to account for the gain instability caused by the temperature and magnetic fluctuations in the room. In order to calculate the absolute temperature, the temperature T has to be solved from the equation below,

$$W(\theta,T) = W(\theta,T) - W(\theta)_e \quad (3)$$

It is not a straightforward task to solve because the equation has exponential terms so a look up graph can be used as shown in figure 1.

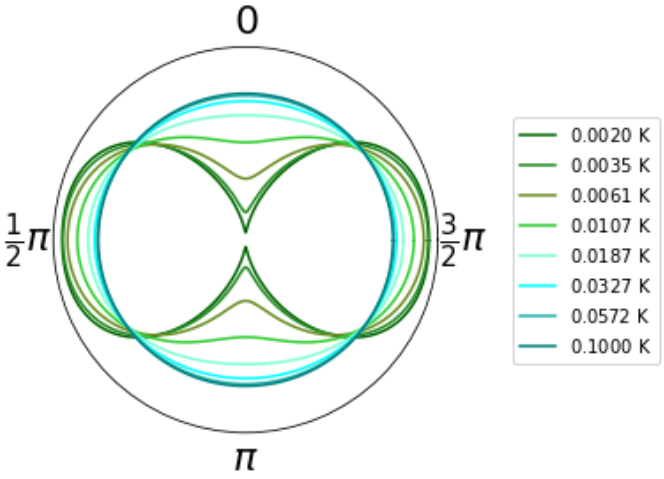


Figure 1. The theoretical $W(\theta)$ distribution as a function of θ plotted for different temperatures [3].

3. Experimental Set-up

The thermometry source is a high-quality cobalt crystal with a purity 99.9%. This cobalt source was cut into size (approx. 1x1x10 mm^3) with the long side cut parallel to the c-axis which is accurately determined [1]. The cobalt was irradiated at NECSA (Nuclear Energy Corporation of South Africa) using the SAFARI-1(South African Fundamental Atomic Research Installation) nuclear reactor. The sample was irradiated using a PTS (pneumatic transport system) or ”rabbits” that is used for precisely-timed

irradiation. The rabbit facility uses a vacuum system to position and remove the sample from the core of the reactor.

The irradiated ^{60}CoCo(hcp) source was positioned next to the mixing chamber on the 10 mK plate which is the coldest point in the dilution refrigerator. The source was placed horizontally in the re-designed sample holder. For the measurements, a NaI detector and a $LaBr_3$ detector were used to capture the energy spectra. These two detectors were positioned at a distance of 47.5 cm away from the source with the NaI detector positioned at 0° and the $LaBr_3$ detector positioned at an angle of 12° with respect to the ^{60}CoCo(hcp)crystals well-defined c-axis. A top-view schematic drawing of the experimental set-up is shown in figure 2.

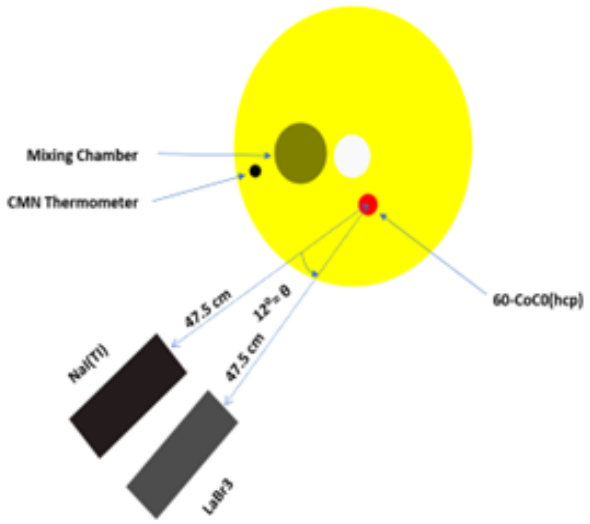


Figure 2. A top view cross section of the UCT Dilution Refrigerator with ^{60}CoCo(hcp) crystal. A top view cross section placed next to the Mixing Chamber.

Figure 3. The NaI scintillation detector positioned at a 0° angle from the ^{60}CoCo(hcp) crystal.

The dilution refrigerator is cooled to 3 K using a pulse tube compressor and the dilution process is used to further cool the dilution refrigerator from 3 K down to 30 mK [4]. Once the dilution refrigerator reaches base temperature (in our case 30 mK), measurements were taken according to the experimental set-up shown in figure 3. Each measurement was taken for 15 minutes once the dilution refrigerator had stabilised to our temperature of interest. The acquisition system used was an USX 30 software, a multi-channel analyser, power supply and amplifier.

4. Detector calibration

Radioactive sources emit electromagnetic radiation in the form of photons which are converted to a current pulse using a Photomultiplier. These current pulses are sorted by a UCS30 ADC (multichannel analyser) into voltage pulse heights. These voltage pulse heights are converted to energy peaks using radioactive sources of known spectrum peak(s). In the case. The radioactive sources used to calibrate the NaI were Barium-133, Sodium-22, Manganese-54, Cadmium-109 and Cobalt-60. In the case of the

$LaBr_3$ detector, it has an intrinsic activity that originates from La-138 and Ac-227. The intrinsic activity pulse height was used to calibrate the $Labr_3$ detector since it is well-known.

5. Analysis considerations/background subtraction

Background noise can originate from numerous sources and a well accounted for background subtraction is important for calculating the areas under the peaks. This can improve the accuracy of our absolute temperature measurement [1]. To ensure the measured spectra is only from our source, the following steps were taken. Repeated 15-minute measurements are taken at every temperature point of interest as mentioned in the experimental section but before the measurements commence, a background measurement was taken for 45 minutes [5]. The background is removed by subtracting the measured background from the measured spectra. This is done by subtracting the background channel-by-channel. The remaining spectra will be background-free.

Lanthanum Bromide background measurement accounts for the background noise in its surrounding environment and its internal intrinsic background noise. The internal intrinsic radioactivity from La-138 and Ac-227 have noticeable peaks on our pulse-height spectrum that contributed to the overall background [6].

Before the area of a peak can be determined, a baseline subtraction has to be done to remove all the unwanted counts. A baseline subtraction is used to subtract all the counts below the peak by fitting a function below the peak in the spectra. Depending on the complexity of the spectra, a polynomial of a certain degree can be used but for our study, a linear fit was used for all baseline subtractions.

6. Results

In this section, we will discuss the activity measurements, the temperature range of a ^{60}CoCo thermometer, and compare the ^{60}CoCo nuclear thermometer to the CMN (Cerium Magnesium Nitrate) thermometer used in the dilution refrigerator. The CMN thermometer is positioned near the mixing chamber as shown in Figure 2. Two detectors with different resolutions were used to compare the measured radiation from the ^{60}CoCo thermometer. Multiple measurements were taken and the average was considered and compared.

6.1. Activity Measurements

The coincidence technique is commonly used for determining the disintegration rates for radioactive materials [7]. Time coincidence is determining how close in time were the first and the following emitted radiation. Since there are multiple nuclei, a case where two emitted gamma-rays originate from different nuclei within the time window is referred to as chance coincidence. Our coincidence set-up consists of 2 detectors, a delay amplifier, a counter and a universal coincidence module.

Table 1. Coincidence measurement summary

	Time(s)	Detector 1	Detector 2	Chance Coincidence	Coincidence
Run	74809	5080646	43265845	4055	5108
Background	666664.7	2100495	2733164	12	112

The background measurement measured over a number of days to account for all background. The activity measurement was taken for approximately 20 hours and an activity of 1321.088±120.093 kBq

was measured. To verify our results, we used an Ionisation chamber and obtained an activity value of 1304.250±195.637 kBq.

6.2. Thermometer temperature range

Nuclear Orientation Thermometers have different temperature ranges which is dependent on the material property and magnetic field. The temperature range of a nuclear thermometer can be determined by determining the partial differential of the directional distribution function. The partial differential of $W(\theta)$ indicates the sensitivity of the nuclear orientation thermometer. For ^{60}CoCo NO, 6.9 mK is the point of high sensitivity and the range of the thermometer is 3 mK to 100 mK [1].

6.3. $LaBr_3$ detector

In figure 4, a comparison between the ^{60}CoCo NO thermometer using a $LaBr_3$ detector and the CMN thermometer has been plotted, where the error bars are the standard uncertainty. The uncertainty of the ^{60}CoCo NO using a $LaBr_3$ was determined by considering the standard deviation of the repeated measurements. As we can see, the uncertainty of the ^{60}CoCo NO $Labr_3$ thermometer increases as the temperature raises whereas the CMN thermometer uncertainty is approximately 1%. At each measured temperature, multiple measurements were taken and the average was considered for both ^{60}CoCo NO thermometer and CMN thermometer. The $LaBr_3$ measurements are comparable to the CMN Thermometer throughout the measured temperature range.

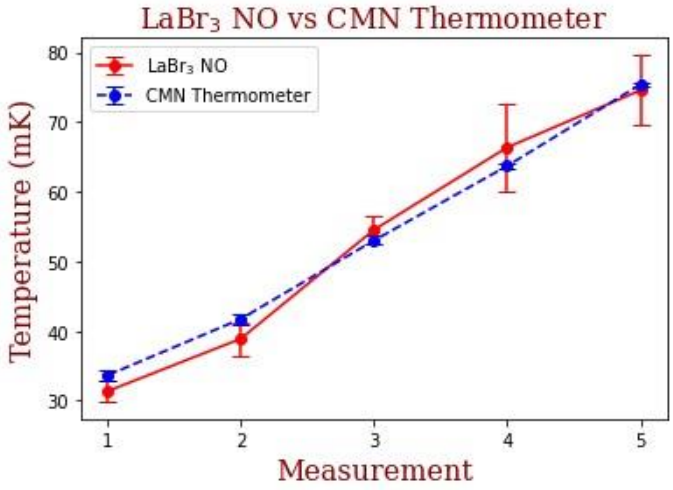


Figure 4. W(θ,T) as a function of temperature at a fixed angle 12°.

6.4. NaI detector

In figure 5, a comparison between the ^{60}CoCo NO thermometer using a NaI detector and the CMN thermometer has been plotted, where the error bars are the standard uncertainty. The uncertainty of the ^{60}CoCo(hcp) NO using a NaI was determined by considering the standard deviation of the repeated measurements. The uncertainty of the ^{60}CoCo NO thermometer increases as the temperature raises. At each measured temperature, multiple measurements were taken and the average was considered for both ^{60}CoCo NO NaI and CMN thermometer. The NaI measurements are comparable to the CMN Thermometer from 30 mK to 50mK the measured temperature range and begins to deviate away from the CMN at higher temperatures.

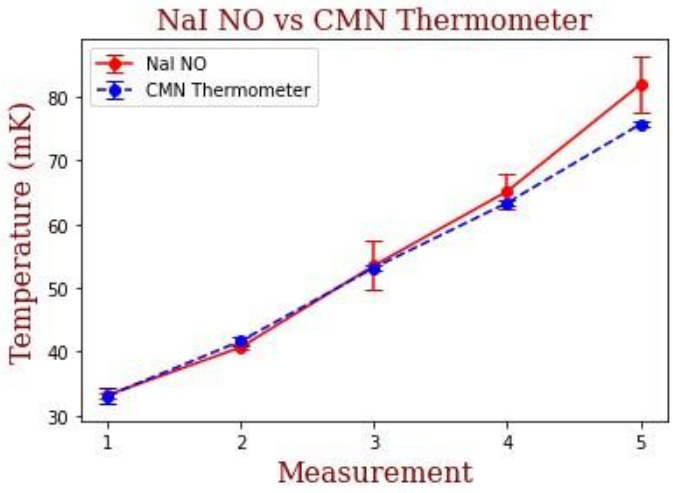


Figure 5. W(θ,T) as a function of temperature at a fixed angle 0°.

7. Conclusion

The NaI detector measurements are comparable to the CMN Thermometer throughout the measured temperature range up to 50 mK. Beyond 50 mK, NaI measurement deviates away from the CMN thermometer whereas the $LaBr_3$ measurement is approximately close to CMN thermometer. The $LaBr_3$ detector has better resolution compared to the NaI(Ti) detector, therefore $W(\theta)_e$ was resolved more accurately above 50 mK [6]. The accuracy of the NaI measurement from 30 mK to 50 mK is superior than that of the $LaBr_3$ detector possibly due to the fact that the intrinsic radioactive radiation will add to the background noise.

References

[1] Marshak H 1983 *Journal of Research of the National Bureau of Standards* **88** 175
[2] Scheck F 2007 *Quantum physics* (Springer)
[3] New T 2016 *^{60}CoCo(hcp) nuclear orientation thermometry (Cape Town: University of Cape Town)*
[4] Chaffin K N 1978 *A He3-He4 dilution refrigerator system* Ph.D. thesis Texas Tech University
[5] Hyslop N *2017 Practical considerations when performing nuclear orientation thermometry for the helium dilution fridge in the physics department of the University of Cape Town (Cape Town: University of Cape Town)*
[6] Knoll G F 2010 *Radiation detection and measurement* (John Wiley & Sons)
[7] Krane K S 1991 *Introductory nuclear physics* (John Wiley & Sons)

Design of a PV power system for grid-connected facilities energy retrofitting: A case study of 15 SAI Battalion, Limpopo Province in South Africa

TS Mulaudzi; D Tinarwo; S Muronga; NE Maluta

University of Venda, Department of Physics, P/Bag X 5050, Thohoyandou, 0950

email: **david.tinarwo@univen.ac.za**

Abstract. The study comprehensively analysed the calculation of an optimally designed grid-connected photovoltaic (PV) system's energy output. Monthly solar radiation data was assessed, and average sunshine hours were generated to design the 15 South Africa Infantry (SAI) Battalion Head Quarters (HQ) building's electric power supply connected to the grid. Individual rooms' appliance inventory with electric parameters and time of use data was done for energy use audit. Due to the high total load power and financial constraints, the power systems were designed to cover part of the load of the building. PVSyst software was used to design and analyse the system. The system overview showed that the annual production is 18130 kWh, whereas the daily normalised output is 4.81 kWh/kWp. Furthermore, the PV array and system losses are 1.23 kWh/kWp/day and 0.14 kWh/kWp/day respectively.

1. Introduction

The increasing population and ageing of the electricity generation stations in South Africa (SA) pose a severe load on the grid. This has led to several load-shedding stages being implemented, and the companies and inhabitants are frustrated. It has become a norm. On the other hand, SA is among those countries that depend heavily on coal as it generates its electricity, hence a high percentage of carbon footprint. According to the National Energy Regulator of South Africa (NERSA), the energy sector alone contributes nearly 80% of total emissions, of which 50% are from electricity generation and liquid fuel production alone [1]. The timing of the transition to a low-carbon economy must be in a manner that is socially just. The demand for renewable energy solutions has surged as the world strives to transition to more sustainable energy sources. Among these solutions, photovoltaic (PV) power systems have emerged as a key player in meeting clean energy targets. The potential of renewable energy sources depends on the country's geographical location and human population [1]. In 2020, electricity generation by solar PV reached approximately 11 581 GWh in Africa [2]. The total capacity of PV technology installed worldwide from 2010 – 2020 has increased from 40 334 to 709 674 MW [3]. According to Statista global data and business intelligence platform and the South Africa Solar Photovoltaic (PV) Market Report from the global data, South Africa leads the continent with a cumulative PV installed capacity of 7021 MW with 4412 MW as industry and residential rooftops as of June 2023, [4]

Grid-connected facilities present an opportunity for energy retrofitting using PV systems. The concept of energy retrofitting involves upgrading existing buildings and infrastructures with energy-efficient

technologies to reduce their environmental footprint and operational costs. By integrating PV power systems into these facilities, the generated electricity offsets a portion of the conventional grid-supplied power. The PV retrofit systems in grid-tied or hybrid forms offer several benefits to the user with reduced electricity bills and, a more reliable supply of electricity and grid load pressure reduction. In its grid-tied mode, the PV system is cheaper with less energy loss by avoiding battery storage and battery inverter. The system is designed to produce more than the local demand and feeds the excess to the grid (virtual storage) for use in the evening when there is no sun.

PVSyst is a widely recognised software tool specifically designed for simulating the performance of PV systems. It allows users to model various PV system design and operation aspects, including solar irradiance, shading, module characteristics, and inverter behaviour. Ghafoor & Munir (2015) used the PVSyst software to design and analyse the economics of an off-grid PV system for household electrification in Faisalabad, Pakistan. Their mathematical modelling results showed that the peak power, area of PV modules, the capacity of the battery backup, and size of the charge controller and inverter were determined to be 1928 Wp and 12.85 m^2, 9640.5 Wh, 56.65 A, and 1020 W, respectively. The economic feasibility analysis of the system was carried out and found that the unit electric energy cost of the off-grid PV system was 26% lower than that of conventional grid electric supply to residential areas [3]. In 2022, Mohammad Baqir and Harpreet Kaur Channi [5] used PVSyst software to design and analyse a 700 kWp in the Daikundi province of Afghanistan. The results showed the system production was 1266 MWh/yr, and the performance ratio was 0.797 [3]. Uwho, Idoniboyeobu, & Amadi (2022) also employed the software for designing and simulating a 500-kW grid-connected PV system with a multilevel inverter that tracks the maximum power from the PV array for the Faculty of Engineering, Rivers State University in Nigeria. Their system was optimised through simulation and achieved an efficiency of 97.7% with a performance ratio of 83.2%. A peak power of 787 kWp with a tilt angle of 5°, [6]. Marais, Kusakana, and Koko (2019) also did a techno-economic analysis using the Homer software for a grid-interactive solar PV system for a South African residential load. Their results revealed that the load demand was adequately met at no shortage and incurred a levelized cost of energy (LCOE) of R0.8546/kWh, South African rand. The software depicted that the allowed load demand that a Solar PV system could meet was 70%, while the grid was used to supply the load when there was no access to sunlight. The system had no storage, so excess energy was sold into the grid during the standard day. [5]

This paper focuses on designing and analysing a 10 kWp PV power system tailored for grid-connected facilities undergoing energy retrofitting in Limpopo province, South Africa. To predict the PV system's performance, it is necessary to design the system by inputting all parameters and environmental factors, such as location, temperature, solar irradiance, configuration, orientation, and inclination.

2. Methodology

2.1. Energy Audit and Load Analysis

The energy consumption of an HQ unit at 15 SAI was investigated since there was no separate electricity data consumption for each building. To have an overview of how much is consumed daily leads to monthly and annual energy consumption. Due to load shedding, retrofitting the PV system was an alternative. Designing a photovoltaic (PV) power system for grid-connected facility energy retrofitting involved several steps to ensure solar energy's efficient and effective integration into the existing infrastructure.

2.1.1. Data Collection

The historical energy bills for an HQ building were unavailable, so a preliminary assessment to understand the building's energy consumption trends over time was conducted. An inventory of

electrical appliances such as lights, HVAC systems, computers, and other energy-consuming elements in the building was done. Collected data were tabulated with their specification, i.e., wattage reading, current and voltage. The daily usage duration of each electrical appliance was recorded to give the full view of the daily consumption and, hence, the monthly consumption. Due to insufficient funds, some devices, such as heating, ventilation, and air conditioning (HVAC) systems, were excluded, and a 10.0 kWp peak power system was considered.

2.1.2. Designing PV system using PVSyst software 7.4.1

The preliminary design of a grid-connected PV system was done as follows: On the PVSyst software, the grid-connected was selected, followed by the geographical site. Since the study site was unavailable on a database, a new site called Manamane was created on a database. Using Meteonorm 8.1, Manamane metadata was downloaded. It comprises annual average parameters such as global horizontal irradiance (kWh/m^2), diffuse horizontal irradiance (kWh/m^2), ambient temperature (degrees), etc. The plane orientation must be optimised for high energy production; thus, the loss should be zero, and it was observed that at $25^{0,}$ the system produced a 0.00 % loss. Pre-sizing of the system consists of planned power (10.0 kWp) and several PV modules; based on the size, ratings, and availability, SunPro power (Si-mono)– 430 Wp, the 2021 model, was chosen. The software helps one to select the correct size of an inverter, so Sungrow with features such as 10 kW and 40 – 560 V was selected. Several parameters and assumptions were considered. The assumptions used are tabulated in table 1: [[7], [8].

Table 1: Parameter Assumptions.

Parameters	Percentage/ ZAR
PV Modules	55 %
Inverters	15 %
Mounting Structures	5 %
Installation	25 %
Fixed Fee-Tariff	R4.04

The feed tariff is R4.04 per kW, while the other parameters' rate was calculated from the total fund needed to build the 10 kWp. Finally, the simulation was run.

3. Results and discussions

Annual energy production is 18 130 kWh, and the system's performance ratio is 77.9%, which is good compared to the one found by Baqir M et al. [3]in Afghanistan. Financial analysis in the form of yearly net profit and cash flow of the system is graphically shown below:

The cost of generating energy using a PV system, the Levelized Cost of Energy (LCOE), is R3.07. As the rate of solar panels keeps decreasing, LCOE will reduce, too. Several parameters, such as investment and charges, financial parameters, and electricity rate, were considered to improve the design of this system. The injection of unused energy generated to the grid increases the annuities. The payback time is 5.6 years.

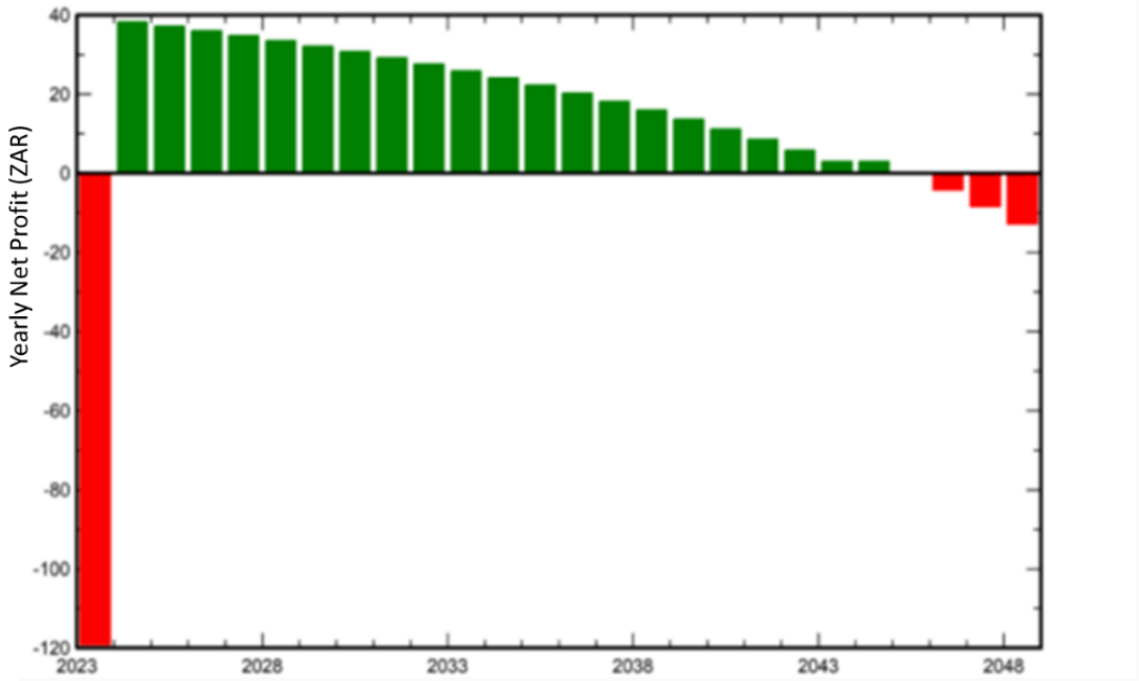


Figure 1: Yearly Net Profit.

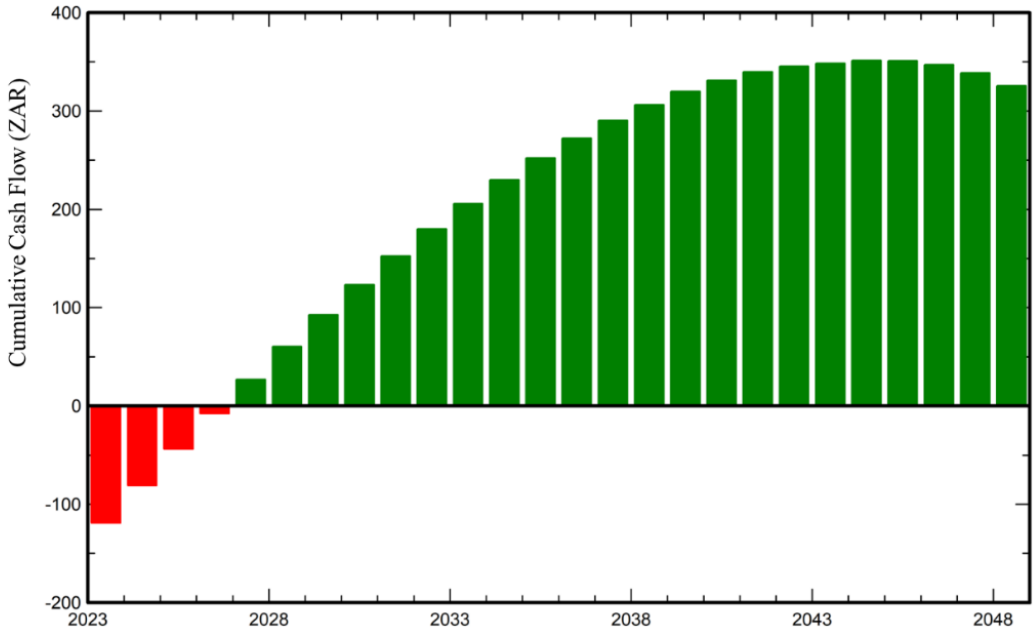
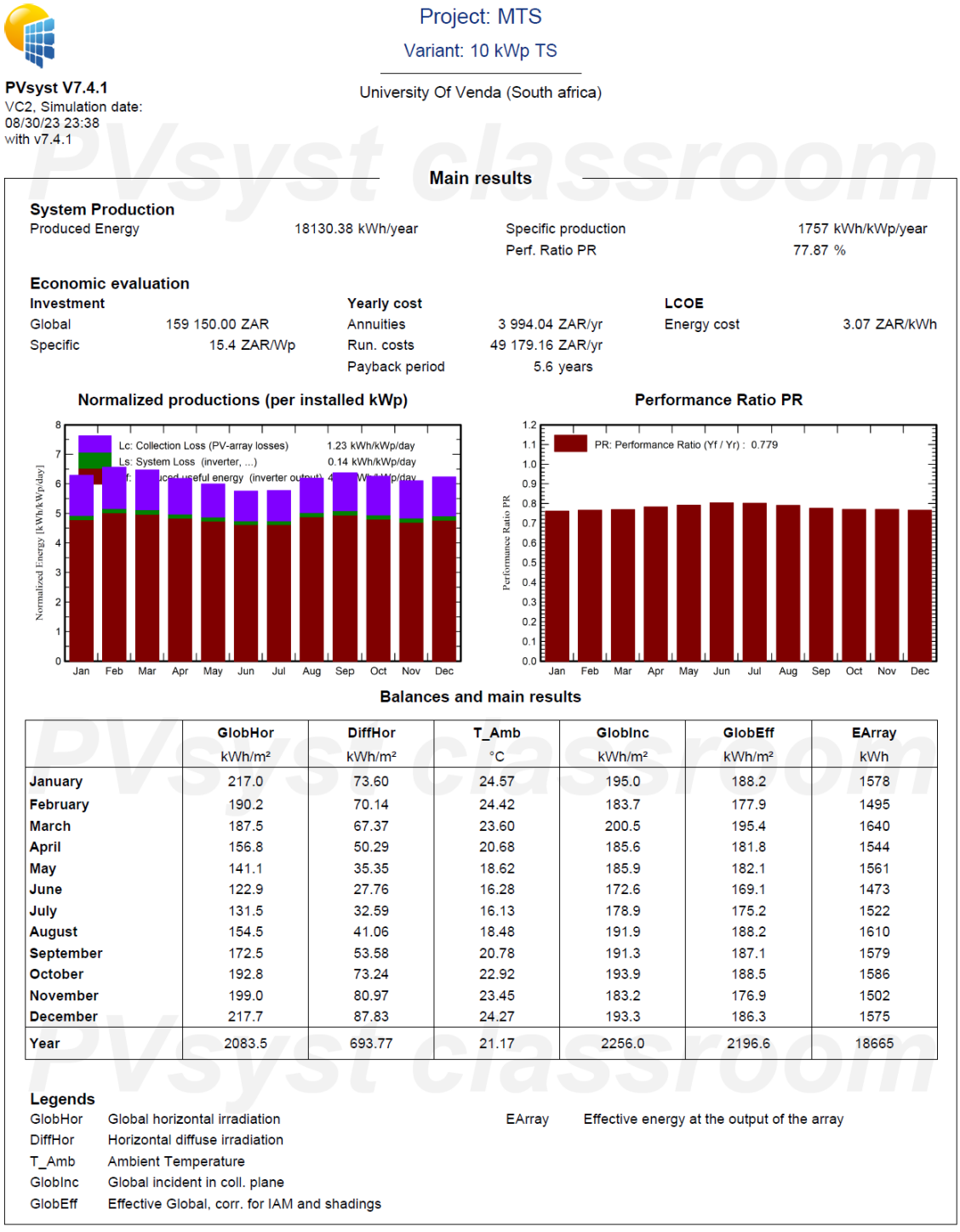


Figure 2: Cumulative Cash Flow.



	GlobHor kWh/m²	DiffHor kWh/m²	T_Amb °C	GlobInc kWh/m²	GlobEff kWh/m²	EArray kWh
January	217.0	73.60	24.57	195.0	188.2	1578
February	190.2	70.14	24.42	183.7	177.9	1495
March	187.5	67.37	23.60	200.5	195.4	1640
April	156.8	50.29	20.68	185.6	181.8	1544
May	141.1	35.35	18.62	185.9	182.1	1561
June	122.9	27.76	16.28	172.6	169.1	1473
July	131.5	32.59	16.13	178.9	175.2	1522
August	154.5	41.06	18.48	191.9	188.2	1610
September	172.5	53.58	20.78	191.3	187.1	1579
October	192.8	73.24	22.92	193.9	188.5	1586
November	199.0	80.97	23.45	183.2	176.9	1502
December	217.7	87.83	24.27	193.3	186.3	1575
Year	2083.5	693.77	21.17	2256.0	2196.6	18665

Figure 3: PV System's Performance Results.

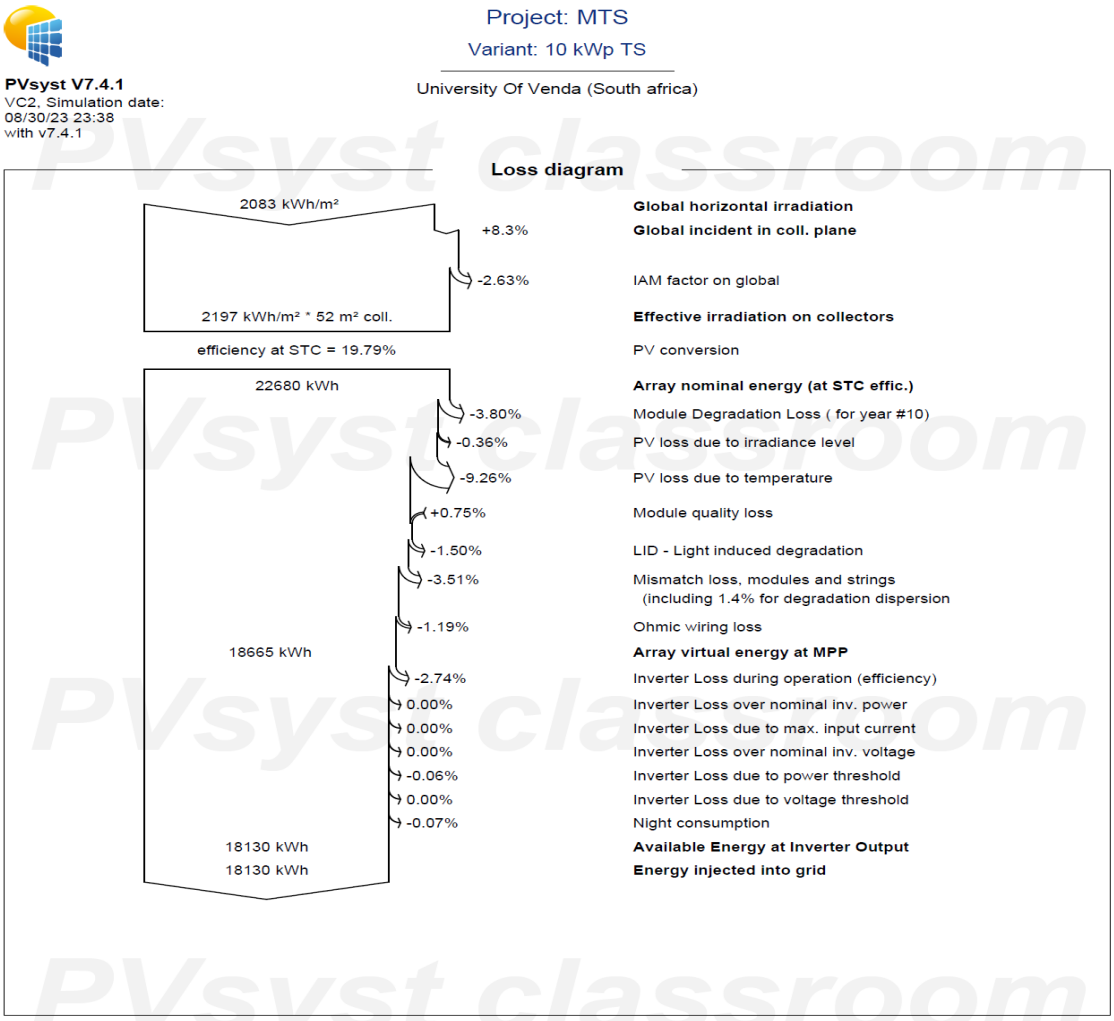


Figure 4: System loss diagram.

Figure 4 displays insight into the causes of the loss of energy. It breaks down the various factors contributing to the overall loss of power and the efficiency reduction of the system. The energy that may be injected into the grid is 18 130 kWh. Noting that energy generation through renewable resources mitigates carbon footprint, is also depicted by the Figure 5.

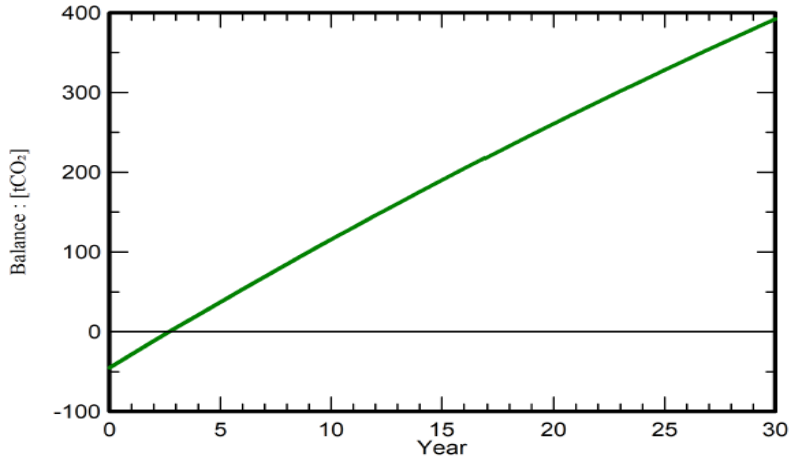


Figure 5: Saves CO_2 Emission vs. Time.

4. Conclusion

The design of a 10 kWp PV system has provided insight into its potential and efficiency. The PV array's tilt and orientation have been set to harness maximum energy. The selected solar panels and inverters are well-suited to the project's requirements, ensuring efficient energy conversion and reliable operation. The loss diagram shown in Figure 4 will help us to take into cognisance the various factors contributing to energy losses in the system. This will enable us to develop strategies for mitigating their impact and optimising the system's overall performance. Financial analysis, the number of solar panels, inverter, and the structure's design had been outlined. A 10 kWp can now be prepared to be installed using the data retrieved from this study.

References

[1] O. M. Akinbami, S. Oke and M. Bodunrin, "The state of renewable energy development in South Africa: An overview.," *Alexandria Engineering Journal,,* vol. 60, no. 6, pp. 5077-5093., 2021.

[2] Government of South Africa, "Government Gazette 42784: Integrated Resource Plan 2019," Government of the Republic of South Africa, Pretoria, South Africa, 2019.

[3] A. O. Maka and J. M. Alabid, "Solar energy technology and its roles in sustainable development.," *Clean Energy,,* vol. 6, no. 3, pp. 476-483., 2022.

[4] GlobalData, "South Africa Solar Photovoltaic (PV) Market Analysis by Size, Installed Capacity, Power Generation, Regulations, Key Players and Forecast to 2035," GlobalData, London, United Kingdom, 2023.

[5] M. Baqir and H. Channi, "Analysis and design of solar PV system using Pvsyst software.," Mohali, India, 2022.

[6] A. Ghafoor and A. Munir, "Design and economics analysis of an off-grid PV system for household electrification," *Renewable and Sustainable Energy Reviews,* vol. 42, pp. 496-502, 2015.

[7] S. Marais, K. Kusakana and S. Koko, "Techno-economic feasibility analysis of a grid-interactive solar PV system for South African residential load," in *2019 International Conference on the Domestic Use of Energy (DUE)*, 2019.

[8] K. Uwho, D. Idoniboyeobu and H. N. Amadi, "Design and Simulation of 500kW Grid Connected PV System for Faculty of Engineering, Rivers State University, Using Pvsyst software," *Iconic Research and Engineering Journals,* vol. 5, no. 8, pp. 2456-8880, 2022.

[9] V. Mlambo, "Living in the Dark: Load Shedding and South Africa's Quest for Inclusive Development," *IAHRW International Journal of Social Sciences Review,* vol. 11, no. 2, pp. 153-160, 2023.

[10] G. Heffner, L. Maurer, A. Sarkar and X. Wang, "Minding the gap: World Bank's assistance to power shortage mitigation in the developing world.," *Energy,* vol. 35, no. 4, pp. 1584-1591, 2020.

[11] B. Adebanji, O. Atoki, T. Fasina, O. Adetan and A. Abe, "Comparative study of off-grid and grid-connected hybrid power system: issues, future prospects, and policy framework," *Indonesian Journal of Electrical Engineering and Computer Science,* vol. 22(2), pp. 752-759, 2021.

[12] S. Asamoah, S. Gyamfi, F. Uba and G. Mensah, "Comparative assessment of a stand-alone and a grid-connected hybrid system for a community water supply system: A case study of Nankese community in the eastern region of Ghana," *Scientific African,* vol. 19, p. e01331, 2022.

[13] GlobalData, "South Africa Solar Photovoltaic (PV) Market Analysis by Size, Installed Capacity, Power Generation, Regulations, Key Players and Forecast to 2035," GlobalData, London, United Kingdom, 2023.

[14] H. Al-Najjar, C. Pfeifer, R. Al-Afif and H. El-Khozondar, "performance Evaluation of a Hybrid Grid-Connected photovoltaic Biogas-Generator Power System," *Energies 2022,,* vol. 15, 2022.

[15] K. Ismaila, A. Sahin, B. Yilbas and A. Al-Sharafi, " Thermo-economic optimization of a hybrid photovoltaic and thermoelectric power generator using overall performance index," *J. Therm. Anal. Calorim. ,* vol. 144, pp. 1815-1825, 2021.

[16] H. Al-Najjar, H. El-Khozondar, C. Pfeifer and R. Al Afif, "Hybrid grid-tie electrification analysis of bio-shared renewable energy systems for domestic application.," *Sustain. Cities Soc.,* vol. 77, 2022.

[17] A. Chand, K. Prasad, K. Mamun, K. Sharma and K. Chand, "Adoption of grid-tie solar system at residential scale," *Clean Technologies,* vol. 1(1), pp. 224-231, 2019.

[18] S. Sen, S. Ganguly, A. Das, J. Sen and S. Dey, "Renewable energy scenario in India: opportunities and challenges," *J Afr. Earth Sci.,* vol. 122, p. 25–31, 2016.

[19] B. Scharfetter and M. Van Dijk, "Legislation governing the implementation of small-scale hydropower projects for rural electrification in South Africa.," *Journal of Energy in Southern Africa,* vol. 28, no. 2, pp. 14-28, 2017.

[20] J. Guo, V. Fast, P. Teri and K. Calvert, "Integrating land-use and renewable energy planning decisions: A technical mapping guide for local government.," *ISPRS International Journal of Geo-Information,* vol. 9, no. 5, p. 324, 2020.

[21] M. Dahmoun, B. Bekkouche, K. Sudhakar, M. Guezgouz, A. Chenafi and A. Chaouch, "Performance evaluation and analysis of grid-tied large scale PV plant in Algeria," *Energy for Sustainable Development,* vol. 61, pp. 181-195, 2021.

[22] M. Adaramola and E. Vågnes, "Preliminary assessment of a small-scale rooftop PV-grid tied in Norwegian climatic conditions," *Energy Conversion and Management,,* vol. 90, pp. 458-465, 2015.

[23] J. Barzola-Monteses and M. Espinoza-Andaluz, "Performance analysis of hybrid solar/H2/battery renewable energy system for residential electrification.," *Energy Procedia,* vol. 158, pp. 9-14, 2019.

[24] V. Tomer and G. Tiwari, "Techno-economic evaluationof grid-connected PV system for households with feed in tariff and time of day tariffs regulation in New Delhi-a sustainable approach," *Renewable Sustainable Energy,* vol. 70, pp. 822-835, 2017.

[25] K. Okedu, R. Uhunmwangho and N. Bassey, "A Comparative Study of on and off Grid Tied Integrated Diesel Solar PV Generation System," *INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGIES,* vol. 1, 2015.

[26] T. Lambert, P. Gilman and P. Lilienthal, "Micropower system modeling with HOMER," *Integration of alternative sources of energy,* vol. 1(1), pp. 379-385, 2006.

[27] R. Dash, L. Behera, B. Mohanty and P. K. Hota, "Cost and sensitivity analysis of a microgrid using HOMER-Pro software in both grid-connected and standalone mode," *ICRIEECE,* pp. . 3444-3449, 2018.

[28] S. Asamoah, S. Gyamfi, F. Uba and G. Mensah, "Comparative assessment of a stand-alone and a grid-connected hybrid system for a community water supply system: A case study of Nankese community in the eastern region of Ghana," *Scientific African,* vol. 19, 2022.

[29] O. Babatunde, J. Munda and Y. Hamam, "Selection of a hybrid renewable energy systems for a low-income household.," *Sustainability, ,* vol. 11, no. 16, p. 4282., 2019.

DIVISION

G

THEORETICAL AND COMPUTATIONAL PHYSICS

Weak Gravity Conjecture for dilaton de Sitter black holes in extra dimension

H Noshad[1], A S Cornell[1], A Chrysostomou[1,2], A Deandrea[2]

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

[2] Institut de Physique des Deux Infinis de Lyon. Université de Lyon, UCBL, UMR 5822, CNRS/IN2P3. 4 rue Enrico Fermi, 69622 Villeurbanne Cedex, France.

E-mail: hnoshad@uj.ac.ir, acornell@uj.ac.za, chrysostomou@ipnl.in2p3.fr, deandrea@ipnl.in2p3.fr

Abstract. In this proceeding, we consider charged black hole solutions of the Einstein-Maxwell-dilaton theory in de Sitter spacetime. Our investigation focuses on examining the location and existence of horizons as a function of the parameters for mass, charge, and dilaton coupling strength (α) in the context of extra-dimensional spacetimes, and investigating possible modifications of the Weak Gravity Conjecture in extra-dimensional spacetimes. In 5-dimensions, for a special value of $\alpha_c^2 = 1$, we observe that there is a new extremality condition, or a new WGC bound. Moreover, we find that, because of the complexity of the metric for $\alpha < \alpha_c$, the extremal case is not achieved for de Sitter spacetime.

1. Introduction

In the ever-expanding landscape of theoretical physics, the swampland criteria has emerged as a crucial guideline for exploring the boundaries of consistent quantum gravity theories [1, 2]. Among these, the weak gravity conjecture (WGC), first proposed in [3], has been one of the earliest and most extensively studied criteria. In a theory coupled to gravity with a U(1) gauge symmetry, the WGC asserts that a state of charge q and mass m must satisfy the condition:

$$g^2 q^2 > \frac{1}{2}\kappa^2 m^2, \tag{1}$$

where g is the U(1) gauge coupling and κ is related to the Newton's constant G and the reduced Planck mass (M_P) through $\kappa^2 = \frac{1}{M_P^2} = 8\pi G$.

The WGC has profound implications, particularly regarding extremal black holes and their decay processes. It ensures that black holes can decay without leaving remnants, aligning elegantly with fundamental principles of charge and energy conservation. When contemplating a black hole in an asymptotically flat spacetime, the compelling arguments presented in [3] provide valuable insights into the dynamics of black hole decays. This process results in the emergence of states with combined masses smaller than the original black hole, driving our understanding of black hole evolution and its consequences.

Notably, the reach of the WGC expands beyond flat spacetimes, encompassing the Einstein-Maxwell-dilaton theory in flat spacetime [4] and even (Anti-)de Sitte spacetime [6, 7]. The

interplay between these theories opens doors to captivating scenarios, including an extension of the Reissner-Nordström de Sitter black hole solution [9, 10] to the Einstein-Maxwell-dilaton theory [5, 8, 11, 12, 13, 14, 15].

In the quest to unite these theories, our work embarks on a profound endeavor: fusing the three elements - Einstein-Maxwell-dilaton theory, de Sitter backgrounds in extra dimensions, and the modifications to the WGC (dS-WGC) in this context. We seek to ascertain the specific adjustments required for equation (1) to hold within the framework of de Sitter backgrounds in extra dimensions. To derive the form of the dS-WGC, we delve into the intricacies of black hole physics, leveraging the same fundamental arguments that led to equation (1) in a flat spacetime [3]. Our investigation reveals that black hole decays must lead to states incapable of forming new black holes so as to avoid the existence of naked singularities, consistent with the principles of the weak cosmic censorship.

As such, the paper is structured as follows: In section 2 we present our black hole solution, along with some generic formulae applicable to all values of α and general dimensions. In section 3 we discuss the horizons of the charged dilatonic de Sitter black hole in five dimensions for three cases: $\alpha = \alpha_c$, $\alpha > \alpha_c$, and $\alpha < \alpha_c$, and we explore the WGC for these cases. Lastly, we summarize our findings in section 4.

2. Einstein-Maxwell-Dilaton Black Holes in n Dimensions and Arbitrary α

Dilatonic gauge theories usually result from the compactification of higher dimensional theories of gravity [16, 11, 17]. We start from the action of an Einstein-Maxwell-scalar theory in n dimensions,

$$S = \int d^n x\sqrt{-g}[R - 2\partial_\mu\phi\partial^\mu\phi - e^{-2\alpha\phi}F^2], \tag{2}$$

where R is the Ricci scalar curvature, $F^2 = F_{\mu\nu}F^{\mu\nu}$ arises from the Maxwell field, and α represents the coupling between the dilaton field and Maxwell field. It is important to observe that when the dilaton field is absent (i.e., $\alpha = 0$), the action described by equation (2) simplifies to the conventional Einstein-Maxwell theory accompanied by a cosmological constant.

The black hole metric solution of the equations of motion is given by:

$$\begin{aligned} ds^2 &= -\left\{\left[1-\left(\frac{r_+}{r}\right)^{n-3}\right]\left[1-\left(\frac{r_-}{r}\right)^{n-3}\right]^{1-\gamma(n-3)} - H^2r^2\left[1-\left(\frac{r_-}{r}\right)^{n-3}\right]^{\gamma}\right\}dt^2 + \\ &\left\{\left[1-\left(\frac{r_+}{r}\right)^{n-3}\right]\left[1-\left(\frac{r_-}{r}\right)^{n-3}\right]^{1-\gamma(n-3)} - H^2r^2\left[1-\left(\frac{r_-}{r}\right)^{n-3}\right]^{\gamma}\right\}^{-1} \\ &\left[1-\left(\frac{r_-}{r}\right)^{n-3}\right]^{-\gamma(n-4)} dr^2 + r^2\left[1-\left(\frac{r_-}{r}\right)^{n-3}\right]^{\gamma} d\Omega^2_{n-2}, \end{aligned} \tag{3}$$

where $H^2 = |\Lambda|/3$ is Hubble parameter and Λ is the cosmological constant. Recall that when $\Lambda = 0$ our space is asymptotically flat, otherwise $\Lambda > 0$ is for dS and $\Lambda < 0$ AdS. In this work we consider only dS spacetime.

Also $\gamma = 2\alpha^2/(n-3)(n-3+\alpha^2)$ and ϕ_0 corresponds to the asymptotic value of $\phi(r)$ as r approaches infinity. The mass and charge of the black hole, in relation to the integration constants r_+, r_-, are described as:

$$\begin{aligned} Q^2e^{2\alpha\phi_0} &= (4\Omega_{n-2})^2\frac{(n-2)(n-3)^2(r_+r_-)^{n-3}}{2(n-3+\alpha^2)}, \\ M &= \frac{\Omega_{n-2}}{16\pi}\left((n-2)r_+^{n-3} + \frac{n-2-p(n-4)}{p+1}r_-^{n-3}\right). \end{aligned} \tag{4}$$

where $\Omega_{n-2} = 2\pi^{\frac{D-1}{2}}/\Gamma\left(\frac{n-1}{2}\right)$ is the volume of the unit $(D-2)$ sphere and the constant p is given by:

$$p = \frac{(2-n)\gamma}{(n-2)\gamma - 2}. \tag{5}$$

Let us express the $0,0$ component of the metric in an alternative form:

$$P(r) = 1 - \left(\frac{r_+}{r}\right)^{n-3} - H^2 r^2 \left(1 - \left(\frac{r_-}{r}\right)^{n-3}\right)^{\alpha_c}, \tag{6}$$

where $\alpha_c = \alpha^2 - (n-3)^2/(n-1)$. It is worth noting that in the case of $\alpha = 0$ we retrieve the Reissner-Nordström solution, where r_+ , r_- correspond to the outer and event horizons, respectively. However, for $\alpha \neq 0$, r_- indicates the position of a singular surface, while r_+ remains as the only event horizon of the black hole.

In order to explore the solutions of $P(r) = 0$, we break $P(r)$ into the sum of two contributions: $W_1 = 1 - (\frac{r_+}{r})^{n-3}$ and $W_2 = H^2 r^2 \left(1 - (\frac{r_-}{r})^{n-3}\right)^{\alpha_c}$. The zeros of $P(r)$ correspond to the intersection points of the two curves defined by W_1 and W_2. Furthermore, it is noteworthy that the situation where the zeros coalesce into one occurs when W_1 and W_2 are tangent, leading to the conditions $P(Z) = 0$ and $P'(Z) = 0$. In fact, we will find two expressions, $Z_{1,2}$ indicate the extremal limits of the black hole. This means that they signify the coincidence of the cosmological horizon and the event horizon, or the singularity surface and the event horizon.

Subsequently, throughout the remainder of this proceeding, we shall investigate the various regions of the WGC in the context of 5 dimensions.

3. The horizons of charged dilatonic de Sitter black hole in 5 dimensions

In 5 dimensions, the values of r_+ and r_- are expressed as follows:

$$r_- = \sqrt{\frac{2}{\pi}} \sqrt{\frac{Q^2(\alpha^2+2)}{4M^2 + \sqrt{6Q^2\alpha^2 + 16M^2 - 12Q^2}}}, \tag{7}$$

$$r_+ = \frac{1}{\sqrt{3\pi}} \sqrt{4M^2 + \sqrt{6Q^2\alpha^2 + 16M^2 - 12Q^2}}. \tag{8}$$

Subsequently, we need to ensure that the values of $r_\pm$ are real, which leads us to the condition:

$$6Q^2\alpha^2 + 16M^2 - 12Q^2 \geq 0. \tag{9}$$

If this condition holds, the metric remains real and physically meaningful. Next, we will consider various values of α and explore the regions of the WGC corresponding to each of them. By analysing different scenarios for α, we aim to gain insights into how the WGC behaves and how it is influenced by the choice of this parameter.

3.1. $\alpha = \alpha_c$

When $\alpha = 1$, we can determine the explicit expressions for the black hole horizons as $P(r) = (H^2 r^4 - r^2 + r_+^2)$, and they can be written as:

$$r_{c,h} = \sqrt{2} \frac{\sqrt{1 \pm \sqrt{1 - 4H^2 r_+^2}}}{2H} \tag{10}$$

where r_c represents the cosmological horizon and r_h corresponds to the event horizon of the black hole. Notably, we have two additional equations that are negative, which is why they have not

been included here. By employing equation (4), we can represent the roots (10) as expressions that depend on the mass and charge of the black holes, denoted as M and Q respectively.

The graphical representation of the various regions for the black hole horizons can be observed in figure 1. In this figure, the red curve represents $r_c - r_h = 0$, and the green curve represents $r_h - r_- = 0$. Moreover, the blue line shows the condition for having an extremal black hole in flat space, which is $r_+ = r_-$ ($M^2 = Q^2/2$). The black line represents the condition given by equation (9), below which the metric becomes complex.

Located at $(M, Q) = \left(3\pi/16, \sqrt{3}\pi/4\sqrt{2}\right)$, the triple point marks the convergence of regions with two, one, and zero solutions. When masses exceed $M > 3\pi/16$, the event horizon fails to form, signifying the threshold beyond which asymptotically de Sitter black hole solutions become unattainable. In the region between $M = 3\pi/16$ and the red curve, no horizons exist. Both the event and cosmological horizons are imaginary, leaving only a naked singularity in this interval. In the region bounded by the red and green curves, two horizons exist - the event horizon and the cosmological horizon. The green line corresponds to the extremal black hole, where the event horizon and singular surface coincide. In the region beneath this line, a single horizon is present, namely the cosmological horizon, while the event horizon is obscured by the singular surface. Hence, the region complying with the WGC is delimited between the green and black lines. At this particular value of α, we notice the emergence of a novel bound for the WGC, distinct from the WGC bound in flat space (blue line) and $r_+ = r_-$ no longer serves as the black hole extremality condition. In order to validate that equation (10) can be regarded as

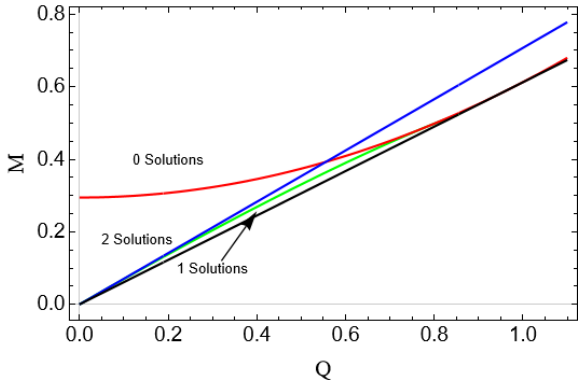


Figure 1. Horizon number for five dimensional $\alpha = 1$ de Sitter Black Hole in (M, Q) for $H = 1$. The red curve represents $r_c - r_-$, the green curve illustrates $r_h - r_-$, the blue line depicts the extremality condition $r_+ = r_-$ ($M^2 = \frac{1}{2}Q^2$) in flat space, and the black line corresponds to equation (9).

a WGC bound, we derive an expansion of $r_h - r_-$ in the flat space limit, $H \to 0$, as follows:

$$M^2 = \frac{Q^2}{2} + \frac{\sqrt{2}Q^3H^2}{4\pi} - \frac{3Q^4H^4}{16\pi^2} + \frac{3\sqrt{2}Q^5H^6}{16\pi^3} + \mathcal{O}(H^8), \tag{11}$$

where the first term corresponds to the WGC in flat space for five dimensions (equivalent to the equation for the blue line), while the subsequent terms represent corrections to this bound.

3.2. $\alpha > \alpha_c$ and $\alpha < \alpha_c$

Our focus now shifts to the scenario $\alpha > 1$. Figure 2 show the findings for $\alpha = 2$, illustrating the distinct regions corresponding to the black hole horizons. In this scenario, an additional

constraint has been introduced to ensure that the singularity remains within the Hubble horizon, depicted by the orange curve. This constraint requires $r_- < 1/H$, as black hole solutions cannot exist when the radius of the singular surface exceeds that of the Hubble horizon. Similar to the $\alpha = 1$ case, the red curve indicates the point where the cosmological and event horizons coincide, while the blue line represents the extremality condition for the black hole in flat space, where $r_+ = r_-$. Within the region between the red curve and the blue line, black hole solutions with two horizons are present. Below the blue curve, solutions entail a naked singularity along with a cosmological horizon. Interestingly, for this specific value of α, the extremality condition for the black hole ($r_h = r_-$) coincides with the flat space case. Consequently, for $\alpha > \alpha_c$, no new WGC bound emerges. Additionally, for comparison, we have examined the case when $\alpha < \alpha_c$

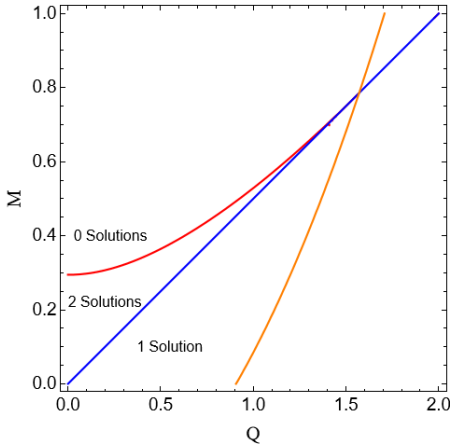


Figure 2. Number of horizons for $\alpha = 2$ as a function of M and Q in five dimensions, for $H = 1$.

in figure 3. In this scenario the upper bound is the red curve where the event and cosmological horizons coincide. The lower bound is indicated by the black line, where the metric is at the brink of becoming complex. In this scenario, the condition for an extremal black hole, $r_h = r_-$, becomes complex and appears to be concealed within the inaccessible region (the complex region for the metric). Consequently, for the anticipated solutions with only the cosmological horizon, there are no clear predictions available.

4. Conclusions

In this proceeding, we explored the properties and behaviour of dilaton black holes in extra dimensions. The action of the Einstein-Maxwell-scalar theory in n dimensions provided a foundation for our investigation, considering the coupling between the dilaton field and Maxwell field as being denoted by α.

The black hole metric solutions were derived, and the mass and charge of the black hole were expressed in terms of integration constants r_+ and r_-. We investigated the various regions of the WGC in five dimensions spacetime, revealing intriguing insights for different values of α. For $\alpha = 1$, we observed the emergence of a novel WGC bound, distinct from the flat space case, and the extremality condition $r_+ = r_-$ no longer held. For $\alpha > \alpha_c$, no new WGC bound was found, and the extremality of the black hole was found to be consistent with flat space. In the case of $\alpha < \alpha_c$, the reality of the metric was examined, and predictions for solutions with only the cosmological horizon were not evident. Our comprehensive analysis enhances our understanding of dilatonic black holes and their connection to the WGC in five dimensions.

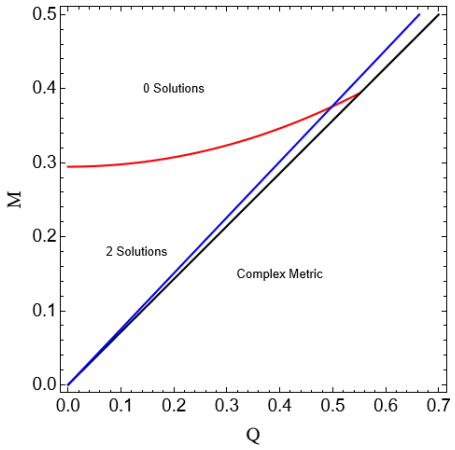


Figure 3. Number of horizons for $\alpha = 0.5$ in (M, Q) coordinate in five dimensions, for $H = 1$. The red curve shows $r_c - r_-$, the blue line is the extremal black hole in flat space $r_+ = r_-$ and the black line shows the bound of a real- valued metric, equation (9).

Acknowledgments

HN is supported by the Faculty of Science at the University of Johannesburg. ASC is supported in part by the National Research Foundation (NRF) of South Africa; AC is supported by the NRF and Department of Science and Innovation through the SA-CERN programme and a Campus France scholarship. The authors extend their thanks to Seong Chan Park for fruitful discussions.

References

[1] Vafa C 2005 The String landscape and the swampland *Preprint* hep-th/0509212.
[2] Ooguri H, Vafa C 2007 *Nuclear physics* B **064007** 21-33.
[3] Arkani-Hamed N, Motl L 2007 *Journal of High Energy Physics* **2007** 060.
[4] Heidenreich B, Reece M 2016 *Journal of High Energy Physics* **2016** 1-41.
[5] Gibbons GW, Maeda K 1988 *Nuclear Physics B* **298** 741-775.
[6] Antoniadis I, Benakli K 2020 *Fortschritte der Physik* **68** 2000054.
[7] Zhang CY, Liu P, Liu Y, Niu C, Wang B 2022 *Physical Review D* **105** 024010.
[8] Poletti SJ, Twamley J, Wiltshire DL 1995 *Physical Review D* **51** 5720.
[9] Romans LJ 1992 *Nuclear Physics B* **383** 395–415.
[10] Ginsparg P, Perry M J 1983 *Nuclear Physics B* **222** 245–268.
[11] Gao Ch J, Zhang Sh N 2005 *Physics Letters B* **605** 185–189.
[12] Hendi SH, Sheykhi A, Panahiyan S, Panah BE 2015 *Physical Review D* **92** 064028.
[13] Yu MH, Ge XH 2022 *The European Physical Journal C* **82** 14.
[14] Okounkova M 2019 *Physical Review D* **100** 124054.
[15] Pani P, Cardoso V 2009 *Physical review D* **79** 084031.
[16] Horowitz G T , Strominger A 1991 *Nuclear Physics B* **360** 197–209.
[17] Duff MJ and Lu JX 1994 *Nuclear Physics B* **416** 301–334.

Quasinormal modes calculated with physics-informed neural networks

A M Ncube[1] and A S Cornell[2]

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

E-mail: [1]ncubeanele4@gmail.com,[2]acornell@uj.ac.za

Abstract. The literature on the computation of black hole quasinormal modes (QNMs) is replete with the use of various approximation methods to solve the "quasi-Sturm Liouville" type problems governing the damped oscillations produced by perturbed black holes. Among the newest techniques is the physics-informed neural network (PINN) algorithm, a machine learning-based, general-purpose differential equation solver that has recently been implemented successfully to compute the QNMs of Kerr black holes perturbed by gravitational fields (of spin-weight, $s = -2$). As a path towards computation of QNM overtones, which are known to be significant in the black hole perturbation theory description of observed gravitational wave signals, QNMs of asymptotically flat Kerr black holes are obtained for the modes $n = 0, \ell = 2, m = 0$, considering spin parameters $a \in [0.0, 0.1, 0.2, 0.49, 0.4999]$. Compared with the QNMs approximations obtained using the continued fraction method, a basic set-up of PINNs is seen to approximate QNMs with increasing accuracy (error $< 1\%$ at best) with decreasing BH spin.

1. Background

The quasinormal modes (QNMs) of asymptotically flat Kerr spacetimes are known to be pertinent in the analysis of gravitational wave signals produced in the aftermath of the coalescence of black holes (BHs), such as in the case of the well-known stellar-mass binary system linked to GW150914 [1]. In general, the level of significance of gravitational wave astrophysics, and by extension the study of QNMs, has broadened with the recent confirmation of nano-hertz gravitational wave signals produced by supermassive BHs [2]. Therefore, the research on BH perturbation theory, which has hitherto been computational and centred on predictions from general relativity (GR), is now a theory that is borne out by analyses [3, 4] showing its precision in describing observed signals.

Within BH perturbation theory, the metric tensor of a BH perturbed by gravitational fields, or direct metric perturbations, is given as (considering a linear approximation) [5]:

$$g_{\mu\nu} = g^0_{\mu\nu} + \delta g_{\mu\nu}, \tag{1}$$

where $g^0_{\mu\nu}$ signifies the background metric subjected to linear perturbations ($\delta g_{\mu\nu}$). In this case, the perturbing gravitational field does not backreact on the background implying that $\delta g^2_{\mu\nu}$ and higher order terms are negligible [5]. Of interest in the astrophysical case is the perturbation of

a Kerr metric, which (in Boyer-Lindquist co-ordinates) is given as [6]:

$$ds^2 = \left(1 - \frac{2Mr}{\Sigma}\right) dt^2 + \frac{4Mar\sin^2\theta}{\Sigma} dt d\phi - \frac{\Sigma}{\Delta} dr^2 - \Sigma d\theta^2 - \frac{\Lambda}{\Sigma}\sin^2\theta d\phi^2, \tag{2}$$

where

$$\Delta = r^2 + a^2 - 2Mr, \quad \Sigma = r^2 + a^2\cos^2\theta, \quad \Lambda = (r^2+a^2)^2 - \Delta a^2 \sin^2\theta. \tag{3}$$

The exterior of the BH is considered to be the space beyond the outer horizon ($r > r_+$), where r_+ is the larger of the roots of Δ. Considering the no-hair conjecture, the Kerr metric is fully described by its mass M and spin angular momentum $J = 2Ma$, within the system of units $G = c = 2M = 1$ (G is Newton's gravitational constant and c is the speed of light in vacuum). In this case, the spin parameter a lies within the range $0 \leq a < \frac{1}{2}$; that is, within the Schwarzschild limit $a = 0$ and the extremal limit of spin $a = \frac{1}{2}$.

The derivation of the Teukolsky equation governing the perturbations of a Kerr BH to produce QNMs (ψ) can be found in Ref. [7]. When the QNMs are in the form given as [6, 7]:

$$\psi = e^{-i\omega t + im\phi}{}_sR_{\ell m\omega}(r){}_sS_{\ell m\omega}(\theta), \tag{4}$$

the Teukolsky equation is separated into radial and angular eigenvalue problems solved by radial and angular eigenmodes $R(r)$ and $S(\theta)$, respectively. They are [6, 7]:

$$\left[\Delta^{-s}\frac{d}{dr}\left(\Delta^{s+1}\frac{d}{dr}\right) + \left(\frac{H^2 - 2is(r-m)H}{\Delta} + 4is\omega r + 2am\omega - A + s(s+1)\right)\right] R(r) = 0, \tag{5}$$

$$\left[\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d}{d\theta}\right) + \left(A - \frac{m^2 + s^2 + 2ms\cos\theta}{\sin^2\theta} - a^2\omega^2\sin^2\theta - 2a\omega s\cos\theta\right)\right] S(\theta) = 0, \tag{6}$$

where $H = (r^2+a^2)\omega - am$, s is the spin-weight of the perturbing field ($s = -2$ for gravitational fields), A is a separation constant and ω signifies the centrepiece QNM frequencies. The astrophysically relevant boundary conditions ensuring that QNMs are ingoing at the event horizon (r_+) and outgoing at spatial infinity are:

$$R(r_* \to -\infty) \sim \frac{e^{ikr_*}}{\Delta^s}; \quad R(r_* \to \infty) \sim \frac{e^{i\omega r_*}}{r^{2s+1}}, \tag{7}$$

and r_* denotes the tortoise co-ordinate:

$$r_* = r_+ \frac{r_+^2 + a^2}{r_+ - r_-}\ln\left(\frac{r - r_+}{r_+}\right) - \frac{r_-^2 + a^2}{r_+ - r_-}\ln\left(\frac{r - r_-}{r_+}\right), \tag{8}$$

where $r_\pm = (1 \pm \sqrt{1 - 4a^2})/2$. These boundary conditions render the differential equations (5) and (6) non-hermitian, such that ω is complex-valued. By contrast, in the special case where $\omega \in \mathbb{R}$ the equations reduce to Sturm-Liouville eigenvalue problems [6].

The solving of the equations (5) and (6) to obtain ω and A is a longstanding challenge within BH perturbation theory that is underpinned by a lack of analytical, closed form solutions. Therefore, it is generally the case that approximation methods are necessary, and in the Kerr case, these include the continued fraction method (CFM) and the WKB method [8, 9]. Considering that many of the extant approximation methods do not extend to all scenarios, the Kerr-Newman BH being the clearest example, the motivation remains for the continued development of novel, semi-analytical or purely numerical techniques. Among the newest of these is a method utilising deep neural networks (DNNs) to solve differential equations; namely, physics-informed neural networks (PINNs). This approach has been shown to achieve the same

level of precision in the computation of the fundamental QNMs of Schwarzschild and Kerr BHs as more established techniques (see Refs. [10, 12]), but it still largely remains to be refined and fully developed to achieve deterministic computation of QNM overtones.

The following sections will elaborate on the use of PINNs to compute the complex-valued eigenvalues ω_n and $A_{\ell m}$. Section 2 describes the implementation of the PINNs algorithm to solve the radial and angular ODEs (equations (5) and (6)). Thereafter, section 3 presents the approximated QNMs and their errors compared to the standard values obtained using the CFM. Lastly, section 4 presents a discussion and outlook.

2. Implementation of PINNs to compute Kerr QNMs

PINNs (the use of DNNs as numerical solvers of partial differential equations (PDEs)) were put forth by and subsequently further developed as described in Refs. [13, 14, 15]. The universal approximation theory that applies to neural networks overall, enables PINNs to operate as general-purpose approximators of PDEs, including the non-hermitian eigenvalue problems found in BH perturbation theory. A key feature of PINNs is the loss function, which for eigenvalue problems takes the form:

$$\mathcal{L} = w_{PDE}\mathcal{L}_{PDE} + w_{reg}\mathcal{L}_{reg}, \tag{9}$$

where $\mathcal{L}$ is the total loss function and $\mathcal{L}_{PDE}$ denotes the PDE "residual", i.e. the mean of the L_2 norm of the PDE evaluated at selected points within the computational domain. The other loss term $\mathcal{L}_{reg}$ represents additional regularisations, while w_{PDE} and w_{reg} represent weights. In general, $\mathcal{L}_{PDE}$ for a d-dimensional, homogeneous second-order eigenvalue problem is given as:

$$\mathcal{L}_{PDE}(\theta) = \frac{1}{N_{points}} \sum_{\mathbf{x}\in points}^{N_{points}} \left\| f(\mathbf{x}; \frac{\partial \hat{u}}{\partial x_1}, \ldots, \frac{\partial \hat{u}}{\partial x_d}; \frac{\partial^2 \hat{u}}{\partial x_1 \partial x_1}, \ldots, \frac{\partial^2 \hat{u}}{\partial x_d \partial x_d}; \hat{\lambda}) \right\|_2^2. \tag{10}$$

In this case, the PINN approximations of the eigenfunction and eigenvalue are $\hat{u} = \hat{u}(\theta)$ and $\hat{\lambda} = \hat{\lambda}(\theta)$ (respectively), with the DNN parameters (i.e. tunable weights and biases) as $\theta = \{\mathbf{W}^\ell, \mathbf{b}^\ell\}_{1\leq\ell\leq L}$. The index ℓ specifies the hidden layer, $\mathbf{W}^\ell$ denotes a weight matrix and $\mathbf{b}^\ell$ is a bias vector. Equation (10) shows that the $\mathcal{L}_{PDE}$ is obtained by evaluating the PDE at N_{points} selected from the d-dimensional spatiotemporal domain.

For the construction of a DNN to act as the eigenfunctions satisfying the PDE and $\mathcal{L}_{PDE}$, successive linear transformations are applied beginning from the input vector $\mathbf{x} \in \mathbb{R}^d$ (of coordinate points) and terminating at the DNN output representing the eigenfunction $\hat{u}$. Overall, the series of mappings that build the DNN is given, generally, by the recurrence relation:

$$\mathcal{N}^\ell = \sigma(\mathbf{W}^\ell \mathcal{N}^{\ell-1} + \mathbf{b}^\ell) \in \mathbb{R}^{N_\ell}, \quad \text{with } \mathcal{N}^\ell = \mathcal{N}^\ell(\theta, \mathbf{x}) \quad 1 \leq \ell \leq L, \tag{11}$$

where σ is some component-wise nonlinear activation function which is typically not present in the output layer $\ell = L$, where $\mathcal{N}^L(\theta, \mathbf{x}) = \hat{u}(\theta, \mathbf{x})$. Essentially, through the minimisation of equations (9) and (10) by subjecting the DNN to training loops within which θ is updated, the DNN finds approximations of the eigenpairs $\hat{u}$ and $\hat{\lambda}$ satisfying the eigenvalue problem.

In solving the Teukolsky equation (equations (5) and (6)), it is necessary to carry out a change in variables (to obtain a finite domain) and, in addition, incorporate the boundary conditions within ansatzes of the eigenfunctions. For the former, we set $x = r_+/r$ and $u = \cos\theta$ such that $x \in [0, 1]$ and $u \in [-1, 1]$, respectively [12]. The ansatzes, derived by Ref. [8] are:

$$R(r) = e^{i\omega r}(r - r_-)^{-1-s+i\omega+i\sigma_+}(r - r_+)^{-s-i\sigma_+} f(x), \quad \sigma_+ = \frac{\omega r_+ - am}{\sqrt{1 - 4a^2}}, \tag{12}$$

$$S(u) = e^{a\omega u}(1 + u)^{|m-s|/2}(1 - u)^{|m+s|/2} g(u), \tag{13}$$

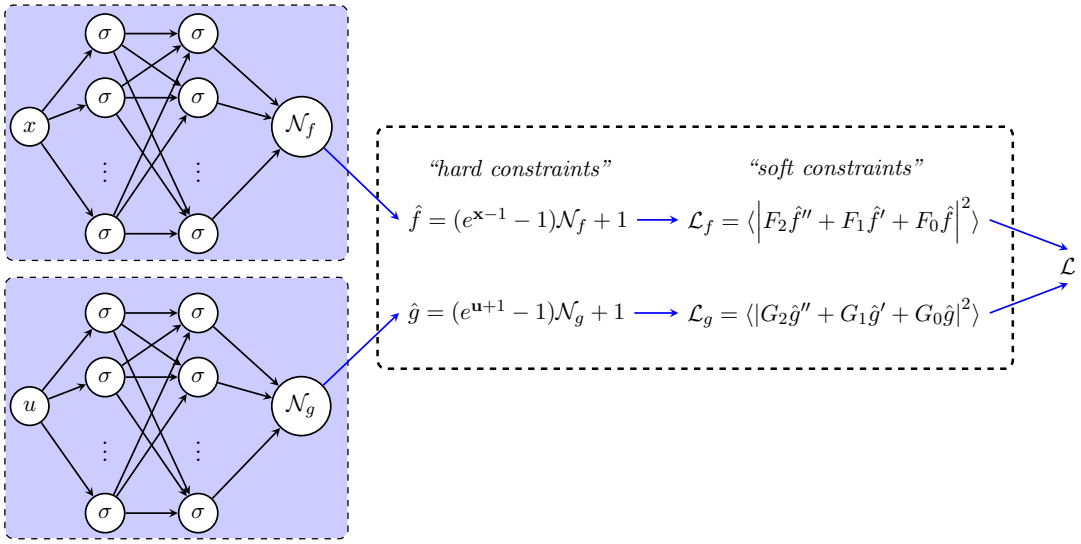


Figure 1. A schematic of the PINN set-up with "hard" and "soft" constraints imposed on the PINN approximations of the radial and angular eigenmodes: $\mathcal{N}_f$ and $\mathcal{N}_g$, respectively.

From these considerations, equations (5) and (6) are, as derived by Ref. [12]:

$$F_2 f'' + F_1 f' + F_0 f = 0, \quad G_2 g'' + G_1 g' + G_0 g = 0, \tag{14}$$

with the coefficients $F_i = F_{0,1,2}$ and $G_i = G_{0,1,2}$ as functions of x and u, respectively. As depicted in figure 1, these are the physics constraints applied to the neural networks approximating f and g. Additionally, the normalisation conditions associated with f and g are imposed, as "hard constraints" (i.e. satisfied exactly by DNNs), on the outputs of the neural networks, $\mathcal{N}_f(\mathbf{x})$ and $\mathcal{N}_g(\mathbf{u})$, which are the approximations of the radial and angular eigenfunctions, respectively [12]:

$$f(\mathbf{x}) = (e^{\mathbf{x}-1} - 1)\mathcal{N}_f(\mathbf{x}) + 1, \quad g(\mathbf{u}) = (e^{\mathbf{u}+1} - 1)\mathcal{N}_g(\mathbf{u}) + 1. \tag{15}$$

Implementation of this set-up of PINNs is done using the functions provided by PyTorch (a Python deep learning library) that has a base class `torch.nn.Module` key to the construction of DNNs. The approximations of the centrepiece eigenvalues $\{\omega_n, A_{\ell m}\}$ and various DNNs set-ups that were tested are given in the following section.

3. QNM frequencies and separation constants for $s = -2, n = 0, \ell = 2, m = 0$.
As is well-known, the setting up of DNN function approximators is not an exact science and involves labouring in "heuristic design"; as such, we set up an arbitrary structure similar to that which we employed in Ref. [10]. That is, for each of the DNNs representing the eigenfunctions of the Teukolsky equation, there are two hidden layers with 50 neurons in each layer. In addition, we used a self-scalable hyperbolic tan activation function: $(1+\beta\mathbf{x})\tanh\mathbf{x}$, and the Adam algorithm [11] in each optimisation step over 5×10^4 training epochs.

Regarding the number of training points, 100 points, where the differential equations are evaluated, are randomly selected from the spatial domains $x \in [0, 1]$ and $u \in [-1, 1]$. In general, the ability to generalise (i.e. to interpolate accurately) is a desirable quality in neural networks which is optimised by choosing representative points to train the neural networks. To obtain representative training sets from the spatial domains of the Teukolsky equation, where the functions $f(x)$ and $g(u)$ are bounded, random sampling of points is one valid approach to

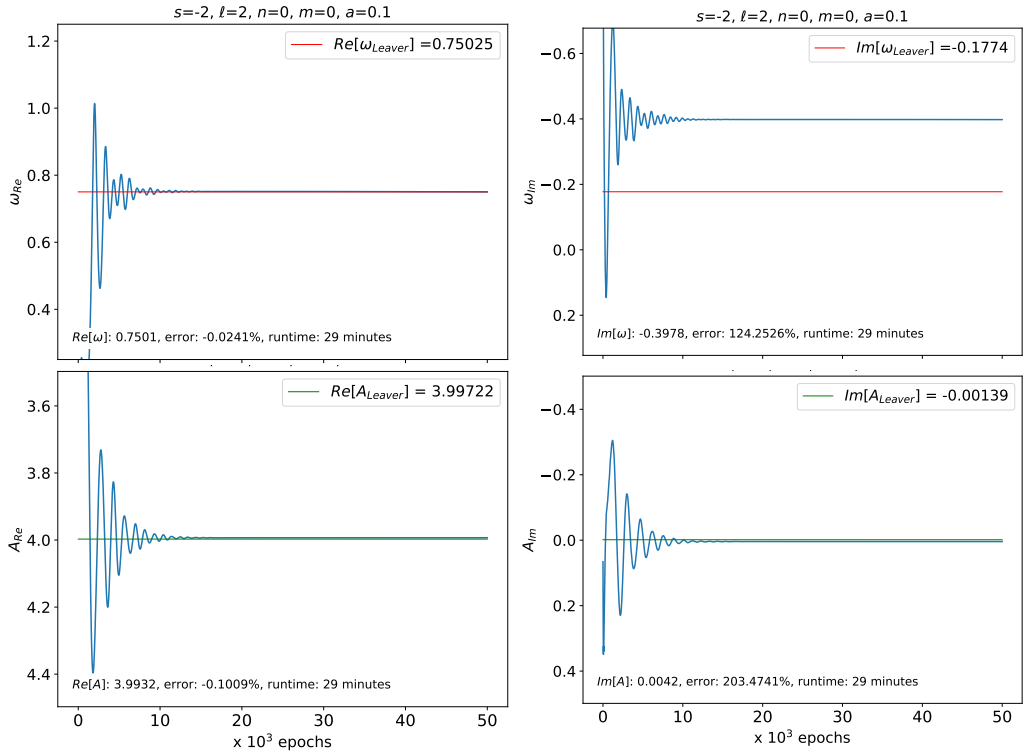


Figure 2. PINN approximations of ω_n and $A_{\ell m}$ for a slowly rotating Kerr black hole (i.e. $a = 0.1$ or $J = 0.2Mc$). Note, the "Leaver" values are those obtained with the CFM [8].

Table 1. Comparing the PINN approximations with the "Leaver" values obtained with the CFM [8] for $a \in [0.0, 0.2, 0.49, 0.4999]$.

	CFM		PINN	
a	ω_n	$A_{\ell m}$	ω_n	$A_{\ell m}$
0.0	0.7473 - 0.1779i	4.0000 + 0.0000i	0.7497 - 0.3953i (0.31%)(122.18%)	3.9954 + 0.0012i (-0.11%)(–%)
0.2	0.7594 - 0.1757i	3.9886 + 0.0056i	0.7667 - 0.4012i (0.97%)(128.41%)	3.9858 + 0.0146i (-0.07%)(161.45%)
0.49	0.8445 - 0.1471i	3.9127 + 0.0315i	0.9076 - 0.5207i (7.46%)(254.07%)	3.9211 + 0.1261i (0.21%)(300.03%)
0.4999	0.8502 - 0.1436i	3.9077 + 0.0323i	0.9680 - 0.5765i (13.85%)(301.35%)	3.9110 + 0.1548i (0.08%)(379.63%)

achieve this. This is corroborated by works such as Ref. [16], where random sampling at every epoch is seen to be a "strong baseline that can achieve close to state-of-the-art performance on many benchmark PDE". Note that no seed values are added, and thus the DNN learn the pair $\{\omega_n, A_{\ell m}\}$ based purely on the Teukolsky equation and the boundary conditions. In addition to the physics-based constraints, we utilise a loss term to penalise trivial solutions to the radial ODE, given as $\mathcal{L} = 1/\langle \hat{f}^2 \rangle$.

Table 1 shows the PINN approximations of ω_n and $A_{\ell m}$ obtained for $s = -2, n = 0, m = 0$. The errors given are percentage deviations from comparing the numerical approximations and the corresponding values obtained with the CFM. An important takeaway from the results pertains to the increase in percentage error with the BH spin parameter a. One conjecture is that this indirectly indicates some association between the physics embodied by the functions, f and g, and the layers and widths of the neural networks (considering that the baseline set-up of PINNs used here works well for the spinless case). Therefore, hyperparameter optimisation, specifically in data-driven learning yet to be done, should be implemented to rigorously find this link.

4. Discussion and Outlook

Using an arbitrary, baseline PINN setup, some PINN approximations of eigenvalues of the Teukolsky equation have been obtained. It is observed in table 1 that the real parts of ω_n and $A_{\ell m}$ are approximated relatively well (error $< 1\%$ for $a = 0$), while the imaginary parts deviate more significantly possibly due to convergence to neighbouring eigenvalues close to the target point. In general, the deviations are exacerbated as $a \to \frac{1}{2}$, which is expected given there is an addition of the azimuthal index m of *spheroidal harmonics*, eliminating the simplicity of the *spherical harmonics* in the case where $a = 0$ and $m = 0$. These results shed light on future steps needed to refine PINNs in the context of Kerr BHs (improving on the accuracy of the approximated ω). Different enhancements of the technique, such as using the supervised learning approach, will be explored to improve QNM computations.

Acknowledgments

ASC is supported in part by the National Research Foundation (NRF) of South Africa; AMN is supported by an SA-CERN Excellence Bursary through iThemba LABS.

References

[1] Abbott B P, Abbott R, Abbott T D, Abernathy M R, Acernese F, Ackley K, Adams C, Adams T, Addesso P, Adhikari R X et al 2016 *Phys. Rev. Lett.* **116** 061102.
[2] Agazie G, Anumarlapudi A, Archibald A M, Arzoumanian Z, Baker P T, Bécsy B, Blecha L, Brazier A, Brook P R, Burke-Spolaor S et al 2023 *Astrophys. J.* 1 **951** L8.
[3] Giesler M, Isi M, Scheel M A and Teukolsky S A 2019 *Phys. Rev. X* **9** 041060.
[4] Isi M, Giesler M, Farr W M, Scheel M A and Teukolsky S A 2019 *Phys. Rev. Lett.* **123** 111102.
[5] Konoplya R A and Zhidenko A 2011 *Rev. Mod. Phys.* **83** 793.
[6] Green S R, Hollands S, Sberna L, Toomani V and Zimmerman P 2023 *Phys. Rev. D* **107** 064030.
[7] Teukolsky S A 1973 *Astrophys. J.* **185** 635.
[8] Leaver E W 1985 *Proc. R. Soc. Lond. A Math Phys. Sci.* 1823 **402** 405.
[9] Seidel E and Iyer S 1990 *Phys. Rev. D* **41** 374.
[10] Cornell A S, Ncube A and Harmsen G 2022 *Phys. Rev. D* **106** 124047.
[11] Kingma D. and Ba J. 2015 *3rd Int Conf on Learning Representations* ed. Y Bengio and Y LeCun (San Diego, CA, USA).
[12] Luna R, Calderón B J, Seoane M J J, Torres-Forné A and Font J A 2023 *Phys. Rev. D* **107** 064025.
[13] Raissi M, Perdikaris P and Karniadakis G E 2019 *J. Comput. Phys.* **378** 686.
[14] Lu L, Meng X, Mao Z and Karniadakis G M 2021 *SIAM Rev.* **63** 208.
[15] Karniadakis G E, Kevrekidis I G, Lu L et al 2021 *Nat. Rev. Phys.* **3** 422.
[16] Daw A, Bu J, Wang S, Perdikaris P and Karpatne A 2022. *Preprint* arXiv:2207.02338 [cs.LG].

Topological Edge States in 2D Su–Schrieffer–Heeger Models

C S van Niekerk[1] and R Warmbier[2]

[1] Department of Physics, University of Johannesburg, Auckland Park, Johannesburg, South Africa

[2] School of Physics and Mandelstam Institute for Theoretical Physics, University of the Witwatersrand, Johannesburg, South Africa

E-mail: csvanniekerk@uj.ac.za[1], robert.warmbier@wits.ac.za[2]

Abstract. Here we present the standard SSH model as well as a new extended version of the SSH model including second-nearest-neighbour interactions. For these models the characteristic topological phases are presented. The standard SSH model presents phases with corner as well as edge states, showing as horizontal lines and piece-wise horizontal lines in the band structures respectively. The zero-energy corner state is suppressed in semi-periodic configurations allowing for the edge state to be the lowest energy state.

1. Introduction

Topological materials have an intriguing 2D quantum state where the bulk of the material is insulating, while the edges carry a current. The mathematical concept of topology has been used to define and predict cases where this occurs [1, 2].

For an topological insulator, the conducting surface states occur on the boundary of the material, i.e. a d-dimensional material can host a boundary state with dimensions $\leq d-1$ [3]. States with dimensions $< d-1$ are classified as higher order topological states [4]. The conducting surface states are defined to be resistant to defects and impurities provided the symmetry of the system is preserved.

The 2D Su-Schrieffer-Heeger(SSH) model is based on a simplified tight-binding approach for rectangular unit cells and can describe topological phases [5, 6, 7, 8, 9, 10, 11]. The ordinary 2D SSH model includes nearest neighbor interactions only, where second nearest neighbor interactions can be included to lower the symmetry of the model [10, 11, 12, 13]. In this work, we outline the effect of different periodicity on the boundary states of a 2D SSH model with and without second nearest neighbor (SNN) interactions.

2. The Model

In our previous work, an extended 2D SSH model with second nearest neighbor interactions was constructed. [10, 11] The model coupled together two 1D SSH dimerized chains in the x and y (nearest neighbor interactions), and the xy directions (second nearest neighbor interactions) forming an $ABCD$ unit cell as shown in Figure 1. The real-space model outlined was used to construct the finite system [10].

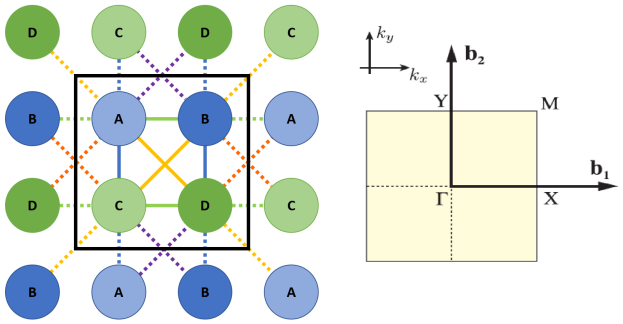


Figure 1. Construction of a 2D system using two different 1D SSH chains coupled together in the y-direction. Lattice sites A, B, C, and D represent one unit cell (outlined in black). Dashed and solid lines show all the possible nearest and next nearest neighbor hoppings from each unit cell. The Brillouin Zone of this unit cell is shown on the right.

Assuming a Bloch wave solutions of the form $\left|\psi(\vec{\mathbf{k}})\right\rangle = \sum_X \sum_{n,m} \left[e^{i\vec{\mathbf{k}}\cdot(n,m)}\phi_X(\vec{\mathbf{k}})\left|X_{n,m}\right\rangle\right]$; with $X \in \{A, B, C, D\}$, and with $E\phi = h(\vec{\mathbf{k}})\phi$, the reciprocal space Hamiltonian in the basis $\hat{\psi}^\dagger_{\vec{\mathbf{k}}} = (\hat{A}, \hat{B}, \hat{C}, \hat{D})$ can be written as:

$$H = \hat{h}_{SSH}(\vec{\mathbf{k}}) = \begin{pmatrix} 0 & s_x(k_x) & s_y(k_y) & s_{xy}(\vec{\mathbf{k}}) \\ s_x^*(k_x) & 0 & s'_{xy}(\vec{\mathbf{k}}) & s_y(k_y) \\ s_y^*(k_y) & s'^*_{xy}(\vec{\mathbf{k}}) & 0 & s_x(k_x) \\ s^*_{xy}(\vec{\mathbf{k}}) & s_y^*(k_y) & s_x^*(k_x) & 0 \end{pmatrix}, \tag{1}$$

where $s_l(k_l) = \gamma_l + \gamma'_l e^{-ik_l}$, $l \in \{x, y\}$ and $s_{xy}(\vec{\mathbf{k}}) = \delta + \delta_x e^{-ik_x} + \delta_y e^{-ik_y} + \delta_{xy} e^{-i(k_x+k_y)}$ and $s'_{xy}(\vec{\mathbf{k}}) = \delta + \delta_x e^{ik_x} + \delta_y e^{-ik_y} + \delta_{xy} e^{-i(-k_x+k_y)}$. The corresponding real-space Hamiltonian can be found via Fourier transform and is given in Ref. [10]. In the model, γ_l and γ'_l; $l \in \{x, y\}$ are the nearest intra-cell and inter-cell interactions, and δ and δ_p, $(p \in \{x, y, xy\})$, are the intra-cell and inter-cell SNN neighbor interactions.

Topological properties are encoded in the model Hamiltonian and its symmetries. The model without SNN interactions has time-reversal, particle-hole, chiral and inversion symmetry placing it in the BDI class of systems. [14] The addition of SNN interactions leads to the loss of chiral and particle hole symmetry and therefore places the model into a lowered symmetry AI class of systems.

When the system has time-reversal and inversion symmetry, a vectored parity invariant can be used to define the topology of the material. The invariant tracks the number of parity changes in the bands at the high symmetry points by calculating an inversion eigenvalue with $\xi(\vec{k}) = \iota H \iota^{-1}$ for the occupied bands, where ι is the unitary inversion operator and $\vec{k}$ corresponds to a point in momentum space. Due to the unitary nature of ι, the inversion eigenvalues can only take on values of ± 1 corresponding to an even or odd parity at each point [15, 16]. The four time reversal invariant momenta [17], i.e. $\Gamma(\vec{\mathbf{k}} = (0, 0))$, $X(\vec{\mathbf{k}} = (\pi, 0))$, $Y(\vec{\mathbf{k}} = (0, \pi))$ and $M(\vec{\mathbf{k}} = (\pi, \pi))$ are shown in Figure 1 on the right. Due to the additional symmetry in the system, the Γ and M points always have the same parity eigenvalue and therefore can be used interchangeably. The vectored parity invariant (ν_x, ν_y) is calculated using $(-1)^{\nu_x} = \prod_{n_{occ}} \xi(X)/\xi(\Gamma)$ and $(-1)^{\nu_y} = \prod_{n_{occ}} \xi(Y)/\xi(\Gamma)$, where n_{occ} refers to the occupied bands.

3. Numerical Results

The SSH model supports both insulating and metallic bulk solutions. For this work, we will only consider the insulating regions in our discussion. Throughout this work, a system size of 30×30 unit cells was used for the real space system. For localisation checks in the open boundary case

the surface was assumed to be 4 unit cells thick.

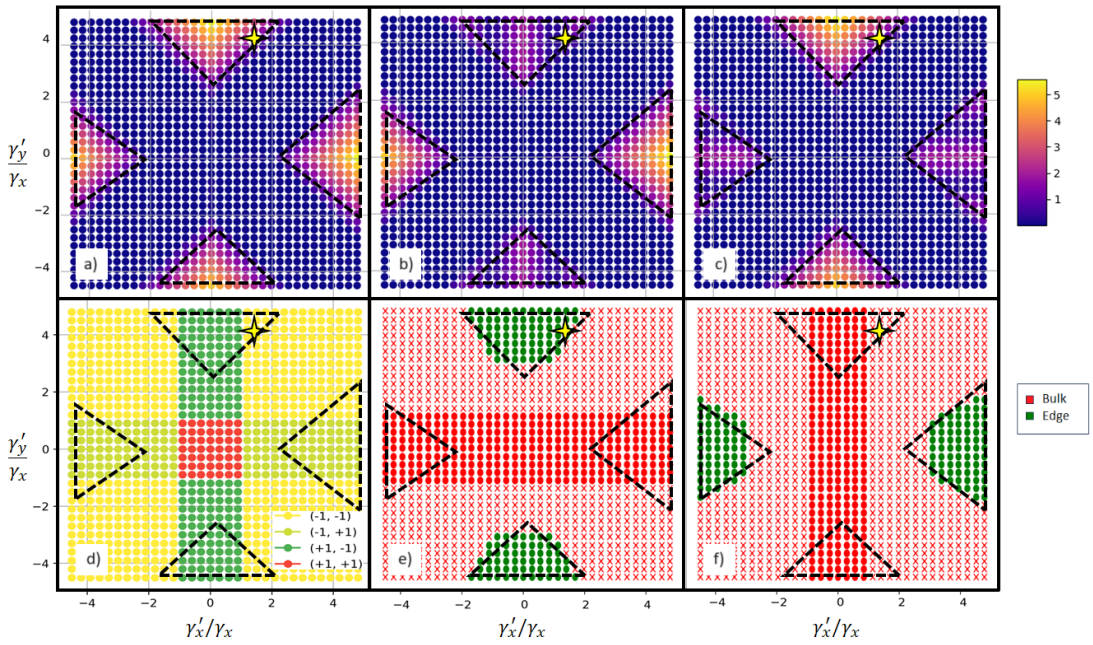


Figure 2. Band gap size for **a)** the completely periodic, **b)** the $x-$, and **c)** the $y-$ periodic real space system for a nearly isotropic SSH model varying with γ_x'/γ and $\gamma_y'\gamma$. Blue(yellow) regions indicate a closed(open) band gap. **d)** The vectored parity invariant for the SSH model with no SNN interactions. An invariant value of -1 indicates a topological phase, while a value of $+1$ indicates a trivial phase. The **e)** $x-$ and **f)** $y-$ periodic wave function localisation of the bands across the parameter space. The different markers refer to a weak ($\times$) and strong ($\bullet$) topological phase. Throughout the paper, we focus on a representative point which is topological in both directions. This point is highlighted by the yellow star where $\gamma_x'/\gamma = 1.5, \gamma_y'/\gamma = 4.2$.

A purely isotropic SSH model ($\gamma_x = \gamma_y$ and $\gamma_x' = \gamma_y'$) is known to always be metallic. Metallic phases are a challenge to characterize topologically due to the large number of states at the Fermi level. A slight modification to the isotropic case can be made to also produce insulating cases - where $\gamma_x = \gamma_y = \gamma$. To further reduce the degrees of freedom, all parameters are scaled by γ, without loss of generality. Figure 2*a*) shows the bulk band gap of this reduced model as a function of the two independent parameters γ_x'/γ and γ_y'/γ, while Figure 2*d*) corresponds to the parity invariant over the same parameter space. This shows that completely trivial phases are all metallic, while partially and completely topological phases can be either metallic or non-metallic. When $\frac{\gamma_l'}{\gamma_l} < 1$, when l is x or y, the system is trivial and when $\frac{\gamma_l'}{\gamma_l} > 1$, the system is topological.

Panels *b*) and *e*) of Figure 2 show the band gap and wave function localisation for the partially periodic systems with periodic boundary conditions (PBCs) in the x direction. Panels *c*) and *f*) show the cases with $y-$PBCs. Partially periodic cases with large γ_y'/γ (γ_x'/γ) exhibit a decreased band gap compared to the periodic case, pointing at edge states in the bulk gap. The localisation of the partially periodic wave functions of the valence band maximum (VBM) show that the wave functions are localised in the bulk or edge in distinct regions of the parameter space. In the metallic region, the VBM is localised in the bulk, meaning that there exist no edge states in the band gap. When the highest valence band bulk state has a larger energy than the highest valence band edge state, and the localisation is on the surface, this refers to a weak

topological phase. Regions in the figure with a × marker indicate these points. For the SSH model with no SNN interactions there are no weak topological surface states. Along the periodic direction, the bulk state valence band maximum has a lower energy than the edge state valence band maximum, but the surface localisation in that direction is destroyed.

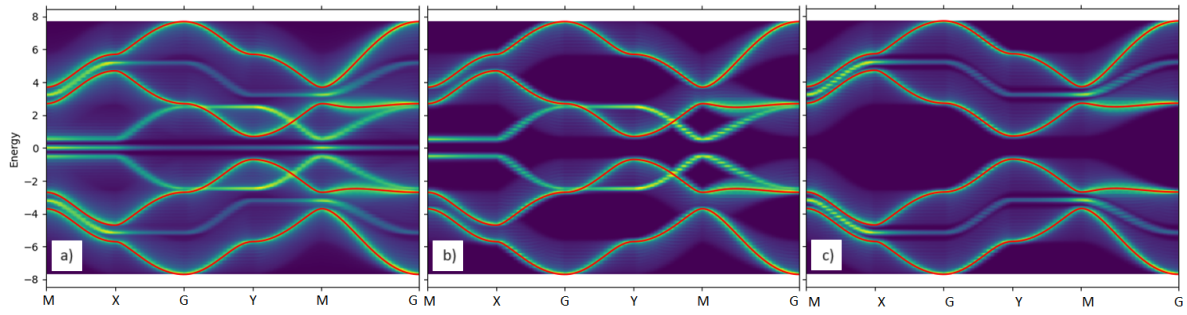


Figure 3. Band structure $\gamma_x'/\gamma = 1.5, \gamma_y'/\gamma = 4.2$ for **a)** open, **b)** $x-$ and **c)** $y-$ PBCs. Band structures were obtained by Fourier transforming the real space wave functions. The red lines correspond to bands obtained from the reciprocal space Hamiltonian.

A completely topological, insulating point in parameter space with $\gamma_x'/\gamma = 1.5, \gamma_y'/\gamma = 4.2$ corresponding to the yellow star in Figure 2. It is selected to further study the behavior of the topological edge states. For the extended SSH model, this point is known to have a zero energy state as well as other surface states. Given the anisotropy in x and y directions we expect different results for partial periodicity in of these directions only. The band structure for this set of parameters is shown in Figure 3 for the open boundary, x and y periodic cases. The red lines show bands obtained using the reciprocal space Hamiltonian. The extended SSH model without SNN interactions is known to have particle-hole symmetry as is evident in the band structure. When an open boundary is present, as in panel a), there is a zero energy corner state which is necessarily removed when enforcing any type of periodicity, as in panels b) and c). When comparing panels b) and c) of Figure 3 the edge state is preserved despite the periodicity, with the exception that for x PBCs, the surface states lay in the bulk band gap, aka between the 2^{nd} and 3^{rd} bulk band, while for y periodicity they are outside of the bulk band gap. This illustrates why one type of periodicity has a localised surface state at that point, while the other does not. SNN interactions break chiral and particle-hole symmetry, thereby changing the topological properties of the system. Depending on the metallicity of the system, this has different consequences. For example, for a metallic system, the variation of the SNN parameters resulted in a gap opening and a phase transition as a result. This case presented the same type of result as the standard SSH model - the surface state was preserved along the non-periodic direction. Sufficiently large SNN interactions close the band gap of a non-metallic system thereby destroying the surface states as well. The same non-metallic representative point is considered as before with $\gamma_x'/\gamma = 1.5, \gamma_y'/\gamma = 4.2$. Figure $4a) - c)$ show band gaps for the different PBCs. There are partial surface states in the case with PBCs in the x direction, but they are destroyed in the y direction. The parity invariant in Figure $4d$) predicts that the entire surface conducting region should be topologically protected. The red stars in Figure 4 indicate a point in the yellow and green regions of topologically protected space.

The addition of SNN interactions shifts the surface states in and out of the band gap. For PBCs in the x direction, the surface state is preserved in part of the region, while the localization in the y periodic case is removed. To illustrate this, Figure $5a)(c)$ and $b)(d)$) show the x and y periodic band structure of the partially(completely) topological region with $\delta = -0.3(0.6)$ and $\delta_{xy} = 0.9(-0.9)$. The bulk bands show the breaking of the particle-hole and chiral symmetries.

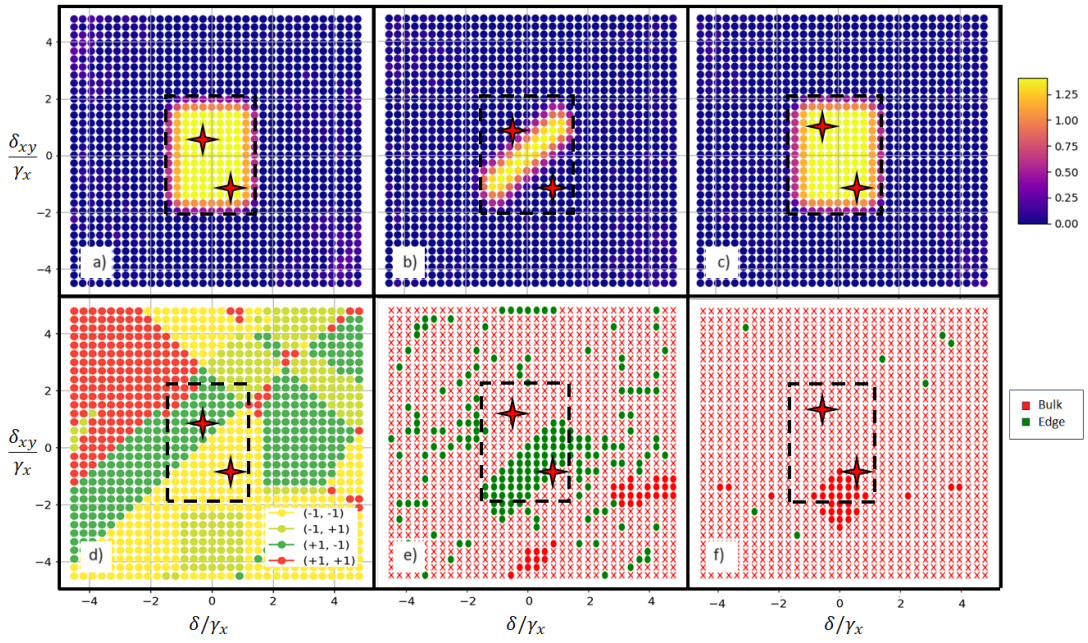


Figure 4. A real space non-metallic system with $\gamma_x'/\gamma = 1.5, \gamma_y'/\gamma = 4.2$ varying over δ/γ and δ_{xy}/γ. Band gap for a system with **a)** completely periodic, **b)** $x-$ and **c)** $y-$ PBCs. **d)** Parity invariant for this case, where ± 1 refers to topological/trivial phases. Band localisation for **e)** $x-$ and **f)** $y-$ PBCs. he markers $\times$ and $\bullet$ refer to a weak/strong topological phase. A point from each of these regions is chosen for further studies and are indicated by the red stars. These points refer to points in parameter space with $\delta = -0.3$, $\delta_{xy} = 0.9$, and $\delta = 0.6$, $\delta_{xy} = -0.9$ respectively.

The band structure of surface states on the other hand preserve the symmetry of these states completely under x PBCs. For one of the two bands in the y periodic case one of these surface states is preserved and has its original symmetry. This is due to the surface states being topologically protected by symmetry.
When calculating the invariant, the parity of the VBM is calculated. In the partially topological state (green region of Figure 4*d*)), the VBM is outside of the band gap making it a weak topological state under x PBCs as shown in Figure 5*a*). The y PBCs (Figure 5*b*)) retain the valence band surface state. For the completely topological case (yellow region of Figure 4*d*)), the VBM shown in Figure 5*c*) is inside the band gap making it a strong topological phase when x periodic boundary condition are present. For y PBCs as plotted in Figure 5*d*), the VBM surface state is destroyed. The behaviour where if the VBM is outside of the gap, the conduction band minimum (CBM) is inside the gap, or where if the states are outside of the gap when the VBM state is present, the CBM state is destroyed and vice versa confirms the difference in parity invariant in the green and yellow regions.

4. Conclusion

In general the 2D SSH model predicts zero energy corner states in the fully topological phase, which cannot exist in systems with partial periodicity. In this case the edge states close to the Fermi level become the primary topological states.
Chiral and particle-hole symmetry preserve the surface states in the general 2D SSH model. Depending on the periodicity, surface states in the band gap are topological, while those completely shifted outside of the gap are not.

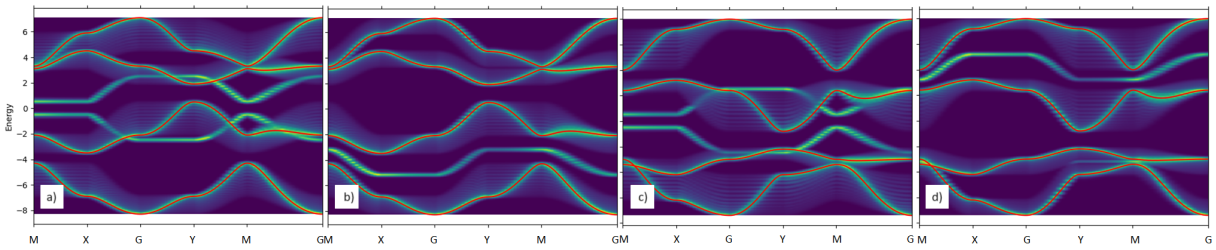


Figure 5. Band structure with $x-$ PBCs when **a)** $\delta = -0.3$, and **c)** $\delta = 0.6$, $\delta_{xy} = -0.9$. The corresponding figures for $y-$PBCs are shown in **b)** and **d)**. The values of $\gamma_x'/\gamma = 1.5$ and $\gamma_y'/\gamma = 4.2$ throughout.

The addition of SNN interactions breaks the chiral and particle hole symmetry of the bulk bands, but are still preserved for the symmetry protected topological states to a large degree. For states which were initially inside of the band gap, both states remain, but the SNN interactions shift either the VBM or CBM outside of the band gap. If the VBM was outside of the gap, the system is referred to as a weak topological insulator. For states that were initially outside of the band gap, either the state belonging to the VBM or CBM is destroyed while the other is preserved. With the knowledge of the effects of periodicity on systems with and without SNN interactions, it is possible to predict and then control the type of localisation of surface conducting states in a given system.

Acknowledgments
The support of the DSI-NRF Centre of Excellence in Strong Materials (CoE-SM) 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-SM. R.W. also acknowledges support by the Mandelstam Institute for Theoretical Physics.

References

[1] Qi X L and Zhang S C 2011 *Rev. Mod. Phys.* **83**(4) 1057–1110
[2] Hasan M Z and Kane C L 2010 *Rev. Mod. Phys.* **82** 3045
[3] Chen Z G, Xu C *et al.* 2019 *Phys. Rev. B* **100**(7) 075120
[4] Langbehn J, Peng Y *et al.* 2017 *Phys. Rev. Lett.* **119**(24) 246401
[5] Liu F and Wakabayashi K 2017 *Phys. Rev. Lett.* **118**(7) 076803
[6] Xie B Y, Wang H F *et al.* 2018 *Phys. Rev. B* **98**(20) 205147
[7] Obana D, Liu F *et al.* 2019 *Phys. Rev. B* **100**(7) 075437
[8] Kim M and Rho J 2020 *Nanophotonics* **9**(10) 3227 – 3234
[9] Bomantara R W, Zhou L *et al.* 2019 *Phys. Rev. B* **99**(4) 045441
[10] van Niekerk C S and Warmbier R 2023 *Phys. Status Solidi B - In Production* (*Preprint* 2311.00427)
[11] van Niekerk C S and Warmbier R 2023 *Phys. Status Solidi B* **260** 2300018
[12] Ma H, Zhang Z *et al.* 2022 *Phys. Rev. B* **106**(24) 245109
[13] Luo X J, Pan X H *et al.* 2023 *Phys. Rev. B* **107**(4) 045118
[14] Ryu S, Schnyder A P, Furusaki A and Ludwig A W W 2010 *New Journal of Physics* **12** 065010
[15] Hughes T L, Prodan E *et al.* 2011 *Phys. Rev. B* **83**(24) 245132
[16] Fang C, Gilbert M J *et al.* 2012 *Phys. Rev. B* **86**(11) 115112
[17] Fu L and Kane C L 2007 *Phys. Rev. B* **76**(4) 045302

Viscous cosmological fluids and large-scale structure

B.G Mbewe[1], R.R Mekuria[2], S. Sahlu[1,3], and A Abebe[1,4]

[1]Centre for Space Research, North-West University, Potchefstroom 2520, South Africa
[2] Ala-Too International University, Faculty of Engineering and Computer Science, Bishkek, Kyrgzstan
[3] Entoto Observatory and Research Center, Space Science and Geospatial Institute, Addis Ababa Ethiopia
[4] National Institute for Theoretical and Computational Sciences (NITheCS), South Africa

E-mail: bonang.mbewe@gmail.com

Abstract. In this paper, we study the viscous fluid cosmological model that when certain conditions are invoked mimics the ΛCDM model. The background equations governing the evolution of viscous interacting fluids in a multifluid system are derived. The Markov Chain Monte Carlo (MCMC) simulation is applied to constrain the best-fit cosmological parameters with Supernova Type 1a data. In addition, linear cosmological perturbations are investigated in a dust-matter-dominated frame using a $1+3$ covariant formalism approach. It is evident from the perturbation results obtained that the model predicts the disintegration of bound structures of large-scale structures in the late-time universe.

1. Introduction

Recent cosmological observations have indicated that the universe is undergoing an accelerated expansion. This accelerated expansion of the universe phenomena associated with dark energy has opened up a plethora of dark energy theories that have been proposed in order to address the behaviour that the current universe is displaying. ΛCDM amongst others, is one of the models proposed to explain the history of cosmic acceleration taking a simple and straightforward form of dark energy as a constant (also known as the cosmological constant Λ and cold dark matter (CDM)) [1, 2, 3]. Although ΛCDM has proved successful in most observational tests, as scientific models often do, it has faced challenges in explaining certain phenomena, and as such more models for dark energy are proposed as an alternative to ΛCDM.

In this paper, we scrutinize a model that, when certain conditions are invoked, mimics ΛCDM. For a multifluid universe with viscous effects taken into consideration [4] and the dark sector components coupled together [5, 6], allowing for exchange of energy from one to the other, and lastly, the equation of state for dark energy is assumed to be given in an inhomogeneous form [3, 7, 8, 9, 10]. In addition, the model will be used to test/study the observational feasibility and implications for the large-scale structure formation of the universe. First, after the background equations governing the model are obtained, the cosmological background parameters are to be constrained using Supernovae Type 1a data by use of MCMC simulation. Thereafter, linear cosmological perturbations will be studied in the $1+3$ covariant formalism for a multifluid system [11, 12, 13, 14]. Moreover, the density perturbation equations and solutions will be derived, and the cosmological implications of the density perturbations on large-scale structure formation will be discussed.

2. Background cosmology of viscous interacting dark fluids (VIDF).
In a homogeneous and isotropic universe, the background equations for a viscous interacting fluid in a flat geometry setting of FLRW metric [15] read as:

$$\begin{aligned}
&\dot{\rho}_r + 3H(\rho_r + p_r) = 0,\\
&\dot{\rho}_d + 3H(\rho_d + p_d) = Q,\\
&\dot{\rho}_\Lambda + 3H(\rho_\Lambda + p_\Lambda) = -Q,\\
&\dot{H} = -\frac{1}{2}(\rho_{tot} + p_{tot}).
\end{aligned} \tag{1}$$

Here a dot denotes a derivative with respect to cosmic time. ρ_r, ρ_d, ρ_Λ and ρ_{tot} refer to the energy density of radiation, dust, vacuum and total fluid while $p_r = \rho_r/3$, $p_d = 0$, p_Λ and p_{tot} refer to isotropic pressure terms of vacuum and total fluid. H is the Hubble parameter, and Q is the interacting parameter. Natural standard units have been adopted, that is $8\pi G = c = 1$, and will be used throughout this manuscript. Taking the source term as taken by Brevik [16]:

$$Q = \delta H \rho_d, \tag{2}$$

where δ is a dimensionless constant. However, it can be noticed from the setup of conservation equations given by equation (1) that when $\delta > 0$, energy exchange flows from dark energy to dark matter and conversely for $\delta < 0$, while when $\delta = 0$ the dark fluids are decoupled from each other [17]. Moreover, taking the equation of state for dark energy to be given by an inhomogeneous form as:

$$p_\Lambda = w_\Lambda \rho_\Lambda - \zeta, \quad \text{where} \quad \begin{cases} w_\Lambda = A_0 \rho_\Lambda^{\alpha-1} - 1, \\ \zeta = \zeta_0 \rho_{\Lambda 0} \left(\frac{\rho_\Lambda}{\rho_{\Lambda 0}} \right)^m. \end{cases} \tag{3}$$

By assuming that the dimensionless constants $\alpha = m = 1$, to simplify the fluid equation for dark energy, the reduced equation of state for dark energy, now reads:

$$p_\Lambda = (A_0 - 1 - \zeta_0)\rho_\Lambda, \tag{4}$$

with A_0 and ζ_0 being dimensionless constants. The solutions that govern the evolution of each fluid species dynamics, in a multi-fluid universe are obtained as follows:

$$\begin{aligned}
&\Omega_r = \frac{\Omega_{r0}}{h^2}(1+z)^4 \quad ; \quad \Omega_d = \frac{\Omega_{d0}}{h^2}(1+z)^{3-\delta} \quad ;\\
&\Omega_\Lambda = \frac{1}{h^2}\Bigg[(1-\Omega_{r0})(1+z)^{3(A_0-\zeta_0)}\\
&+ \frac{\Omega_{d0}}{3(1+\zeta_0-A_0)-\delta}\left\{\delta(1+z)^{3-\delta} - 3(1+\zeta_0-A_0)(1+z)^{3(A_0-\zeta_0)}\right\}\Bigg].
\end{aligned} \tag{5}$$

Here Ω_{r0}, Ω_{d0} refer to the fractional energy densities for radiation and dust evaluated at present cosmic time. Furthermore, $h = H/H_0$ denotes a normalised Hubble parameter.
It is observed from Fig. 1, that the viscous interacting model compared to the ΛCDM model has the following features: the radiation-dust matter equality occurs earlier, while the dust matter-dark energy equality epoch occurs later than ΛCDM. The dust-matter domination era is longer than that predicted by the ΛCDM model. Furthermore, the viscous interacting model violates all the energy conditions by having a dark energy density that is negative. The Hubble parameter governing the expansion of viscous dark-energy model is given as:

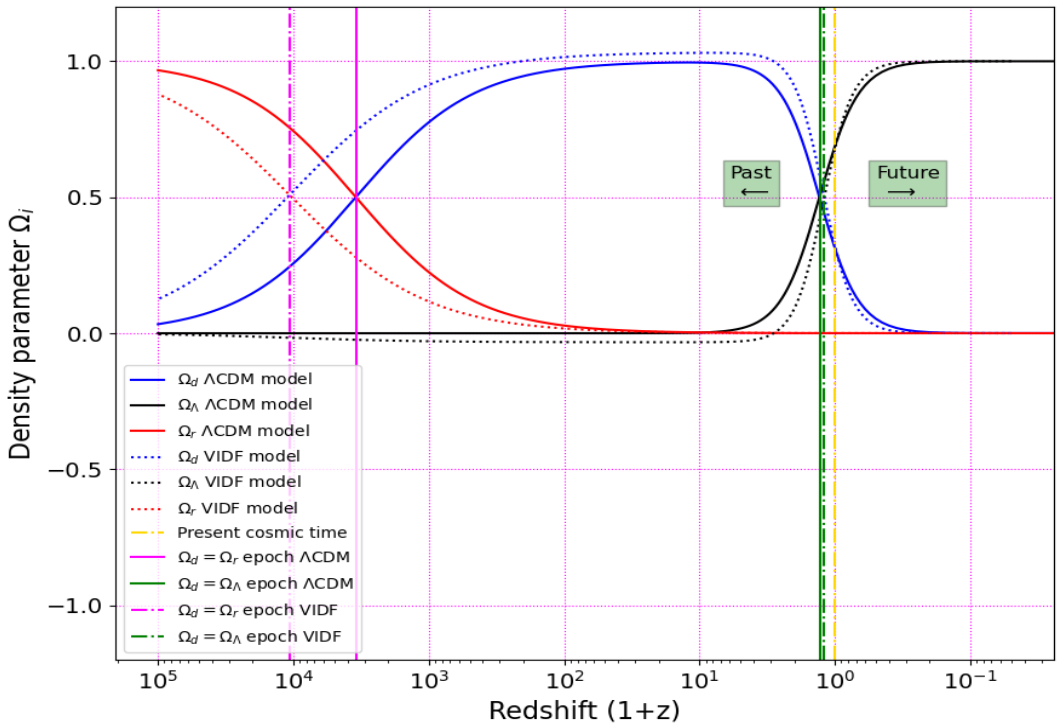


Figure 1: The evolution of energy density parameter for a multi-fluid universe.

$$
\begin{aligned}
H =& H_0\Bigg\{\Omega_{r0}\Big[(1+z)^4-(1+z)^{3(A_0-\zeta_0)}\Big] \\
&+\Omega_{d0}\left[\frac{3(1+\zeta_0-A_0)}{3(1+\zeta_0-A_0)-\delta}\right]\Big[(1+z)^{3-\delta}-(1+z)^{3(A_0-\zeta_0)}\Big]+(1+z)^{3(A_0-\zeta_0)}\Bigg\}^{\frac{1}{2}}.
\end{aligned}
\tag{6}
$$

The deceleration parameter can be obtained from Eq. (6) and it yields as:

$$
\begin{aligned}
q(z) =& \frac{1}{2h^2}\Bigg[\Omega_{r0}\Big\{4(1+z)^4-3(A_0-\zeta_0)1+z)^{3(A_0-\zeta_0)}\Big\} \\
&+\Omega_{d0}\left(\frac{3(1+\zeta_0-A_0)}{3(1+\zeta_0-A_0)-\delta}\right)\Big\{(3-\delta)(1+z)^{3-\delta}-3(A_0-\zeta_0)(1+z)^{3(A_0-\zeta_0)}\Big\} \\
&+3(A_0-\zeta_0)(1+z)^{3(A_0-\zeta_0)}\Bigg]-1.
\end{aligned}
\tag{7}
$$

From Eq. (6), the luminosity distance can be expressed as [14, 18]:

$$
\mu = 25+5\log_{10}\left[3000\bar{h}^{-1}(1+z)\int_0^z\frac{dz'}{H(z')/H_0}\right]. \tag{8}
$$

In Fig. 3, we see the transition from a decelerating universe to an accelerating universe occurring much later than that of the ΛCDM model, which was made evident in Fig. 1 when the dust-dark energy equality epoch was later, unlike the predicted ΛCDM. Moreover, the two models seem to both predict the fact that the universe will indeed be in an accelerated regime indefinitely.

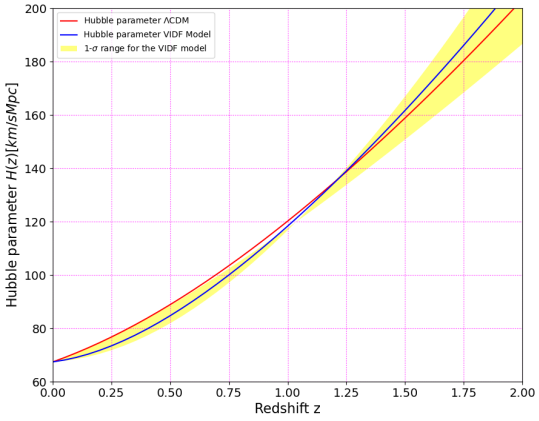


Figure 2: The evolution of Hubble parameter

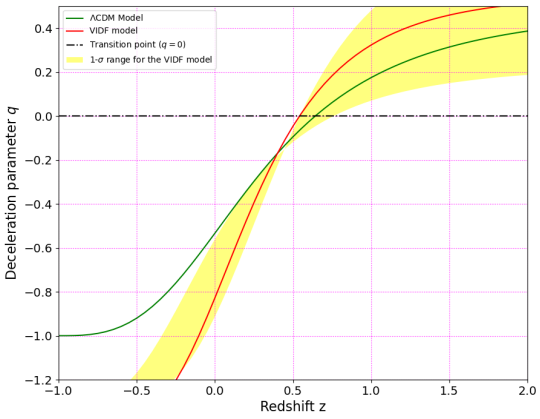


Figure 3: The evolution of deceleration parameter

3. MCMC results of the viscous interacting fluids

By applying the same methodology as detailed in [13], we constrain the best-fit cosmological parameter using Eq.(8) with SNIa data, and the results are presented in Figs. 4 and 5 for ΛCDM and VIDF models, respectively.

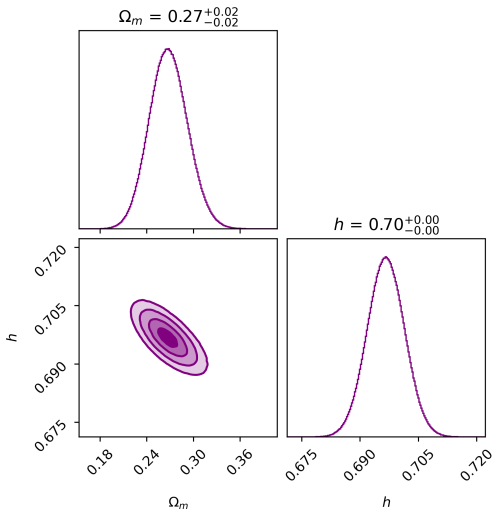


Figure 4: Best fit parameters of ΛCDM model using MCMC.

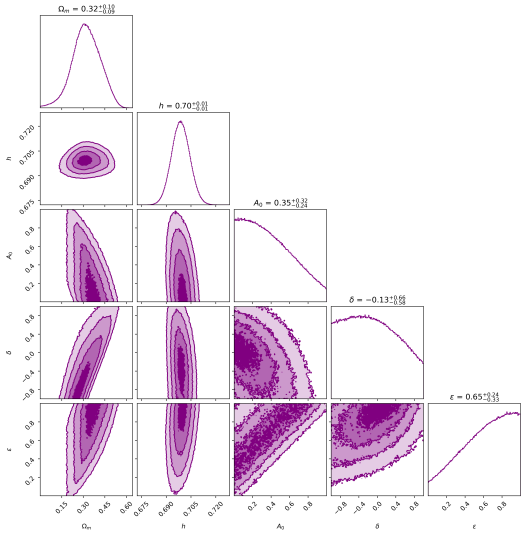


Figure 5: Best fit parameters of VIDF model using MCMC.

In Fig. 5 the constrained value of dust matter is in fairly good agreement with that obtained from Planck's 2018 data [19]. However, as for the Hubble parameter, there exist some discrepancies as the value obtained is not within the 1σ deviation. In these MCMC simulation results, radiation was taken into account; however, it was treated as a constant, since when it is taken as a parameter to be constrained, the simulation fails to resolve it, and, as such, constraining it yields a uniform distribution over its entire search space, as outlined in detail by [13].

4. Linear cosmological perturbations in VIDF model

Taking the constrained cosmological parameters together with the background VIDF cosmology to study the growth of matter density fluctuations using a $(1+3)$ linear covariant cosmological perturbation to observe the contribution that the VIDF model has on large-scale structure formation. Using $(1+3)$ linear covariant formalism, we start by defining the spatial gradient variables [12, 14]as:

$$D_a^d = \frac{a}{\rho_d}\nabla_a\rho_d \quad , \quad D_a^\Lambda = \frac{a}{\rho_\Lambda}\nabla_a\rho_\Lambda \quad , \quad Z_a = a\nabla_a\Theta, \tag{9}$$

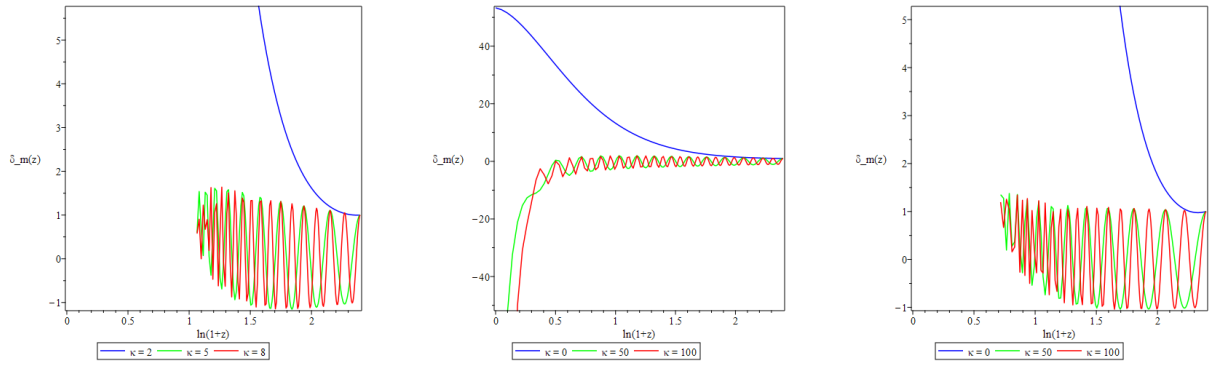
where D_a^i is the inhomogeneity of the energy density of the species $(i = d, \Lambda)$, Z_a is the volume expansion of the fluids, and $\Theta = 3H$. Taking the cosmic time derivative, scalar decomposition and harmonic decomposition of the spatial gradient variables, we obtain the following:

$$\begin{aligned}
\Delta_d' =& \left(\frac{3-\delta}{3(1+z)}\right)\frac{\mathcal{Z}}{h} + \left(\frac{3-\delta}{(1+z)}\right)\left(\frac{w_d\Omega_d}{(1+w_t)\Omega_t}\right)\Delta_d + \left(\frac{3-\delta}{(1+z)}\right)\left(\frac{(w_\Lambda-\zeta_0)\Omega_\Lambda}{(1+w_t)\Omega_t}\right)\Delta_\Lambda,\\
\Delta_\Lambda' =& -\left[\frac{1}{3}\left(3(\zeta_0 - A_0) - \frac{\delta\Omega_d}{\Omega_\Lambda}\right)\right]\frac{\mathcal{Z}}{(1+z)h}\\
&-\left\{\left[3(\zeta_0 - A_0) - \frac{\delta\Omega_d}{\Omega_\Lambda}\right]\left[\frac{w_d\Omega_d}{(1+w_t)\Omega_t}\right] - \frac{\delta\Omega_d}{\Omega_\Lambda}\right\}\frac{1}{(1+z)}\Delta_d\\
&-\left\{\left[3(\zeta_0 - A_0) - \frac{\delta\Omega_d}{\Omega_\Lambda}\right]\left[\frac{(w_\Lambda-\zeta_0)\Omega_\Lambda}{(1+w_t)\Omega_t}\right] + \frac{\delta\Omega_d}{\Omega_\Lambda}\right\}\frac{1}{(1+z)}\Delta_\Lambda,\\
Z' =& \frac{2}{(1+z)}\mathcal{Z}\\
&-\left[\left\{\frac{k^2(1+z)^2}{3h^2} - 1 - \frac{1}{2}\left((1+3w_d)\Omega_d + \big[1+3(w_\Lambda-\zeta_0)\big]\Omega_\Lambda\right)\right\}\frac{w_d\Omega_d}{(1+w_t)\Omega_t}\right.\\
&\left.\quad - \frac{1}{2}(1+3w_d)\Omega_d\right]\left(\frac{3h}{(1+z)}\right)\Delta_d\\
&-\left[\left\{\frac{k^2(1+z)^2}{3h^2} - 1 - \frac{1}{2}\left((1+3w_d)\Omega_d + \big[1+3(w_\Lambda-\zeta_0)\big]\Omega_\Lambda\right)\right\}\frac{(w_\Lambda-\zeta_0)\Omega_\Lambda}{(1+w_t)\Omega_t}\right.\\
&\left.\quad - \frac{1}{2}\big[1+3(w_\Lambda-\zeta_0)\big]\Omega_\Lambda\right]\left(\frac{3h}{(1+z)}\right)\Delta_\Lambda.
\end{aligned} \tag{10}$$

Here, Δ_m, Δ_Λ and $\mathcal{Z}$ are the scalar perturbations for dust matter, dark energy, and volume expansion. The prime denotes the derivative with respect to the redshift. It is seen from Eq. (10) that conditions required for the ΛCDM model retrieval, that indeed even on the perturbation level is also the case [1].

Fig. 6, gives different extrema displayed by the VIDF model. In Fig. 6a, the central values obtained from MCMC simulation were used to obtain the density contrast of dust matter, and it can be seen for long wavelengths that the energy density fluctuations growths faster, however, there seems to be a singularity at a redshift of (~ 2), while for short wavelengths the energy density seems to oscillate with growing energy density fluctuations, that is the same behaviour

[1] All the derivations of the equations found in this proceedings paper can be obtained in this Google Drive link `https://drive.google.com/drive/folders/1h9K_JW36rUv8Yti6befMkf4U-PKqWxyO`



(a) Central values taken of MCMC.

(b) Upper boundary values taken of MCMC.

(c) Lower boundary values taken of MCMC.

Figure 6: The density contrast of VIDE model as a function of redshift.

noticed when considering the lower boundary values obtained from the simulation. However, in Fig. 6b, there exist no singularity which was experienced by the model in Figs. 6a and 6c and we thus see the growth of matter density fluctuation throughout redshift for longer wavelength while a decaying one when considering short wavelengths.

5. Conclusion

In this paper the VIDF model has been explored in the FLRW universe where homogeneity and isotropy are implied and compared with the ΛCDM model. In the background cosmology of the VIDF model, it is noted that the era of dust matter-dominated epoch is much longer in comparison to the ΛCDM and also that the radiation-dust matter equality occurs much earlier, while dust-dark energy density equality occurs at a later cosmic time. Furthermore, the transition point from decelerated expansion to accelerated expansion occurs much later than the one anticipated by the standard model. Both models seem to agree when it comes to accelerating expansion of the universe prediction.

MCMC simulation was adopted in constraining the model's background cosmological parameters using supernova data. The results obtained by constraining the cosmological parameters of the model were more in agreement with the recent cosmological data given by Planck's 2018 data. Although, the were more models studied, we did not consider them for the proceeding since some of them were not given results at all while others did not give comparable results to that of Plank's 2018 by overestimating. However, the results obtained are preliminary, as more data (such as BAO, CMB, OHD, R22 ect.) have to be tested on the model in order to see how the model fits the observation. The linear cosmological perturbations were also investigated in a dust matter-dominated frame using a $1+3$ covariant approach. It was evident from the perturbation results obtained that the model predicts the rip of large-scale structures at the late-time universe, and this comes mainly from the interaction of dark-fluid components. Future work will be to use different cosmological data to constrain the background parameters and large-scale structure even better.

References

[1] Bamba K, Capozziello S, Nojiri S and Odintsov S D 2012 *Astrophysics and Space Science* **342** 155–228
[2] Copeland E J, Sami M and Tsujikawa S 2006 *International Journal of Modern Physics D* **15** 1753–1935
[3] Shi K, Huang Y and Lu T 2012 *Monthly Notices of the Royal Astronomical Society* **426** 2452–2462
[4] Fabris J C, Goncalves S and Ribeiro R d S 2006 *General Relativity and Gravitation* **38** 495–506
[5] Bolotin Y L, Kostenko A, Lemets O A and Yerokhin D A 2015 *International Journal of Modern Physics D* **24** 1530007
[6] Chimento L P, Jakubi A S and Pavón D 2000 *Physical Review D* **62** 063508
[7] Astorga-Moreno J, Chagoya J, Flores-Urbina J and García-Aspeitia M A 2019 *Journal of Cosmology and Astroparticle Physics* **2019** 005
[8] Garcia-Aspeitia M A, Hernandez-Almada A, Magana J, Amante M H, Motta V and Martínez-Robles C 2018 *Physical Review D* **97** 101301
[9] García-Aspeitia M A, Martínez-Robles C, Hernández-Almada A, Magaña J and Motta V 2019 *Physical Review D* **99** 123525
[10] Hernandez-Almada A, Magaña J, García-Aspeitia M A and Motta V 2019 *The European Physical Journal C* **79** 1–9
[11] Abebe A, Abdelwahab M, De la Cruz-Dombriz A and Dunsby P K 2012 *Classical and quantum gravity* **29** 135011
[12] Ellis G F, Maartens R and MacCallum M A 2012 *Relativistic cosmology* (Cambridge University Press)
[13] Hough R, Sahlu S, Sami H, Elmardi M, Swart A M and Abebe A 2021 *arXiv preprint arXiv:2112.11695*
[14] Sahlu S, Ntahompagaze J, Elmardi M and Abebe A 2019 *The European Physical Journal C* **79** 1–31
[15] Brevik I, Obukhov V and Timoshkin A 2015 *Astrophysics and Space Science* **355** 399–403
[16] Brevik I, Myrzakulov R, Nojiri S and Odintsov S D 2012 *Physical Review D* **86** ISSN 1550-2368 URL `http://dx.doi.org/10.1103/PhysRevD.86.063007`
[17] van der Westhuizen M A and Abebe A 2023 Interacting dark energy: clarifying the cosmological implications and viability conditions (*Preprint* `2302.11949`)
[18] Chanda A, Roy B C, Bamba K and Paul B C 2023 *arXiv preprint arXiv:2309.09158*
[19] Benetti M, Borges H, Pigozzo C, Carneiro S and Alcaniz J 2021 *Journal of Cosmology and Astroparticle Physics* **2021** 014

First-principles study on the effect of Pt addition on the stability of B2 $Ti_{50}Ru_{50}$ – a supercell approach

BS Ngobe[1, 2], MJ Phasha[1], MP Molepo[2] and Y Yamabe-Mitarai[3]
[1] Advanced Materials Division, MINTEK, Randburg, South Africa
[2] School of Physics, University of the Witwatersrand, Johannesburg, South Africa
[3] Graduate School of Frontier Sciences, University of Tokyo, Kashiwanoha, Japan

E-mail: **Bonganing@mintek.co.za and 2623288@students.wits.ac.za**

Abstract. Innovative structural materials that are capable of meeting the constantly changing demands in engineering can be accomplished by integrating computational thermodynamics with validated experimental techniques to enhance existing materials. One example of such material is Nitinol ($Ti_{50}Ni_{50}$), a well-known shape memory alloy (SMA). The distinctive aspect of Nitinol emanates from its ability to undergo martensitic phase transformation (MPT), from a high-temperature B2 phase to a low-temperature monoclinic (B19') martensite phase. Despite the outstanding shape memory effect (SME) and excellent super elasticity, Nitinol is only used in operating temperatures below 373 K. This demonstrates the need for innovative SMAs able to overcome Nitinol's application temperature limitations. Such SMAs are classified as high-temperature shape memory alloys (HTSMAs). The B2 Ti-PGM-based compounds are known for their high martensitic transformation temperatures (MTT). Consequently, in this study, the effect of Pt addition on the stable B2 $Ti_{50}Ru_{50}$ is investigated using first-principles calculations to stimulate the martensitic transformation. This study showed that the systematic replacement of Ru with Pt atoms can induce a martensitic phase transformation from the B2 $Ti_{50}Ru_{50-x}Pt_x$. In particular, B2 $Ti_{50}Ru_{50-x}Pt_x$ with Pt composition ≥18.75 at. % where $C' \lesssim 0$ has been identified as ternary alloy compositions with possible martensitic phase transformation.

1. Introduction

The increasing interest in shape memory alloys (SMAs) containing platinum group metals (PGMs) can be attributed to their versatile applications in high-temperature areas such as aerospace, automotive, power plants, and chemical industries [1]. This talks about the need for new and futuristic structural materials to meet the revolution of the 21st century. For instance, the expensive Ti-PGMs base compounds like TiPt and TiPd, have emerged as promising materials for the development of high-temperature shape memory alloys (HTSMAs) [2-3]. The outstanding structural behaviour of SMA is driven by the martensitic phase transformation that occurs between the high-symmetry austenite phase to low-symmetry martensitic phase(s) on cooling from high to low-temperature region [1, 4-5]. This is due to their ability to transform from a higher temperature phase cubic B2 to lower temperature phases such as $L1_0$/B19/B19' at temperatures higher than 373 K [4-7], surpassing the transformation temperature of well-known SMA, the TiNi alloys [1, 6]. Nitinol's low transformation temperature has adversely hindered the wider applications of this advanced smart alloy at high temperatures.

The aforementioned SMA features are derived from an ordered CsCl-type intermetallic phase, which forms instantly from the molten state of nearly equiatomic chemical composition [1, 6-7] and transforms to martensite phase upon cooling to lower temperatures [2, 4, 7-8].

Unlike compounds with shape memory effect, certain Ti-PGM-based compounds such as B2 of the TiRu system [4], exhibit similar ordered B2 phase at high temperatures but remain stable even at lower temperatures. Thus hindering it from undergoing phase transition to low-temperature phases. Amongst Ti-PGM-based compounds, the B2 of TiRu should be comparatively less expensive as Ru is the cheapest among PGM metals.

In this study, we report work carried out using first-principles calculations to pin-track B2-$Ti_{50}Ru_{50-x}Pt_x$ ternary compositions with the prospect of stimulating SME on stable B2-$Ti_{50}Ru_{50}$ by substituting some of Ru atoms with Pt atoms through supercell approach. In pursuit of such advanced structural materials with SME, we evaluated the thermodynamic, mechanical, electronic and lattice dynamical stability.

2. Computational methods

The computational predictions reported here were carried out using first-principles calculations based on the density functional theory (DFT) formalism as implemented in Cambridge Serial Total Energy Package (CASTEP) code [9-10]. The ultra-soft pseudopotentials (USSPs) were used to model the electron-ion interactions [10]. The electron-exchange correlation was described by the Perdew-Burke-Ernzerhof (PBE) functional of the generalized gradient approximation (GGA) [11]. A satisfactory level of convergence for the total energy of the reference binary $Ti_{50}Ru_{50}$ and $Ti_{50}Pt_{50}$ B2 unit cells was achieved with an energy cut-off of 500 eV and a 13×13×13 k-point mesh. Meanwhile, for the B2 ternary $Ti_{50}Ru_{50-x}Pt_x$ system in a 2×2×2 supercell, we utilized a 7×7×7 k-point grid. This 2×2×2 supercell comprised of 16 atoms represents the crystal structure of B2 $Ti_8Ru_{8-x}Pt_x$ equivalent to $Ti_{50}Ru_{50-x}Pt_x$.

All the equilibrium B2 crystal structures were obtained through geometry optimization in the Brayden-Fletcher-Goldfarb-Shanno (BFGS) minimization scheme [12]. The convergence criterion of less than 1×10^{-5} eV/atom, the maximum residual forces of 0.03 eV/Å, the maximum residual bulk stress of 0.05GPa and the maximum atomic displacement of 1×10^{-3} Å were utilised to achieve maximum accuracy.

Figure 1 (a) – (c) shows the B2 crystal structures of (a) $Ti_{50}Ru_{50}$ unit cell, 2×2×2 supercell comprised of 8 unit cells to represent (b) $Ti_{50}Ru_{50}$ and (c) $Ti_{50}Ru_{50-x}Pt_x$ compositions, respectively, that were used to investigate the structural and thermodynamic, mechanical, electronic and vibrational stability of the crystal compounds reported in this research work.

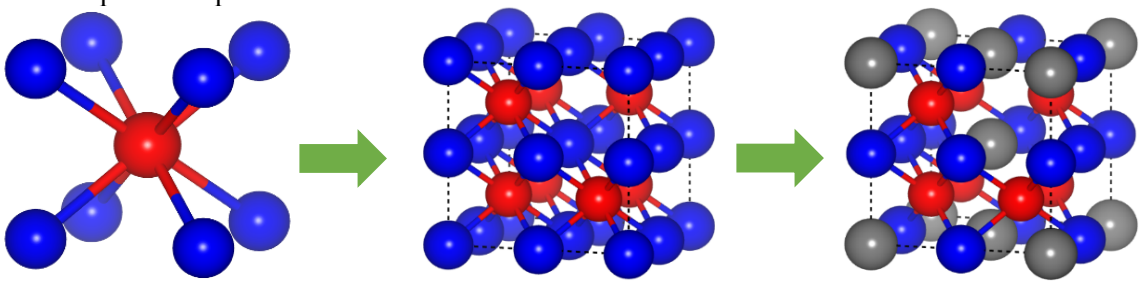


Figure 1. Schematic representation of B2 (a) unit cell crystal geometry of TiRu, its (b) 2×2×2 binary and resultant (c) ternary supercell structures used in this study work, with red, blue and grey depicting Ti, Ru and Pt atoms respectively.

3. Results and discussion

3.1. Structural and thermodynamic properties

The obtained structural equilibrium lattice parameters and thermodynamic characteristics of the investigated compounds are presented in Table 1. The presented ground state results for the benchmark binary alloys were found to agree well with those reported by other researchers [2, 5, 13-14]. Additionally, the outcomes depicted in Table 1 indicate that there is an increase in lattice parameters as

the concentration of Pt increases. This is due to the atomic radius of platinum (1.39 Å) which is larger than that of ruthenium (1.34 Å). The lattice parameters of the B2 compounds reported in this study are consistent with those reported by other researchers, namely 3.08 and 3.19 Å for $Ti_{50}Ru_{50}$ and $Ti_{50}Pt_{50}$ [2, 5], respectively.

Table 1. Structural parameters of the investigated B2 crystal geometry.

Crystal structure	Lattice parameters, a (*Å*)		Heats of formation, (ΔH_F) eV.atom^{-1}	
	This work	**Literature**	**This work**	**Literature**
$\mathbf{Ti_{50}Ru_{50}}$	3.08	3.09[2]	-0.745	-0.770[13], -0.743[14]
$\mathbf{Ti_{50}Ru_{43.75}Pt_{6.25}}$	3.09	-	-0.750	-
$\mathbf{Ti_{50}Ru_{37.5}Pt_{12.5}}$	3.11	-	-0.740	-
$\mathbf{Ti_{50}Ru_{31.25}Pt_{18.75}}$	3.12	-	-0.739	-
$\mathbf{Ti_{50}Ru_{25}Pt_{25}}$	3.13	-	-0.770	-
$\mathbf{Ti_{50}Ru_{18.75}Pt_{31.25}}$	3.15	-	-0.761	-
$\mathbf{Ti_{50}Ru_{12.5}Pt_{37.5}}$	3.16	-	-0.789	-
$\mathbf{Ti_{50}Ru_{6.25}Pt_{43.75}}$	3.17	-	-0.827	-
$\mathbf{Ti_{50}Pt_{50}}$	3.18	3.19[5]	-0.862	-0.825[4]

The ability of a compound to exist in a stable phase can be inferred from the heats of formation also referred to as enthalpies of formation, as described in Equation 1 [14]. These enthalpies indicate a compound's thermodynamic potential to undergo chemical formation. In the case of a phase being stable, ΔH_F should have the lowest value. Conversely, it becomes unstable with a positive ΔH_F at 0 K [15].

$$\Delta H_F = (E_C - \sum_i x_i E_i), \quad (1)$$

where E_C represent the total energy of the investigated compound and E_i represents the elemental total energies of their ground-state crystal structures at x_i at. % fraction concentration. As presented in Table 1, all the investigated compounds were found to be thermodynamically stable ($\Delta H_F < 0$), with pure $Ti_{50}Pt_{50}$ (-0.83 eV/atom) found to be the most stable. Initially, the stability was found to decrease with the addition of Pt atoms less than 25 at. %, and increases thereafter with reference to $Ti_{50}Ru_{50}$. Again, the enthalpies results reported in this study were found to be in accordance with results reported by other researchers [5, 14], -0.825 and -0.743 eV/atom for $Ti_{50}Pt_{50}$ and $Ti_{50}Ru_{50}$, respectively.

3.2. Mechanical stability

The mechanical and dynamic behaviour of any crystal can be deduced from the elastic constants (C_{ij}), which are fundamental parameters of first-principles calculations [16], in which, both mechanical and physical properties of materials can be estimated. Cubic crystal consists of the simplest cubic form of a

stiffness matrix, where the number of the independent elastic constants is reduced to three (C_{11}, C_{12} and C_{44}) in the Voigt-Reuss-Hill (VRH) notation, and the shear elastic coefficient C' for assessing the compound's prospect to undergo a phase transition (to ascertain stability or instability) at lower temperatures can be expressed as shown in Equation 2:

$$C' = \left(\frac{C_{11}-C_{12}}{2}\right). \tag{2}$$

As per the dynamical theory proposed by Born-Huang [17], the mechanical stability criteria applicable to B2 compounds can be established by comprehensively meeting the criteria outlined by the set of expressions in Equation 3:

$$C_{11} > 0, C_{44} > 0, C_{11} - C_{12} > 0 \;\; C_{11} > |C_{12}|, C_{11} + 2C_{12} > 0. \tag{3}$$

The crystal's mechanical stability weakens as it approaches the phase transition point [18]. This occurs when the tetragonal shear modulus (C') becomes smaller than the monoclinic shear constant (C_{44}), which characterizes the crystal's resistance against shearing. Figure 2 presents the elastic parameters of the investigated compounds. It can be noted in Figure 2 for pure $Ti_{50}Ru_{50}$ and $Ti_{50}Ru_{50-x}Pt_x$ ternaries with Pt compositions less than 18.75 at. % satisfied all the mechanical stability criteria as detailed in Equation 3, while ternaries with Pt composition greater and equal to 18.75 at. % adheres to the criteria ($C_{11} < C_{12}$). The aforementioned signifies that pure $Ti_{50}Ru_{50}$ and ternaries with Pt < 18.75 at. % do not undergo MPT, and subsequently cannot be regarded as SMA. While ternaries with Pt ≥ 18.75 will undergo MPT and are likely to be regarded as HTSMA due to their negative C'. This is similar to what has been observed on the $Ti_{50}Pt_{50}$ compound ($C'<0$) [19] which is inclined to undergo a phase transition at much higher temperatures [3]. Figure 2, further shows as Pt composition ≳ 12.5 at. %, C_{44} become larger than C', signifying phase transition is approached.

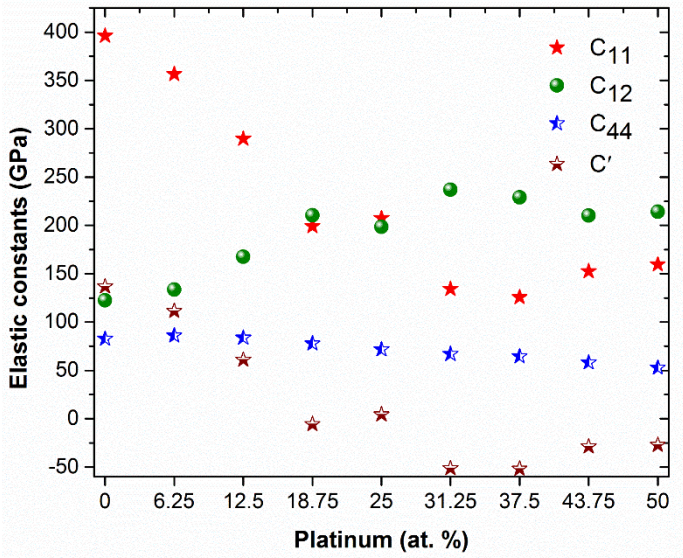


Figure 2. The elastic constants as well as their calculated modulus of elasticity of the investigated B2 compounds

3.3. Electronic and vibrational stability

Figures 3 and 4 represent the electronic properties of the studied compounds, the total density of states and their corresponding phonon dispersion curves plotted along the selected Brillouin zone (BZ), respectively, computed at 0K.

The total density of states (TDOS) is used to study the electronic stability by examining the Fermi level's position ($E-E_F = 0$) relative to the pseudogap. To achieve phase stability, the valence and the conduction bands should coincide perfectly at the Fermi level by cutting at the centre of the pseudogap and contrary to that results in phase instability and subsequent phase transition [20-23].

The phonon dispersion curves provide a vital analysis of the lattice vibrational stability of compounds at their ground state. A compound is considered vibrational stable at 0 K if there are no soft or negative vibrational frequencies along high symmetry directions in the BZ [24]. On the contrary, the presence of negative vibrational frequencies indicates the lattice vibrational instability of the crystal, an indication of the likelihood to undergo a phase transition, driven by the displacement of atoms from original lattice positions in the austenite high-temperature phase [24-25].

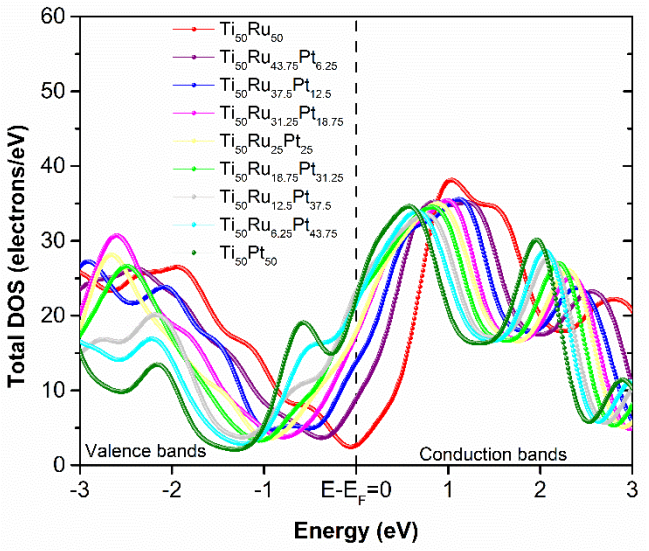
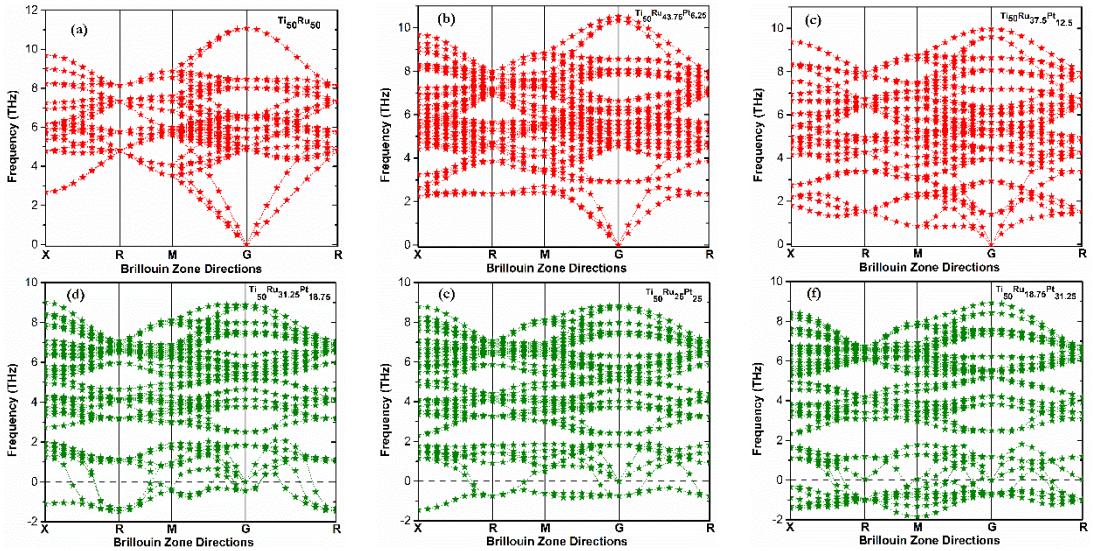


Figure 3. Total density of states (TDOS) of the investigated B2 compounds.



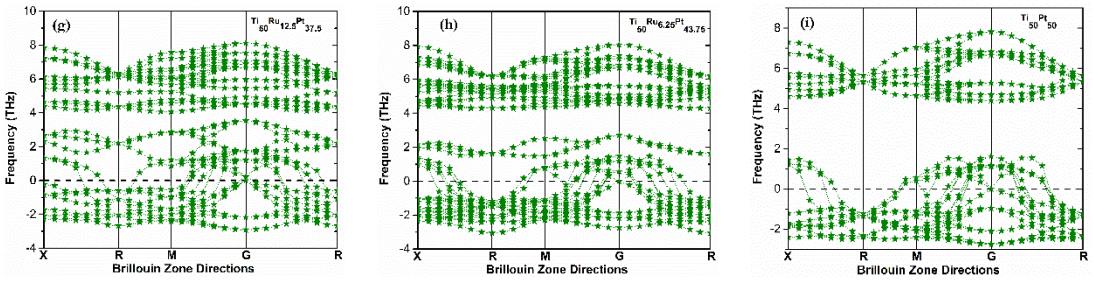


Figure 4. Phonon dispersion curves of the benchmark B2 binary and $Ti_{50}Ru_{50-x}Pt_x$ ternary compounds plotted along selected Brillouin zone.

Figure 3 shows that B2 $Ti_{50}Ru_{50}$ remain stable, as indicated by its Fermi level cutting the TDOS in the pseudogap (deep valley). It is noted that partial substitution of Ru atoms with Pt atoms results in shifting the pseudogap towards the valence bands (left region), resulting in the instability of such resulting compositions at 0 K, similar to that of pure $Ti_{50}Pt_{50}$.

Figure 4 (a) – (i) shows the selected high symmetry points, i.e., X-R-M-G-R-X, in the BZ of the investigated B2 compounds. Figure 4 shows that ternary compounds with less than 18.75 at. % Pt (b and c) was found to possess only positive frequencies similar to that of the $Ti_{50}Ru_{50}$ (a) compound, indicating no phase transformation. This positive vibrational frequency on these compounds is an indication of rigidness observed only on the gamma (G) point. The phonon dispersion curves of ternaries with 18.75 at. % Pt and above were found to have both positive and negative vibrational frequencies, this vibrational frequencies can be observed on all the selected BZ symmetry-points. Thereby making the lattice vibrational dynamics of these compounds unstable at 0 K, an indication of possible phase transformation.

4. Conclusions

This study successfully determined the thermodynamic, electronic, mechanical, and lattice dynamic characteristics of binary and ternary B2 $Ti_{50}Ru_{50-x}Pt_x$ compositions using the first-principles technique.

The predicted lattice parameters and heats of formation for binary B2 $Ti_{50}Ru_{50}$ and $Ti_{50}Pt_{50}$ are in excellent agreement with the results reported in the literature, attributed to the accuracy of the computational parameters used. The calculated heats of formation reveals that the addition of Pt increased the thermodynamic stability of the B2 $Ti_{50}Ru_{50}$.

From a mechanical stability perspective, all compounds with Pt composition <18.75 at. % satisfy the Born-Huang criteria ($C' > 0$). For Pt compositions ≥ 18.75 at. %, C_{44} surpasses C', signifying possible phase transition.

Results from the total density of states show that the substitution of Ru with Pt atoms shifted the position of the Fermi level towards the conduction band region. While the phonon spectrum of the stable B2 $Ti_{50}Ru_{50}$ consists of only positive frequencies, the ternary compounds with Pt content ≥ 18.75 at. % were found to have both positive and negative vibrational frequencies, indicating a potential martensitic transformation.

In summary, this study has shown that the systematic replacement of Ru with Pt atoms can induce martensitic phase transformation from the B2 $Ti_{50}Ru_{50-x}Pt_x$. In particular, B2 $Ti_{50}Ru_{50-x}Pt_x$ with Pt composition ≥18.75 at. % at points where $C' \lesssim 0$, has been identified as ternary alloy compositions with possible martensitic phase transformation.

Acknowledgements

This paper is published with the permission of MINTEK and The University of the Witwatersrand. The authors would like to thank the Advanced Metals Initiative (AMI) of the Department of Science and Innovation (DSI) for financial support. In addition, Dr Maje Phasha is grateful for financial support from the National Research Foundation (NRF) of South Africa - JSPS GRANT No: 148782. Lastly, the

gratitude is also extended to the Centre for High-Performance Computing (CHPC) in Cape Town for allowing us to carry out the calculations using their remote computing resources.

References

[1] Wadood A 2016 *Adv. Mate. Sci. Eng*. **2016** 9
[2] Kong ZZ, Duan YH, Peng MJ, Qu DY and Bao LK 2019 *Physica. B: Cond. Matter*. **573** 21
[3] Yamabe-Mitarai Y, Hara T, Miura S and Hosoda H 2006 *Mater. Trans.* **47** 657
[4] Bao W, Liu D, Duan Y and Peng M 2019 *Philos. Mag*. **99** 22
[5] Jung J, Gosh G, and Olson G 2003 *Acta Mater.* **51** 6341
[6] Donkersloot H and Vucht JV 1970 *J. Less-Common. Met*. **20** 91
[7] Xing WW, Chen XQ, Li DZ, Li LY, and Fu CL 2012 *Intermetallics* **28** 24
[8] Jain E, Pagare G, Chouhan SS and Sanyal SP 2014 *Intermetallics* **54** 184
[9] Clark SJ, Segall MD, Philip JDL, Pickard CJ, Hasnip PJ, Probert MJ, Refson K and Payne MC 2005 *Z. Kristallogr*. **220** 567
[10] Blöchl PE 1994 *Phys. Rev*. B **50** 17953
[11] Perdew JP and Wang Y 1992 *Phys. Rev. B* **45** 13249
[12] Fischer T and Almlof J 1992 *J. Phys. Chem*. **96** 9774
[13] Semenova E 1997 *Powder. Metall. MCeram*. **36** 404
[14] John R and Ruben H 2011 *Mater. Sci. Appl*. **2** 1366
[15] Gachon JG, Selhaoui N, Aba B and Hertz J 1992 *J. Phase Equilibria* **13** 511
[16] Wu SC, Fecher GH, Naghavi SS and Felser C 2019 *J. Appl. Phys*. **128** 18
[17] Gomisa O, Manjónb FJ, Rodríguez-Hernándezc P and Muñozc A 2019 *J. Phys. Chem. Solids*, **124** 120
[18] Sedlák P, Janovská M, Bodnárová L, Heczko O and Seiner H 2020 *Metals* **10** 1383
[19] Mahlangu R, Phasha MJ, Chauke HR and Ngoepe PE 2013 *Intermetallics* **33** 32
[20] Ngobe BS, Phasha MJ, and Mwamba IA 2021 *SAJST* **40** 205
[21] Ravindran P and Asokamani R 1997 *Bull. Mater. Sci*. **20** 613
[22] Kulkova SE and Valujsky DV 2001 *J. Phys. IV* **11** 58
[23] Kulkova SE, Valujsky DV, KPim JS, Lee G and Koo YM 2001 *Solid State Commun*. **119** 623
[24] Si L, Jiang Z-Y, Zhou B, and Chen W-Z 2012 *Phys. B* **407** 351
[25] Togo A and Tanaka I 2015 *Scri. Mater*. **108** 5

Geometric bound states in ϕ^4 theory

JF Du Plessis[1,*] **and WA Horowitz**[2]

[1] Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701, South Africa
[2] Department of Physics, University of Cape Town, Rondebosch 7701, South Africa

E-mail: *dpljea028@myuct.ac.za

Abstract. We investigate previously discovered geometric bound states, found when considering massive scalar ϕ^4 theory in a compactified spacetime with 2 spatial dimensions finite and 1 infinite. Such geometric bound states have also been observed in other finite geometries, and are expected to be a generic feature of finite volume field theories. We investigate the different behaviours of the NLO $2 \to 2$ scattering amplitude and the effective coupling in the vicinity of such geometric bound states. We find that as the geometric bound state is approached both the effective coupling and NLO amplitude decreases. We however also find that the NLO amplitude, in contrast to the effective coupling, diverges extraordinarily close to the bound state, where we make precise the meaning of "extraordinarily close".

1. Introduction

The formation of quark-gluon plasma (QGP) has been a central topic in heavy-ion physics. While traditionally associated with large nucleus-nucleus (A+A) collisions, recent observations indicate its potential formation in high multiplicity proton-proton (p+p) and proton-nucleus (p+A) collisions. These findings, stemming from data at the Large Hadron Collider (LHC) [1, 2, 3, 4], have been supported by the effective description of low momentum particle distributions in these systems using relativistic, nearly inviscid hydrodynamics [5, 6]. The employed (effectively infinite system size) equation of state, derived from lattice Quantum Chromodynamics (QCD) [7], suggests that the medium seemingly formed in these collisions is akin to the nearly inviscid QGP observed in larger systems.

Recent studies have highlighted the significance of finite system size effects. In particular, a massless free scalar thermal field theory with Dirichlet boundary conditions revealed that these effects can introduce substantial corrections to thermodynamic properties, potentially mimicking the temperature-dependent behavior of full QCD [8]. These corrections are non-trivial, with values reaching up to 40% for p+p collisions and around 10% for mid-central nucleus-nucleus collisions. Quenched lattice QCD calculations with periodic boundary conditions have underscored the significance of finite system size effects, particularly evident in systems exhibiting asymmetric finite lengths [9].

The equation of state, encompassing properties such as the speed of sound and the trace anomaly, is fundamental for hydrodynamic simulations of relativistic collisions. A notable discovery is that a free massless scalar field theory, even with the breaking of conformal symmetry due to Dirichlet boundary conditions, yields a traceless energy-momentum tensor [10]. When finite-system-size-like corrections were incorporated into the QCD coupling, a significant

reduction in the trace anomaly was observed [10]. This has implications for the extraction of shear and bulk viscosities from hydrodynamic simulations. This also underscored the importance of understanding the system size dependence of the coupling in order to calculate finite system size corrections to the EOS.

In previous work the effective coupling in ϕ^4 theory in a spacetime with periodic boundary conditions on different numbers of the spatial dimensions was investigated [11]. The existence of geometric bound states, where the geometry of the compactified spacetime induces a pole in the S-matrix was discovered. The existence of the geometric bound states were found to be coupling-independent, but depended on the number of spacetime dimensions compactified, as well as their (potentially asymmetric) characteristic lengths. It was observed that, although the NLO amplitude was infinite *at* the geometric bound states, the effective coupling decreases near it, and is identically 0 at the bound state. We note that we will use "at the geometric bound state" to mean evaluated on arguments corresponding to a bound state.

This seeming contradiction is explored in this work. We show that in a spacetime with 2 finite spatial dimensions of equal characteristic lengths the amplitude does indeed decrease near the geometric bound state, and only diverges to extraordinarily close to the momentum needed to form the bound state.

2. Effective coupling

By a resummation of bubble diagrams the ϕ^4 effective coupling was shown to be

$$-i\lambda_{\text{eff}}(s,t,\{L_i\}) = \frac{-i\lambda(\mu)}{1-\lambda(\mu)\big(\overline{V}(s,\{L_i\})+\overline{V}(t,\{L_i\})+\overline{V}(u,\{L_i\})\big)} \tag{1}$$

in [11], where, suppressing summation bounds over i,

$$\overline{V}_n(p^2,\{L_i\}) = \\ -\frac{1}{32\pi^2}\int_0^1 dx\left\{\ln\frac{\mu^2}{m^2-x(1-x)p^2-i\varepsilon}+2\sum_{\vec{k}\in\mathbb{Z}^n}{}' K_0\left(2\pi\sum L_i k_i\sqrt{m^2-x(1-x)p^2-i\varepsilon}\right)\right\} \tag{2}$$

in a finite system with n periodic boundary conditions of length scales $\{L_i\}$ [12]. We denote the sum excluding the origin by $\sum'$ and K denotes the modified Bessel function of the second kind. With two finite dimensions, one finds that the s channel contribution is numerically ill-behaved, and as such another approach is needed. Using Eqs. (A.5) and (A.17) with a square lattice we can then write

$$\overline{V}_2(s,L;\mu) = -\frac{1}{32\pi^2}\left[\ln\left(\frac{\mu^2}{m^2}\right)+a_2(Lm) \right. \\ \left. +\frac{1}{\pi}\sum_{l=0}^{\infty} r_2(l)\left(\frac{\frac{4}{L}\arcsin\left(\frac{L}{2}\sqrt{\frac{s+i\varepsilon}{(Lm)^2+l}}\right)}{\sqrt{s+i\varepsilon}\sqrt{4l-L^2(s-4m^2+i\varepsilon)}}-\frac{1}{l+(Lm)^2}\right)\right]. \tag{3}$$

As found in [11], one finds the existence of geometric bound states. When $L^2(s-4m^2)/4$ is close to an l^* with $r_2(l^*)\neq 0$, then one can see that Eq. (3) must diverge. Let us then introduce the dimensionless $\sigma \equiv L^2(s/4-m^2)$, such that we get $L^2 s = 4(\sigma+(Lm)^2)$. We can also consider then the dimensionless $\alpha \equiv Lm$, giving

$$\overline{V}_2(\sigma,\alpha;\mu) = -\frac{1}{32\pi^2}\left[\ln\left(\frac{\mu^2}{m^2}\right)+a_2(\alpha)+\frac{1}{\pi}\sum_{l=0}^{\infty} r_2(l)\left(\frac{\arcsin\left(\sqrt{\frac{\alpha^2+\sigma}{\alpha^2+l}}\right)}{\sqrt{\sigma+\alpha^2}\sqrt{l-\sigma}}-\frac{1}{l+\alpha^2}\right)\right], \tag{4}$$

where we now implicitly require $0 < \mathrm{Im}\,\sigma \ll 1$. If we then consider some $|\delta| \ll 1$ with $0 < -\mathrm{Im}\,\delta \ll |\delta|$ and l^* with $r_2(l^*) \neq 0$. The imaginary part of δ is simply the pole prescription, so outside of determining branch cuts we can safely treat δ as a real number. We can then take $\sigma = l^* - \delta$, which gives

$$\overline{V}_2(l^* - \delta, \alpha; \mu) \approx \text{const.} - \frac{1}{64\pi^2} \frac{r_2(l^*)}{\sqrt{l^* + \alpha^2}\sqrt{\delta}}. \tag{5}$$

In the perturbative (in infinite volume) regime we can then see that by recognising that the other contributions times λ must be small, giving

$$\lambda_{\text{eff}} \approx \frac{64\pi^2\lambda}{64\pi^2 + \lambda \frac{r_2(l^*)}{\sqrt{l^*+\alpha^2}\sqrt{\delta}}}, \tag{6}$$

and similarly the NLO amplitude will have

$$64\pi^2\mathcal{M} \approx -\lambda \left[64\pi^2 - \lambda \frac{r_2(l^*)}{\sqrt{l^* + \alpha^2}\sqrt{\delta}}\right]. \tag{7}$$

We can see that for positive δ, both $|\lambda_{\text{eff}}|$ and $|\mathcal{M}|$ will decrease as δ decreases, until

$$\delta \approx \frac{\lambda^2}{2^{12}\pi^4} \frac{r_2^2(l^*)}{l^* + \alpha^2}, \tag{8}$$

where the magnitude of the amplitude will start diverging, while the magnitude of the effective coupling will continue to go to zero. Close, but not too close, to the geometric bound state, we therefore get that the NLO amplitude also decreases, as the effective coupling was observed to do in [11]. In more quantitative terms we then have that the amplitude will diverge only when $L^2|p_*^2 - p^2| \lesssim \frac{\lambda^2}{2^{12}\pi^4} \frac{r_2^2(l^*)}{l^*+\alpha^2}$ where p^* denotes the geometric bound state momentum. Since $r_2^2(l^*)/l^*$ acquires its maximum of 16 at $l^* = 1$, this quantity will be $< 10^{-4}\lambda^2$.

3. Conclusions

We calculated the asymptotic form of the s channel divergence near a geometric bound state. This asymptotic result was then further used to investigate the behaviour of the NLO amplitude and effective coupling near the geometric bound state. We found that the NLO amplitude decreases as the geometric bound state is approached, and only diverges extraordinarily close to the momentum needed to form the bound state. The effective coupling decreases as the geometric bound state is approached, and is identically zero at the bound state. We note that the amplitude diverges only in the region where the one-loop Feynman diagrams would have an order one contribution, and as such the amplitude is not expected to be well described by the perturbative expansion in this region. This work therefore showed that the decrease in the effective coupling near the geometric bound state is not in contradiction with a seemingly infinite amplitude *at* the geometric bound state. To better understand these geometric bound states an exact exploration akin to that captured in the Bethe-Salpeter equation is expected to be necessary. This is left for future work. We note that the NLO s channel divergence is not of the form assumed in the derivation of the Bethe-Salpeter equation, and as such it is expected that the Bethe-Salpeter equation will need to be modified to account for this.

Acknowledgments

WAH wishes to thank the South African National Research Foundation and the SA-CERN Collaboration for support and New Mexico State University for its hospitality, where part of the work was completed. JFDP wishes to thank the SA-CERN Collaboration for financial support. The authors wish to thank Matt Sievert, Alexander Rothkopf, Bowen Xiao and Jonathan Shock for valuable discussions.

Appendix A. s Channel dispersion relation

We are interested in evaluating the s channel contribution to the NLO amplitude

$$\overline{V}_2(s,\{L_i\},\mu) = -\frac{1}{32\pi^2}\int_0^1 dx\Bigg\{\ln\frac{\mu^2}{m^2-x(1-x)s-i\varepsilon} + 2\sum_{\vec{k}\in\mathbb{Z}^2}{}' K_0\left(2\pi\sum L_i k_i\sqrt{m^2-x(1-x)s-i\varepsilon}\right)\Bigg\}, \tag{A.1}$$

which in its current form is numerically ill-behaved. For simplicity, let us define

$$f(\sigma,\varepsilon) = \int_0^1 dx\ln\left(\frac{1}{1-x(1-x)\sigma-i\varepsilon}\right) \tag{A.2}$$

$$g_2(\Lambda,\sigma,\varepsilon) = f(\sigma,\varepsilon) + 2\sum_{\vec{k}\in\Lambda}{}'\int_0^1 dx K_0\left(2\pi\|\vec{k}\|\sqrt{1-x(1-x)\sigma-i\varepsilon}\right) \tag{A.3}$$

where Λ is an two-dimensional lattice. We note that using the dominated convergence theorem [13] one can show that, as long as there are no $\vec{\ell}\in\Lambda^*$ (where Λ^* is the dual lattice) with $\ell^2=\frac{\sigma}{4}-1$, one can safely take both the $\sigma\to\infty$ or lattice spacings of Λ to ∞ limits using that $K_0\sim e^{-x}/\sqrt{x}$ to show that $g_2(\Lambda,\sigma,\varepsilon)\sim f(\sigma,\varepsilon)$ in these limits. Furthermore, it is possible to show that $f(\sigma,0^+)\sim 2+i\pi-\ln(\sigma)$ for large σ. Eqs. (A.2) and (A.3) then allow us to write Eq. (A.1) as

$$\overline{V}_2(s,\{L_i\},\mu) = -\frac{1}{32\pi^2}\left[\ln\left(\frac{\mu^2}{m^2}\right)+g_2\left(\Lambda(\{L_i m\}),\frac{s}{m^2},\varepsilon\right)\right], \tag{A.4}$$

where $\Lambda(\{L_i m\})$ is the 2-dimensional rectangular lattice with lattice spacings given by the $L_i m$. In the special case of all finite lengths being equal,

$$\overline{V}_2(s,L,\mu) = -\frac{1}{32\pi^2}\left[\ln\left(\frac{\mu^2}{m^2}\right)+g_2\left(\Lambda(Lm),\frac{s}{m^2},\varepsilon\right)\right], \tag{A.5}$$

where $\Lambda(Lm)$ is then the two-dimensional square lattice with all lattice spacings equal to Lm.

Imaginary part

We will try to exploit the complex structure of these functions to find a numerically well behaved equivalent form for Eq. (A.1). We have chosen to self-consistently take the arg, log, K_0, and the square root branch cuts along the negative real axis. This straightforwardly gives that both f and g_2 have a branch cut along the positive real axis from $\sigma=4$ to $\sigma=\infty$. The real part of both f and g_2 is continuous across the branch cut, and the imaginary part changes sign. We further need to make the assumption that there are no poles for g_2 in the complex σ plane off the positive real axis. This assumption seems reasonable and is supported by numerics. We do however note that there are physical situations where the S-matrix has poles violating this assumption [14]. Something we can easily calculate and evaluate is the imaginary part of f as follows:

$$\begin{aligned}\operatorname{Im} f(\sigma,\varepsilon) &= \int_0^1 dx \operatorname{Im}\ln\left(\frac{1}{1-x(1-x)\sigma-i\varepsilon}\right)\\ \operatorname{Im} f(\sigma,0^+) &= \pi\sqrt{1-\frac{4}{\sigma}},\end{aligned} \tag{A.6}$$

where ε helped us avoid the branch cut and choose the sign of the imaginary part. We will use 0^+ to denote that our ε is some infinitesimal positive real number, but since ε only selects a branch cut we will neglect to write it. This does mean we need to be continually mindful of

branch cuts. We can then continue onward to calculate

$$I \equiv \mathrm{Im} \int_0^1 dx K_0 \left(2\pi\|\vec{k}\|\sqrt{1-x(1-x)\sigma - i\varepsilon}\right)$$

We again use the ε to avoid the branch cut and indicate the sign of the imaginary part, giving us the following after dropping $\mathcal{O}(\varepsilon)$

$$I = \pi \int_{\frac{1}{2}}^{\frac{1}{2}+\frac{1}{2}\sqrt{1-\frac{4}{\sigma}}} dx J_0 \left(2\pi\|\vec{k}\|\sqrt{x(1-x)\sigma - 1}\right), \tag{A.7}$$

where J denotes the Bessel function of the first kind. We can further simplify Eq. (A.7) by using the substitution $y \equiv (2x-1)/\sqrt{1-4/\sigma}$ as well as $\int_0^1 dy\, J_0(a\sqrt{1-y^2}) = \mathrm{sinc}(a)$ to find

$$I = \frac{\pi}{2}\sqrt{1-\frac{4}{\sigma}}\,\mathrm{sinc}\left(2\pi\|\vec{k}\|\sqrt{\frac{\sigma}{4}-1}\right). \tag{A.8}$$

We can combine Eqs. (A.6) and (A.8) to find

$$\mathrm{Im}\, g_2(\Lambda,\sigma,0^+) = \pi\sqrt{1-\frac{4}{\sigma}}\sum_{\vec{k}\in\Lambda}\mathrm{sinc}\left(2\pi\sqrt{\frac{\sigma}{4}-1}\|\vec{k}\|\right) \tag{A.9}$$

which now allows us to use the Poisson summation formula in order to write Eq. (A.9) as

$$\mathrm{Im}\, g_2(\Lambda,\sigma,0^+) = \frac{|\Lambda^*|}{\sqrt{\sigma}}\sum_{\vec{\ell}\in\Lambda^*}\frac{\theta\left(\frac{\sigma}{4}-1-\ell^2\right)}{\sqrt{\frac{\sigma}{4}-1-\ell^2}}, \tag{A.10}$$

as detailed in [11, 12].

For asymptotically large $|\sigma|$, we find that $\left|g_2(\Lambda,\sigma,\varepsilon)/\sigma^2\right|$ falls of faster than $|\sigma|^{-1}$. We can use this as follows. Consider the Cauchy integral formula

$$\frac{d}{d\sigma}\left(g_2(\Lambda,\sigma,0^+)\right) = \frac{1}{2\pi i}\int_\gamma dz \frac{g_2(\Lambda,z,0^+)}{(z-\sigma)^2}, \tag{A.11}$$

where γ is some small counter-clockwise contour around the pole at $z=\sigma$.

Looking at the complex structure as discussed above and shown in Fig. A1, we can see that we can deform our contour around the branch cut just under the real axis (offset due to $\varepsilon = 0^+$). Then since $g_2(\Lambda,\sigma,0)/\sigma^2$ falls off faster than σ^{-1}, the integral along γ_1 does not contribute. It's also possible to show that the integral along γ_3 does not contribute. This leaves us with γ_2, γ_4.

Using the previously discussed analytic structure we then find that:

$$\frac{d}{d\sigma}\left(g_2(\Lambda,\sigma,0^+)\right)\Big|_{\sigma=\sigma_0+i\eta} = \frac{1}{\pi}\int_4^\infty dz \frac{\mathrm{Im}\, g_2(\Lambda,z,0^+)}{(z-\sigma_0-i\eta)^2} \tag{A.12}$$

$$= |\Lambda^*|\sum_{\ell\in\Lambda^*}\int_4^\infty dz \frac{\theta\left(\frac{z}{4}-1-\ell^2\right)}{\sqrt{z}\sqrt{\frac{z}{4}-1-\ell^2}}\frac{1}{(z-\sigma_0-i\eta)^2}. \tag{A.13}$$

Integrating both sides w.r.t. σ_0 and then evaluating the integral over z then gives

$$g_2(\Lambda,\sigma+i\eta,0^+) - g_2(\Lambda,\sigma_0+i\eta,0^+)$$
$$= \frac{4|\Lambda^*|}{\pi}\sum_{\ell\in\Lambda^*}\left(\frac{\arcsin\left(\frac{1}{2}\sqrt{\frac{\sigma+i\eta}{1+\ell^2}}\right)}{\sqrt{\sigma+i\eta}\sqrt{4\ell^2-(\sigma+i\eta-4)}} - \frac{\arcsin\left(\frac{1}{2}\sqrt{\frac{\sigma_0+i\eta}{1+\ell^2}}\right)}{\sqrt{\sigma_0+i\eta}\sqrt{4\ell^2-(\sigma_0+i\eta-4)}}\right). \tag{A.14}$$

We need to carefully split up the terms, where an appropriate and natural term to add and

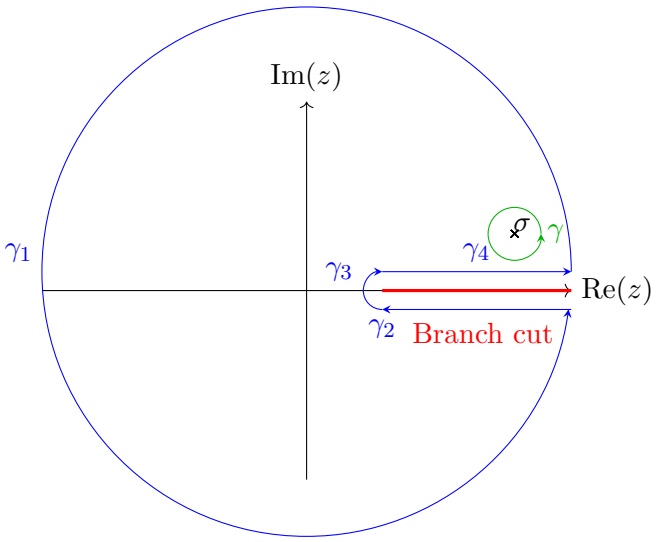


Figure A1. The analytic structure of Eq. (A.11).

subtract is $1/1+\ell^2$, which after rearranging gives

$$g_2(\Lambda,\sigma,0^+) = \left[g_2(\Lambda,\sigma_0,0^+) - \frac{|\Lambda^*|}{\pi}\sum_{\ell\in\Lambda^*}\left(\frac{2\arcsin\left(\frac{1}{2}\sqrt{\frac{\sigma_0}{1+\ell^2}}+i\eta\right)}{\sqrt{\sigma_0}\sqrt{\ell^2+1-\frac{\sigma_0}{4}-i\eta}} - \frac{1}{1+\ell^2}\right)\right] + \frac{|\Lambda^*|}{\pi}\sum_{\ell\in\Lambda^*}\left(\frac{2\arcsin\left(\frac{1}{2}\sqrt{\frac{\sigma}{1+\ell^2}}+i\eta\right)}{\sqrt{\sigma}\sqrt{\ell^2+1-\frac{\sigma}{4}-i\eta}} - \frac{1}{1+\ell^2}\right). \tag{A.15}$$

Since the left hand side is independent of σ_0, we know the first term must be a function of only the lattice. We know the asymptotic form for g_2 for large σ_0, allowing us to define

$$a_2(\Lambda) \equiv 2 + i\pi + \lim_{\sigma_0\to\infty}\left[-\ln(\sigma_0) - \frac{|\Lambda^*|}{\pi}\sum_{\ell\in\Lambda^*}\left(\frac{2\arcsin\left(\frac{1}{2}\sqrt{\frac{\sigma_0}{1+\ell^2}}+i\eta\right)}{\sqrt{\sigma_0}\sqrt{\ell^2+1-\frac{\sigma_0}{4}-i\eta}} - \frac{1}{1+\ell^2}\right)\right] \tag{A.16}$$

in order to obtain

$$g_2(\Lambda,\sigma+i\eta,0^+) = a_2(\Lambda) + \frac{|\Lambda^*|}{\pi}\sum_{\ell\in\Lambda^*}\left(\frac{2\arcsin\left(\frac{1}{2}\sqrt{\frac{\sigma}{1+\ell^2}}+i\eta\right)}{\sqrt{\sigma}\sqrt{\ell^2+1-\frac{\sigma}{4}-i\eta}} - \frac{1}{1+\ell^2}\right). \tag{A.17}$$

References

[1] Abelev B *et al.* (ALICE) 2013 *Phys. Lett. B* **719** 29–41 (*Preprint* 1212.2001)
[2] Aad G *et al.* (ATLAS) 2013 *Phys. Rev. Lett.* **110** 182302 (*Preprint* 1212.5198)
[3] Adare A *et al.* (PHENIX) 2013 *Phys. Rev. Lett.* **111** 212301 (*Preprint* 1303.1794)
[4] Aad G *et al.* (ATLAS) 2016 *Phys. Rev. Lett.* **116** 172301 (*Preprint* 1509.04776)
[5] Bzdak A, Schenke B, Tribedy P and Venugopalan R 2013 *Phys. Rev. C* **87** 064906 (*Preprint* 1304.3403)
[6] Weller R D and Romatschke P 2017 *Phys. Lett. B* **774** 351–356 (*Preprint* 1701.07145)
[7] Borsanyi S, Fodor Z, Hoelbling C, Katz S D, Krieg S and Szabo K K 2014 *Phys. Lett. B* **730** 99–104 (*Preprint* 1309.5258)
[8] Mogliacci S, Kolbé I and Horowitz W A 2020 *Phys. Rev. D* **102** 116017 (*Preprint* 1807.07871)
[9] Kitazawa M, Mogliacci S, Kolbé I and Horowitz W A 2019 *Phys. Rev. D* **99** 094507 (*Preprint* 1904.00241)

[10] Horowitz W A and Rothkopf A 2021 *50th International Symposium on Multiparticle Dynamics* (*Preprint* 2109.01422)
[11] Horowitz W A and Plessis J F D 2023 *arXiv* (*Preprint* 2308.08651)
[12] Horowitz W A and Du Plessis J F 2022 *Phys. Rev. D* **105**(9) L091901 URL https://link.aps.org/doi/10.1103/PhysRevD.105.L091901
[13] Du Plessis J F and Horowitz W A 2022 *International Nuclear Physics Conference Proceedings*
[14] Moroz A and Miroshnichenko A E 2019 *New Journal of Physics* **21** 103035 URL https://dx.doi.org/10.1088/1367-2630/ab484b

Deep Inelastic Scattering in Nuclear Collisions

W. A. Horowitz

Department of Physics, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa

E-mail: wa.horowitz@uct.ac.za

Abstract. Deep inelastic scattering (DIS) is a key measurement in nuclear physics. One of the massive advantages of DIS is that powerful mathematical theorems—known as factorization—have been proven. We discuss the application of these theorems and effective field theory to collisions between electrons and protons and the potential extension of these theorems to collisions between electrons and nuclei, which will be relevant at the future eRHIC facility.

1. Introduction

A microsecond after the Big Bang, the universe cooled rapidly to around a trillion degrees, transforming space into a new state called the quark-gluon plasma (QGP). As cosmic expansion persisted, the intricate dynamics of this initial era left their mark on the universe's history, ultimately shaping its large-scale structure.

Many body quantum field theory is a vibrant area of research. Even seemingly "simple" systems, governed solely by the electromagnetic force, exhibit behaviors that defy full understanding from fundamental principles. Examples include water's phase behavior [1] and high-temperature superconductivity [2]. This motivates us to explore the non-Abelian generalization of many-body dynamics in quantum field theories, in order to compare with their Abelian counterparts, because there's even the possibility that the non-Abelian theories display simpler aspects, and in order to understand better the physics of the early universe [3–7].

Remarkably, collisions of heavy ions at the Relativistic Heavy Ion Collider (RHIC) and the Large Hadron Collider (LHC) offer a glimpse into these intricate non-Abelian dynamics and early universe physics. "Hard probes," observables related to particles with a scale large compared to the natural scale of the nuclear force, $\Lambda_{QCD} \sim 200$ MeV, are key tools for the exploration of the QGP formed in these collisions [8].

To relate measured particle distributions to QGP characteristics, theoretical calculations are crucial. For hard probes, one approach assumes weak coupling between probes and a QGP medium. Expressions have been derived for the energy transferred from probes to the QGP during propagation [10–18]. These energy loss-based models excel in describing experimental data [19–22]. A current thriving area of research aims to establish a firmer theoretical basis for energy loss derivations, especially in understanding how corrections to leading-order results scale with probe energy and system size [23–26].

In Deep Inelastic Scattering (DIS), high-energy collisions between electrons and hadronic objects occur. The initial kinematics of both particles are known, and the final electron kinematics are measured; all other final state particles are integrated over.

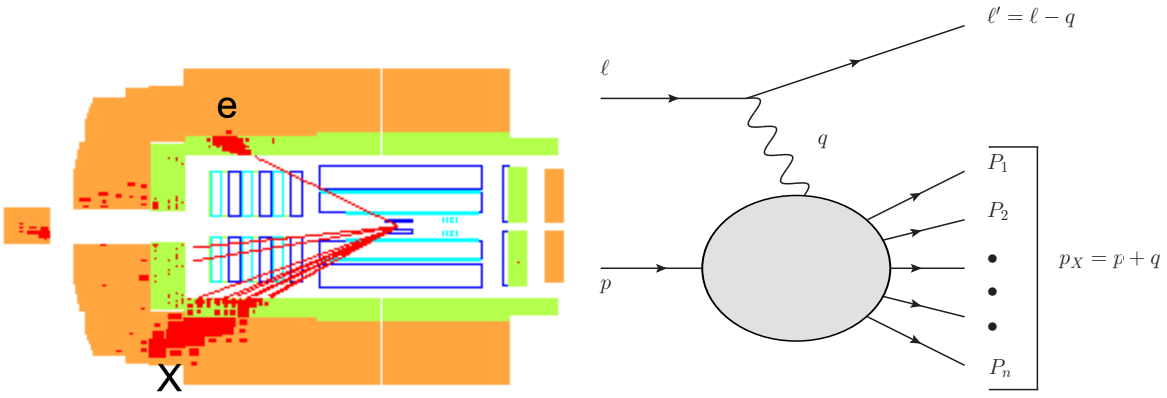


Figure 1. (Left) A DIS event measured by the H1 collaboration at DESY [9]. In this frame, the electron enters from the left and the proton from the right. The deflected electron is tagged as the single track in the top half of the detector and a jet of hadronic particles is identified in the bottom half of the detector, approximately back-to-back in azimuth. $Q^2 = 25030$ GeV^2 for this event. (Right) Scattering kinematics of a DIS event.

We show a typical DIS event, observed by H1 at DESY [9], Fig. 1. Electrons enter from the left at 27.5 GeV, protons from the right at 920 GeV, yielding $\sqrt{s} = 320$ GeV. The deeply-inelastic events involve the scattered electron's large-angle deflection due to significant momentum transfer. A hadronic jet is labeled X. We focus on DIS cross sections inclusive of all final state particles.

Choosing inclusive observables dramatically simplifies the theory. For factorizable processes [27], such as DIS, the prediction is an expansion in $1/Q^2$, also known as the twist expansion. In DIS, leading twist scales with Q^0, termed Bjorken scaling [28]. In general, twist contributions scale as $1/Q^{\tau-2}$, with twist τ. DIS's leading twist contribution is then twist-2. Twist labels connect with the Operator Product Expansion (OPE) [29].

We would like to find a way to extend the factorization framework into the formalism of energy loss in QGP, allowing for a rigorous foundation for jet tomography in heavy ion collisions.

2. Introduction to Factorization

We would like to work with the natural kinematic variables, x and Q^2. In order to understand these variables and the connection to the physical lab frame, let's take a closer look at the kinematics of a DIS process. Please see Appendix A for our lightcone conventions.

Let's begin by clearly defining our variables. We consider the collision of an incoming electron of momentum ℓ with an incoming hadron of momentum p. The electron is deflected by the collision to a final momentum ℓ', so that the momentum transfer to the target hadron is $q \equiv \ell - \ell'$. Consequently, the total momentum of the hadronic fragments is $p_X = p + q$. These kinematics are illustrated in Fig. 1.

The cross section for *fully inclusive* deep inelastic scattering is kept differential only in the final-state momentum ℓ' of the scattered electron, while summing over the particle content and final-state kinematics of all the hadronic debris represented by X. As a result, we seek to write all our expressions in terms of the momenta of the initial electron ℓ, initial hadron p, and scattered electron ℓ'. We are thus motivated to define the Lorentz-invariant quantities

$$Q^2 \equiv -q^2; \qquad x \equiv \frac{Q^2}{2p \cdot q}, \tag{1}$$

where Q^2 is the virtuality of the photon and $x = x_{\rm Bj}$ is known as Bjorken-x. Since $Q^2 = -q^2 = -t$ with t the Mandelstam invariant, we have that $Q^2 \geq 0$ and the virtual photon is spacelike. (Excluding the case of perfectly forward scattering $q = 0$, one has $Q^2 > 0$.)

The kinematic range of the variables x and Q^2 are:

$$1 \geq x \geq 0; \qquad xs \geq Q^2 \geq 0; \qquad Q^2 \gg m^2_{proton} \sim k^2_T \sim \Lambda^2_{QCD}; \tag{2}$$

i.e. there's only one dimensionless quantity, $m/Q \ll 1$.

We wish to determine the "leading twist" contribution, which is the leading contribution in m/Q. Higher order corrections are "higher twist."

It's extremely useful to work in the natural frame for the problem, the Breit frame, which is otherwise known as the "brick wall" frame. In this frame,

$$q^\mu = (0, \mathbf{0}, -Q)^\mu = [-Q, Q, 0]^\mu; \quad p^\mu_{proton} = [p^+, 0, \mathbf{0}]^\mu; \quad Q = x\, p^+; \quad p^\mu_{parton} = [\xi p^+, p^-, \mathbf{p}]^\mu,$$

where ξ is known as "Feynman-x." ξ is the amount of momentum carried by the parton within the proton. At leading order, $\xi = x$: remarkably, a macroscopic measurement provides femtoscopic insight into the nature of the proton.

To leading order in the electromagnetic coupling, one may write the DIS cross section in terms of hadronic currents:

$$E_{\ell'} \frac{d\sigma^{e^\lambda h^s \to e^{\lambda'} X}}{d^3\ell'} = \frac{2}{s} \frac{\alpha^2_{EM}}{Q^2} L^{\lambda\lambda'}_{\mu\nu} W^{\mu\nu}_s = \frac{1}{2s} \frac{\alpha^2_{EM}}{Q^2} L^{\lambda\lambda'}_{\mu\nu} \frac{1}{\pi} \int d^4z e^{iz\cdot q}{}_{in} \langle p^s | J^{\dagger\mu}(z) J^\nu(0) | p^s \rangle_{in}, \tag{3}$$

which decomposes the inclusive DIS cross section, differential only in the final-state electron momentum ℓ', into the *leptonic tensor*

$$L^{\lambda\lambda'}_{\mu\nu} \equiv \Big[\bar{u}_\lambda(\ell)\gamma_\mu u_{\lambda'}(\ell')\Big]\Big[\bar{u}_{\lambda'}(\ell')\gamma_\nu u_\lambda(\ell)\Big] \tag{4}$$

describing the scattering of the electron and the *hadronic tensor*

$$W^{\mu\nu}_s \equiv \frac{1}{4\pi} \int d^4z e^{iz\cdot q}{}_{in} \langle p^s | J^{\dagger\mu}(z) J^\nu(0) | p^s \rangle_{in}. \tag{5}$$

Note that λ (λ') are the incoming (outgoing) electron polarizations. The QED physics is entirely contained within the leptonic tensor while the hadronic physics is entirely contained in the hadronic tensor.

Simple dimensional analysis implies that the hadronic tensor must scale as Q^0, known as Bjorken scaling [28]. Bjorken scaling was observed in the 1960's in SLAC-MIT data [30]; see Fig. 2.

The relevant question, then, is: can Bjorken scaling be broken anomalously by higher order corrections? The answer is a definitive yes. In order to determine rigorously the higher order corrections, one first performs a Libby-Sterman region analysis [27, 32, 33]. Such an analysis shows where an amplitude might diverge in the UV or the IR. There are generically three potential regions in a diagram where these divergences might occur (and thus contribute to anomalously breaking a symmetry or scaling): the hard, the collinear, and the soft. These regions are categorized by the size of their momentum components:

$$\begin{aligned}
\text{Hard}: \quad & p^\mu \sim [Q\,,\, Q\,,\, Q]^\mu \quad , \quad p^2 \sim Q^2\,, \\
\text{Collinear}: \quad & p^\mu \sim [Q\,,\, \tfrac{M^2}{Q}\,,\, M]^\mu \quad , \quad p^2 \sim M^2 \ll Q^2\,, \\
\text{Soft}: \quad & p^\mu \sim [M\,,\, M\,,\, M]^\mu \quad , \quad p^2 \sim M^2 \ll Q^2\,.
\end{aligned} \tag{6}$$

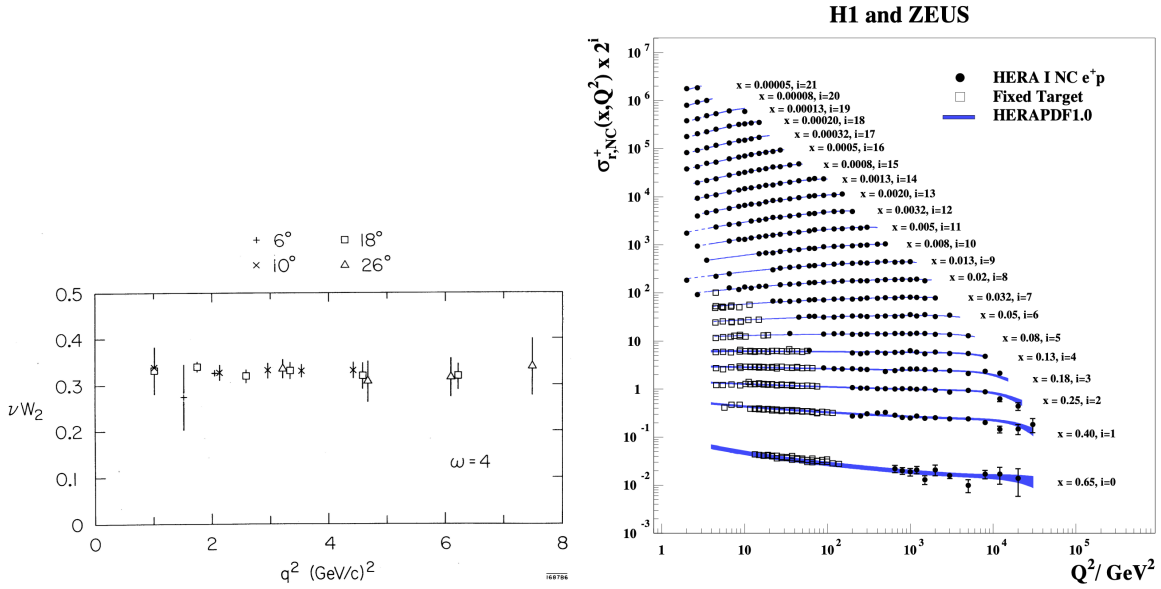


Figure 2. (Left) An early observation of Bjorken scaling [30]. (Right) Measurement of anomalous breaking of scaling compared to QCD predictions [31].

When a Libby-Sterman analysis is performed for a DIS process, one finds that there are only two regions: a hard region and a collinear region; see Fig. 3, where the logical progression from the most general diagram containing an incoming electron and proton and an outgoing electron and proton to the leading QED contribution to the leading twist contribution is shown[1]. Since all propagators in the hard region are off-shell by order Q^2, this part of the QCD diagram is perturbatively calculable with the coupling evaluated at $\alpha_s(Q^2) \ll 1$. There are potentially an unlimited number of collinear gluons that connect the hard and collinear parts of the diagram; these gluons exponentiate into a Wilson line that lives in the parton distribution function (PDF), the object that gives a sense of the number of partons of a particular type in a proton [27]. Although the parton distribution function itself is not calculable perturbatively, one may use renormalization group flow to derive the famous DGLAP evolution equations, which give the Q^2 evolution of the PDFs. The H1 and ZEUS experiments showed that these higher orders in QCD corrections to the leading order Bjorken scaling give precise, quantitative description of data [31]; see Fig. 2.

Interesting recent work has tried to apply factorization to jet broadening in semi-inclusive deep inelastic scattering (SIDIS) in nuclei [34, 35]. The idea is to explicitly compute the twist-4 contribution to the process, which is enhanced by the large size of the nucleus. What's not yet entirely clear is if the collinear factorization framework, which assumes that one can integrate out completely the transverse momentum of the radiated gluon, ignoring the finite kinematics of the problem, is appropriate for this observable [36]. Nonetheless, this is an interesting new avenue of research that may prove to be very important for future studies at an electron-ion collider such as eRHIC.

[1] Note that this analysis is actually connecting the *in/in* result Eq. 5 to the usual *in/out* formalism via the Optical Theorem. Then

$$\int d^4z e^{iz\cdot q}{}_{in}\langle p^s|J^\mu(z)J^\nu(0)|p^s\rangle_{in} = 2\,\mathrm{Im}\, i\int d^4z e^{iz\cdot q}{}_{out}\langle p^s|T\{J^\mu(z)J^\nu(0)\}|p^s\rangle_{in}, \tag{7}$$

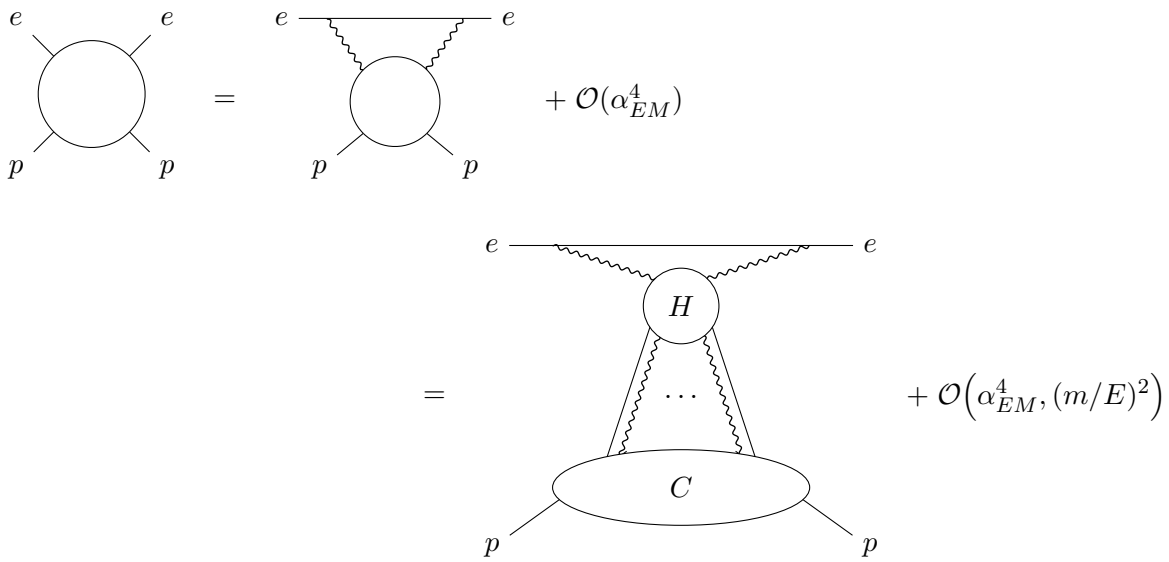


Figure 3. (Left) The most generic Feynman diagram corresponding to forward e+p scattering. (Centre) The leading order in α_{EM} contribution. (Right) The leading twist contribution.

3. Conclusions

In the aftermath of the Big Bang, the universe rapidly transitioned to the quark-gluon plasma state, leaving an indelible imprint on its subsequent evolution and current large-scale structure. Many body quantum field theory is a fascinating, active area of research, with implications across many orders of magnitude of size and of epochs in time.

The Relativistic Heavy Ion Collider and the Large Hadron Collider enable the study of non-trivial, emergent, many body, non-Abelian dynamics and also early universe physics through heavy ion collisions. Utilizing "hard probes," particles with significant scale emerging from these collisions, researchers gain insights into the properties of the QGP. Theoretical calculations involving weak coupling assumptions provide expressions for energy loss from hard probes to the QGP during their passage through the medium. These energy loss-based models have shown remarkable success in describing experimental data. Work is now ongoing to try to place these energy loss calculations on a more rigorous theoretical footing.

In the realm of Deep Inelastic Scattering, high-energy collisions between electrons and hadronic objects provide a deeper understanding of the quark substructure. Fully inclusive DIS cross sections focus solely on the final-state momentum of the scattered electron while summing over hadronic debris. Kinematic variables like x and Q^2 offer insights into the nature of the scattering process, with Bjorken scaling providing a pivotal principle. However, higher order corrections can break this scaling, and a Libby-Sterman analysis reveals hard and collinear regions crucial for perturbative calculations.

Recent research has attempted to apply these rigorous theoretical techniques of factorization from electron-proton collisions to larger hadronic systems such as electron-ion collisions and ion-ion collisions. There are many interesting avenues of future research that can be taken going forward.

Acknowledgments

The author gratefully acknowledges generous support from the South African National Research Foundation and the SA-CERN Collaboration. The author further gratefully acknowledges extensive valuable discussions with Matthew D. Sievert and Antonio Garcia Vallejo.

Appendix A. Light-Front Coordinates

Throughout this proceedings we employ natural units $\hbar = c = 1$. Consider a generic Lorentz four-vector p^μ with components

$$p^\mu = \left(p^0\,,\,\vec{p}\right)^\mu = \left(p^0\,,\,\mathbf{p}\,,\,p^3\right)^\mu \tag{A.1}$$

where $\mathbf{p} = (p^1\,,\,p^2)$ denotes the 2D vector in the transverse plane. We define *light-front* coordinates in the following asymmetric way,

$$p^\pm \equiv p^0 \pm p^3 \qquad \Leftrightarrow \qquad \begin{array}{rcl} p^0 & = & \frac{1}{2}(p^+ + p^-) \\ p^3 & = & \frac{1}{2}(p^+ - p^-). \end{array} \tag{A.2}$$

References

[1] Kapil V, Schran C, Zen A, Chen J, Pickard C J and Michaelides A 2021 The first-principles phase diagram of monolayer nanoconfined water (*Preprint* 2110.14569)

[2] Aprahamian A *et al.* 2015 Reaching for the horizon: The 2015 long range plan for nuclear science

[3] Gyulassy M and McLerran L 2005 *Nucl. Phys. A* **750** 30–63 (*Preprint* nucl-th/0405013)

[4] Jacak B and Steinberg P 2010 *Phys. Today* **63N5** 39–43

[5] Accardi A *et al.* 2016 *Eur. Phys. J. A* **52** 268 (*Preprint* 1212.1701)

[6] Abdul Khalek R *et al.* 2022 *Nucl. Phys. A* **1026** 122447 (*Preprint* 2103.05419)

[7] Casalderrey-Solana J, Liu H, Mateos D, Rajagopal K and Wiedemann U A 2014 *Gauge/String Duality, Hot QCD and Heavy Ion Collisions* (Cambridge University Press) ISBN 978-1-139-13674-7 (*Preprint* 1101.0618)

[8] Majumder A and Van Leeuwen M 2011 *Prog. Part. Nucl. Phys.* **66** 41–92 (*Preprint* 1002.2206)

[9] Meyer J Hera e-p scattering events observed in the h1 detector http://www-h1.desy.de/pictures/H1-short.event.tutorial2.pdf accessed: 2021-01-07

[10] Bjorken J D 1982 FERMILAB–PUB–82–059–THY

[11] Baier R, Dokshitzer Y L, Mueller A H, Peigne S and Schiff D 1997 *Nucl. Phys. B* **484** 265–282 (*Preprint* hep-ph/9608322)

[12] Gyulassy M, Levai P and Vitev I 2001 *Nucl. Phys. B* **594** 371–419 (*Preprint* nucl-th/0006010)

[13] Wang X N and Guo X f 2001 *Nucl. Phys. A* **696** 788–832 (*Preprint* hep-ph/0102230)

[14] Djordjevic M and Gyulassy M 2004 *Nucl. Phys. A* **733** 265–298 (*Preprint* nucl-th/0310076)

[15] Salgado C A and Wiedemann U A 2003 *Phys. Rev. D* **68** 014008 (*Preprint* hep-ph/0302184)

[16] Armesto N, Salgado C A and Wiedemann U A 2004 *Phys. Rev. D* **69** 114003 (*Preprint* hep-ph/0312106)

[17] Arnold P B, Moore G D and Yaffe L G 2003 *JHEP* **05** 051 (*Preprint* hep-ph/0302165)

[18] Wicks S, Horowitz W, Djordjevic M and Gyulassy M 2007 *Nucl. Phys. A* **784** 426–442 (*Preprint* nucl-th/0512076)

[19] Armesto N *et al.* 2012 *Phys. Rev. C* **86** 064904 (*Preprint* 1106.1106)

[20] Horowitz W A and Gyulassy M 2011 *Nucl. Phys. A* **872** 265–285 (*Preprint* 1104.4958)

[21] Horowitz W A 2013 *Nucl. Phys. A* **904-905** 186c–193c (*Preprint* 1210.8330)

[22] Cao S *et al.* (JETSCAPE) 2021 *Phys. Rev. C* **104** 024905 (*Preprint* 2102.11337)
[23] Kolbe I and Horowitz W A 2019 *Phys. Rev. C* **100** 024913 (*Preprint* 1511.09313)
[24] Horowitz W A and Du Plessis J F 2022 *Phys. Rev. D* **105** L091901 (*Preprint* 2203.01259)
[25] Faraday C, Grindrod A and Horowitz W A 2023 (*Preprint* 2305.13182)
[26] Horowitz W A and Du Plessis J F 2023 (*Preprint* 2308.08651)
[27] Collins J 2013 *Foundations of perturbative QCD* vol 32 (Cambridge University Press) ISBN 978-1-00-940184-5
[28] Bjorken J D 1969 *Phys. Rev.* **179** 1547–1553
[29] Peskin M E and Schroeder D V 1995 *An Introduction to quantum field theory* (Reading, USA: Addison-Wesley) ISBN 978-0-201-50397-5
[30] Bloom E D *et al.* 1970 *15th International Conference on High Energy Physics*
[31] Aaron F D *et al.* (H1, ZEUS) 2010 *JHEP* **01** 109 (*Preprint* 0911.0884)
[32] Sterman G F 1995 *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 95): QCD and Beyond* pp 327–408 (*Preprint* hep-ph/9606312)
[33] Sterman G F 2004 *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics in $D \geqq 4$* pp 67–145 (*Preprint* hep-ph/0412013)
[34] Kang Z B, Wang E, Wang X N and Xing H 2014 *Phys. Rev. Lett.* **112** 102001 (*Preprint* 1310.6759)
[35] Kang Z B, Wang E, Wang X N and Xing H 2016 *Phys. Rev. D* **94** 114024 (*Preprint* 1409.1315)
[36] Clayton H, Sievert M D and Horowitz W A 2022 *Eur. Phys. J. C* **82** 437 (*Preprint* 2110.14737)

Assumption Breakdown in Radiative Energy Loss

Coleridge Faraday

Department of Physics, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa

E-mail: frdcol002@myuct.ac.za

W. A. Horowitz

Department of Physics, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa

E-mail: wa.horowitz@uct.ac.za

Abstract. We show that an integral assumption in DGLV radiative energy loss—the large formation time assumption—is violated at high-p_T for phenomenologically relevant parameters. We further investigate the phenomenological impact of placing a new kinematic bound on the radiated gluon transverse momentum, which ensures that there are no contributions to the energy loss from regions of parameter space that violate the large formation time assumption. We find that this places a large sensitivity on the exact kinematic cutoff used, similar to the known collinear cutoff sensitivity, indicating the theoretical need for a rederivation of DGLV radiative energy with the large formation time assumption relaxed in order to make rigorous predictions. We additionally find that this large formation time cutoff dramatically reduces the size of a short pathlength correction to the DGLV radiative energy loss, which is of phenomenological interest in predicting suppression in small $p + A$ systems. We compute the phenomenological predictions utilising this large formation time cutoff in both $p + A$ and $A + A$ collisions at the LHC, in a convolved radiative and elastic energy loss model.

1. Introduction

Studying high-p_T particle spectra offers crucial insights into the many-body dynamics of QCD in high-energy collisions. Significant suppression of high-p_T particles in $A + A$ collisions has been observed in RHIC and LHC experiments [1–4], attributed to parton energy loss in the QGP. This phenomenon aligns well with predictions from pQCD-based models [5–8]. Recent findings in $p + A$ and $p + p$ collisions, including strangeness enhancement [9, 10], quarkonium suppression [11], and collective behavior [12, 13], additionally support QGP formation in these smaller collision systems. Non-trivial modifications of high-p_T particles have also emerged in small collision systems [14–16], necessitating theoretical explanations.

Applying successful $A+A$ models to $p+A$ and $p+p$ collisions presents challenges due to various large system size assumptions. In our previous work [17, 18], we addressed the removal [19] of the *large pathlength assumption* $L \gg \mu^{-1}$ in the Djordjevic-Gyulassy-Levai-Vitev (DGLV) radiative energy loss model [20, 21]. The correction consists of $\mathcal{O}(e^{-\mu L})$ terms, which were previously assumed to be small, and results in the following novel effects: reduction of energy loss, linear growth with partonic energy, and disproportionate size for incident gluons (cf. usual C_A/C_F

color factor scaling). This reduction in energy loss could explain the rapid rise of the charged hadron nuclear modification factor with p_T [22] and the enhancement above unity in $p + A$ collisions [23, 24].

The derivation of this short pathlength correction, and to a lesser extent the original DGLV derivation, benefited significantly from a *large formation time* (LFT) assumption—in addition to the usual collinear, Eikonal, and soft assumptions. In [17] we used a numerical energy loss weighted average of dimensionless ratios which are assumed to be small in the derivation of the energy loss single emission kernel, to show that the large formation time assumption was explicitly violated at high-p_T for both the DGLV radiative energy loss and the DGLV radiative energy loss which receives a short pathlength correction.

The LFT assumption is utilised in all opacity expansion based approaches, for instance DGLV [21], and as such an understanding of the LFT assumption's impact is crucial for interpreting predictions from these models. If such energy loss models receive large contributions from regions of phase space where the LFT assumption is invalid, then their predictions should not be trusted. The fact that the LFT assumption is violated at high-p_T or small pathlengths in the DGLV radiative energy loss model, then implies that theoretical control over the LFT assumption is required in order to make quantitative predictions at high-p_T or in small systems. A rigorous way in which such control may be achieved is through a full rederivation of the radiative energy loss kernels in [19–21], where the LFT assumption is relaxed. In this work, we will explore an alternative method for obtaining theoretical control over the LFT assumption by implementing a kinematic cut on the radiated transverse gluon transverse momentum $\mathbf{k}$ integral. This ensures that the matrix element (modulus squared) is never evaluated in regions of phase space which violate the LFT assumption.

We will see that implementing such a cutoff dramatically reduces the size of the short pathlength correction for high-p_T pions in $A + A$ collisions which had previously received a $\sim$100% negative correction at $p_T \sim \mathcal{O}(100)$ GeV, confirming that the short pathlength correction receives a large contribution at high energies from regions of phase space where the large formation time assumption is invalid. We additionally find that both the short pathlength corrected and original DGLV radiative energy loss incur a large sensitivity as a result of this large formation time cutoff. This is not surprising as there exists a similar sensitivity to the standard collinear cutoff $|\mathbf{k}|_{\max} = 2xE(1-x)$ [25], however, we find that the sensitivity to the large formation time cutoff is significantly larger than that of the collinear cutoff.

One reason why such a sensitivity is important, is in the utilisation of energy loss models as a tomographic probe of the QGP. If one wishes to interpret fitted quantities such as the temperature of the plasma as physical, then an understanding of the sensitivity of the energy loss model to the various assumptions utilised is necessary for deriving a theoretical uncertainty for these quantities. It may also be more difficult for various energy loss models to be ruled out by data, due to the large theoretical uncertainties present in the model as a result of this cutoff sensitivity.

2. Discussion and Results

The radiative energy loss is calculated according to DGLV [21] with short path length corrections as derived in [19]. The number of radiated gluons N^g differential in the momentum fraction

radiated away x is given to first order in opacity L/λ by

$$\frac{\mathrm{d}N^g}{\mathrm{d}x} = \frac{C_R\alpha_s L}{\pi\lambda}\frac{1}{x}\int\frac{\mathrm{d}^2\mathbf{q}_1}{\pi}\frac{\mu^2}{\left(\mu^2+\mathbf{q}_1^2\right)^2}\int\frac{\mathrm{d}^2\mathbf{k}}{\pi}\int\mathrm{d}\Delta z\,\rho(\Delta z)$$
$$\times\left[-\frac{2\left\{1-\cos\left[\left(\omega_1+\tilde{\omega}_m\right)\Delta z\right]\right\}}{\left(\mathbf{k}-\mathbf{q}_1\right)^2+m_g^2+x^2M^2}\left[\frac{\left(\mathbf{k}-\mathbf{q}_1\right)\cdot\mathbf{k}}{\mathbf{k}^2+m_g^2+x^2M^2}-\frac{\left(\mathbf{k}-\mathbf{q}_1\right)^2}{\left(\mathbf{k}-\mathbf{q}_1\right)^2+m_g^2+x^2M^2}\right]\right. \tag{1}$$
$$+\frac{1}{2}e^{-\mu_1\Delta z}\left(\left(\frac{\mathbf{k}}{\mathbf{k}^2+m_g^2+x^2M^2}\right)^2\left(1-\frac{2C_R}{C_A}\right)\left\{1-\cos\left[\left(\omega_0+\tilde{\omega}_m\right)\Delta z\right]\right\}\right.$$
$$\left.\left.+\frac{\mathbf{k}\cdot\left(\mathbf{k}-\mathbf{q}_1\right)}{\left(\mathbf{k}^2+m_g^2+x^2M^2\right)\left(\left(\mathbf{k}-\mathbf{q}_1\right)^2+m_g^2+x^2M^2\right)}\left\{\cos\left[\left(\omega_0+\tilde{\omega}_m\right)\Delta z\right]-\cos\left[\left(\omega_0-\omega_1\right)\Delta z\right]\right\}\right)\right],$$

where the first two lines of the above equation are the original DGLV result [21] and the last two lines are the correction [19]. Variable names are the same as in [17, 19, 21]. Importantly, the radiated gluon has 4-momentum (in lightfront coordinates) $k = \left[2xE, m_g^2+\mathbf{k}^2/2xE, \mathbf{k}\right]$ and the medium exchanged particle $q = [q^+, q^-, \mathbf{q}]$, where x is the radiated (plus) momentum fraction.

The *large formation time assumption* requires that both $\mathbf{k}^2/2xE \ll \mu_1$ and $(\mathbf{k}-\mathbf{q})^2/2xE \ll \mu_1$, and is utilised in GLV [20], DGLV [21], and—to a larger extent—in the short pathlength correction to DGLV [19]. Conventionally, and in our previous work [17, 19, 21], the radiated gluon, incident parton, and outgoing parton are all enforced to be collinear by setting the upper bound on the $\mathbf{k}$ integral as $|\mathbf{k}|_{\max} = 2xE(1-x)$. Requiring that our matrix element (modulus squared) is not integrated over regions of invalidity for either the large formation time assumption or the collinear assumption, leads to the prescription $|\mathbf{k}|_{\max} = \mathrm{Min}(\sqrt{2xE\mu_1}, 2x(1-x)E)$. Note that we enforce only the first of the two large formation time assumptions, as the second assumption defines a far more complicated region of parameter space and we observe numerically that enforcing the first assumption leads to the second holding consistently. We additionally note that at high energies the large formation time cutoff occurs earlier in $\mathbf{k}$ than the collinear cutoff, which can be understood through the different asymptotic scalings of the cutoffs, $\sim E^{1/2}$ vs $\sim E^1$ respectively.

A study of the collinear cutoff in the DGLV radiative energy loss was conducted in [25], where a large sensitivity to the specific choice of cutoff was found. Similarly we show in Fig. 1 the sensitivity of the short pathlength corrected and original DGLV radiative energy loss models to both the large formation time and collinear cutoffs, by varying the respective cutoffs by factors of two. Figure 1 shows that by enforcing the large formation time cutoff, one dramatically reduces the size of the short pathlength correction to the radiative energy loss. Further, the uncorrected DGLV radiative energy loss has an increased sensitivity to the specific cutoff used once the large formation time cutoff is imposed (in comparison to only the collinear cutoff), particularly at higher energies $E \gtrsim 100$ GeV. This sensitivity can be interpreted as the existence of a large weight of the distribution close to the large formation time cutoff. This is further indication that the large formation time assumption breaks down at high p_T for both the short pathlength corrected and original DGLV radiative energy loss, as first discussed in [17].

In principle one should propagate this theoretical sensitivity to the level of the R_{AA}, by computing the R_{AA} which results from the radiative energy loss with $|\mathbf{k}|_{\max}$—set according to the collinear or collinear and large formation time assumption—scaled by some factor $\alpha \sim \mathcal{O}(1)$. Particularly for the short pathlength corrected energy loss this will be extremely numerically intensive, as the short pathlength corrected energy loss kernel is *not* monotonic in $\mathbf{k}$; meaning that the R_{AA} must be calculated for a full range of α. We instead provide a qualitative argument for how the error at the level of $\Delta E/E$ translates to the error at the level of R_{AA}

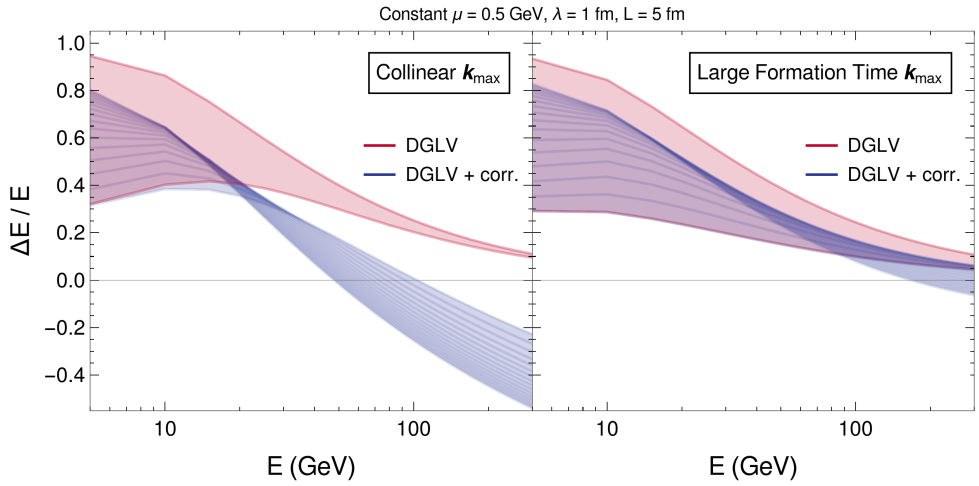


Figure 1. The fractional radiative energy loss calculated according to DGLV and short pathlength corrected DGLV. In the left pane the upper bound on the radiated gluon momentum $\mathbf{k}$ integral is set by the collinear assumption $|\mathbf{k}|_{\max} = 2xE(1-x)$, while in the right pane the upper bound is taken as $|\mathbf{k}|_{\max} = \mathrm{Min}(\sqrt{2xE\mu_1}, 2x(1-x)E)$. Bands are calculated by varying the upper bound by factors of two. Calculations are done with $\mu = 0.5GeV$, $\lambda = 1$ fm, and $L = 5$ fm.

The R_{AA} can be approximated [17] as $R_{AA} \approx \int \mathrm{d}x(1-x)^{n-1}P_{\mathrm{tot}}(x)$ where $P_{\mathrm{tot}}(x)$ is the total probability of a losing a fraction x of the incident parton's energy, and n is the power in the approximately power-law parton production spectra with $n \sim 6$ for gluons at the LHC. If one further assumes that the elastic energy loss is negligible, and take $P_{\mathrm{tot}}(x) = \delta(x-\epsilon)$ where $\epsilon \equiv \Delta E/E = \bar{\epsilon} \pm \Delta\epsilon$, then we find $R_{AA} \sim 1-(n-1)(\bar{\epsilon} \pm \Delta\epsilon) \sim \bar{R}_{AA} \mp (n-1)\Delta\epsilon$. Under this naive approximation, one sees that the sensitivity to the kinematic $\mathbf{k}$ cutoff at the level of $\Delta E/E$ are enhanced by $(n-1) \sim 5$ at the level of the R_{AA}. Note that this is likely an overestimate of the error as we saw qualitatively in [17] that changes at the level of $\Delta E/E$ are softened by the geometry averaging and Poisson convolution procedures in the full energy loss model.

The phenomenological implications of such a large formation time cutoff are investigated using the convolved radiative and elastic energy loss model described in [17], based on the Wicks-Horowitz-Djordjevic-Gyulassy (WHDG) model. The left pane of Fig. 2 shows the predicted R_{AA} for pions in central Pb + Pb collisions as a function of final transverse momentum p_T versus data. Including the large formation time cutoff dramatically reduces the size of the correction, as is expected from Fig. 1. The effect on the R_{AA} calculated with the uncorrected WHDG is negligible. We note however that Fig. 1 indicates that there will likely be a large sensitivity at the level of the R_{AA} to the exact kinematic cutoff on the $\mathbf{k}$ integral used.

The right pane of Fig. 2 shows a similarly dramatic decrease of the size of the short pathlength correction once the large formation time cutoff is enforced, as well as the same negligible effect on the uncorrected DGLV energy loss. Note that in Fig. 2 we only include the radiative energy loss since, as was previously shown in [17, 18], the central limit theorem approximation to the elastic energy loss breaks down in small systems. In the future, this will need to be corrected by the removal of the central limit theorem approximation in the elastic energy loss as well as potential short pathlength corrections to the elastic energy loss.

3. Conclusion

We provided an argument for the breakdown of the large formation time assumption in both DGLV [21] and the short pathlength correction to DGLV [19]. We showed that one can

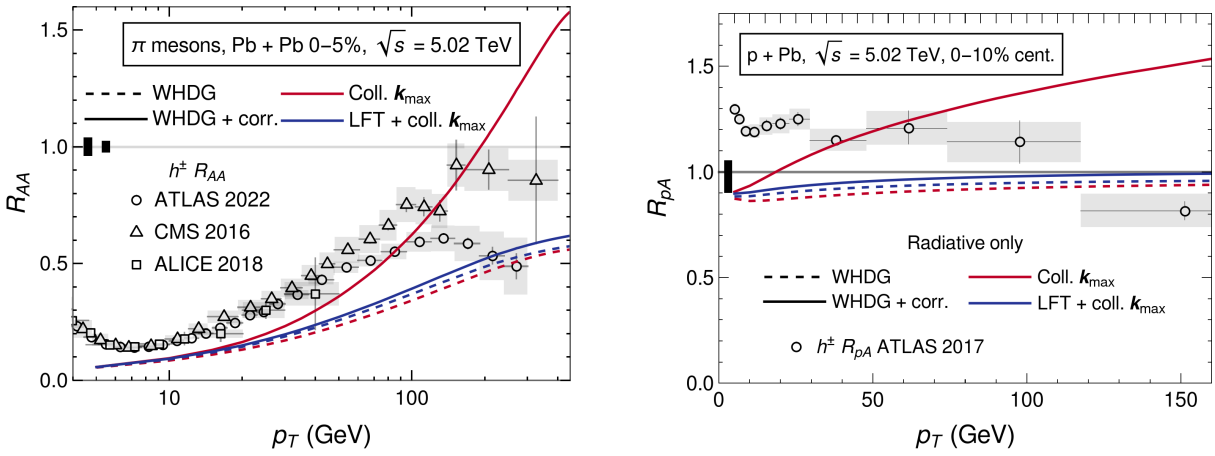


Figure 2. The nuclear modification factor in A+A collisions (left pane) and p+A collisions (right pane) is plotted as a function of the measured transverse momentum p_T for pions. Theoretical predictions are computed for the WHDG model (solid) and the WHDG model which receives a short pathlength correction to the radiative energy loss (dashed). The upper bound of the radiated gluon transverse momentum $\mathbf{k}$ integral is prescribed both by enforcing the collinear assumption (red) and by enforcing both the collinear and large formation time assumptions (blue). In the left pane all plotted curves are computed using only the radiative energy loss and neglecting the elastic energy loss. Data are from ATLAS [23, 26], CMS [27], ALICE [28]

.

artificially impose a cutoff on the radiated gluon momentum $\mathbf{k}$ integral which ensures that the matrix element (modulus squared) is not integrated over regions where the large formation time assumption is invalid. Doing so dramatically reduces the size of the short pathlength corrected DGLV radiative energy loss, indicating that the short pathlength correction to the DGLV radiative energy loss receives an increasingly large contribution from regions where the large formation time assumption is invalid at high p_T. The effect on the uncorrected DGLV radiative energy loss is small $\sim 10\%$. We explore the sensitivity of the energy loss to the exact kinematic bound imposed, by varying kinematic cutoff by factors of two. Doing this we find that imposing the large formation time cutoff, as opposed to the usual collinear cutoff, increases the sensitivity of both the DGLV and short pathlength corrected DGLV to the exact kinematic cutoff chosen.

We further calculate both the R_{AA} and R_{pA} for pions produced in central collisions at the LHC. We find that the dramatic decrease of the correction at the level of $\Delta E/E$ translates to a dramatic decrease in the predicted R_{AA} and R_{pA} as is expected. We note that the sensitivity incurred on the $\Delta E/E$ will likely translate to a similar sensitivity on the R_{AA} and R_{pA}. We argue therefore that GLV, DGLV, and the short pathlength correction to DGLV all need to be rederived, with the large formation time assumption relaxed, for rigorous quantitative predictions to be made for *either* $A + A$ or $p + A$ collisions.

Possible future investigations into small systems could involve establishing a more robust foundation for energy loss calculations [29], exploring adjustments to thermodynamics due to system size [30], analyzing the equation of state in small systems [31], and calculating small system size corrections to the effective coupling [32].

Acknowledgements
CF and WAH thank the National Research Foundation and the SA-CERN collaboration for their support.

References

[1] Adcox K *et al.* (PHENIX) 2002 *Phys. Rev. Lett.* **88** 022301 (*Preprint* nucl-ex/0109003)
[2] Adams J *et al.* (STAR) 2003 *Phys. Rev. Lett.* **91** 072304 (*Preprint* nucl-ex/0306024)
[3] Adler S S *et al.* (PHENIX) 2006 *Phys. Rev. Lett.* **96** 202301 (*Preprint* nucl-ex/0601037)
[4] Adler S S *et al.* (PHENIX) 2007 *Phys. Rev. Lett.* **98** 172302 (*Preprint* nucl-ex/0610036)
[5] Dainese A, Loizides C and Paic G 2005 *Eur. Phys. J. C* **38** 461–474 (*Preprint* hep-ph/0406201)
[6] Schenke B, Gale C and Jeon S 2009 *Phys. Rev. C* **80** 054913 (*Preprint* 0909.2037)
[7] Horowitz W A 2013 *Nucl. Phys. A* **904-905** 186c–193c (*Preprint* 1210.8330)
[8] Wicks S, Horowitz W, Djordjevic M and Gyulassy M 2007 *Nucl. Phys. A* **784** 426–442 (*Preprint* nucl-th/0512076)
[9] Abelev B B *et al.* (ALICE) 2014 *Phys. Lett. B* **728** 25–38 (*Preprint* 1307.6796)
[10] Adam J *et al.* (ALICE) 2016 *Phys. Lett. B* **758** 389–401 (*Preprint* 1512.07227)
[11] Adam J *et al.* (ALICE) 2016 *JHEP* **06** 050 (*Preprint* 1603.02816)
[12] Khachatryan V *et al.* (CMS) 2015 *Phys. Rev. Lett.* **115** 012301 (*Preprint* 1502.05382)
[13] Aad G *et al.* (ATLAS) 2016 *Phys. Rev. Lett.* **116** 172301 (*Preprint* 1509.04776)
[14] Aad G *et al.* (ATLAS) 2015 *Phys. Lett. B* **748** 392–413 (*Preprint* 1412.4092)
[15] Adare A *et al.* (PHENIX) 2016 *Phys. Rev. Lett.* **116** 122301 (*Preprint* 1509.04657)
[16] Acharya S *et al.* (ALICE) 2018 *Phys. Lett. B* **783** 95–113 (*Preprint* 1712.05603)
[17] Faraday C, Grindrod A and Horowitz W A 2023 (*Preprint* 2305.13182)
[18] Faraday C and Horowitz W A 2023 *11th International Conference on Hard and Electromagnetic Probes of High-Energy Nuclear Collisions: Hard Probes 2023* (*Preprint* 2307.08355)
[19] Kolbe I and Horowitz W A 2019 *Phys. Rev. C* **100** 024913 (*Preprint* 1511.09313)
[20] Gyulassy M, Levai P and Vitev I 2001 *Nucl. Phys. B* **594** 371–419 (*Preprint* nucl-th/0006010)
[21] Djordjevic M and Gyulassy M 2004 *Nucl. Phys. A* **733** 265–298 (*Preprint* nucl-th/0310076)
[22] Horowitz W A and Gyulassy M 2011 *Nucl. Phys. A* **872** 265–285 (*Preprint* 1104.4958)
[23] Balek P (ATLAS) 2017 *Nucl. Part. Phys. Proc.* **289-290** 281–284 (*Preprint* 1802.02071)
[24] Acharya S *et al.* (ALICE) 2018 *JHEP* **10** 174 (*Preprint* 1804.09083)
[25] Horowitz W A and Cole B A 2010 *Phys. Rev. C* **81** 024909 (*Preprint* 0910.1823)
[26] Aad G *et al.* (ATLAS) 2023 *JHEP* **07** 074 (*Preprint* 2211.15257)
[27] Khachatryan V *et al.* (CMS) 2017 *JHEP* **04** 039 (*Preprint* 1611.01664)
[28] Sekihata D (ALICE) 2019 *Nucl. Phys. A* **982** 567–570 (*Preprint* 1807.11240)
[29] Clayton H, Sievert M D and Horowitz W A 2022 *Eur. Phys. J. C* **82** 437 (*Preprint* 2110.14737)
[30] Mogliacci S, Kolbé I and Horowitz W A 2020 *Phys. Rev. D* **102** 116017 (*Preprint* 1807.07871)
[31] Horowitz W and Rothkopf A 2022 *SciPost Phys. Proc.* **10** 025 (*Preprint* 2109.01422)
[32] Horowitz W A and Plessis J F D 2022 *Phys. Rev. D* **105** L091901 (*Preprint* 2203.01259)

First principles characterization of nitrogen substitutional point defects in graphane (CH)

H Mapingire and R E Mapasha

Department of Physics, University of Pretoria, Pretoria, 0002 Republic of South Africa

Email: **edwinmapasha@up.ac.za**

Abstract. We present *ab initio* characterization of nitrogen substitutional point defects and their different charge states in the two-dimensional material graphane. Our present first principles calculations were carried out within the framework of density functional theory implemented in the quantum espresso code. We derived the formation energies and thermodynamic charge transition levels for the nitrogen substitutional point defects of the form N_H , N_C and N_{CH} . We also determined the U-parameters for these point defects as it is crucial as far as defect mobility is concerned. The defect formation phenomena are all endothermic, hence the point defects are energetically expensive to form. N_C and N_{CH} yielded a positive effective U-parameter showing their stability should single negative and single positive charge states form. The point defects in this contribution induced spin-polarised defect states within the graphane bandgap and they can thus be exploited for various nano-magnetic and nanoelectronics applications.

1. Introduction

The isolation and characterization of the two-dimensional material (2-D) graphene in 2004 at the University of Manchester by Novoselev and Geim [1] opened floodgates of interest in different kinds of 2-D materials [2,3,4]. Graphene is an sp^2 hybridized material composed of a network of carbon atoms that are covalently bonded to each other to form a stable chicken-wire like hexagonal structure [5]. The isolation of a graphene monolayer was done using the now well known micro-mechanical cleavage or the skotch-tape method [6]. Graphene has a number of interesting properties that have attracted research interest far and wide. However, the absence of a band-gap in graphene has limited its direct application in nano-electronics [7]. A number of material researchers are seeking apt ways of creating a band-gap in graphene in order to open new ground of application for graphene derived 2-D materials [8]. In their work, Sahin *et al* proposed dimensionality reduction and the use of functionalization [9]. Alternatively, one can adsorb atoms or molecules or functional groups onto a graphene layer. These two methods can give rise to charge doping and the opening up of an electronic band-gap. Two-dimensional materials that are synthesised from graphene in more or less similar ways as discussed above are referred to as graphene derivative materials [8]. Among others, we have graphane which is hydrogenated graphene [10], fluorographene which is fluorinated graphene as well as chlorographene or chlorinated graphene. In 2007 Sofo *et al* [11] used first-principles derivation of total energy to predict the existence and stability of graphane (CH) as a semi-conductor hydrocarbon which is completely saturated.

Graphene hydrogenation give rise to the formation of a direct bandgap at the Brillouin zone centre of graphane. The different properties of graphane come into play because of the different ways in which

hydrogen atoms can be intertwined with the carbon atoms on a graphene monolayer. The structures of the six main allotropes of graphane- chair, stirrup, boat-1, boat-2, tricycle and twist-boat- depends on the ways in which the hydrogen atoms are attached to the carbon atoms in each hexagonal carbon ring [12]. The chair conformer is the most stable with the periodic up and down hydrogenation pattern. The other graphane stereo isomers that have been proposed in research articles includes: the armchair; the twist-boat; the twistboat-chair and the tricycle which comes from the chair and stirrup combination [13]. Whitener [14] discussed a number of methods that can be used to synthesise graphane. The prominent ones are: the exposure of graphene to hydrogen plasma and plasma-assisted chemical vapour deposition (CVD). Our current study of the 2-D graphane is motivated by its various apt applications that have been reported in literature. These include piezoelectricity, thermoelectricity, nano-electronics, transistors, nano-composites and electron-phonon superconductors [12]. In our current contribution, we do a comprehensive study of nitrogen substitutional point defects in graphane. We do an indepth study of nitrogen substitutional point defects of the form N_C, N_H and N_{CH} as shown by figure 1. The N_C substitutional point defect is formed when a nitrogen atom substitutes a carbon atom from its lattice site on a graphane monolayer as shown by figure 1(a). The substituted carbon atom has four valence electrons that enables it to covalently bond to other three nearest neighbour carbon atoms as well as to a hydrogen atom giving rise to sp^3 orbital hybridization. The N_H substitutional point defect is formed by inserting a nitrogen atom on a hydrogen atom lattice site as depicted by figure 1(b). The nitrogen atom's five valence electrons will cause the N_H defect in the neutral charge state to have four unpaired electrons since one of the five electrons will be involved in the single covalent bond between the nitrogen atom and the nearest carbon atom. The simultaneous substitution of a carbon atom together with its covalently bonded hydrogen atom by a nitrogen atom forms an N_{CH} point defect on a graphane monolayer as illustrated by figure 1(c). This defect has two unpaired electrons in the neutral charge state, N_{CH}^{0} , while N_{CH}^{-1} and N_{CH}^{+1} have three and one unpaired electrons respectively. Our current study is motivated by a number of factors, which among others include the fact that the knowledge of point defects in graphane is still not yet complete.

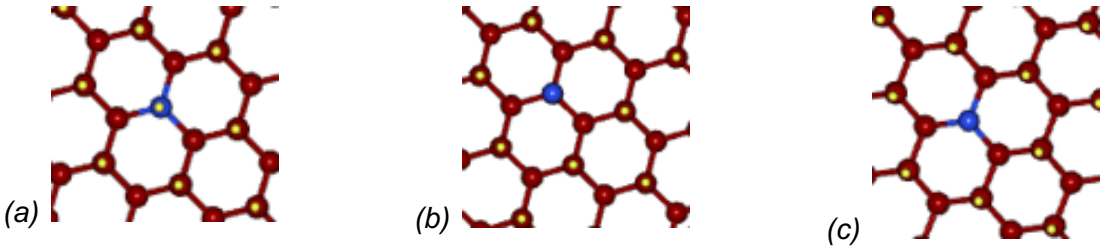


Figure 1: Nitrogen substitutional point defects in graphane: (a) N_C (b) N_H and (c) $N_{CH.}$ The red, blue and yellow spheres represent carbon, nitrogen and hydrogen atoms respectively.

2. Method

Our first principles calculations were performed within the confines of density functional theory (DFT) using the Kohn-Sham methodology [15]. We employed the Quantum espresso code (QE) [16] to run our calculations. Using QE plane wave basis set, we derived fundamental properties of the unit cell of graphane which comprises two carbon and two hydrogen atoms. In deriving these properties, we utilised the Heyd, Scuseria and Ernzerhof (HSE) Hybrid functional [17]. We used the ultra-soft pseudopotentials to model the interaction between the ions and the electrons [18]. We employed the supercell approach to run our entire calculations. The utilisation of the supercell method was motivated by the realisation that this approach facilitates the use of mathematical concepts that exploit the periodicity of the system under considerations [19]. Furthermore, this method can describe the band structure of the crystal system reasonably well and the results yielded by this approach are relatively easy to interpret as compared to other approaches that may present challenges in these aspects [20]. After the derivation of the requisite properties of the unit cell of graphane, we expanded

this cell to a $6 \times 6 \times 1$ supercell comprising 144 atoms - 72 carbon atoms and 72 hydrogen atoms. We created a separation of 15 Å in order to eliminate spurious interlayer interactions which can be a source of unwanted errors for this method. Convergence testing convinced us to use 450 eV and 4500 eV as the cut-off energy values of the wave function and charge density respectively. We used a $3 \times 3 \times 1$ Monkhorst –Pack [21] k-point sampling for the Brillouin zones. To ensure reliable results, the threshold value of the force convergence was set at 1.0×10^{-6} eV/Å.

To understand the properties of any given material, the point of departure is the derivation of the formation energy. As outlined by Van de Walle as well as Zhang and Northrup [22][23] , we derived the defect formation energy, $E^f[D^q]$, using the equation 1.

$$E^f[D^q] = E_T[D^q] - E_T[pure] - \sum_i n_i \mu_i + qE_F + E_{corr} \tag{1}$$

where $E_T[D^q]$ represents the total energy of a supercell in which the defect , D, of charge state q is embeded. $E_T[pure]$ is the total energy of the pristine supercell of the material under consideration and n_i is an integer that represents the number of $i-$type atoms that have been added to (for $n_i > 0$) or removed from (for $n_i < 0$) the supercell in order to create the defect. The term μ_i represents the chemical potentials of the type of atomic species at play while E_F denotes the Fermi energy of the electrons calculated relative to the valence band edge. E_{corr} represents the correction term that takes care of errors that may be introduced by the utilisation of the supercell approach.

We define thermodynamic level of transition for charge states q_a and q_b as the position of the Fermi-level whereby the formation energy of the two charge states are the same [24][25][20]. We determined the charge transition levels using equation 2, whereby $E^f(D^q;E_F = 0)$ represents the value of the formation energy for defect D which is in the q charge state when $E_F = 0$, i.e the Fermi level is located precisely at the valence band maximum.

$$\varepsilon(q_a|q_b) = [E^f(D^q{}_a;E_F = 0) - E^f(D^q{}_b;E_F = 0)]/(q_b - q_a) \tag{2}$$

The thermodynamic transition levels are of extreme importance because it has been shown experimentally that charge state q_a is stable for the Fermi level positions that are below $\varepsilon(q_a|q_b)$ and q_b is stable for the Fermi level positions above $\varepsilon(q_a|q_b)$ [25][20][24]. One other paramount concept we considered in this contribution is the U-parameter which is the energy difference between two consecutive thermodynamic energy transition levels [20]. We determined the U-parameter using the equation 3. The physical quantities E_f^{q+1}, E_f^{q-1} and E_f^q denote the respective formation energies of the defect in charge states $q + 1$, $q - 1$ and q taken in that order.

$$U = E_f^{q+1} + E_f^{q-1} - 2E_f^q \tag{3}$$

3. Results

Table 1 shows an overview of our first principles results for the three point defect configurations. In the next sub-section we systematically analyse our findings in this contribution.

3.1 Nitrogen substituting a carbon atom [N_C]

The dopant nitrogen atom has five valence electrons in its outermost shell. The adsorbed nitrogen atom is expected to also form sp^3 hybridization with three nearby carbon atom and one hydrogen atom. The additional electron brought by the nitrogen impurity is anticipated to substantially increase the electrical conductivity of the graphane material. The N_C point defect is an n-type point defect because

Table 1: Formation energies for nitrogen substitutional point defects in graphane

Defect	Charge State	Formation Energy(eV)	U − parameter (eV)	Fermi − level Position(eV)	Electron Difference With Pristine	Defect Type
N_C	0	8.91	0.14	+ 1.10	− 1	n − type
	−	8.64		+ 1.45	− 2	n − type
	+	5.66		− 2.21	0	−
N_H	0	2.74	−	− 1.15	− 4	n − type
	−	−		−	−	−
	+	2.37		− 2.67	− 3	n − type
N_{CH}	0	7.23	3.72	− 1.93	0	−
	−	7.20		+ 1.39	−1	n − type
	+	6.88		− 2.71	+ 1	p − type

of the presence of this one unpaired electron on the dangling bond. We expect nitrogen-doped graphane to depict paramagnetism or ferro-magnetism because of the presence of this unpaired electron. Our first principles calculations show that N_C defect formation process is non-spontaneous and endothermic with a high positive formation energy of 8.91 eV as shown by table 1. It is energetically expensive to form this defect. We tried to alter the formation energy by populating N_C^0 with an additional electron to form N_C^{-1}. The formation energy slightly decreased to 8.64 eV. Charging the defect by a single positive charge significantly lowers the formation energy albeit positive to 5.66 eV. We deduced that N_C point defect transforms according to the C_s(m) symmetry constraints while both N_C^{-1} and N_C^{+1} are subject to C_{3v}(3m) symmetry group transformations. Since C_{3v}(3m) point group is subject to the presence of spinor states, we suggest that this defect can be exploited for spintronics applications. We also noted that the formation of N_C^0 may be subject to the Jahn-Teller effect [26] because of the lowering of the symmetry from C_{3v}(3m) for pristine graphane to C_s(m) which is depicted in the neutral state of this point defect. Analysis of the DOS plots shows that N_C^0 and N_C^{-1} induces non-spin polarised defect states considerably away from the Fermi level. We noted spin-up states and spin-down states close to the conduction band minimum as shown by figure 2. The DOS graph shows that the majority spin states are symmetric to the minority spin states. We predict that N_C^0 and N_C^{+1} finetune the bandgap in such a way as to widen it while N_C^{-1} seemingly has no effect on the size of the bandgap. We again have observed that both N_C^0 and N_C^{+1} induces shallow donor levels as depicted by figure 2. We also noted that for N_C^0 and N_C^{-1} the Fermi level is pushed to a position very close to the conduction band edge while for N_C^{+1} it is positioned close to the valence band edge. As shown by table 1, N_C yielded a positive U-parameter value of 0.14 eV showing the stability of this defect in terms of allowing the formation of the negative and positive charged states.

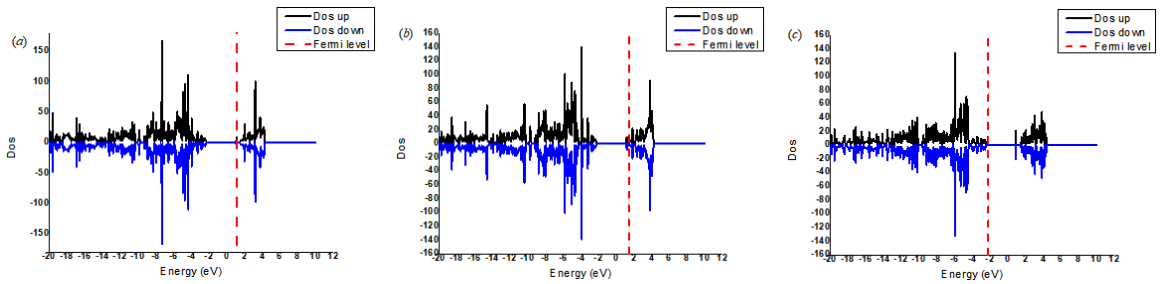


Figure 2: Total density of states (DOS) for a nitrogen atom substituting a carbon atom in graphane: (a) N_C^0 (b) N_C^{-1} (c) N_C^{+1}

3.2 Nitrogen substituting a hydrogen atom [N_H]

The N_H^0 and N_H^{+1} are expected to be n-type defects because of the presence of four and three unpaired electrons respectively. The formation energy of the N_H^0 defect is relatively low at a value of 2.74 eV hence the defect formation phenomenon is non-spontaneous. The withdrawal of a single electron to form N_H^{+1} decreased the formation energy slightly to 2.37 eV as depicted by table 1. In our calculations, the N_H point defect exhibited two charge states and hence has a single thermodynamic charge transition level. There is a lowering of symmetry from the C_{3s}(3m) to the C_S(m) symmetry group for the N_H point defect modified graphane. This seems to show the manifestation of the Jahn-Teller distortion [27]. The C_S(m) symmetry transformations are preserved for the neutral and positive charge states. However, the absorption of an atom on a hydrogen site causes a heavy geometric distortion of the lattice structure around the vicinity of the point defect. The N_H^{+1} defect is electrically active since there are clear spin-polarised defect states between the bandgap. If there are any defect states for N_H^0 and N_H^{-1}, we propose that these states are intertwined with bulk states close to the bandgap edges, hence they are not clearly distinct. The DOS plots shown by figure 3 shows that there is a possibility of this defect widening the bandgap. N_H^{-1} pushes the Fermi-level to the conduction band edge while both N_H^0 and N_H^{+1} moves the Fermi-level to positions close to the valence band edge.

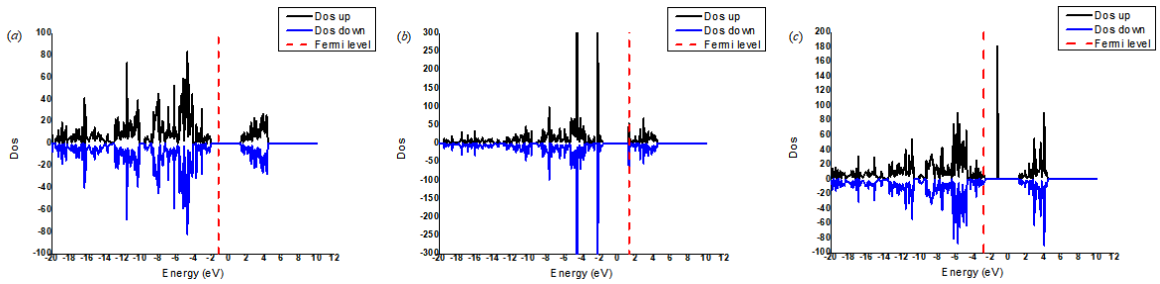


Figure 3: Total density of states (DOS) for a nitrogen substituting a hydrogen atom in graphane: (a) N_H^0 , N_H^{-1} and (c) N_H^{+1}

3.3 Nitrogen substituting both a carbon and hydrogen atom [N_{CH}]

The N_{CH} defect formation process is endothermic and non-spontaneous. This defect is energetically expensive to form with a high positive formation energy of 7.23 eV. Charging the defect negatively and positively slightly reduces the formation energies to 7.20 eV and 6.88 eV respectively as recorded in table 1. The presence of this defect lowers the symmetry of the graphane monolayer from the high symmetry point group C_{3v}(3m) to the point group C_s(m) that has lower symmetry transformations. This change in symmetry may be due to the pseudo Jahn-Teller effects [26]. The withdrawal or the addition of a single electron restores the symmetry to C_{3v}(3m). Spinor functions are inherent to the C_{3v}(3m) point group, hence there is a likelihood of utilising this defect for spintronics applications. A careful analysis of the DOS plots for N_{CH}^0 shows that the majority spin states are symmetric to the minority spin states proving that the material is non-magnetic.

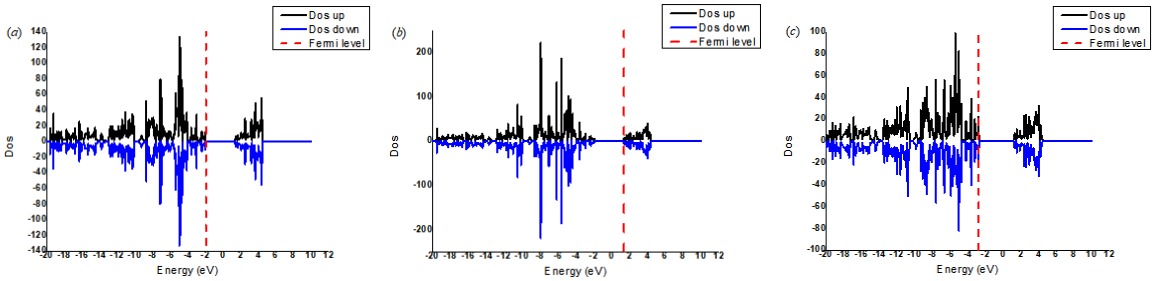


Figure 4: Total density of states (DOS) for a nitrogen atom substituting both a carbon and hydrogen atom simultaneously in graphane: (a) N_{CH}^{0} (b) N_{CH}^{-1} and (c) N_{CH}^{+1}

We have two dangling bonds from the dopant nitrogen atom and we propose that the two electrons on these broken bonds are likely to form a pair. Our DOS results as shown by figure 4 seem to show that this point defect doesn't induce states in the middle of the bandgap. If there are any states introduced by this defect in the bandgap, they are coupled with the bulk states and are therefore not well separated from the band edges. This finding seems not to be in line with Pujari and Kanhere's [28] finding that this type of defect in graphane induces spin-polarised states in the middle of the bandgap. The use of the tetrahedra method to deduce the DOS plots may have caused this uncertainty as far as our current contribution is concerned. However, we do not hasten to add that the presence of N_{CH} defect in a graphane monolayer modifies the bulk states close to the band edges and the N_{CH}^{-1} seem to show some magnetic traits since some states are asymmetric. N_{CH} defect is more stable in the neutral state than the singly positive charge state. The positive U-parameter value of 3.72 eV for N_{CH} shows that this defect does not show instability if positive and negative charged states form. However, the transition from one charge state to another requires a relatively greater amount of energy as compared to N_C point defect.

4. Conclusion

In this contribution we systematically investigated the thermodynamic stability as well as the electronic properties of point defect modified graphane monolayer. We characterised point defects in the form of nitrogen substitutional dopants. We found out that all the point defects considered in this report have relatively high positive formation energies. We also found out that charging the defects increases the formation energy when the Fermi level is fixed at 0 eV. If the Fermi level is not fixed, charging the defects generally decreases the formation energy. All the defects need activation energy to form. We also noted that these point defects are generally subject to the Jahn-Teller effect which lowers the energy of a given system by removing degeneracy. One of our findings is that the n-type point defects do not induce magnetism or defect states in the middle of the bandgap. The findings by Wang *et. al*. [29] was that the n-type defects generally do not induce magnetism in graphane. Our results are considerably consistent with this report. N_C fine-tuned the graphane bandgap and induced localised spin-polarised defect energy states. Therefore these defects can be exploited for tailored applications such as magneto-optical sensing, quantum metrology and nano-electronic uses. N_C and N_{CH} gave rise to positive effective values of the U-parameter. These defects thus exhibit stability should single negative and single positive states form. The arrangement of the thermodynamic transition states are not inverted and this explains the stability of the defect. To the best of our knowledge, this has not yet been reported. Moreover, our current contribution has armed us with the critical knowledge that we required for us to be on a good platform to intensively study the resultant complex formations of these point defects in our next contribution in future. This is worth investigating in order to determine whether a combination of the native defects and the substitutional extrinsic defects can lower the activation energy.

Acknowledgements

The authors are greatly indebted to the university of Pretoria post graduate funding and the Centre for High Perfomance Computing (CHPC) for providing the simulation resources for this project. REM thanks NRF and NiTheCS for supporting the project financially.

References

[1] Novoselov K S, Geim A K, Morozov S V, Jiang D, Zhang Y, Dubonos S V, Grigorieva I V and Firsov A 2004 *Science* **306** 666-69
[2] Araujo P T, Terrones M and Dresselhaus M S 2012: *Materials today* **15** 98-109
[3] Zhang Y G, Cheng G D, Peng W and Tang Z 2014 *Comp. Mat. Sci*. **95** 316-19
[4] Dreyer C E, Alkauskas A, Lyons J L, Janotti A and Van de Walle C G 2018 *Annu. Rev. of Mater. Res*. **48**, 2.1-2.26
[5] Podlivaev A I and Openov L A 2017 *Condensed Matter* **106** 110-15
[6] Geim A K N and Novoselov K S 2007 *Nature Materials* **6** 183-91
[7] Mapasha R E, Molepo M P and Chetty N 2016 *Physica E* **79** 52-58
[8] S. Haldar S and Sanyal B 2016 *Defects in Graphene and its Derivatives*(IntechOpen:Nayak P K)
[9] Sahin H, Leenaerts O, Singh S K and Peeters F M 2015 *Cond-mat. mtrl-sci*. 1-14,
[10] Samarakoon D K and Wang X Q 2011 *Structural and Electronic Properties of Hydrogenated Graphene, Physics and Applications of Graphene - Theory* (InTech)
[11] Sofo J O, Chaudhari A S and Barber G D 2007 *Phys. Rev. B 75*, 153401
[12] Zhou C, Chen S, Lou J, Wang J, Yang Q, Liu C and Huang D 2014 *Nanoscale Research Letters* **9** 26
[13] Reshak A H and Auluck S 2014 **4**(70) 37411-18 *The Royal Society of Chemistry*
[14] Whitener K E 2018 *J. Vac. Sci. Technol. **A*** **36**, 05G401, 1-15
[15] Kohn W and Sham L J 1965 *Phys. Rev.*, **140**, A1133
[16] Giannozzi P *et al* 2009 *Condens. Matter* **21** 395502
[17] Heyd J, Scuseria G E and Ernzerhorf M 2003 *Journal of Chemical Physics* **118** 8207-8215.
[18] Vanderbilt D 1990 *Phys. Rev*. B, **41**, 7892-7895
[19] Koch W and Holthausen M C 2001 *A Chemist's Guide to Density Functional Theory* (Wiley-Vch)
[20] Freysoldt C, Grabowski B, Hickel T, Neugebauer J, Kresse G, Janotti A and Van de Walle C G 2014 *Rev. Mod. Phys.* **86**. 253-305
[21] Monkhorst H J and Pack J D 1976 *Phys. Rev. B* **13** 5188-5192
[22] Van de Walle C G, Laks D B, Neumark G F and Pantelides S T 1993 *Phys. Rev. B*. **47** 9425
[23] Zhang S B and Northrup J E 1991 *Phys. Rev. Lett.***67** 2339
[24] Gali A 2019 *Nanophotonics* **8** (11)
[25] Thiering G and Gali A 2016 *Phys. Rev. B* **94** 125202
[26] Bersuker I B 2001*Chem. Rev*.101, **4** 1067-1114
[27] Gehring G A and Gehring K A 1975 *Rep. Prog. Phys.***38** 1-89
[28] Pujari B S and Kanhere D G 2009 *J. Phys. Chem*. **113**(50). 21063-21067
[29] Wang Y, Ding Y, Shi S and Tang W 2011 *Appl. Phys. Lett*, 98. 163104

The study of the properties of Titanium alloys for biomedical applications employing the first-principles approach

Kobe Mabeba[1*], Enoch Sithole[1], and Rosinah Modiba[1,2]

[1]Department of physics, Sefako Makgatho Health Sciences University, P.O. Box 94, Medunsa Ga-Rankuwa, Pretoria, 0204, South Africa
[2]Materials Science and Manufacturing, Council of Scientific and Industrial Research (CSIR) P.O Box 395 Pretoria 0001, South Africa
E-mail: **mabebak.2@gmail.com**

Abstract. Titanium and its alloys have been widely used as implant materials for replacement of damaged hard tissues. This was due to their outstanding mechanical characteristics and biocompatibility. Some of the applications includes orthopaedic, and endoderm surgeries etc. However, there is a great concern regarding the difference in the implant material's Young's modulus and that of natural bone. The prolonged use of Ti-alloys resulted in the failure of the implants due to stress-shielding effect, hence the need to improve the Young's moduli of these alloys. The aim of the study is to investigate the structural and mechanical stability of $Ti_{12.5}$-$Mg_{87.5-x}$-Si_x alloys using first principle approach. Interestingly, the Young's modulus of the $Ti_{12.5}$-$Mg_{87.5-x}$-Si_x is improved by alloying with Si. Furthermore, the density of states for the alloys suggests that 1.25 at. % Si stabilises the $Ti_{12.5}$-$Mg_{87.5}$-Si_x alloys.

1. Introduction

With the growth of the world's older population, demand for biomedical implants for the replacement of failed hard tissues such as bones, knee implants, and teeth is increasing [1]. A biomedical implant is a foreign substance placed within the human body in order to repair broken hard tissues such as ligaments, tendons or teeth for the purpose of improving the human life over a long period of time [2].

Hamidi et al (2017) noted the work done by Lane in the 1800s, where the metal plates were used as biomedical implants to replace fractures in the human body and began using biomedical implants [3]. Metal implants presented the risk of harmful ion secretions when placed inside of a human body [3]. Ti-alloys presented superior biocompatibility and corrosion resistance as compared to metal implants, CoCr, and stainless-steel biomedical implants [4]. Beta (ß)-Ti alloys presented the lowest young's modulus [5]. Ti-alloys are biomedical implants made up of Titanium as a substrate and metallic elements as components, such as Al, V, Ni and Nb etc., as displayed by for example, $Ti_xAl_6Nb_7$ and $Ti_xAl_6V_4$ alloys [6].

The most popular used ß-Ti alloy is $Ti_xAl_6V_4$ which amounted to 75-80% of the overall manufactured Ti-alloys [3]. Recent investigated alloys included $TixAl_6Nb_7$ and $Ti_xAl_5Fe_{2.5}$ of which they carried greater mechanical abilities than $Ti_xAl_6V_4$ alloy with both Al and V showing the toxicity trait of harmful secretions [6]. An elastic modulus investigation for $Ti_xAl_6Nb_7$ and $Ti_xAl_5Fe_{2.5}$ alloys has shown that the Young's modulus value of the two alloys was greater than that of the human bone [7], which resulted in stress shielding effect. Therefore, there is a need to improve the Ti-based alloys for use in biomedical applications by replacing the toxic elements with those that are safe for use in the human body.

Young's modulus is one of the most important properties of an alloy to look at as it relates to the comfort and the adhesiveness of an alloy. The rise in the modulus value of a material over the human

bone, generally results in the loosening of the material and this leads to the formation of a stress shielding effect.

$(Ti_{5-x}Mg_x)$-Si_3 intermetallic structures were investigated in a recent article using a hexagonal prototype structure of Mn_5Si_3 [8]. Hexagonal crystal structures are discovered to have a higher young's modulus as compared to the cubic structures [9].

In this study, first principle method was used to determine the ternary alloying of a cubic $Ti_{12.5}Mg_{87.5}$ with Si. The stability of the alloys was investigated by determining the structural properties, elastic constants and density of states of the Ti alloys.

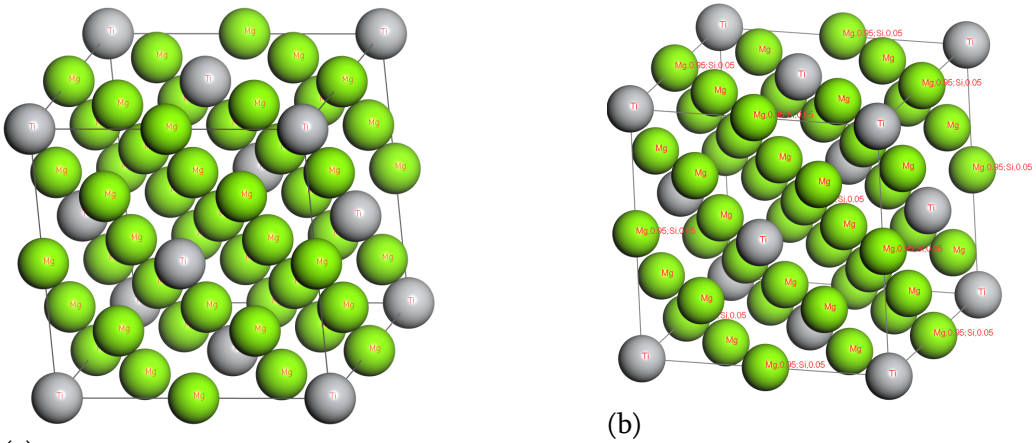
2. Methodology

The calculations were performed with the ab initio density functional theory (DFT) [10, 111] formalism implemented in CASTEP [12] under the projector augmented wave (PAW) [13]. The energy cut-off was 320 eV, which was sufficient to combine the whole energy of the structure. For the exchange correlation function, Perdew and Wang's generalized gradient approximation (GGA-PBE) [14] was chosen. According to Monkhorst and Pack, the Brillouin zone integration is carried out for a sufficiently large set of k points [15]. The solid solution approach (disordered) was used to replace Mg with Si atoms, using visual crystal approximation (VCA) [16] embedded in CASTEP. This is a much simpler and cheaper computational approach that investigates crystals composed of fictitious "virtual" atoms, but with primitive periodicity but interpolated between the behaviour of the atoms in the parent compound. This technique is widely used in band structure calculation [17-19]. VCA module provides an advantage to study smaller concentrations of Si. The structural properties, elastic constants, and density of states for the structures are calculated. The optimization of structural parameters (atomic positions and lattice parameters) is achieved by minimizing the stress tensor and stress tensor forces.

3. Results and discussions

3.1. Structural properties and elastic constants

Figure 1 shows the cubic Fm-3m $Ti_{12.5}Mg_{87.5}$ structure with 32 atoms [20]. The structure was optimized and alloyed with Si as shown in Figure 1 (b). In Table 1, the calculated lattice parameters and elastic constants of the $Ti_{12.5}Mg_{87.5-x}Si_x$ alloy with x varied from 0,625 to 1.875 atomic percentage employing the disordered approach are shown. The lattice parameter of the $Ti_{12.5}Mg_{87.5}$ decreases in the addition of 0.625 at. % Si on the Ti side and increases minimally as Si is increased.



(a) (b)

Figure 1: Structures of (a) $Ti_{12.5}Mg_{87.5}$ and (b) alloyed with Si to form $Ti_{12.5}Mg_{87.5-x}Si_x$ alloy with x varied from 0.625 to 1.875 atomic percentage.

The accuracy of the calculation of elasticity is important for determining the mechanical stability and elasticity of any system. The cubic structure has three independent elastic constants (C_{11}, C_{12} and C_{44}). The mechanical stability criteria system, as described in the other parts of the article [21], are as follows:

$$C_{44} > 0,\ C_{11} > |C_{12}| \text{ and } C_{11} + 2C_{12} > 0 \quad (1)$$

$$C' = \frac{(C_{11} - C_{12})}{2} \quad (2)$$

For structural stability to be considered, the elastic constant stability criterion must be met. In principle, the positive C′ indicates the mechanical stability of the crystal. Interestingly, all independent elastic constants C_{11}, C_{12} and C_{44} are positive throughout the concentration range, indicating the stability of the alloy.

Table 1: Lattice parameters a (Å) and the elastic properties (GPa) of the $Ti_{12.5}Mg_{87.5-x}Si_x$ (x:0.625, 1.25, 1.875) ternaries and their anisotropy A obtained using the disordered approach.

Structure	a (Å)	C_{11}	C_{12}	C_{44}	C’	A
$Ti_{12.5}Mg_{87.5}$	8.995	30	20	21	5	4.2
$Ti_{12.5}Mg_{86.875}Si_{0.625}$	8.901	52	34	37	9	4.1
$Ti_{12.5}Mg_{86.25}Si_{1.25}$	8.919	50	33	38	8	4.7
$Ti_{12.5}Mg_{85.625}Si_{1.875}$	8.950	45	31	37	7	5.2

3.2. Mechanical properties

In addition, the mass (B), cutting (G) and Young modules (E) of structures have been calculated using the Voigt–Reuss–Hill (VRH) average scheme [22- 24] presented in Figure 2. The bulk module generally characterizes the resistance to deformation under applied stress, whereas the cutting module is a measure of the capacity to deform resistant under deformation stress [25]. The higher the value, the greater its resistance to deformation. According to Pugh, the conductivity or fracture behaviour of solids is estimated by the ratio of slicing module to volume module [26]. The authors proposed empirical criteria for ductile metals with G/B below 0.57, and brittle metals with greater than 0.57.

In Table 2, G/B for all the structures was found to be less than 0.57 except for the 1.875 at. % Si which was found to be more. This suggests that the structures are ductile. Interestingly, the Young’s modulus was increased with an increasing the Si. These values are closer to the value of human bone Young’s modulus.

Table 2: The bulk modulus (GPa), shear modulus (GPa), Young’s modulus (GPa) of the $Ti_{12.5}Mg_{87.5-x}Si_x$ (x:0.625, 1.25, 1.875) structures. The Pugh ratio G/B is also calculated.

Structure	B	G	E	G/B
$Ti_{12.5}Mg_{87.5}$	24	12	31	0.5
$Ti_{12.5}Mg_{86.875}Si_{0.625}$	40	22	55	0.55
$Ti_{12.5}Mg_{86.25}Si_{1.25}$	39	21	53	0.54
$Ti_{12.5}Mg_{85.625}Si_{1.875}$	35	20	50	0.57

3.5. Density of states

Total DOS at Fermi level (E_f) is an important indication of the stability of alloys where the high number of DOS suggests the instability whereas the lower number indicates the stability of the structure.

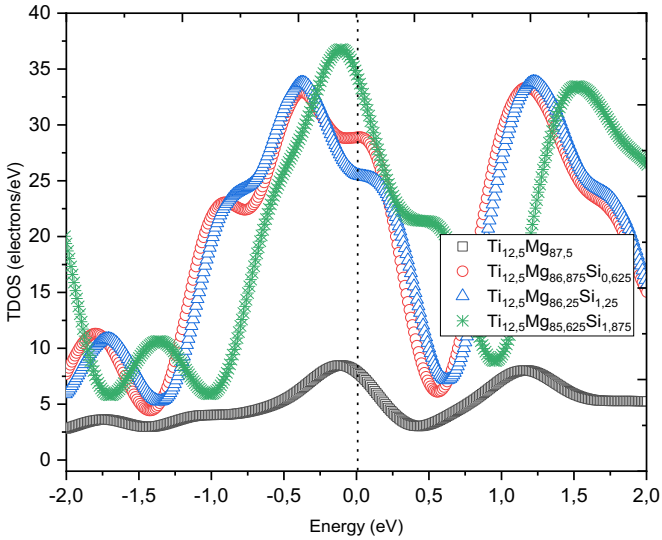


Figure 2: TDOS results for the $Ti_{12.5}Mg_{87.5-x}Si_x$ (x:0.625, 1.25, 1.875) alloys.

To compare its stability more accurately, the complete electronic DOS superimposed on the structure is shown and compared in Figure 2. In agreement with the mechanical properties, the 1.875 at. % Si is the most unstable and the DOS is the highest at Fermi level and $Ti_{12.5}Mg_{87.5}$ alloy the most stable with the lowest DOS. Interestingly, the 1.25 at. % Si is found to be the most stable ternary.

4. Conclusions

First principle method has been employed to determine the ternary alloying of $Ti_{12.5}Mg_{87.5-x}Si_x$ with Si. The addition of 0.625 at. % Si decreased the lattice parameters. As the Si content was increased above 0.625 at %, the lattice parameters then increased slowly. It has been found that this structure is stable, and all elastic constants are positive. The Young modulus of $Ti_{12.5}Mg_{87.5-x}Si_x$ increased with increasing the concentration of Si atoms. Furthermore, after the addition of 1.25 at. % of Si atoms, the density of states for $Ti_{12.5}Mg_{87.5}$ was found to be lower at the Fermi level, which indicates the stability of the alloy.

Acknowledgements

This project was financially supported by the physics department at SMU. The APRDP funding from CSIR is acknowledged. The CHPC at Cape Town provided the computational support.

References

[1] Li Y, Yang C, Zhao H, Qu S, Li X and Li Y 2014 *Matter* **3** 1709
[2] Xiao M, Chen Y M, Biao M N, Zhang X D and Yang B C 2017 *Mater. Sci. Eng. C* **70** 1057
[3] Hamidi H F, Harun W S, Samykano M, Ghani S A, Ghazalli Z, Ahmad F and Sulong A B 2017 *Mater. Sci. Eng. C*. **78** 1263
[4] Hussein A H, Gepreel M A H, Gouda M K, Hefnawy A M and Kandil S H 2016 *Mater. Sci. Eng. C*. **61** 574
[5] Zhuravleva K, Chivu A, Teresiak A, Scudino S, Calin M, Schultz L, Eckert J and Gebert A 2013 *Mater. Sci. Eng. C*. **33** 2280
[6] Niinomi M, Nakai M and Hieda J 2012 *Acta Biomater* **8** 3888
[7] Pan Y and Guan M M, Ceramics. Int. **45** 15649
[8] Chen X J, Zeng M X, Wang R N, Mo Z S, Tang B Y, Peng L M and Ding W J 2012 *Comput. Mater. Sci.* **54** 287
[9] Pan Y 2023 *Int. J. Refract. Hard. Met*. **115** 106277
[10] Hohenberg P and Kohn W 1964 *Phys. Rev. B*. **136** 864
[11] Kohn W and Sham L J 1965 *Phys. Rev. A*. **140** 1133
[12] Clark S J, Segall M D, Pickard C J, Hasnip P J, Probert M J, Refson K and Payne M C, *Z. Kristallogr.* **220** 567

[13] Blöchl P E 1994 *Phys. Rev. B*. **50** 17953
[14] Perdew J P and Wang Y 1992 *Phys. Rev. B*. **45** 13244
[15] Monkhorst H J and Pack J D 1976 *Phys. Rev. B*. **13** 5188
[16] Nordheim L 1931 *Ann. Phys.* **401** 607
[17] De Gironcoli S, Giannozzi P and Baroni S 1991 *Phys. Rev. Lett*. **66** 2116
[18] Marzari N, de Gironcoli S and Baroni S 1994 *Phys. Rev. Lett.* **72** 4001
[19] Ramer N J and Rappe A M 2000 *J. Phys. Chem. Solids.* **61** 315
[20] Kyoi D, Sato T, Rönnebro E, Kitamura N, Ueda A, Ito M, Katsuyama S, Hara S, Noréus D, and Sakai T 2004 *J. Alloys. Compd*. **372** 213
[21] Wu Z J, Zhao E J, Xiang H P, Hao X F, Liu X J and Meng J 2007 *J. Phys Rev. B*. **76** 054115
[22] Ballato A 2009 *Ieee International frequency control symposium: Conf. Series*, **154** 150
[23] Reuss A X and Angew Z 1929 *Math. Mech*. **9** 49
[24] Hill R 1952 *Proc. Soc.* **65** 464
[25] XCII P S F 2009 *Mag. J. Sci.* **45** 823
[26] Pugh S 1954 *Philos. Mag.* **45** 823

The study of Zr and Nb alloyed on the beta-Ti for bio-medical applications: First principle approach

Thabiso Mathews[1*], Enoch Sithole[1], and Rosinah Modiba[1&2]
[1]Department of Physics, SMU, P.O. Box 94, Medunsa, Ga-Rankua, 0204, South Africa.
[2]Department of Material Science and Manufacturing, CSIR, P.O. Box 395, Pretoria, 0001, South Africa.
E-mail: **mathewsthabisoabel@gmail.com**

Abstract. Titanium alloys have been used in many areas such as aviation and bio-medical industry due to their high strength and resistance to fatigue deterioration. Alloying elements such as Nb and Ta have been the primary focus as they are effective β-phase stabilising elements and can effectively reduce the modulus of Ti alloy and are less toxic. However, these dopants tend to increase the elastic modulus of the alloy. As of late, there has been a crucial need to develop non-toxic, more biocompatible, and allergy-free components for bio-medical applications. The purpose of this study is to use first principle approach to investigate the effect of alloying Ti with Nb and Zr to improve the mechanical properties of Ti alloys for use in human implantations.

1. Introduction

Titanium alloys have been widely employed in the aviation, chemical, and medical industries. In the medical industry it is clinically used for orthopaedic implants [1]. Titanium alloys used in the bio-medical industry are required to be both corrosion resistant and biocompatible. Commercially, pure titanium and titanium alloys have been used frequently in biomedical applications since the 1950s [2]. This is because of its low elastic modulus, excellent fatigue strength, high specific strength, good corrosion resistance, and consistent performance, just to name a few [3].

Recently, non-toxic, allergy-free components have been used to create Ti alloys, and their biocompatibility has been taken into consideration. The two primary alloys, Ti-6Al-4V and so-called commercially pure titanium (cp Ti), have clinical success rates of up to 98% after 10 years [4]. Furthermore, both alloys can attach firmly and are biocompatible when in touch with bone and gingival tissues [4]. The metals deterioration is caused by the corrosive environment in humans that contains sodium, water, and protein plasma. This susceptibility to corrosion and compatibility with blood is the main reason why metallic implants are rejected [5]. Furthermore, the long-term, serious health issues could result from the release of harmful aluminum (Al) and vanadium (V) ions. Particularly, cytotoxic V ions and Al have the potential to cause adverse reactions in human tissues and to contribute to neurological disorders like Alzheimer's disease [6].

Numerous studies have undoubtedly focused on β alloys due to their outstanding corrosion resistance, low stiffness, and super-elastic qualities utilising Ta, Mo, Zr, Nb as alloy additives [7]. Nb and Zr components appear to be more biocompatible and less likely to have negative effects like cytotoxicity or neurological disorders. However, Ti alloys' abrasion resistance is rather low [7].

The use of $(MoTa)_xNbTiZr$ alloys with medium and high entropy in biomedical implants was carefully examined [7].Under compression, alloys containing up to 0.4 mole fractions of Mo and Ta displayed a plastic strain of greater than 30%. The findings of the Nano indentation experiment demonstrated that the system's elastic modulus increased linearly with the addition of Mo and Ta. From

a biological standpoint, it was assumed that the inclusion of Ta and Mo over a certain concentration (mole fraction = 0.4) was unfavourable since it increased the system's brittleness and elastic modulus while reducing its ductility [7].

This study aims to investigate the effects of Zr and Nb on mechanical properties of titanium material for bio-medical applications employing the first principle approach. The stability of the alloys with respect to thermodynamic, elastic properties and density of states were studied.

2. Methodology

This work made use of the Perdew-Burke-Eruzerhof (PBE) scheme and the generalized gradient approximation [8] in the Cambridge Sequence Total Energy Package (CASTEP) package [9]. The Brillouin-zone (BZ) integration Pack has been used with a 500 eV energy cut-off and 5×5×5 Monkhorst-Pack k-point grids [10] for the TiZr and TiNb systems. Using ultrasoft pseudopotentials and Koeling-Harmon relativistic effects. The Broydene-Fletchere-Goldarbe-Shanno (BFGS) approach [11], based on total energy minimization, and was used to optimize crystals. The optimization and energy calculations' convergence criteria were carried out with fine quality and the energy convergence of 10^5 eV/atom were used for all calculations. The convergence tolerance for maximum force was (0.03 eV/atom) and maximum displacement at (0.001 atom).

3. Results and discussion

Figure 1 (a) depicts the structure of Ti, with im-3m space group symmetry of a cubic B2 structure also referred to as a high temperature beta phase with the lattice parameter of 3.26 Å. Ti was alloyed with 6.25 %, 12.5%, 18.75% atomic dopant of Nb and Zr on a supercell of 16-atoms as shown in Figure 1 (b).

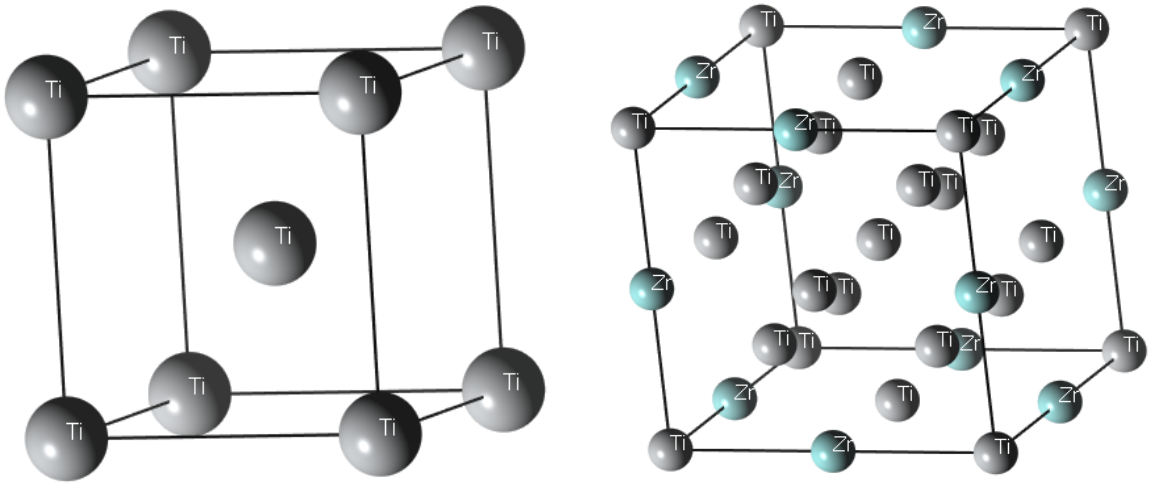


Figure 1. The beta B2 Ti (a) pure structure (b) and alloyed structure.

3.1 Lattice parameters

Table 1 shows dopant at different concentrations and lattice constants for the alloys ($Ti_{100-x}Nb_x$) and ($Ti_{100-x}Zr_x$) for this study and are compared with the available theoretical data [12] and [13]. It is evident that the lattice constant increases gradually as the alloying component concentration rises. This can be attributed to Nb and Zr having somewhat larger atomic radii than Ti. The $Ti_{81.25}Nb_{18.75}$ is greater (3.262 Å) than that of $Ti_{93.75}Nb_{6.25}$ and $Ti_{87.5}Nb_{12.5}$ with (3.258 Å) and (3.257 Å), respectively. Interestingly, at 12.5% Nb the lattice parameters of the TiNb structure decreases slightly and this can be attributed to the stabilising ability of two atoms of Nb. However, the lattice parameters of TiZr increase rapidly in agreement with theoretical data.

Table 1: Equilibrium lattice parameters (Å) for $Ti_{100-x}Zr_x$ and $Ti_{100-x}Nb_x$ alloys

Structures	a=b=c
Ti_{100}	2.94
$Ti_{50}Zr_{50}$[a]	(3.42)
$Ti_{93.75}Zr_{6.25}$	3.278
$Ti_{87.5}Zr_{12.5}$	3.297
$Ti_{81.25}Zr_{18.75}$	3.323
$Ti_{0.75}Nb_{0.25}$[b]	(3.27)
$Ti_{93.75}Nb_{6.25}$	3.258
$Ti_{87.5}Nb_{12.5}$	3.257
$Ti_{81.25}Nb_{18.75}$	3.262

[a][12]
[b][13]

3.2 Thermodynamic Stability and elastic properties.
Calculations of formation enthalpies allows us to determine and examine thermodynamic stability of materials. Table 2 depicts the correlation between the formation enthalpy, elastic constants, and the alloying element content. The energy of formation per atom where determined using equation (1),

$$E_{formation} = \frac{E_{products}}{16} - \left[\left(\frac{(16-x)E_{Ti}}{16}\right) + \left(\frac{(x)E_{M(x)}}{16}\right)\right], \quad (1)$$

where *E* is energy per atom and *M* is Nb or Zr and *x* is the number of atoms. In general, compounds with the least heats of formation entail greater stability whereas highest value indicates instability. The addition of Nb on the Ti atom decreases the formation energy of the compounds, however, the positive heats of formation (compounds required energy) for the binary $Ti_{16-x}Zr_x$ and $Ti_{16-x}Nb_x$ alloys indicates an endothermic process where the products are less stable than the reactants. $Ti_{81.25}Nb_{18.75}$ has the lowest formation energy. The C_{11}, C_{12}, and C_{44} of $Ti_{16-x}Zr_x$ and $Ti_{16-x}Nb_x$ alloys vary with alloying element content. The stability criterion for a cubic [14] structure are given by the equations (2) and (3).

$$C_{11} + 2C_{12} > 0 \,, (C_{11} - C_{12}) > 0 \text{ and } C_{44} > 0 \quad (2)$$

$$C' = \frac{(C_{11}-C_{12})}{2} \quad (3)$$

It can be seen in Table 2 that as Zr content is added to Ti, the alloy becomes mechanically unstable due to C_{11} being less than C_{12} these resulting in a negative C'. Interestingly, when Nb is added on the Ti, the structure is found to be mechanically stable with all the C_{ij}'s being positive. With an increase in dopant concentration, both C_{11} and C_{12} of $Ti_{16-x}Nb_x$ alloys rise, whilst C_{11} and C_{12} of $Ti_{16-x}Zr_x$ decrease and result in failing to meet the Born Huang Stability criteria.

3.3 *Mechanical Properties*
The mechanical properties of a material, such as its shear modulus, Poisson's ratio G/B, bulk modulus, and Young's modulus, can be studied using the elasticity parameters of matter [14]. Table 3 shows the mechanical properties of $Ti_{100-x}Zr_x$ and for $Ti_{100-x}Nb_x$. The Young's moduli of the $Ti_{100-x}Zr_x$ (x: 6.25, 12.5, 18.75) and shear modulus were found to be unstable with the negative values. Interestingly, the Young's Modulus of Ti increases with an increase in Nb dopant, and the value is closer to that of the

Table 2: Formation energy and elastic constants for $Ti_{100-x}Zr_x$ and for $Ti_{100-x}Nb_x$ (x: 6.25, 12.5, 18.75) structures.

Structures	ΔE_f (eV)	C_{11} (GPa)	C_{12} (GPa)	C_{44} (GPa)	C' (GPa)
Ti_{100}		187	17	47	85
$Ti_{93.75}Zr_{6.25}$	0.113	88	111	43	-12
$Ti_{87.5}Zr_{12.5}$	0.133	86	114	39	-14
$Ti_{81.25}Zr_{18.75}$	0.116	79	104	37	-12
$Ti_{50}Zr_{50}$[a]		(92)	(96.8)	(37)	-
$Ti_{0.75}Nb_{.25}$[b]		(128)	(115)	(15)	
$Ti_{93.75}Nb_{6.25}$	0.102	108	41	108	34
$Ti_{87.5}Nb_{12.5}$	0.094	123	116	40	4
$Ti_{81.25}Nb_{18.75}$	0.084	129	110	27	9.41

[a][12]
[b][13]

bone's Young's modulus value (10-30 GPa) [6]. This is consistent with existing data reported by Gao *et al*., [16] where DFT calculations of BCC Ti, Hf and Zr predicted concentrations of 5-15 % of the dopants to be insufficient to reach a stable $Ti_{100-x}Zr_x$ alloy.

$Ti_{87.5}Nb_{12.5}$ shows the highest resistance to deformation with a bulk modulus of 118 (GPa). Pugh's ratio was employed to determine whether a solid was brittle or ductile. The empirical standards stipulate that for brittle metals, G/B must be greater than 0.57 and less than 0.57 for ductile metals [17]. The Ductility of $Ti_{100-x}Nb_x$ compounds decrease with increasing amount of doping atoms, with the most ductile being $Ti_{93.75}Nb_{6.25}$.

Table 3: The Young's (Y), shear (G) and bulk moduli (B) of the $Ti_{100-x}Zr_x$ and for $Ti_{100-x}Nb_x$. The Pugh's ratio (G/B) of the alloys is also calculated.

Structure	*B* (GPa)	*G* (GPa)	*Y* (GPa)	*G/B* (GPa)
Ti_{100}	134	51	113	0.38
$Ti_{93.75}Zr_{6.25}$	103	-15	-46	-
$Ti_{87.5}Zr_{12.5}$	105	-29	-97	-
$Ti_{81.25}Zr_{18.75}$	100	-28	-94	-
$Ti_{93.75}Nb_{6.25}$	108	12	36	0.11
$Ti_{87.5}Nb_{12.5}$	118	17	48	0.14
$Ti_{81.25}Nb_{18.75}$	117	18	51	0.15

3.4 Density of states

Figure 2 and 3 show the partial (PDOS) and total (TDOS) density of states of mater for both $Ti_{100-x}Zr_x$ and $Ti_{100-x}Nb_x$ systems, respectively. The densities of states provide information on the nature of bonds, state hybridization, and the relationship of orbitals to the band structure of materials.

The TDOS of structures with identical constituents can be used to simulate the stability trend in terms of how they behave at the Fermi level, where the structure with the greatest and smallest density of states at E_f is regarded as the least and most stable, respectively. TDOS for $Ti_{100-x}Zr_x$ and $Ti_{100-x}Nb_x$ structures is made up of low but broad s- derived bands, intermediate p-derived bands and narrow but high d-derived bands as can be seen in Figure 2. The TDOS are mostly contributed by the d-states with s-orbitals contributing minimally. It can be seen in Figure 3 that $Ti_{87.5}Nb_{12.5}$ is found to be the most stable with the lowest energy at E_f with $Ti_{93.75}Zr_{6.25}$ being the least stable with the highest DOS at Fermi. The stability trend of the TDOS is in good agreement with the heats of formation results were $Ti_{87.5}Nb_{12.5}$ was found to be thermodynamically stable.

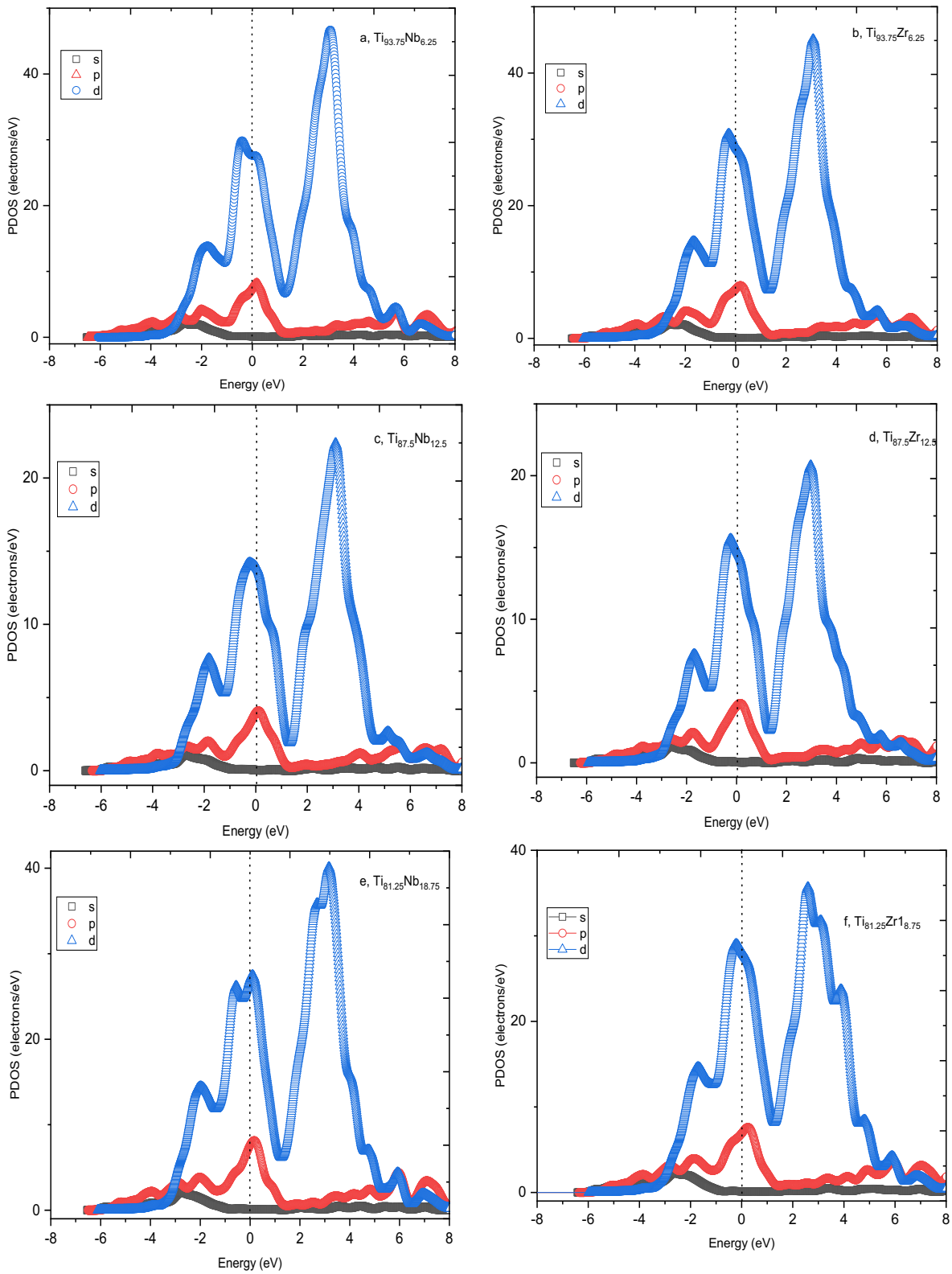


Figure 2 (a-f). PDOS for $Ti_{100-x}Zr_x$ and $Ti_{100-x}Nb_x$ structures.

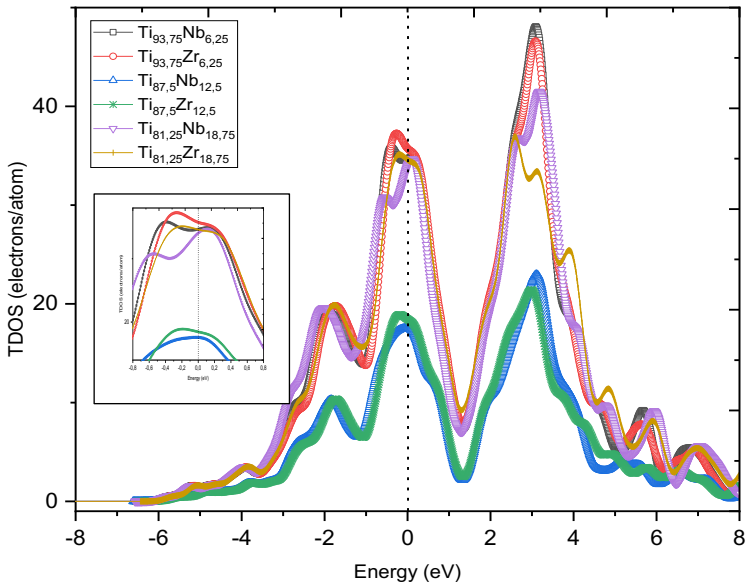


Figure 3. TDOS for $Ti_{100-x}Zr_x$ and $Ti_{100-x}Nb_x$ structures.

4. Conclusions

This paper presents the first-principle calculations performed to optimize the structures of $Ti_{100-x}Zr_x$ and $Ti_{100-x}Nb_x$ unit cells and to investigate their lattice parameters, heats of formation, elastic constants and density of states. The results show that the lattice parameters increas with increasing elements concentrations. Furthermore, the $Ti_{87.5}Nb_{12.5}$ structure was found to be the most stable, with the lowest heats of formation and density of states as compared to TiZr structures. These properties legitimized TiNb alloys to be applicable in biomedical industry.

Acknowledgements

The author appreciates the assistance from Sefako Makgatho Health Sciences University. The Council for Scientific and Industrial Research's primary responsibility and Accelerated Principal Researcher Development Program APRDP funding for the financial assisatnce. High-performance computing centre (CHPC) in Cape Town was used for calculations and simulations.

References

[1] Zhang L C and Chen L Y 2019 *Adv. Eng. Mater*. **21** 1801215
[2] Sarraf M, Ghomi E R, Alipour S, Ramakrishna S and Sukiman N L 2022 *Bio. Des. Manuf.* **5** 371
[3] Sasikumar Y, Indira K and Rajendran N 2019 *J. Bio. Tribo. Corros.* 5(36): 1-25 (2019).
[4] Nicholson W J 2020 *Prosthesis* **2** 100
[5] John A A, Jaganathan S K, Supriyanto E and Manikandan A 2016 *Current. Sci.* **111** 1003
[6] Li P, Ma X, Tong T and Wang Y 2020 *J. Alloys. Compd.* **815** 152412
[7] Bansal P, Singh G and Sidhu H S 2012 *J. Mater. Res.* **36** 431
[8] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
[9] Segall M D, Lindan P J, Probert M A, Pickard C J, Hasnip P J, Clark S J and Payne M C 2002 *J. Phys. Condens. Matter*. **14** 2717
[10] Monkhorst H J and Pack J D 1976 *Phys. Rev. B.* **13** 5188
[11] Pfrommer B G, Cote M, Louie G and Cohen M L 1997 *J Comp. Phys.* **131** 233
[12] Wang B T, Zhang P, Liu H Y, Li W D and Zhang P 2011 *J. Appl. Phys.* **109** 063514
[13] Ikehata H, Nagasako N, Furuta T, Fukumoto A, Miwa K and Saito T 2004 *Phys. Rev. B.* **70** 174113
[14] Kanoun M B, Goumri-Said S and Jaouen M 2007 *Phys. Rev. B.* **76** 134109
[15] Rahman A, Mousumi, K, Ali L, Khatun R, Rahman Z, Hasan S S, Hasan W and Hasan Z 2023 *Results. Phys.* **44** 106141

[16] Gao M C, Gao P, Hawk J A, Ouyang L, Alman D E and Widom M 2017 *J. Mater. Res.* **32** 3627
[17] Gxowa-Penxa Z, Daswa P, Modiba R, Mathabathe M N and Bolokang A S 2021 *Mater. Chem. Phys.* **259** 124027

Surface properties of Al_2O_3 and Fe_2O_3 (100) surfaces using First-Principle method

PM Dibakwana[1], DM Tshwane[2,4] ME Sithole[1] and R Modiba[1,3]

[1]Department of Physics, SMU, P.O. BOX 94, Medunsa, Ga-Rankuwa, 0204, South Africa
[2]Next Generation Enterprises and Institutions Cluster, CSIR, P.O Box 395. Pretoria 0001, South Africa
[3]Advanced Material and Engineering, CSIR, P.O Box 395. Pretoria 0001, South Africa
[4]National Institute for Theoretical and Computational Sciences (NITheCS), Gauteng 2000, South Africa

E-mail: **patriciadibakwana.27@gmail.com**

Abstract. Al_2O_3 and Fe_2O_3 are oxide materials that are widely used in a variety of applications, including catalysis, ceramic-based semiconductors, corrosion protection, and thermal barrier coatings. However, most of the applications of material strength depend on its surface properties. The purpose of the study was to investigate the structural stability between the Al_2O_3 and Fe_2O_3 surfaces. In this study First Principle calculations based on density functional theory were performed with Cambridge Serial Total Energy Package to investigate the surface stability between Al_2O_3 and Fe_2O_3 surfaces. The results show that the surface energy of Al_2O_3 is lower than that of Fe_2O_3 at different slab sizes, however, both materials are thermodynamically stable since they exhibit low surface energy. The results suggest that Al_2O_3 is more stable and energetically favourable since it exhibits a low surface energy in all slab sizes than Fe_2O_3 surfaces. In addition, partial density of states and work function were computed and analysed.

1. Introduction

The demand to improve the currently used materials for thermal barrier coatings has arisen as a result of daily technological advancement. Transition metal oxides have captured the attention of both experimental and theoretical researchers due to their diverse physical and chemical properties, making them highly versatile for various applications like catalysis and ceramic-based semiconductors, etc. [1]. However, most of the material's application depends on its surface properties. The study of surface properties is crucial for numerous scientific, technological, and design advances. Surface properties, including roughness, hardness, adhesion, and conductivity, play vital roles in determining material functionality. The material surfaces are significant in applications such as thermal barrier coatings for high temperatures, gas turbines, engines, heterogeneous catalysts [2], microchip packaging, and corrosion prevention [3]. Understanding surface properties is important for optimizing material performance, developing coatings, and surface treatment. It enables advancements across industries and enhances the functionality and overall value of materials used for thermal barrier coating in automobile and airline industries.

Neufeld *et el.* [4] investigated the role of an Al_2O_3 overlayer on Fe_2O_3 for the splitting of water using the density functional theory (DFT) method. The results showed that catalysis occurs primary on the α-Fe_2O_3 exposed areas of -Fe_2O_3 rather than directly over α-Al_2O_3. The presence of α-Al_2O_3 coverage was found to decrease the overpotential needed to oxidize water on α-Fe_2O_3. Another study by Tamijani *et al*. [5] compared clean, hydrated, and defect α-Al_2O_3 and α-Fe_2O_3 (110) using first-principles

characterization and molecular simulation. The results were compared with experimental characterizations of hydrated alumina and hematite. The outcomes indicated that the introduction of defects upon hydration and hydroxylation led to thermodynamically unfavorable structures, yet these structures exhibited stability similar to that of the hydrated ideal terminations. Tamijani *et al*. [5] observed an enhancement in water oxidation when α-Al_2O_3 and α-Fe_2O_3 were employed as coverings. This improvement was attributed to the reduction in the work function of α-Fe_2O_3 upon α-Al_2O_3 coverage, which facilitates the extraction of electrons during the water oxidation reaction. Consequently, these results suggest that choosing an overlayer with a lower work function than that of the catalyst as a strategy for the future development of a better catalyst.

Surface chemistry unquestionably affects a solid's catalytic activity and growth rate because the crystal surface is the main interface through which it interacts with its surroundings. Surface energy (s), which is defined as the energy needed per unit area of the new surface created to divide the material into two along a Miller plane, is a major determinant of surface chemistry [6]. Previous studies addressing the surface free energy of alumina and hematite surfaces as a function of water vapor pressure demonstrate the theoretical conclusions produced by ab initio thermodynamics method using a gas phase particle reservoir [7]. Furthermore, as the demand for better steel material quality grows, a better understanding of the mechanism of steel refining reactions is required to optimize the process. In general, refining reactions occur at the surface-metal interface. Surface energy, work function, and density of state are all important surface properties, and studying the surface properties of Al_2O_3 and Fe_2O_3 is important for gaining fundamental insight into the behaviors of oxide materials, enabling their effective utilization and advancement in various scientific, technological, and industrial applications.

2. Computational setup

Calculations reported here were computed using BIOVIA Material Studio Software, which uses density functional theory (DFT) [8]. The DFT is contained in Cambridge Serial Total Energy Package (CASTEP) code [9]. The Perdew-Burke-Ernzernhof (PBE) functional of the generalized gradient approximation (GGA) [10, 11] was used to explain the electron-exchange correlation. Crystal structures of Al_2O_3 and Fe_2O_3 were built using Material Studio software with the following structural properties: For Al_2O_3, the theoretical lattice parameters were a=b=4,78 Å, c=13,01 Å, space group 3Rc, a kinetic energy cutoff of 450 eV, and 4×4×1 k-point meshes, which proved to be sufficient to converge for Al_2O_3. The theoretical lattice parameters for Fe_2O_3 were as follows: a=b=5,04 Å and c=13,752 Å with a space group of 3Rc. The lattice parameters, atomic arrangement, structural phases and space group are used to describe the crystal structure. The calculation's k-point set used in a calculation defines the accuracy of the Brillouin zone sampling. Using Broyden-Fletcher-Goldfarb-Shanno (BFGS) [12] algorithm as implemented in CASTEP [13], Geometry optimization was performed for both materials. The displacement and force tolerance were set at 10^{-3} and 0.03, respectively. Surface energy (γ) was calculated as follows:

$$\gamma = \frac{1}{2}(E_{surf} - NE_{bulk})$$

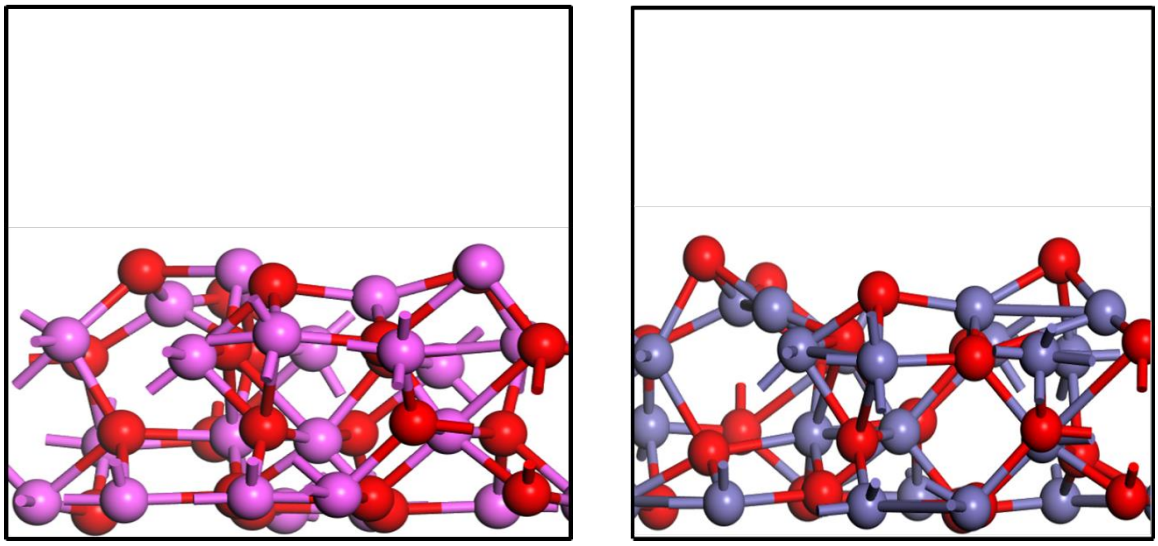
Where E_{surf} and E_{bulk} represent the total energies of the slabs and bulk, respectively. The area and number of atoms are represented by A and N, respectively.

3. Results and discussion

3.1 Structural properties and surface energies

Figure 1 illustrates the simulated crystal structure of Al_2O_3 and Fe_2O_3 at 4 and 5 atomic layers respectively. Al_2O_3 exhibits trigonal structure containing a space group of R3c and Fe_2O_3 exhibit rhombohedral crystal structure having R3c space group. The lattice parameter of Al_2O_3 is a=b= 4,759 Å, c=12,991 Å and for Fe_2O_3 a=b=5,035 Å, c=13,72 Å. The calculated parameter for both materials are in good agreement with the results of previous simulations [14]. The structural stability of Al_2O_3 and

Fe_2O_3 was analyzed using surface energy, which is defined as the energy associated with the presence of a crystal surface [15]. It is expressed per unit area. The surface energy of both material's varies depending on the surface terminations and how the atoms are arranged. Low surface energy is an important factor for thermodynamic stability and contributes to the stability of the surface of a material. Surface energy of both materials were calculated using equation 1 and the results are shown in table 1. The surface energy of Al_2O_3 (100) calculated ranges from 0.07 to 0.05 which is lower than the surface energy value of Fe_2O_3 (100) at 4 and 5 layers 0.13 to 0.32 respectively. According to research, the Al_2O_3 structure has the lowest surface energy, which suggest that the 5 layers structure is more stable than the 4 layers structure. Additionally, it was discovered that the surface energy strength varies depending on the thickness of the surface slab. On both surfaces, it was discovered that 4 atomic layers have more energy than 5 layers. Similar observation was reported by Tshwane *et al* [16] during their study on properties of Ti_2AlV (100) and (110) employing the first-principle method.



(a) (b)

Figure 1. Atomic structures of (a) Al_2O_3 and (b) Fe_2O_3 (100) surfaces.

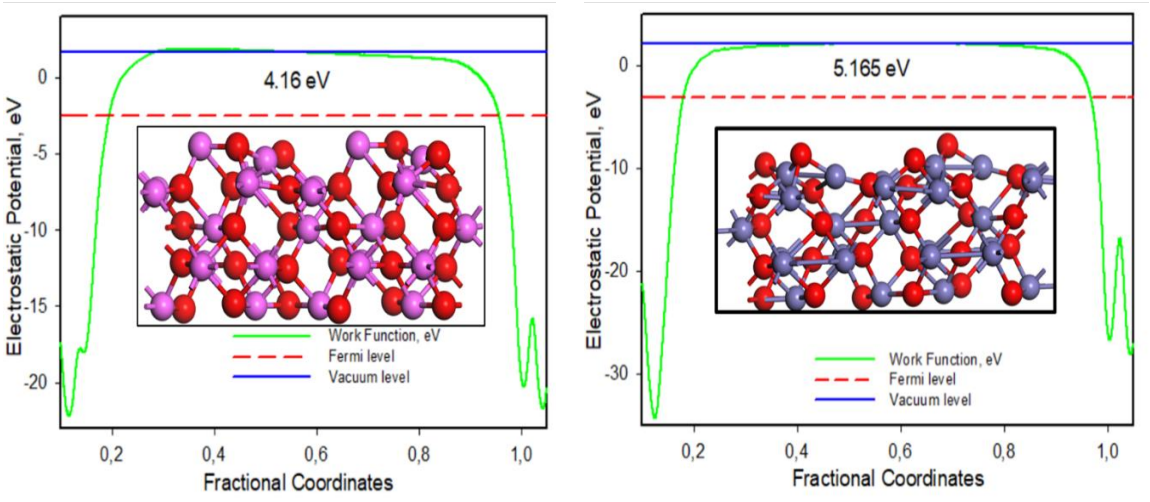
Table 1. Calculated energies of Al_2O_3 and Fe_2O_3 (100) surfaces.

Surface	Structure	Layers	γ(eV/Å^2)
100	Al_2O_3	5	0.05
		4	0.07
	Fe_2O_3	5	0.13
		4	0.32

3.1. Work function

The work functions are the basic crystal solid surface parameters for understanding various physical, chemical and structural surface conditions [17]. Studies have shown that the work function tends to vary depending on the surface termination and crystallographic orientation [16, 18]. Generally, Al_2O_3 surfaces tend to have lower work functions compared to Fe_2O_3 surfaces. The calculated work function of Al_2O_3 and Fe_2O_3 (100) surfaces were calculated at 4 and 5 layers. Figure 2 (a) and (b) shows the plot of the computed work function and table 2 shows the calculated work function at 4 and 5 layers. A difference in work function results from the redistribution of electrons and the rearranging of atomic charges. The values for the Al_2O_3 (100) and Fe_2O_3 (100) surfaces at 4 and 5 layers vary from 4.16 to 4.74 and 5.01 to 5.17 eV, respectively at different slab size. A lower work function indicates that

electrons are more easily emitted from the material's surface, implying higher electron emission. Al_2O_3 is more likely to emit electrons from the surface compared to Fe_2O_3 when subjected to external energy, such as heat or light.



(a) (b)

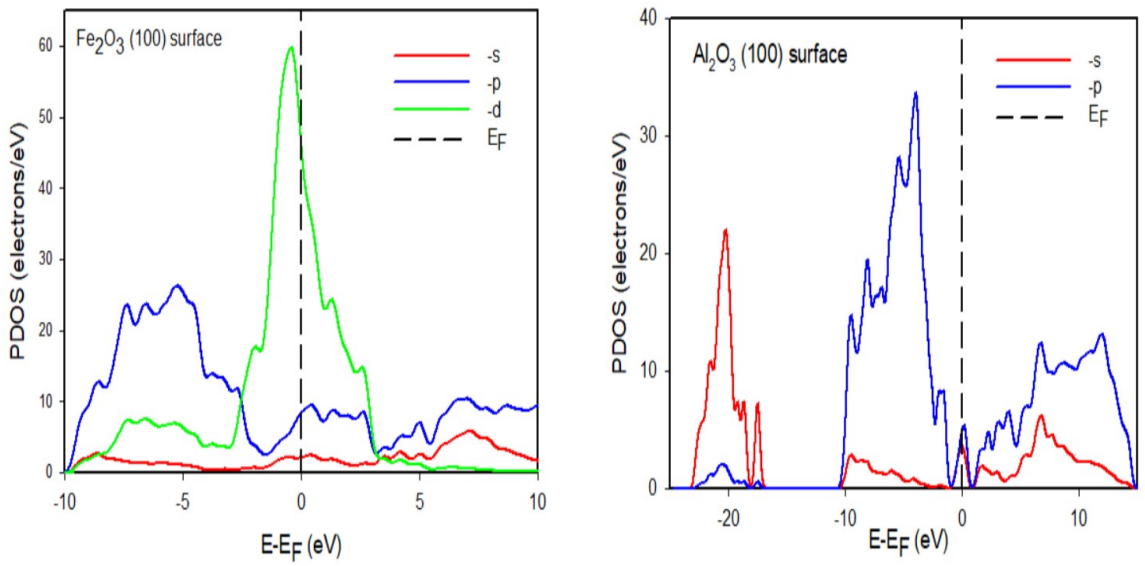
Figure. 2. Work functions (Φ) of (a) Al_2O_3 5 layers and (b) Fe_2O_3 (100) 4 layers.

Table 2: Calculated work function (eV) of Al_2O_3 and Fe_2O_3 (100) surface in 4 and 5 layers.

Surface	Structure	Layers	Φ(eV)
100	Al_2O_3	5	4.16
		4	4.73
	Fe_2O_3	5	5.17
		4	5.01

3.2. Density of state

The projected density of states (PDOS) is an analysis that chemically intuitively describes the charge density as calculated by DFT. To investigate the electron interactions between the atomic orbital systems and the atomic bonding between Al-O and Fe-O the DOS was determined. Figure 3 displays the PDOS graphs of the Al_2O_3 and Fe_2O_3 (100) surfaces with the Fermi energy (EF) indicated on the x-axis by zero. The PDOS) was investigated to better understand the electronic properties and interaction of Al_2O_3 and Fe_2O_3. The electronic for Al_2O_3 mainly consists of -s and -p orbitals that is mostly contributed by Al and O states. Furthermore, the hybridization of -s p orbital was observed at approximately -20 eV while near the fermi level is mainly consist of -p orbital from Al states. In addition, for Fe_2O_3 it was seen that the -d states predominately contribute more than -s and -p states on the fermi level. This indicates clear the metallic behaviour and d electrons from Fe atoms.



(a) (b)

Figure 3. PDOS of (a) Al_2O_3 (100) 5 layers and (b) Fe_2O_3 (100) 4 layers.

4. Conclusions

The study intended to compare Al_2O_3 and Fe_2O_3 slabs as they are different materials that behaves differently, to know which material is better as far as their stabilities are concern. It was found that both Al_2O_3 and Fe_2O_3 (100) surfaces structures were discovered to be thermodynamically stable. Moreover, the results revealed that Al_2O_3 (100) is more stable than Fe_2O_3 (100) surface and the results revealed that surface stability is determined by the number of layers, with a 5 layer of Al_2O_3 (100) being the most stable amongst all surfaces evaluated in addition, according to the work function analysis, the value of the work function changes from one surface to the next and rises as the slabs increased.

Acknowledgements

The authors acknowledge CHPC and Sefako Makgatho Health Science University for the computational resources provided. Department of Physics for the opportunity to carry out this work.

References

[1] Niu J, Jiang Y, He Y, Tao A, Chen C, Ma X J and Ye H 2022 *Compt. Mater. Sci*. **210** 111442
[2] Haynes J A, Pint B A, More K L, Zhang Y and Wright I G 2002 *Oxid. Mater.* **58** 513
[3] Hou P P Y 2008 *Annu. Rev. Mater. Res.* **38** 275
[4] Neufeld O, Yatom N and Toroker M C 2022 *J. Phys*. **15** 7237
[5] Tamijani A A, Augustine L J, Bjorklund J L and Catalano J G 2022 *Mol. Simul.* **48** 247
[6] Natalie M and Sohlberg S K 2023 *Phys. Chem.* **25** 13351
[7] Wang X G, Chaka A and Scheffler M 2000 *Phys. Rev. Lett*. **84** 3650
[8] Mattsson A E, Schultz P A, Desjarlais M P, Mattsson T R and Leung K 2005 *Simul. Mater. Sci. Eng.* **13** R1
[9] Clark S J, Segall M D, Pickard C J, Hasnip P J, Probert M I, Refson K and Payne MC 2005 Z. Kristallogr. **220** 570
[10] Perdew P J and Wang Y 1992 *Phys. Rev.* B. **45** 13249
[11] Perdew P J, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett*. **77** 3865
[12] Dai Y H 2002 *SIAM. J. Optim.* **13** 693
[13] Segall M D, Lindan P J D, Probert M J, Hasnip C J, Clark J and Payne MC 2002 *J. Condens. Matter. Phys*. **14** 2717

[14] Xue H, Wei X, Guo W and Zhang X 2002 *J. Alloys. Compd.* **820** 53070
[15] Packham D E 2003 *Int. J. Adhes. Adhes* **23** 4378
[16] Tshwane D M and Modiba R 2022 *RAPDASA-RobMech-PRASA-CoSAAMI: Conf. Series* **370** 09005
[17] Hoffmann R 1988 *Rev. Mod. Phys*. **60** 601
[18] Wang J and Wang S G 2015 *Appl. Surf. Sci*. **35** 1046