

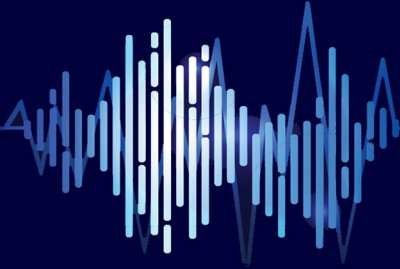


Iranian Association of Electrical
and Electronics Engineers



Shahid Chamran
University of Ahvaz

Journal of Applied Research in Electrical Engineering



Vol. 2, No. 2, Summer and Autumn 2023



PUBLISHER: SHAHID CHAMRAN UNIVERSITY OF AHVAZ

E-ISSN: 2783-2864

P-ISSN: 2717-414X



**Iranian Association of
Electrical and Electronics
Engineers**

Journal of Applied Research in Electrical Engineering

E-ISSN: 2783-2864

P-ISSN: 2717-414X



**Shahid Chamran
University of Ahvaz**

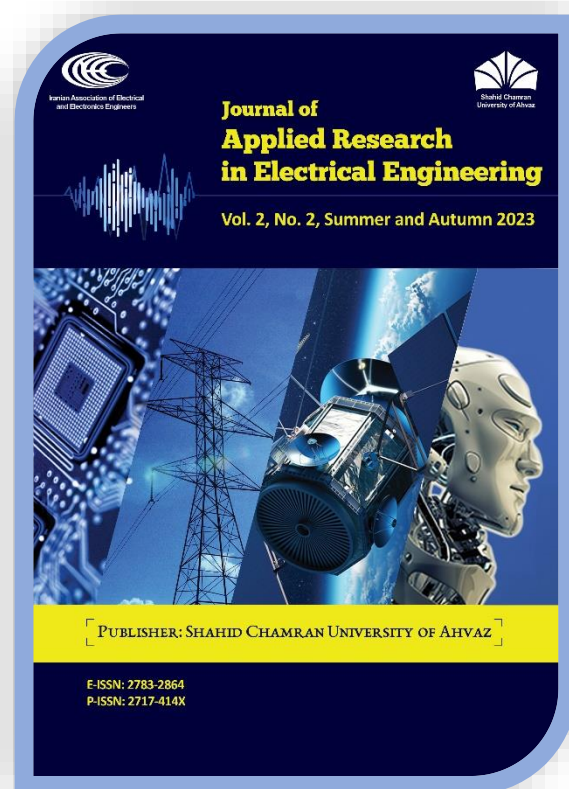
**Journal of Applied Research in Electrical Engineering (JAREE),
Volume 2, Number 2, Summer and Autumn 2023**

Publisher: Shahid Chamran University of Ahvaz, Iran

This magazine is the result of a formal partnership of **Shahid Chamran University of Ahvaz** and **Iranian Association of Electrical and Electronics Engineers**

Website: <https://jaree.scu.ac.ir>

E-mails: jaree@scu.ac.ir; jaree.scu@gmail.com



Address: Department of Electrical Engineering, Faculty of Engineering, Shahid Chamran University of Ahvaz, Golestan Street, Ahvaz, Iran

P.O. Box: 61357-85311






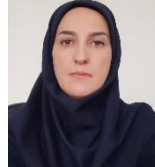
Tel: +986133226782, +989122876375

Fax: +986133226597

Editorial Board

	Editor-in-Chief		Director-in-Charge
	<p>Prof. Mahmood Joorabian Shahid Chamran University of Ahvaz, Ahvaz, Iran <i>Power System Planning, Renewable Energy and Smart grid, FACTS Devices</i> mjoorabian@scu.ac.ir</p>		<p>Prof. Seyed Ghodratalah Seifossadat Shahid Chamran University of Ahvaz, Ahvaz, Iran <i>Power System Protection, Power Electronics, Power Quality</i> seifossadat@yahoo.com</p>
	Managing Editor		Executive Assistant
	<p>Dr. Alireza Saffarian (Associate Professor) Shahid Chamran University of Ahvaz, Ahvaz, Iran <i>Power System Protection, Power System Stability, Power Quality, Distribution Systems</i> a.saffarian@scu.ac.ir</p>		<p>Dr. Hossein Farzin (Assistant Professor) Shahid Chamran University of Ahvaz, Ahvaz, Iran <i>Reliability of power systems, Microgrids and distributed energy sources, Smart grids</i> farzin@scu.ac.ir</p>
	Associate Editor		Associate Editor
	<p>Dr. Yousef Seifi Kavian (Associate Professor) Shahid Chamran University of Ahvaz, Ahvaz, Iran <i>Electronics and Telecommunications</i> y.s.kavian@scu.ac.ir</p>		<p>Dr. Mohsen Saniei (Associate Professor) Shahid Chamran University of Ahvaz, Ahvaz, Iran <i>Power and Control</i> m.saniei@scu.ac.ir</p>
	Editorial Board		Editorial Board
	<p>Prof. Saeedallah Mortazavi Shahid Chamran University of Ahvaz, Ahvaz, Iran <i>Intelligent Control Systems, Power system Control, Automation, Fuzzy logic, Neural Networks</i> mortazavi_s@scu.ac.ir</p>		<p>Prof. Abdolnabi Kosarian Shahid Chamran University of Ahvaz, Ahvaz, Iran <i>Solid State Electronic Devices, Solar Cell Fabrication Technology</i> a.kosarian@scu.ac.ir</p>
	Editorial Board		Editorial Board
	<p>Prof. Ebrahim Farshidi Shahid Chamran University of Ahvaz, Ahvaz, Iran <i>Analog and Digital Integrated Circuits, Data Converters, Microelectronics</i> farshidi@scu.ac.ir</p>		<p>Prof. Naser Pariz Ferdowsi University of Mashhad, Mashhad, Iran <i>Nonlinear Control, Hybrid Systems, Aeronautics, Industrial Control, Applied Mathematics</i> n-pariz@um.ac.ir</p>
	Editorial Board		Editorial Board
	<p>Dr. Yousef Seifi Kavian (Associate Professor) Shahid Chamran University of Ahvaz, Ahvaz, Iran <i>Digital Circuits and Systems, Communication Networks and Distributed Systems</i> y.s.kavian@scu.ac.ir</p>		<p>Dr. Mohsen Saniei (Associate Professor) Shahid Chamran University of Ahvaz, Ahvaz, Iran <i>Power System Dynamics, High Voltage Engineering, Electricity Market, Microgrid</i> m.saniei@scu.ac.ir</p>
	Editorial Board		Editorial Board
	<p>Prof. Abbas Zarifkar Shiraz University, Shiraz, Iran <i>Optical Electronics</i> zarifkar@shirazu.ac.ir</p>		<p>Dr. Edris Pouresmaeil (Associate Professor) Aalto University, Espoo, Finland <i>Integration of renewable energies into the power grid</i> edris.pouresmaeil@aalto.fi</p>
	Editorial Board		Editorial Board
	<p>Prof. Reza Ghaderi Shahid Baheshti University, Tehran, Iran <i>Control Theory, System Identification, Control Systems, Fuzzy Engineering</i> r_ghaderisbu.ac.ir</p>		<p>Prof. Fushuan Wen Tallinn University of Technology, Tallinn, Estonia <i>Power systems and power economics</i> fushuan.wen@taltech.ee</p>
	Editorial Board		Editorial Board
	<p>Prof. Majid Sanaye-pasand University of Tehran, Tehran, Iran <i>Power Systems Protection, Digital Protective Relays, Power Systems Automation, Power System Transients</i> msanaye@ut.ac.ir</p>		<p>Prof. Zabih (Fary) Ghassemlooy North Umbria University, Newcastle upon Tyne, United Kingdom <i>Optical Communications, Visible Light, Communication Systems</i> z.ghassemlooy@northumbria.ac.uk</p>

Editorial Board (Continued)

	<p style="text-align: center;">Editorial Board</p> <p>Prof. Mahdi Tavakoli University of Alberta, Alberta, Canada <i>Robotics and Telerobotics, Haptics and Teleoperation Control, Surgical and Therapeutic Robotics, Image-Guided Surgery</i> mahdi.tavakoli@ualberta.ca</p>		<p style="text-align: center;">Editorial Board</p> <p>Prof. Mohamad Hassan Modir Shanechi Illinois Institute of Technology, Chicago, USA <i>Nonlinear and intelligent systems, Power system dynamics and security, Power system planning and maintenance scheduling</i> shanechi@iit.edu</p>
	<p style="text-align: center;">Editorial Board</p> <p>Prof. Mohammad Shahidehpour Illinois Institute of Technology, Chicago, USA <i>Power Systems, Microgrids, Power System Operation, Power System Planning</i> ms@iit.edu</p>		<p style="text-align: center;">Language Editor</p> <p>Majid Sadeghzadeh Hemayati <i>English Language Editing</i> m_s_hemayati@yahoo.com</p>
	<p style="text-align: center;">Page Designer</p> <p>Dr. Mahyar Abasi (Assistant Professor) Arak University, Arak, Iran <i>Power</i> m-abasi@araku.ac.ir</p>		<p style="text-align: center;">Executive Director</p> <p>Masoomesh Shaygan Shahid Chamran University of Ahvaz, Ahvaz, Iran <i>Electronics</i> masoomesh.shaygan@gmail.com</p>

About Journal

Journal of Applied Research in Electrical Engineering (J. Appl. Res. Electr. Eng.) is a single-blind peer-reviewed, **open access** and **free of charge** international journal published by Shahid Chamran University of Ahvaz in cooperation with Iranian Association of Electrical and Electronics Engineers (IAEEE). The JAREE is a medium for global academics to exchange and spread the latest discoveries and advances in their applied research in electrical engineering. The JAREE aims at presenting important results of analytical, computational and experimental works within all specialized fields of electrical engineering (electronics, power, control and telecommunications). It welcomes high quality original research papers from contributors throughout the world. All papers are subject to a peer reviewing procedure. Submission, processing and publication of the papers in JAREE is **free of charge**.

Types of accepted papers include:

- Research articles
- Review articles
- Applied articles

Research papers are expected to present innovative solutions, novel concepts, or creative ideas that can help to address existing or emerging technical challenges in electrical engineering.

Application papers are expected to share valuable industry experiences on dealing with challenging technical issues, developing/adopting new standards, applying new technologies or solving complex problems. JAREE welcomes application papers that can have a significant impact on industry practices in the coming years.

Review papers are expected to provide insightful and expert reviews, tutorials, or study cases on an important, timely and widely-interested topic in electrical engineering.

All researchers in the fields of electrical science are invited to publish their scientific and research achievements in this journal. Interested authors can submit their manuscripts in the journal's website. More information is available in the website on how to prepare and submit the manuscripts.

Amis and Scope

The *Journal of Applied Research in Electrical Engineering* aims to provide a medium for dissemination of innovative and consequential papers that present analytical, computational and experimental works within all specialized fields of electrical engineering (electronics, power, control and telecommunications). The scope of the journal includes, but is not limited to, the following topic areas:

Electronics:

- Optical electronics, photonics and solar cells
- Integrated analog circuits and mixed signals
- Integrated radio frequency circuits
- Digital electronics (VLSI)
- Semiconductor devices
- Sensor technology

Power:

- Dynamics and stability of the power systems
- Power system protection
- Electric power quality
- Operation and planning of the power systems
- High voltage insulation technology
- Flexible AC Transmission Systems (FACTS)
- Electric power distribution systems
- Smart grids, micro-grids, renewable energies and distributed generation
- Reliability of electrical energy systems
- Energy management and electricity market
- Electric machines and transformers
- Power electronic and electric drives

Control:

- Linear and non-linear control systems
- Adaptive, optimal and stochastic control
- Fuzzy systems, neural networks and intelligent control
- Robotic and mechatronic
- Modeling, Identification and optimization of systems
- Guidance and navigation systems
- Automation, industrial control and instrumentation

Telecommunications:

- Signal and image processing
- Wireless and cellular communication systems
- Telecommunication networks
- Radar and sonar
- Information theory and coding
- Cognitive radio
- Antenna design
- Microwave devices
- Wave propagation and electromagnetic compatibility

Indexing Databases and Social Networks

Iran Ministry of Science, Research and Technology Scientific Journals

<https://journals.msrt.ir/home/detail/16058/>

Directory of Open Access Journals (DOAJ):

<https://doaj.org/toc/2783-2864>

Google Scholar:

<https://scholar.google.com/citations?user=F7KQPtYAAAAJ&hl=en&authuser=1>

Directory of Open Access Scholarly Resources (ROAD):

<https://portal.issn.org/resource/ISSN/2783-2864>

LinkedIn:

<https://www.linkedin.com/in/journal-of-applied-research-in-electrical-engineering-jaree-7540871b2/>

Academia:

<https://shahidchamranahwaz.academia.edu/JournalAppliedResearchinElectricalEngineering>

Mendeley:

<https://www.mendeley.com/profiles/journal-of-applied-res-in-electrical-engineer/>

Twitter:

https://twitter.com/jaree_scu

Facebook:

<https://www.facebook.com/jaree.scu>

Researchgate:

https://www.researchgate.net/profile/Jaree_Engineering

Telegram:

<https://t.me/jareescu>

Instagram:

<https://www.instagram.com/jaree.scu/>

Journal homepage:

<http://jaree.scu.ac.ir>

Journal emails:

jaree@scu.ac.ir, jaree.scu@gmail.com

Guide for Authors

How to submit a manuscript

For the initial submission, the authors have to just send the main manuscript file and the signed [Copyright Form](#) of the journal. While preparing manuscripts for initial submission, authors are kindly requested to follow the guidelines, described below:

- The manuscript should be written in a Microsoft Word file (.doc or .docx).
- The file should include text (preferably in 10 points, “Times New Roman” font) and all figures (figures can be placed within the text at the appropriate point or at the end of the text).
- The manuscript pages should be prepared either using a double-column single-line spacing layout or a single-column double-line spacing layout. A margin of at least 1.5 cm on each side is required.
- All papers should be composed of Title, Author Name, Affiliation, Corresponding author email, Abstract, Keywords, Body, and References.
- The manuscript should be written in good English. It should have been carefully checked for clarity, conciseness, the correctness of grammar, and typographical errors.
- The corresponding author should sign the journal copyright form on behalf of any and all co-authors and upload it to the Journal’s Submission System when submitting the manuscript. The journal copyright form can be downloaded from [here](#).
- The corresponding author can use the [JAREE Template for Cover Letter](#) as a default text for the cover letter when submitting the manuscript.
- It is recommended that the title of the paper does not contain abbreviations or formulae.
- The abbreviations used in the abstract should be introduced both in the abstract and again on first use in the body.
- References should be numbered in the order they are mentioned in the text.

Manuscript Submission

Submission to this journal proceeds totally online and you will be guided stepwise through the creation and uploading of your files. All correspondence, including notification of the Editor's decision and requests for revision, takes place by e-mail. To submit your manuscript, click on the [Submit Manuscript](#) link on the journal's homepage. Then, click on [Register](#) to create an author account. A message is sent to your email address containing your username and password. Then, login to the Journal’s Submission System at the [User's login](#) page using the username and password to submit your new manuscript. Once you have logged in, you can change your password by clicking on the My Home link at the top menu.

Copyright and Open Access License

An author submitting a paper should ensure that he or she has the right to publish the paper and that it contains nothing defamatory. The JAREE will assume that all co-authors have agreed to the submission of any paper received. The corresponding author should sign the journal copyright form on behalf of any and all co-authors and upload it to the Journal’s Submission System when submitting the manuscript.

Contents

Article Title and Authors	Page No.
New Approach in Minerals Cracking Using Electromechanical Breakdown Theory Asaad Shemshadi, and Mohamad Reza Khojaste	103
Combined Economic Emission Dispatch in a Grid-Connected Microgrid Using an Improved Mayfly Algorithm Nicholas Kwesi Prah II, Elvis Twumasi, and Emmanuel Asuming Frimpong	112
Improving the Diagnosis of Sudden Cardiac Death by Using Non-Linear Features of the ECG Signal and Hybrid RBF Afshin Koliji, Sara MihanDoost, Nematollah Ezzati, and Ehsan Mostafapour	120
Selection and Tuning Propagation Path Loss Model for Hawassa City, Ethiopia at 1800 MHz Frequency Tamirat Yenealem, and Robel Getachew	127
Different Types of Distributed Optimal Leader-Follower Consensus Protocol Design for a Class of High-Order Multi-Agent Systems Farideh Azadmanesh, and Reza Ghasemi	136
Morphological and Statistical Filter-Based Multi-Direction Building Detection for SAR Images Fateme Amjadipour, Maryam Imani, and Hassan Ghassemian	145
Design of a Tunable 4th Order OTA-C Band-Pass Filter for use in Front-End of ADC Mohsen Ghaemmaghani, and Shahbaz Reyhani	152
Nonlinear Observer-Based Control Design for a Three-Axis Inertial Stabilized Platform Mohammad Mehdi Zohrei, and Hamidreza Javanmardi	158
Optimal Sizing and Placement of Series Capacitors in Distribution Networks Using Modified Elephant Herding Optimization Algorithm Elvis Twumasi, Abdul-Fatawu Seini Yussif, and Emmanuel Asuming Frimpong	173
Re-examination of the Doubly Fed Induction Machine (DFIM) Model Taking Into Account the Rotor Reactive Power Mohammad Naser Hashemnia	182
Full-Duplex Device-to-Device Relays in a Novel Hybrid and Adaptive Joint Relaying Network: Symbol Error Analysis and Optimum Power Allocation Mehdi Basiri Abarghouei, and Reza Saadat	194
Performance Analysis of a Steep-Slope Bi-channel GaSb-GaAs Extended Source Tunnel Field Effect Transistor with Enhanced Band-to-Band Tunneling Current Zahra Ahangari	206

Copyrights

© 2024 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).





Iranian Association of
Electrical and Electronics
Engineers

Journal of Applied Research in Electrical Engineering

E-ISSN: 2783-2864

P-ISSN: 2717-414X

Homepage: <https://jaree.scu.ac.ir/>



Research Article

New Approach in Minerals Cracking Using Electromechanical Breakdown Theory

Asaad Shemshadi *, and Mohamad Reza Khojaste

Arak University of Technology, Arak, Iran

* Corresponding Author: shemshadi@arakut.ac.ir

Abstract: One type of electrical breakdown in solid insulation is electromechanical failure. In mineral processing, crushing rocks is energy-intensive. Rock crushing using high voltage has many advantages, including high stone-breaking efficiency, and is a new and efficient way to break the stone. The shape of the electrode, the amount of applied voltage, and the selection of drilling process parameters are the main obstacles to using this method. In this study, based on the equivalent circuit of high voltage electro pulse failure, a mathematical model of high voltage electro pulse discharge in rock has been developed. Then, a high-voltage simulation model is developed based on the coaxial cylindrical electrode structure. This paper investigates the use of electromechanical failure phenomena for crushing minerals. High voltage pulses are used to crush the rock, then by simulating the relevant circuit, the necessary voltage for crushing three minerals is obtained and the feasibility of using this method is discussed. Finally, using the simulation and the obtained results, the possibility of using this method for crushing minerals has been investigated. This study provides a scientific basis for quantifying and predicting rock crushing using high-voltage technology to improve drilling efficiency and reduce energy loss.

Keywords: Electromechanical failure, crushing rocks, electrical breakdown in solids.

Article history

Received 15 May 2022; Revised 11 December 2022; Accepted 20 February 2023; Published online 11 July 2023.

© 2023 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)

How to cite this article

A. Shemshadi, and M. R. Khojaste, "New approach in minerals cracking using electromechanical breakdown theory," *J. Appl. Res. Electr. Eng.*, vol. 2, no. 2, pp. 103-111, 2023. DOI: [10.22055/jaree.2023.40836.1059](https://doi.org/10.22055/jaree.2023.40836.1059)



1. INTRODUCTION

The mining industry is a major consumer of energy, accounting for 5-7% of global consumption. High mineral demand in the world has led many mining companies to process lower grade ores. Processing these ores requires a lot of energy to release minerals. With increasing energy consumption in the world and declining first-class ores, mining companies are forced to find alternative ways to reduce total energy consumption. Due to the phenomenon of electromechanical failure in solid insulation, if we can implement a mechanism that can be used to crush minerals, the needs of the mining industry can be met by reducing energy consumption. EPB is one of the new technologies of stone breaking [1, 2]. The advantages of this method compared to other methods of breaking the rock are controllable energy level, no pollution, no stone particles and its low cost. In the EPB process, short voltage pulses with high amplitude and current are required. There is a problem with the delay between the maximum voltage and current [3, 4]. The bumps and depressions in the rocks strongly affect the distribution of the electric field and reduce the efficiency of this method [5,

6]. A lot of work has been done by different people on breaking stones with electro pulse. Among them, Zhang developed a model of crushing rock by a plasma channel based on the principles of motion transfer. According to the classical blast theory, rock was considered as a homogeneous, isotropic and incompressible material in the model [7]. Due to the non-uniformity characteristics of rock, pores and electrical distortions, a failure model has been developed based on the theory of failure weaknesses and the cylindrical channel model. The proportional coefficient of electric field distribution is also included in the model. Breaking rock with electrical energy can be a cost-effective and better way to break up rock mass, in oil and gas extraction, tunnel construction and the like. Qiong Hu [8] has proposed plasma pulse discharge technology in order to achieve effective separation of ore and thin-layer bedrock and reduce the impact on the deep ocean environment. The energy transfer model is based on the equivalent circuit of the discharge channel to simulate the energy injection and power consumption of the discharge channel. A pulsed power supply is designed and developed for experimental research. The breakdown voltage of artificial shells and artificial

bedrock have been tested. The results show that the breakdown voltage is apparently different. The energy consumption is calculated and the results show that energy injection is useful for crushing. C. Li et al. [9] developed a complete high-pressure EPB damage model in granite, which includes a shock wave model and a high-pressure EPB damage model in granite. In this study, the use of the EPB damage model captures the complete process of high-voltage EPB, from discharge to shock wave generation, and thus the rock crushing through electric pulse can be simulated and calculated. The time-varying waveform of shock waves with different electrical parameters has been simulated and calculated based on the model. In the EPB granite damage geometry model, different shock wave forms are loaded to the surface and internal rock. Then the fracture process of the rock surface and inside and the rock fracture mechanism have been analyzed using EPB.

2. PRINCIPLES OF BREAKING ROCK WITH HIGH VOLTAGE PULSE

Breaking a rock using a high voltage pulse is possible in two ways: 1- Hydroelectric 2- Electro-pulse, which is more effective in the latter case.

In breaking rock using high voltage pulse, heterogeneous and non-conductive feature of most ores is exploited. The ore is placed in a dielectric fluid, usually water, and high voltage pulses are applied to it. As shown in Fig. 1, when the high voltage pulse reaches its maximum in less than 500 ns, the breakdown voltage of the water is higher than most solids. Therefore, when the applied voltage is higher than the breakdown voltage, discharge occurs in the solids first. Thus, the rock is broken by the first voltage pulse and a plasma channel is created. Then, the energy stored in the power supply enters the plasma channel. As the plasma channel expands, tension is created. When this stress exceeds the strength of the rock, the rock breaks. The circuit schematic is shown in Fig. 2.

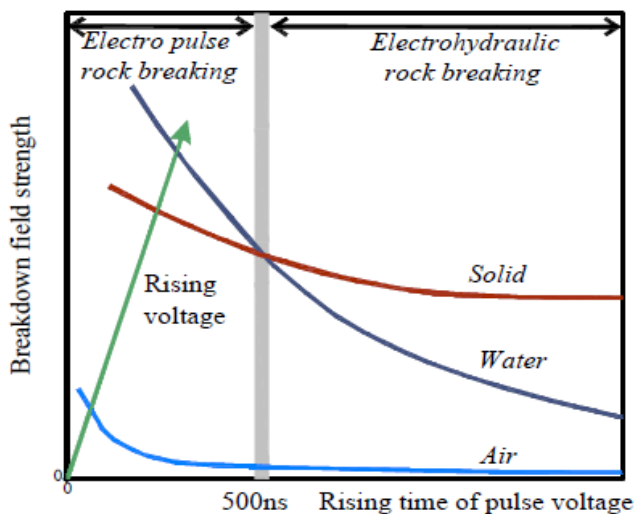


Fig. 1: The breakdown voltage of different materials relative to the pulse rise time [10].

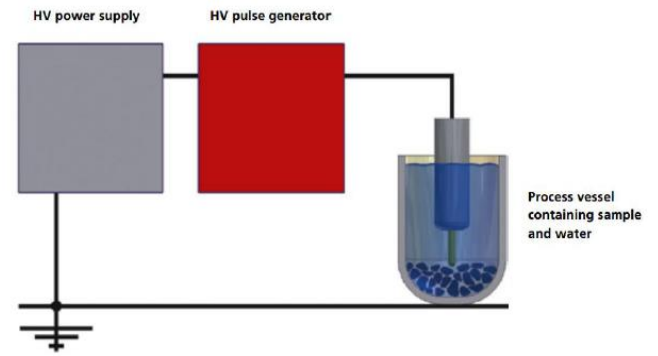


Fig. 2: The circuit schematic [10].

The process of breaking a rock with a voltage pulse can be divided into three parts: 1- Electromechanical failure 2- Injection of energy into the plasma channel 3-shock-wave fracturing [11, 12].

2.1. Electromechanical Breakdown Modelling

When an insulator with a dielectric constant ϵ is placed in a strong electric field, the field of impact of the mechanical gravitational force is applied to the insulation and depending on its mechanical strength, the initial length of the insulation reaches from d_0 to d . If d_0 is the initial thicknesses of the part with Yang constant (Y) and the insulation thickness is reduced to d by applying voltage V , then if the electrostatic force is tolerated by the solid insulator, the compressive stress due to the electrostatic force will be in equilibrium with the mechanical strength:

$$\frac{1}{2} \epsilon_0 \epsilon_r \frac{V^2}{d^2} = Y \ln \frac{d_0}{d} \quad (1)$$

$$V = \sqrt{\frac{2Yd^2}{\epsilon} \ln \frac{d_0}{d}} \quad (2)$$

Derivative of equation (2) with respect to d , voltage that leads to mechanical instability insulation is achieved in the following conditions:

$$\frac{d}{d_0} = e^{-1/2} \quad (3)$$

Therefore, the maximum electric field strength that can be tolerated for insulation is equal to:

$$V_{max} = 0.6d_0 \sqrt{\frac{Y}{\epsilon}} \quad (4)$$

$$E_{max} = 0.6 \sqrt{\frac{Y}{\epsilon}} \quad (5)$$

It is observed that the electromechanical breakdown field of insulation depends on the electrical (ϵ) and mechanical (Y) characteristics of the insulation simultaneously. According to the theory that breakdown occurs at weak points [10], the breakdown of solid dielectrics begins at weak points such as cracks or holes. According to [11] for the plasma channel model in electromechanical failure, the Kratal cylindrical channel model is used. As shown in Fig. 3, l_{ch} is the length of the plasma channel, V_{ch} is the volume of the plasma channel, and r is the radius of the plasma channel ($=10\mu\text{m}$). In the field of electrostatics, the plasma channel contains electrostatic and electromechanical energy [13]. Wes electrostatic energy can be expressed as follows:

$$W_{es} = \frac{1}{2} \epsilon_0 \epsilon_r E V_{ch} \tag{6}$$

where, E is the electric field, V_{ch} is the volume of the plasma channel, ϵ_0 is the vacuum dielectric constant and ϵ_r is the relative dielectric of dielectric. The electromechanical energy of the plasma channel, according to the above, is as follows [13]:

$$W_{em} = \frac{\epsilon_0 \epsilon_r E^4}{8Y} V_{ch} \tag{7}$$

where, E is the electric field and Y is the modulus of elasticity. The total energy of the plasma channel is equal to the sum of the electrostatic and electromechanical energies [13]:

$$W_{tot} = \left(\frac{1}{2} \epsilon_0 \epsilon_r E + \frac{\epsilon_0 \epsilon_r E^4}{8Y} \right) V_{ch} \tag{8}$$

As the applied electric field increases and the plasma channel expands, the surface energy of the rock must be overcome. When the total energy inside the rock holes is greater than the rock surface energy (W_{ec}), the plasma channel expands until the high and low voltage electrodes are connected (Similar to Fig. 4). Failure occurs at this point [13]:

$$W_{ec} = 2G\pi r l \tag{9}$$

$$W_{tot} \geq W_{ec} \tag{10}$$

where W_{ec} is the rock surface energy and G is the surface free energy of rock.

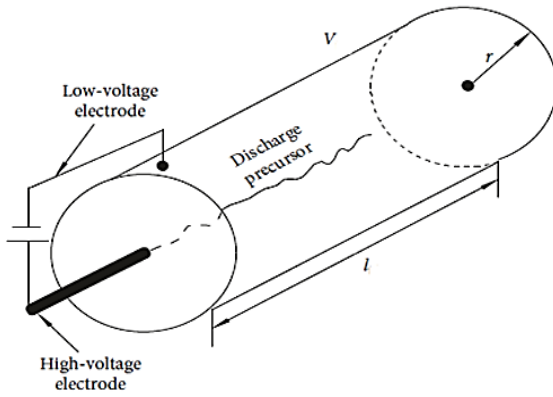


Fig. 3: Plasma channel model [10].

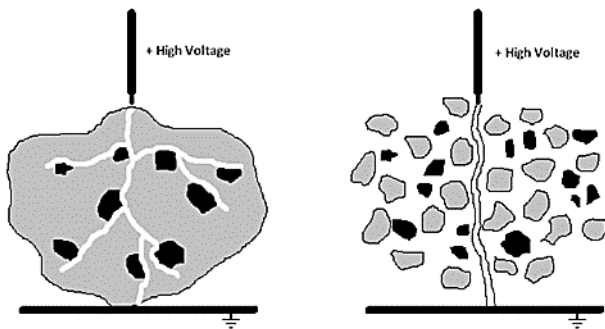


Fig. 4: Plasma channel expansion and failure [10].

According to the mentioned equations we have:

$$E = \left(\frac{2}{r \epsilon_0 \epsilon_r} (\sqrt{r^2 Y^2 + 4GrY} - rY) \right) \tag{11}$$

Once the high and low voltage electrodes are connected, the rock heat up rapidly along the plasma current, resulting in increased internal pressure. A sharp increase in pressure causes an explosion in the ore particle, resulting in a pressure pulse that shatters the rock. Under the influence of an external electric field, an inductive charge is generated at the end of the plasma channel, where there is a free charge. The charge changes the intensity of the electric field at the end of the channel. The relationship between the external electric field and the electric field at the end of the channel is as follows:

$$E_0 = hE \tag{12}$$

where h is the coefficient of proportionality of the electric field distribution, which is expressed as follows:

$$h = \frac{k \epsilon_r}{1 + (k-1)\epsilon_r} \tag{13}$$

That k is a dimensionless constant and corresponds to the size and direction of the plasma channel. According to [14], assuming that the inner hole is a normal sphere, the value of k is 36.

According to Equations (11)-(13), the electric field required to crush the rock will be as follows:

$$E_0 = \frac{1+2\epsilon_r}{3\epsilon_r} \left(\frac{2}{r \epsilon_r \epsilon_0} (\sqrt{r^2 Y^2 + 4GrY} - rY) \right)^{\frac{1}{2}} \tag{14}$$

The electric field required to break a rock depends on the mechanical properties of the material according to Equation (14). Changes in the electric field relative to the mechanical properties of the rock are shown in Figs. 5- 8.

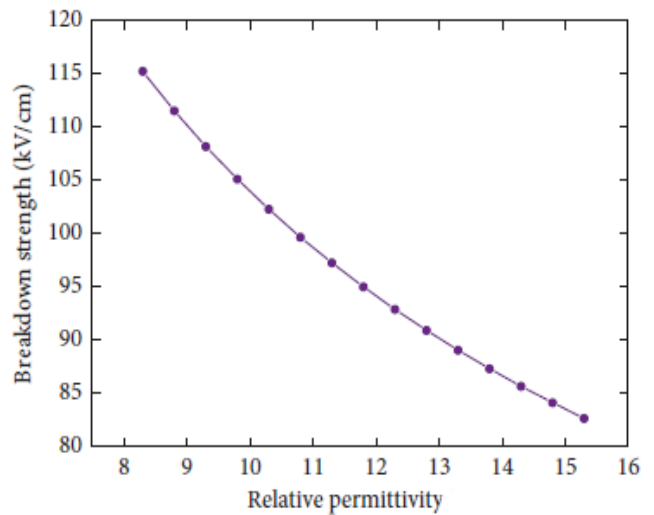


Fig. 5: Changes in the electric field relative to relative permittivity.

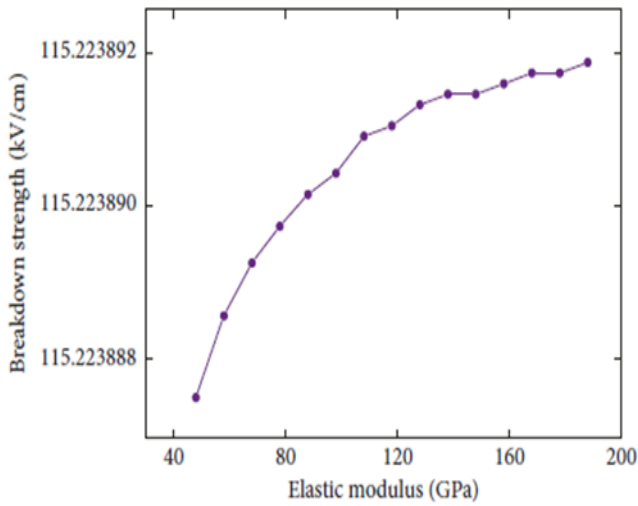


Fig. 6: Changes in the electric field relative to elastic modulus.

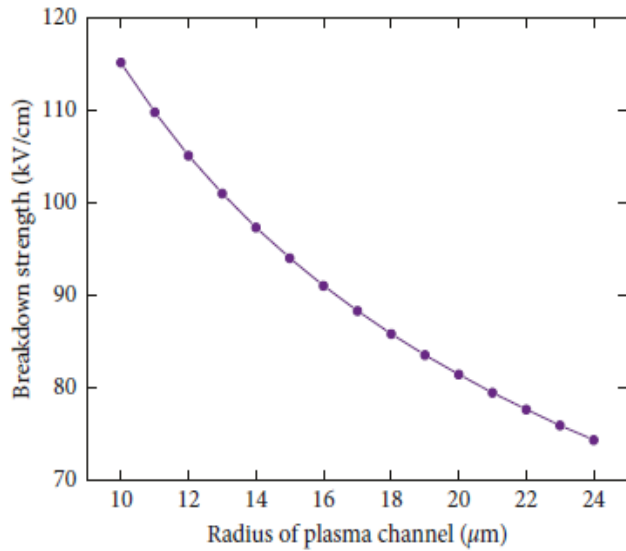


Fig. 7: Changes in the electric field relative to radius of plasma channel.

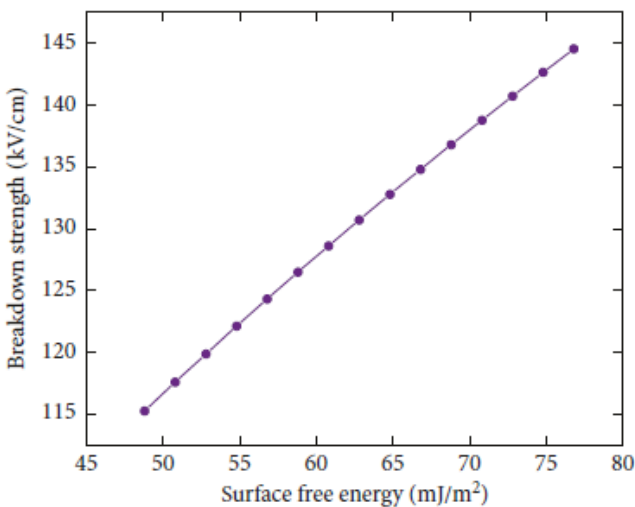


Fig. 8: Changes in the electric field relative to surface free energy.

In Fig. 8 surface free energy or interfacial free energy or surface energy quantifies the disruption of intermolecular bonds that occurs when a surface is created. In the physics of solids, surfaces must be intrinsically less energetically favorable than the bulk of a material (the atoms on the surface have more energy compared with the atoms in the bulk of the material), otherwise there would be a driving force for surfaces to be created, removing the bulk of the material. The surface energy may therefore be defined as the excess energy at the surface of a material compared to the bulk, or it is the work required to build an area of a particular surface. Cutting a solid body into pieces disrupts its bonds and increases the surface area, and therefore increases surface energy. If the cutting is done reversibly, then conservation of energy means that the energy consumed by the cutting process will be equal to the energy inherent in the two new surfaces created. The electric field required to crush several mineral samples has been calculated (The table numbers are based on [15]).

2.2. Modelling the Circuit Required to Break the Rock

A circuit similar to Fig. 9 is used to create the electric field needed to break the rock. Capacitive pulse generators are mostly used in electro discharge technologies. Capacitor C is used to store energy from the source and create the necessary electric field. As mentioned, water acts as an insulating material. The resistance of the R_z circuit includes the resistance of the capacitors, wires and switches, while the inductor is intended for induction in the connecting wires and discharge channel. When the switch is closed, the capacitor pulse generator generates a high pressure pulse to strike the rock and produce the plasma channel. The energy stored in the capacitor is injected into the plasma channel. This energy increases the pressure of the channel, creating mechanical pressure that causes the rock to break. Rock impedance extends from electrical insulation to electrical breakdown, in which a plasma channel is created, which is associated with discharge time and current. In this paper, the Weizel-Rompe model is used for the plasma channel [10], which expresses impedance as a current integral:

$$R_{td}(t) = K * l(\int_0^t i^2(t)dt)^{-1/2} \tag{15}$$

where K is the coefficient of resistance (611V.S1/2/m) [11] and l is the length of the plasma channel (equal to the distance between the electrodes). The distance between the electrodes is 10 mm, and i is the circuit current. According to Kirchhoff equations, the following equation is obtained for the circuit of Fig. 9:

$$L \frac{di}{dt} + (R_z + R_{td}) i(t) + U_c(t) = 0 \tag{16}$$

In this equation, i is the current of the circuit, L is the inductance of the circuit, and U_c is the instantaneous voltage of the capacitor, which is obtained from the following equation:

$$i(t) = C \frac{dU_c}{dt} \tag{17}$$

By placing equations (15) and (17) in equation (16), the following equation is obtained:

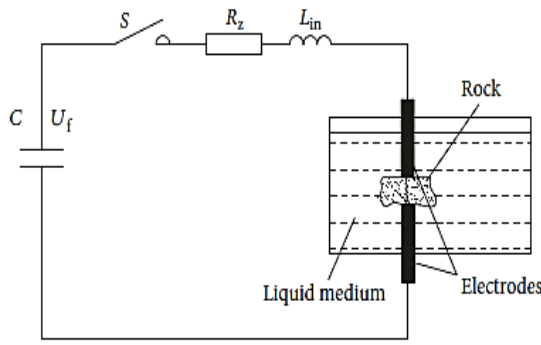


Fig. 9: The equivalent circuit creates a high voltage pulse to break the rock.

$$\frac{d^2i(t)}{dt^2} + \left[\frac{R_z}{L} + \frac{Kl}{L} \left(\int_0^t i^2(t) dt \right)^{\frac{1}{2}} \right] \frac{di(t)}{dt} - \frac{Kl}{2L} \left(\int_0^t i^2(t) dt \right)^{\frac{3}{2}} i^3(t) + \frac{1}{LC} i(t) = 0 \quad (18)$$

$$\begin{cases} x_1 = i(t) \\ x_2 = \frac{di}{dt} \\ x_3 = \int_0^t i^2(t) dt \end{cases} \quad (19)$$

Second-order differential equations (19) can be turned into a first-order differential equation system. From the system of equations (19) we can conclude:

$$\begin{cases} \frac{dx_1}{dt} = x_2 \\ \frac{dx_2}{dt} = \frac{Kl}{2L} x_3^{-1.5} x_1^3 - \left(\frac{R_z}{L} + \frac{Kl}{L} x_3^{-1.5} \right) x_2 - \frac{1}{LC} x_1 \\ \frac{dx_3}{dt} = x_1^2 \end{cases} \quad (20)$$

The initial conditions of the system of equations (20) will be as follows:

$$\begin{cases} x_1(0) = 0 \\ x_2(0) = \frac{U_0}{L} \\ x_3(0) = 0 \end{cases} \quad (21)$$

According to the equivalent circuit of Fig. 9, the plasma channel energy can be expressed as follows:

$$W(t) = \int_0^t i^2(t) R_{td}(t) dt \quad (22)$$

According to the law of energy stability, the energy from the power supply is injected into the plasma channel and when the plasma channel expands, it is converted into heat wave energy and mechanical stress. The energy stability equation is as follows [7]:

$$\frac{dW}{dt} = \frac{P*dV}{dt} + \frac{1}{\gamma-1} \frac{d(PV)}{dt} \quad (23)$$

where γ refers to the entropy index by which the energy distribution of the shock wave and heat in the plasma channel is determined. For condensate materials, its value is from 1.05 to 1.25 and its value in this article is 1.1 [7]. Impulse wave pressure (P), which is caused by the expansion of the plasma channel can be expressed as follows:

$$P = \frac{\gamma-1}{V*\gamma} W \quad (24)$$

For rocks, the generated plasma channel can be thought of as an expanding cylindrical piston. According to the plasma cylindrical channel model and the Rankine-Hugoniot

relationship, the relationship between the volume of the plasma channel and the energy injected into the plasma channel is as follows:

$$\frac{dV}{dt} = \left(\frac{\sqrt{7} \alpha^{14}}{3\sqrt{\rho_0}} \right) * \left[\left(\frac{W}{V} * \frac{\gamma-1}{\gamma} + \beta \right)^{\frac{3}{7}} - \beta^{\frac{3}{7}} \right] * \left(\sqrt{\frac{V}{\pi l}} * 2\pi l \right) \quad (25)$$

where α and β are the shock wave coefficients ($3.001*108$ Pa and $3*108$ Pa, respectively) [15] and ρ_0 is the density of the substance in question. Using the variable change and the Euler's method for the common solution of equations (15) – (25), we obtain the changes in the discharge voltage, the discharge current, the plasma channel energy and the shock wave pressure.

3. SIMULATION AND ANALYSIS OF RESULTS

The equivalent circuit of Fig. 9 is simulated in MATLAB with all the items intended for it. Simulations are performed for three minerals, copper ore, chalcopyrite and hematite. In the following, a case study and feasibility study of using the phenomenon of electromechanical breakdown for crushing minerals will be discussed. In this simulation, a shock wave is used to create the electric field needed to break the rock. The shock wave is 1.2/50 microseconds. The value of the inductor in the circuit is equal to $5\mu\text{H}$ and the resistance R_z is equal to 1Ω . In the simulation for all three minerals, the parameter of Table 1 is used. By performing the simulation, the breakdown voltage required to crush the three studied minerals is obtained as shown in Table 2. The capacitive generator can be used to generate a shock wave. The simulation is performed for 1 cm^3 pieces of rock samples and the required voltage and power as well as the current drawn from the source are calculated.

According to the voltage and power diagrams, it is observed that with increasing the yang modulus of the ore, the amount of voltage and power of the source also increases and may limit the use of this method. By increasing the constant dielectric value, the breakdown voltage decreases and consequently the energy required to crush the rock also decreases.

4. FEASIBILITY STUDY OF USING ELECTROMECHANICAL BREAKDOWN METHOD FOR CRUSHING ROCK

One of the factors that can limit this method is the dimensions of the sample to be crushed, so for this purpose, the effect of stone dimensions on the voltage required to crush the stone is investigated. According to Fig. 19 and Fig. 20, with increasing rock thickness, the voltage required to break the rock decreases and with increasing rock area, the required voltage increases. According to studies, the greatest effect on the voltage required for crushing rock is related to the Young's modulus.

5. DISCUSSION AND CONCLUSION

In this paper, the use of electromechanical breakdown phenomenon for crushing minerals was investigated. This method consumes less energy than other mechanical methods for crushing rock, and since energy consumption in the mining industry and in the world in general is one of the main concerns, this method can be a good option. According to the simulation, it is concluded that with increasing the relative

dielectric constant and the radius of the plasma channel, the breakdown electric field decreases and with increasing the Young's modulus and the surface free energy, the breakdown electric field increases. Once the breakdown is complete, energy consumption can be improved by using a high-capacity current source. The energy required to crush the mentioned mineral samples is calculated from the following equation in joules:

$$W = \frac{1}{2} C (V_{max})^2 \tag{26}$$

where C is the capacitance of the sample, V_{max} is the maximum voltage required, and W is the energy discharged by the electrodes in joules.

According to Table 1, the energy required to crush each 1cm^3 of the samples is less than the mechanical energy required. According to Fig. 11, Fig. 14, and Fig. 17 After the rock is crushed, the electrodes are shorted and according to Fig. 10, Fig. 13 and Fig. 16 a strong current is drawn from the power supply, so it is necessary to install relays to cut off the power supply. This method can be used to weaken rocks and create cracks in them that mechanical crushing machine requires less energy. That is, a device consisting of an electric crusher and a mechanical crusher is made to weaken the rock in the first stage under electric field stress and then crush it with a mechanical crusher, although considering the possibility of using this method in three samples of ore, the amount of electric field required is in the right range and it is possible to supply the required energy. In this method, the breakdown electric field strength depends on relative permittivity, surface free energy, radius of channel and Young's modulus, which is the effect of each factor according to Figs. 5- 8.

The field required for breakdown in this method depends on the mechanical and electrical properties of the ore. If multiple pulses are used instead of one pulse to crush the rock, the performance of the system will be improved as the energy required will be reduced. By using the modelling, time-varying waveforms of voltage, current, and power can be obtained numerically according to different electrical parameters. As mentioned, according to Fig. 19 and Fig. 20, as the thickness of the stone increases, the voltage required to break the stone decreases, and as the surface of the stone increases, the required voltage increases. This may cause limitations in the use of this method and it is not possible to use it in some samples. This study is important for the effect of using high voltage to improve rock breaking efficiency, reduce energy losses and select drilling process parameters.

The mathematical model and numerical simulation model used in this study are useful in analysing the effect of different factors on this process under different conditions. The results of this article show the possibility of better optimization of the rock breaking process using high voltage.

Table 1: Physical properties of studied minerals.

Mineral	Y (GPa)	ϵ_r	G (j/m ²)	E (kV/cm)
Copper ore	128	5.6	1.65	1847
Chalcopyrite	115	1.6	1.45	1906
Hematite	220	1.9	1.9	1790

Table 2: Sample breakdown voltage.

Mineral	Supply Voltage (kV)
Copper ore	426
Chalcopyrite	345
Hematite	1360

Table 3: Energy required to crush minerals.

Sample	Energy (mj)
Copper ore	44.96
Chalcopyrite	8.42
Hematite	155.5

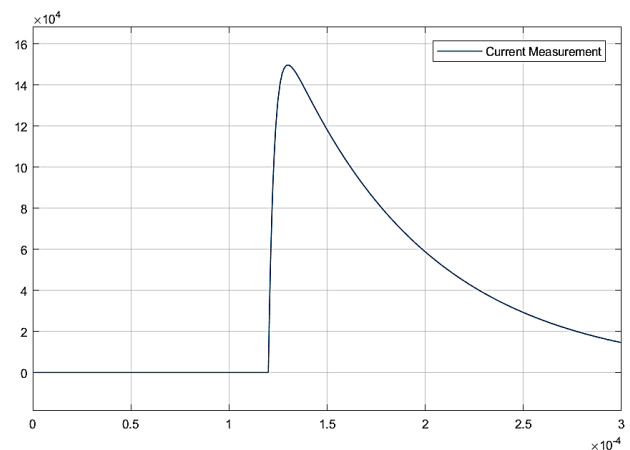


Fig. 10: Current drawn from the source to break the chalcopyrite ore.

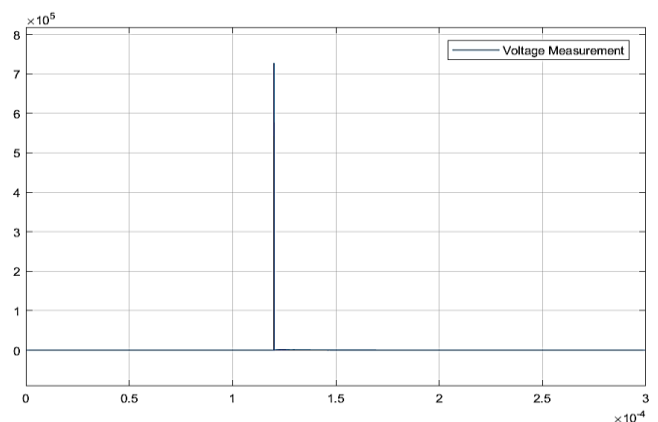


Fig. 11: Voltage required to break chalcopyrite ore.

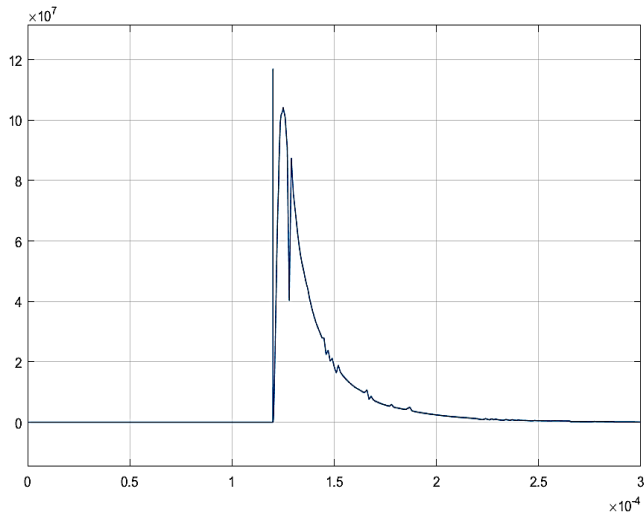


Fig. 12: Source power required to break chalcopyrite ore.

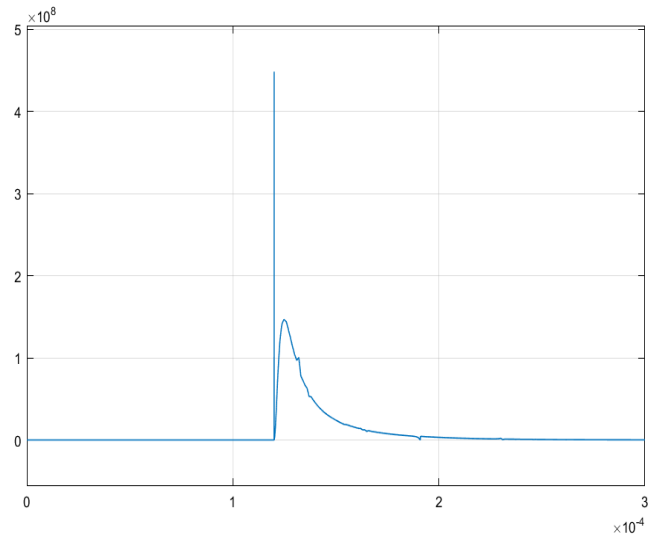


Fig. 15: Source power required to break copper ore.

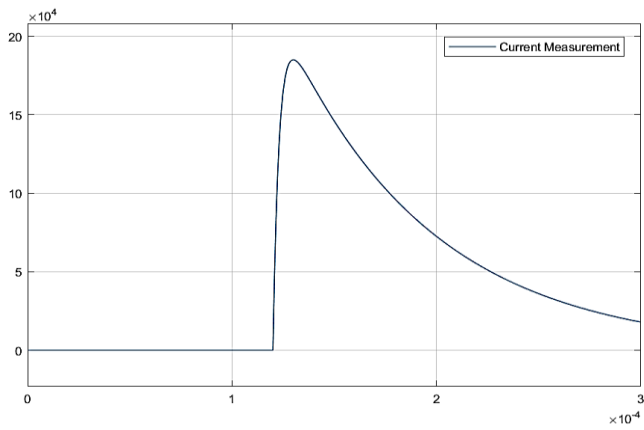


Fig. 13: Current drawn from the source to break the copper ore.

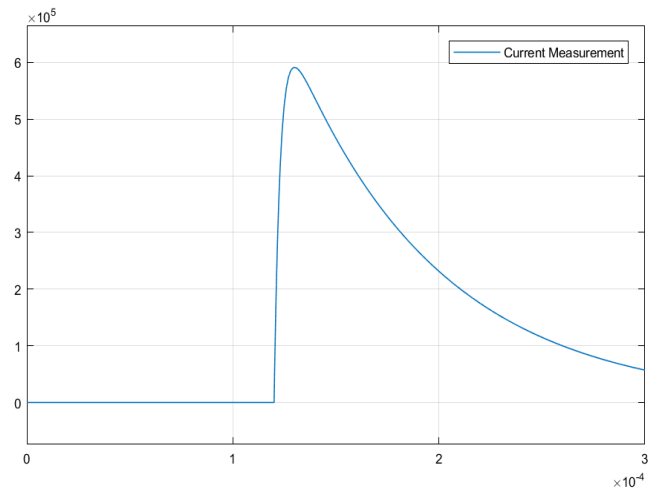


Fig. 16: Current drawn from the source to break the hematite ore.

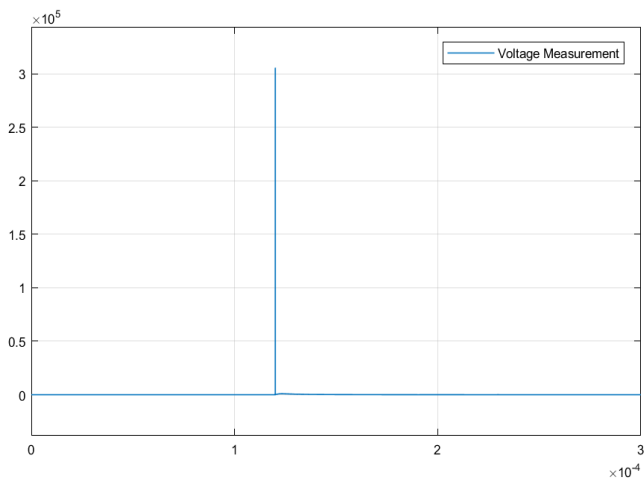


Fig. 14: Voltage required to break copper ore.

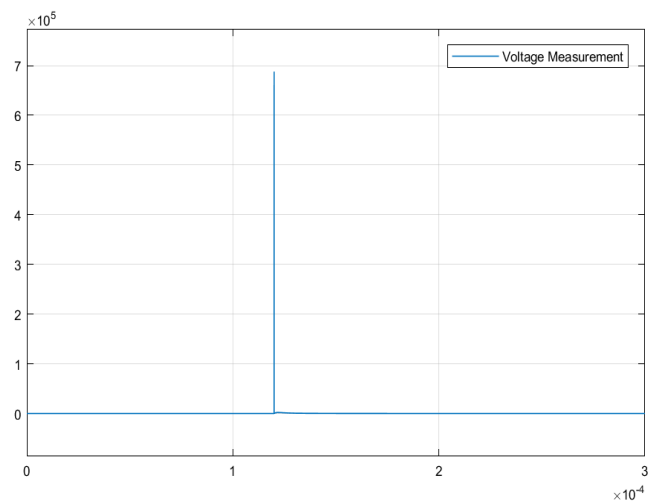


Fig. 17: Voltage required to break hematite ore.

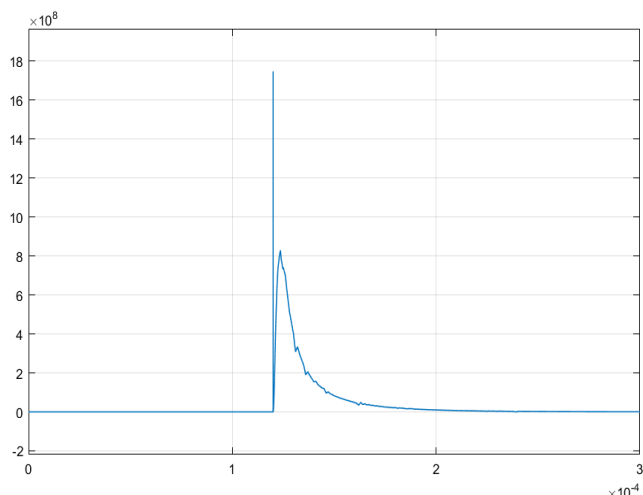


Fig. 18: Source power required to break hematite ore.

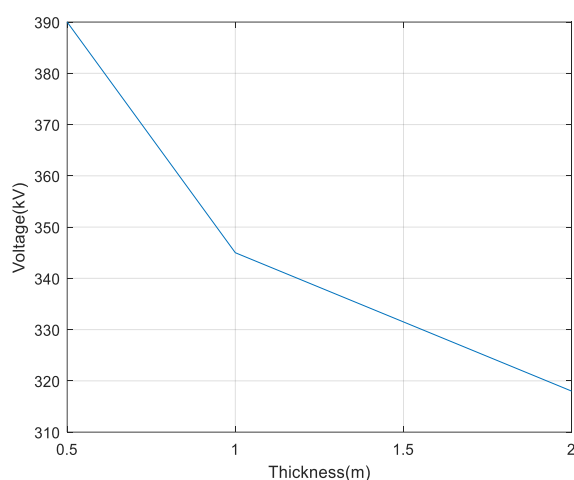


Fig. 19: The change in voltage required to crush the rock relative to the change in rock thickness.

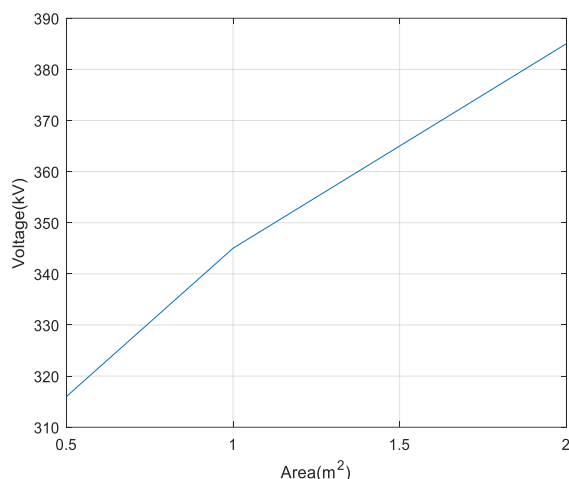


Fig. 20: The change in voltage required to crush the rock relative to the change in rock area.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Asaad Shemshadi: Conceptualization, Formal analysis, Funding acquisition, Project administration, Supervision, Validation, Roles/Writing - original draft. **Mohamad Reza Khojaste:** Data curation, Investigation, Methodology,

Resources, Software, Visualization, Writing - review & editing.

DECLARATION OF COMPETING INTEREST

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, have been completely observed by the author.

REFERENCES

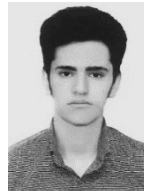
- [1] B. F. Lehmann, M. Reich, M. Mezzetti, E. Anders, and M. Voigt, "The future of deep drilling-A drilling system based on electro impulse technology," *Oil Gas European Magazine*, vol. 43, pp. 187–191, 2017.
- [2] S. M. Razavian, B. Rezai, and M. Irannajad, "Investigation on pre-weakening and crushing of phosphate ore using high voltage electric pulses," *Advanced Powder Technology*, vol. 25, no. 6, pp. 1672–1678, 2014.
- [3] M. Ito, S. Owada, T. Nishimura, and T. Ota, "Experimental study of coal liberation: electrical disintegration versus roller crusher comminution," *International Journal of Mineral Processing*, vol. 92, no. 1-2, pp. 7–14, 2009.
- [4] W. Zuo, F. Shi, K. P. van der Wielen, A. Weh, and W. Alexander, "Ore particle breakage behaviour in a pilot scale high voltage pulse machine," *Minerals Engineering*, vol. 84, pp. 64–73, 2015.
- [5] V. I. Kurets, V. V. Lopatin, and M. D. Noskov, "Influence of local heterogeneities on discharge channel trajectory at electric pulse destruction of materials," *Journal of Mining Science*, vol. 36, no. 3, pp. 268–274, 2000.
- [6] E. Wang, F. Shi, and E. Manlapig, "Experimental and numerical studies of selective fragmentation of mineral ores in electrical comminution," *International Journal of Mineral Processing*, vol. 112-113, pp. 30–36, 2012.
- [7] Z. C. Zhang, "Rock fragmentation by pulsed high voltage discharge and drilling equipment development," Ph.D. Thesis, Zhejiang University, Hangzhou, China, 2013, in Chinese.
- [8] K. P. Van der Wielen, R. Pascoe, A. Weh, F. Wall, and G. Rollinson, "The influence of equipment settings and rock properties on high voltage breakage," *Minerals Engineering*, vol. 46-47, pp. 100–111, 2013.
- [9] J. Biela, C. Marxgut, D. Bortis, and J. Kolar, "Solid state modulator for plasma channel drilling," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 16, no. 4, pp. 1093–1099, 2009.
- [10] C. Li, L. Duan, S. Tan, V. Chikhotkin, and W. Fu, "Damage model and numerical experiment of high-voltage electro pulse boring in granite," *Energies*, vol. 12, no. 4, p. 727, 2019.

- [11] N. Kuznetsova, V. Lopatin, V. Burkin, V. Golovanevskiy, D. Zhgun, and N. Ivanov, “+eoretical and experimental investigation of electro discharge destruction of non-conducting materials,” in *Proceedings of the IEEE Pulsed Power Conference*, pp. 267–273, Chicago, IL, USA, June 2011.
- [12] J. C. Fothergill, “Filamentary electromechanical breakdown,” *IEEE Transactions on Electrical Insulation*, vol. 26, no. 6, pp. 1124–1129, 1991.
- [13] I. W. Mcallister, “Partial discharges in spheroidal voids,” *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 4, pp. 456–461, 2002.
- [14] M. F. Ashby, and D. R. H. Jones, *Engineering materials*. Pergamon press, 1980.
- [15] J. L. Jiang, S. Lan, and J. X. Yang, “Numerical analysis of the plasma characteristics of pulsed discharge in water,” *Journal of Harbin University of Science and Technology*, vol. 13, pp. 107–111, 2008.

BIOGRAPHY



Asaad Shemshadi was born on Nov 1, 1979. He received the B.Sc. degree from the Shiraz University, in 2003, the M.Sc. degree from the Kashan University, Iran in 2007, and PhD degree from Khaje Nasir Toosi University of Technology in 2014, all in Electrical engineering. His research interests are: Vacuum Interrupters design and analysis, high voltage simulations, thermal plasma modeling, high voltage equipments design, transients in vacuum arc quenching and pulsed power.



Mohammad Reza Khojasteh was born on November 26, 1998 in Arak. He received his B.Sc degree from Arak University of Technology, in 2021, in the field of electrical engineering. Interests are: power systems, simulation of high voltage phenomena, protection of power systems.

Copyrights

© 2023 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).





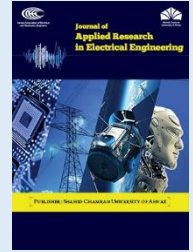
Iranian Association of
Electrical and Electronics
Engineers

Journal of Applied Research in Electrical Engineering

E-ISSN: 2783-2864

P-ISSN: 2717-414X

Homepage: <https://jaree.scu.ac.ir/>



Research Article

Combined Economic Emission Dispatch in a Grid-Connected Microgrid Using an Improved Mayfly Algorithm

Nicholas Kwesi Prah II, Elvis Twumasi , and Emmanuel Asuming Frimpong

Department of Electrical and Electronic Engineering, College of Engineering, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

* Corresponding Author: etwumasi.coe@knust.edu.gh

Abstract: The Combined Economic Emission Dispatch (CEED) is an important consideration in every power system. In this paper, a modified Mayfly Algorithm named Modified Individual Experience Mayfly Algorithm (MIE-MA) is used to solve the CEED optimization problem. The modified algorithm enhances the balance between exploration and exploitation by utilizing a chaotic decreasing gravity coefficient. Additionally, instead of the MA relying solely on the best position, it calculates the experience of a mayfly by averaging its positions. The CEED problem is modeled as a nonlinear optimization problem constrained with four equality and inequality constraints and tested on a grid-connected microgrid that consists of four dispatchable distributed generators and two renewable energy sources. The performance of the MIE-MA on the CEED problem is compared to Particle Swarm Optimisation (PSO), an MA variant that incorporates a levy flight algorithm named IMA and Dragonfly Algorithm (DA) using the MATLAB R2021a software. The MIE-MA achieved the best optimum cost of 11306.6 \$/MWh, compared to 12278.0 \$, 12875.8\$, and 17146.4\$ of the DA, IMA, and PSO respectively. The MIE-MA also achieved the best average optimum cost over 20 runs of 12163.48 \$, compared to 12555.36 \$, 13419.67 \$, and 17270.08 \$ of the DA, IMA, and PSO respectively. The hourly cost curve of the MIE-MA was also the best compared to the other algorithms. The MIE-MA algorithm thus achieves superior optimal values with fewer iterations.

Keywords: MIE-MA, mayfly algorithm, swarm intelligence, economic dispatch.

Article history

Received 22 May 2023; Revised 12 July 2023; Accepted 02 August 2023; Published online 25 November 2023.

© 2023 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)

How to cite this article

N. K. Prah II, E. Twumasi, and E. A. Frimpong, "Combined economic emission dispatch in a grid-connected microgrid using an improved mayfly algorithm," *J. Appl. Res. Electr. Eng.*, vol. 2, no. 2, pp. 112-119, 2023. DOI: [10.22055/jaree.2023.43803.1070](https://doi.org/10.22055/jaree.2023.43803.1070)



1. INTRODUCTION

Optimal operations and effective planning of electric power generation systems are critical elements in electric power systems. Problems relating to controlling, operating, and economic dispatch of power generation resources are still being addressed [1]. Economic dispatch power generation resources are useful in finding the right balance between system performance and cost of operation [2, 3]. Traditionally, the effects of emissions due to power system operations were not considered in the economic dispatch problem. However, due to environmental concerns such as global warming and pollution, it has become necessary to add an emission cost component to the economic dispatch problem. This is collectively known as the Combined Economic Emission Dispatch (CEED), which aims to mitigate emission levels from all generating units and promotes the transition

towards a sustainable and environmentally friendly power generation paradigm [3]. The addition of the emission component, and consideration of intermittent renewable sources in the CEED problem increases the complexity of the problem due to the increase non-convexity, increased constraints, and decision variables, increased computational time and the need for trade-off between emission levels and economic cost. The CEED presents a highly nonlinear optimization problem with constraints such as link capacity limits, ramp rate limits, load-demand constraints, and the generator's maximum capacity limit constraints and the non-convexity significantly introduces multiple local optima, making it difficult to find the optimal solution. Due to this complexity, it is inefficient to employ conventional mathematical and numerical methods to solve the CEED problem, due to their high tendency to get stuck at local optimum [4]. More sophisticated nature-inspired metaheur-

istic algorithms have been adopted to solve complex problems due to their effectiveness in achieving near-global or global solutions for complex nonlinear problems [5].

Metaheuristic algorithms, a computational intelligence paradigm, prove valuable in addressing such complex problems. They provide benefits including effortless application to continuous and discrete problems, convex and non-convex problems, reduced mathematical complexity, and effective search for global optimal solutions [6]. Over the years, metaheuristic algorithms have been applied to economic dispatch problems in microgrids. In [7] PSO was used to solve the CEED problem while considering demand response, uncertainty in load and RESs on a microgrid with four dispatchable DGs. In [8], Grasshopper Algorithm was used to solve the CEED problem in an IEEE system which has three thermal generators, the algorithm was shown to outperform FA and BA in achieving the CEED. In [9], Gravitational Search Algorithm (GSA) was used as an optimization tool to solve the CEED problem. The algorithm was shown to outperform the BBO algorithm in achieving the optimal CEED. In [10], Economic dispatch combined with emission dispatch was performed using the Crow Search Algorithm (CSA). CSA was shown to outperform PSO and GA in achieving the optimal CEED. The above algorithms still have issues with local stagnation and so researchers are coming up with new and improved algorithms to address these issues [11].

The Mayfly algorithm, one of the latest Swarm Intelligence Optimization Algorithms (SIOAs) is considered a promising algorithm and, it draws inspiration from the movement and mating behaviour of mayflies [12, 13]. It shows great promise with its improved capabilities in exploration and exploitation. It has been shown to outperform other algorithms such as PSO, Invasive Weed Optimisation (IWO), Bees algorithm (BA), GA, Differential Evolution (DE), Firefly Algorithm (FA), Harmony Search Algorithm (HSA), etc. MA has been successfully applied by researchers to tackle complex problems [13]. For instance, in [14], for agricultural unmanned aerial vehicles (UAVs), it was used to solve a 2D path planning problem. In reference [15], MA was applied to tackle the optimal power flow problem in regulated electricity markets. Moreover, the algorithm has demonstrated advancements in Maximum Power Point Tracking (MPPT) for photovoltaic systems [11].

However, MA does have certain limitations, including premature convergence and stagnation. As a result, further contributions are required to address these issues. Other researchers have made some attempts to address the above problems. A modified version of MA known as PGB-IMA was presented in reference [13]. This variant selected the global best from the entire mayfly population, including both males and females, to enhance exploration capabilities. While this enhancement proved effective for unimodal functions, it exhibited slower convergence or local optima entrapment on multimodal functions. In [12], levy flight was employed to improve exploration in the MA for solving the CEED problem. However, this approach occasionally caused mayflies to fly out of smaller search spaces. Researchers introduced ModMA in reference [14]. This approach incorporated various techniques, including adaptive Cauchy

method, exponent decreasing inertia weight, an improved crossover operator and decreasing inertia weight, to achieve a balance between exploitation and exploration within the MA. ModMA improved the convergence rate but still faced the challenge of local optima trapping. Consequently, further improvements are necessary to address the issues of local optima entrapment and premature convergence, thus enhancing the overall performance of MA.

Therefore, this work proposes a modified MA that addresses the aforementioned deficiencies of the MA and consequently applies it to the CEED problem to obtain optimal results. The modification introduced in the algorithm focuses on individual experience, resulting in MIE-MA. This algorithm enhances mayfly movement to improve convergence rate and overcome local optima entrapment. This is achieved by replacing the personal best (pbest) in the original Mayfly Algorithm with personal experience (Pexp). Pexp is calculated as the average of all positions visited by a mayfly, ensuring equal contribution from all visited positions. This enables the mayflies to effectively explore their search spaces, avoiding premature stagnation and overlooking optimal solutions. As a result, the MIE-MA approach yields optimal solutions with fewer iterations. Additionally, the algorithm incorporates a chaotic decreasing gravity coefficient to strike a balance between exploration and exploitation. The innovation of this work lies in the modification of the mayfly algorithm by modifying the individual experience formula of the MA, adopting the chaotic decreasing gravity coefficient to the mayfly and applying it to a CEED problem in a grid-connected microgrid. The rest of the paper is structured as follows: [Section 2](#) describes the problem formulation, [Section 3](#) describes the mayfly algorithms, [Section 4](#) contains the proposed MIE-MA, [Section 5](#) illustrates the implementation of the MIE-MA on the CEED, [Section 6](#) contains results and analysis and finally, [Section 7](#) concludes the paper.

2. PROBLEM FORMULATION

The CEED problem is modeled as a nonlinear constrained optimization problem with both equality and inequality constraints [7].

2.1. Decision Vector

For each hour, the decision vector consists of the power output of the dispatchable generators (DG) and the power exchanged with the upstream grid. i.e.

$$X = [P_{DG(1)}, \dots, P_{DG(N)}, P_{grid}]$$

where, P_{DG} is power output of dispatchable generators, N is the total number of dispatchable generators, and P_{grid} is power exchange with upstream grid. P_{grid} is positive when power is bought from the grid and negative when power is transferred to the upstream grid.

2.2. Objective Function

The cost of dispatchable DGs is denoted by equation (1), while their emission treatment cost is described by equation (2).

$$C_i(P_{DG,i}) = \alpha_i(P_{DG,i})^2 + b_i P_{DG,i} + c_i \quad \$/h \quad (1)$$

$$em_i(P_{DG,i}) = d_i(P_{DG,i})^2 + e_i P_{DG,i} + f_i \quad \$/h \quad (2)$$

α_i , b_i , and c_i are operational cost coefficients, and d_i , e_i , and f_i are emission treatment cost coefficients for CO₂, SO₂ and NO_x, respectively.

The objective function includes the operation and emission costs, i.e.,

$$OF = \sum_{i=1}^{NG} \sum_{t=1}^{24} C_i (P_{DG,i}(t)) + \sum_{i=1}^{NG} \sum_{t=1}^{24} em_i (P_{DG,i}(t)) + \sum_{t=1}^{24} \rho(t) P_{grid}(t) + \sum_{t=1}^{24} \gamma P_{grid}(t) \quad (3)$$

where $\rho(t)$ is the market price at t^{th} time, and γ is the emission factor of power purchased from upstream grid at the t^{th} time.

2.3. Constraints

The constraints are power balance constraints, output limit constraints, ramp rate limits and power transfer limits.

2.3.1. Power balance

$$P_d(t) = P_{grid}(t) + \sum_{i=1}^{NG} P_{DG,i}(t) + \sum_{p=1}^{Nren} P_{renew,p}(t)$$

where $P_{renew,p}(t)$ is the power output of P^{th} renewable DG unit at the t^{th} time, and $Nren$ is the total number of renewable DGs.

2.3.2. Output limits

The power output of each dispatchable DG must be within its output limits, i.e.

$$P_{DG,i,min} \leq P_{DG,i}(t) \leq P_{DG,i,max}$$

2.3.3. Ramp rate limits

The gradient of change of the output of the DG must be limited, i.e.,

$$P_{DG,i}(t) - P_{DG,i}(t - 1) \leq RU_{DG,i}$$

$$P_{DG,i}(t - 1) - P_{DG,i}(t) \leq RD_{DG,i}$$

2.3.4. Power transfer limits

To comply with the power flow limit of the link connecting the microgrid and the upstream grid, it is necessary to ensure that the following equation is met:

$$-P_{transf,max} \leq P_{grid}(t) \leq P_{transf,max}$$

3. MAYFLY ALGORITHM

The MA mimics the way mayflies fly and mate. It is a combination of the major pros of PSO [16], FA [17], and GA [18]. The MA consists of the following 6 phases [4]:

3.1. Initialisation

During this phase, a random set of male and female mayflies is generated. Each mayfly is assigned a current velocity (v_i) and position (x_i), denoted as $v_i = (v_{i1}, v_{i2}, \dots, v_{in})$ and $x_i = (x_{i1}, x_{i2}, \dots, x_{in})$, respectively. The positions of the mayflies are then adjusted using information from their best position ($pbest$) and the best position in the entire population ($gbest$).

3.2. Male Mayfly Movement

The position is updated by adding to the position, a velocity v_i^{t+1} . For male mayflies, the velocity is expressed as

$$v_{ij}^{t+1} = v_{ij}^t + a_1 e^{-\beta r_p^2} (pbest_{ij} - x_{ij}^t) + a_2 e^{-\beta r_g^2} (gbest_j - x_{ij}^t)$$

In the given equation, the visibility coefficient is denoted as β , and a_1 and a_2 represent positive constants representing attraction refers to the best position attained by the male mayfly indexed as "i" in dimension "j". The variables " r_p " and " r_g " represent the Euclidean distances between the position " x_i " and the individual best position ($pbest$) and the global best position ($gbest$), respectively. g is the gravity coefficient and is typically a fixed number ranging from 0 to 1.

3.3. Female Mayfly Movement

Female mayflies move towards male mayflies i.e. $y_i^{t+1} = y_i^t + v_i^{t+1}$, y denotes position of the female mayfly. For male mayflies, the velocity is expressed as:

$$v_{ij}^{t+1} = \begin{cases} g * v_{ij}^t + a_2 e^{-\beta r_{mf}^2} (x_{ij}^t - y_{ij}^t), & \text{if } f(y_i) > f(x_i). \\ g * v_{ij}^t + f_l * r, & \text{if } f(y_i) \leq f(x_i). \end{cases}$$

The velocity and position of the i -th female mayfly in dimension j at iteration t are respectively represented by v_{ij}^t and y_{ij}^t . The attraction constant is denoted as " a_2 " and the visibility coefficient is represented by " β ". The variable " r_{mf} " indicates the Euclidean distance between the female mayfly indexed as " i " and the male mayfly indexed as " i ". The random walk coefficient, " f_l ", signifies that a female may not be attracted to a male. Additionally, r represents a random value ranging from -1 to 1.

3.4. Mating Mayflies

This enhances exploration through communication between mayflies by producing offsprings off_1 and off_2 . The crossover operator is used, i.e., $off_1 = r * m + (1-r) * f$ and $off_2 = r * f + (1-r) * m$, where m , f , and r are male mayfly, female mayfly, and a random number between 0 and 1, respectively

3.5. Mutation of Mayflies

This enhances the exploitation of the MA. This is expressed as

$$offspring_n = offspring_n + \sigma N_n(0,1).$$

In this equation, σ represents the standard deviation, while N_n denotes the standard normal distribution.

3.6. Reduction of Nuptial Dance and Random Walk

This aids in striking a balance between exploitation and exploration. This is expressed as: $d_t = d_o \delta^t$, $0 < \delta < 1$ and $fl_t = fl_o \delta^t$, $0 < \delta < 1$, respectively.

In this equation, t represents the iteration counter, and δ is a fixed value in the range of (0, 1).

3.7. Pseudocode of the Algorithm

- The pseudocode of the algorithm is shown below:
- Define the objective function as $f(x)$, where x represents the vector $(x_1, \dots, x_d)^T$.
 - Set the initial positions and velocities for the male population of mayflies.
 - Set the initial positions and velocities for the female population of mayflies.
 - Evaluate the solutions and determine the global best solution ($gbest$).

- **Do while** iteration < maximum iterations
 - Adjust the velocities and positions of both sets of mayflies.
 - Solve their objective function values.
 - Separate mayflies based on their objective function value.
 - Enable mating amongst them.
 - Use crossover operator to produce offspring.
 - Assign genders (male and female) to offspring randomly.
 - Do a replacement of worst solutions with best solutions.
 - reassign the personal best (*pbest*) and global best (*gbest*) solutions with their latest values.
- **End while**
- Show the final results.

4. MODIFIED INDIVIDUAL EXPERIENCE MAYFLY ALGORITHM

In the original Mayfly Algorithm (MA), each mayfly's position is adjusted based on its individual experience (*pbest*) and the experience of its neighbours. However, this approach may limit the contribution of mayflies that are consistently moving at a better rate than the global best (*gbest*), potentially leading to stagnation if the *gbest* is trapped in a local optimum.

To address this issue, a modification is proposed. In this modified approach, the experience of a mayfly is calculated as the average of the positions it has visited in the search space. This enhanced representation provides a better understanding of how the mayflies are approaching the global optimum and ultimately yields optimal values in the search space. The formulation for this modification is as follows:

$$P_{exp,i}^t = \frac{\sum_{t=1}^{iter} x_i^t}{iter} \quad (4)$$

where, $P_{exp,i}^t$ is mayfly *i*-th experience at step *t*, *iter* is iteration number at current step, and x_i^t is mayfly *i*-th position at step *t*.

To enhance the balance between exploration and exploitation in the Mayfly Algorithm (MA), a strategy of adopting a chaotic random decreasing gravity coefficient is implemented. This strategy is motivated by a study [5] that explored different weight strategies in Particle Swarm Optimization (PSO) and their impact on the algorithm's performance. The study found that a chaotic random decreasing inertia weight strategy yielded the highest accuracy. This strategy enables alternating between rough and fine search in all evolutionary processes [6].

In this work, a slightly modified version of the chaotic random decreasing inertia weight strategy is utilized. The formulation for this strategy is as follows:

$$g = (gmax - gmin) * \left(\frac{MaxIt - iter}{MaxIt} \right) + gmin * z \quad (5)$$

$$z = 4 * z * (1 - z)$$

where *gmax* and *gmin* represent the maximum and minimum inertia weights, respectively, and *z* is a number randomly picked between 0 and 1.

5. IMPLEMENTATION OF THE MIE-MA ON THE CEED

The studied grid connected microgrid consists of four dispatchable DGs (MTs) and two renewable energy sources as shown in Fig. 1 [7, 19].

The data used in this research consists of the operational characteristics of four dispatchable DGs (micro-turbines), day ahead power forecasts of the two renewable energy sources and the grid price for each hour. The data was obtained from [7]. Table 1 and Table 2 in the Appendix represent the utilized data.

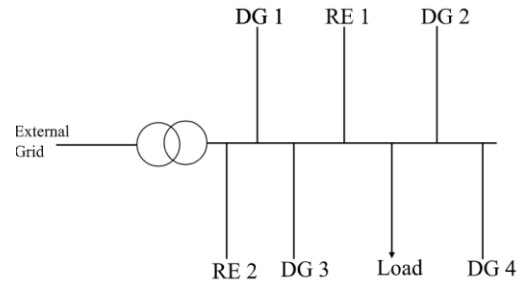


Fig. 1: Diagram of the studied system.

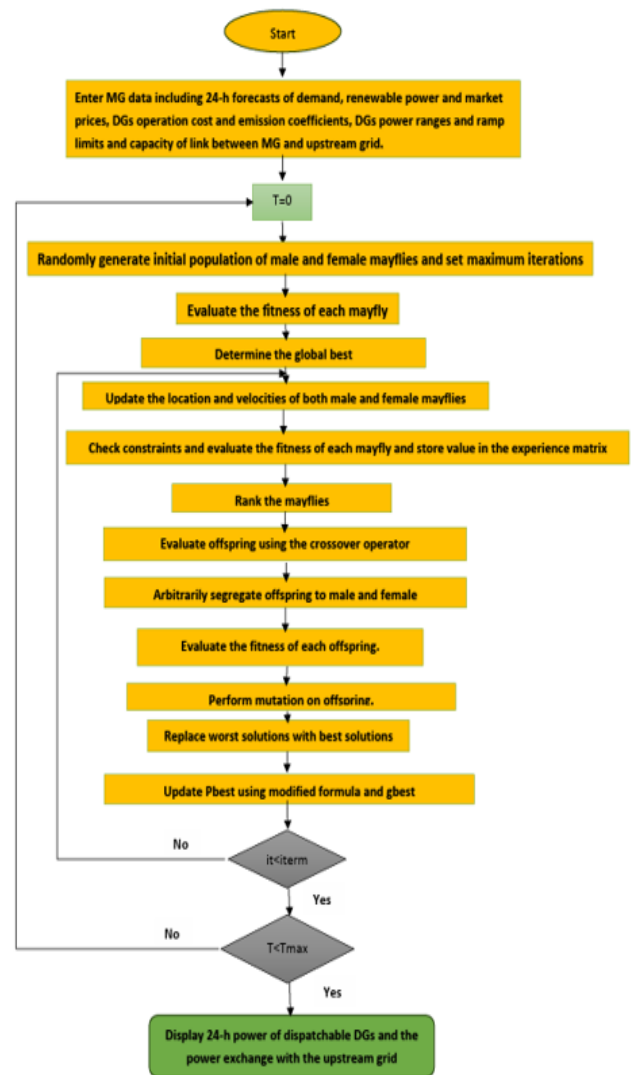


Fig. 2: Flowchart of the Implementation method.

According to [19], for Micro-Turbines (MT), the emission of CO₂, SO₂ and NO_x are 0.7239, 0.0036 and 0.1994 kg/MWh respectively. For a macro grid, the emissions of CO₂, SO₂ and NO_x are 0.8891, 1.8016 and 1.6021 kg/MWh respectively. The treatment cost for CO₂, SO₂ and NO_x are 0.0311, 2.1999 and 9.3332 \$/kg, respectively [7]. Therefore $d_i = f_i = 0$, $e_i = 1.8620$, and $\gamma = 18.9161$.

The optimal parameters for MIE-MA include maximum iterations=100, male and female population = 200 (each one), $a_1 = 1$, $a_2 = 1.5$, $\beta = 2$, $d = 0.1$, $f_i = 0.1$, $g = 0.8$, $\delta = 0.77$, $gmax = 0.9$, and $gmin = 0.2$.

6. RESULTS AND DISCUSSION

This section contains the results from the application of the MIE-MA on the CEED problem compared to a variant of the MA (IMA) [11] which incorporates levy flight algorithm, PSO and DA. The comparison parameters are optimum cost, average cost and standard deviation over 20 individual runs and hourly cost curves.

6.1. Optimum Cost Value

Table 3 shows the lowest value of the objective function obtained by each algorithm. It can be observed from the results above that the MIE-MA achieved the best optimal cost compared with PSO, DA and IMA, this is due to the superior exploration and exploitation abilities of the MIE-MA.

6.2. Mean Cost Values Standard Deviation

Table 4 shows the average optimum cost over 20 runs obtained by the various algorithms. It can be observed from the table above that the MIE-MA achieved the best average cost and standard deviation over 20 independent runs. This further shows the accuracy and consistency of the MIE-MA over the other algorithms.

6.3. Hourly Cost Curves and Convergence Curve

Fig. 3 illustrates the optimum cost of the objective function at each hour. It can be observed from the curves that for each hour, the MIE-MA consistently achieves a lower objective function value than the other algorithms. The convergence curve in Fig. 4 also shows the optimal cost achieved for each iteration in the final hour. The MIE-MA converged to its final value before the 700th iteration.

6.4. Dispatch of DGs and Exchanged Power

Table 5 shows the 24-hour dispatch results for the grid connected microgrid after the application of the modified individual experience mayfly algorithm. The negative values indicate a sale of power to the external grid. It can be observed from the dispatch that the power balance constraints, link capacity constraints, ramp rate limits and DG capacity limits are all respected. Furthermore, the dispatch indicates an effective sale of about 10.6MW over 24 hours to the external grid and for hours where the prices were higher (hours 12 to 24- see appendix), the dispatch indicates an effective sale of about 18.5 MW. This represents extra money for the MG to offset some of its operation and emission treatment costs.

Table 3: Optimum cost values of the various algorithms.

Algorithm	Optimum Value (\$)
MIE-MA	11306.6
ModLMA	12875.8
PSO	17146.4
DA	12278.0

Table 4: Average optimum cost and standard deviation over 20 runs.

Algorithm	Average Value (\$)	Standard Deviation
MIE-MA	12163.48	235.85
IMA	13419.67	323.69
PSO	17270.08	445.62
DA	12555.36	277.75

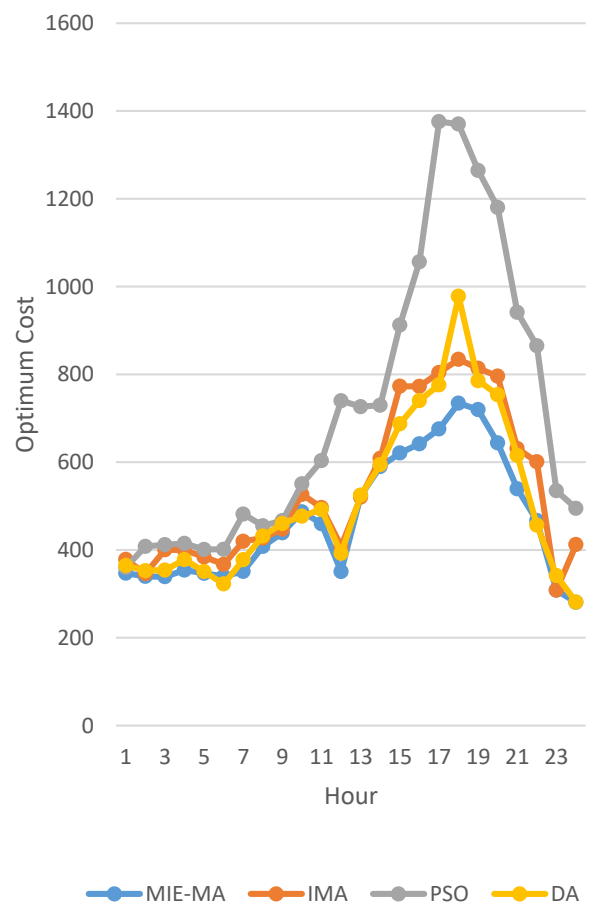


Fig. 3: Hourly cost curves.

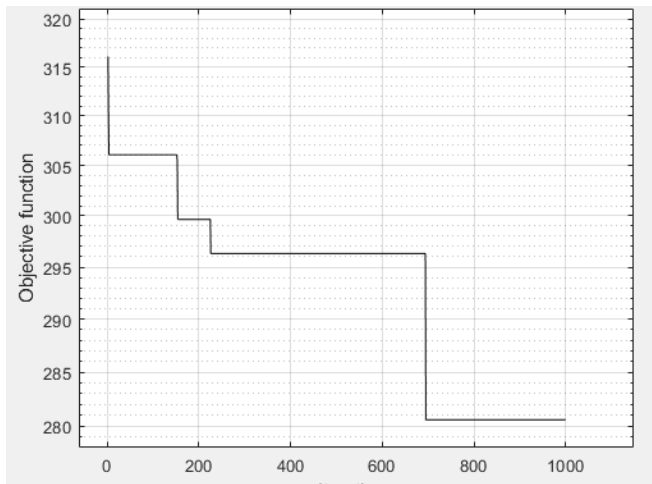


Fig. 4: Convergence curve of the MIE-MA in solving the CEED.

Table 5: Dispatch of DGs.

T	P _{DG,1} (MW)	P _{DG,2} (MW)	P _{DG,3} (MW)	P _{DG,4} (MW)	P _{Grid} (MW)
1	3.852	1.2083	1.0626	0.8104	1.7969
2	1.2509	1.3581	0.9457	1.0152	3.9703
3	1.9208	1.2048	0.8068	0.9873	3.5505
4	4.3311	1.6524	0.8028	0.9074	1.3366
5	5	1.6675	1.1183	1.1321	-0.4604
6	4.813	3.6753	1.401	1.0374	-2.7628
7	4.8629	4.9925	1.496	0.9214	-2.7726
8	4.9768	3.9284	1.0741	0.8172	-0.5763
9	3.8179	1.8088	1.0172	0.9024	3.0957
10	4.5615	3.2682	1.6244	1.0409	0.9612
11	4.7876	4.9833	0.9646	0.9939	-0.2693
12	4.8209	5	2.6222	2.1431	-3.5097
13	4.6665	4.553	2.9348	1.9861	-1.4302
14	4.7749	4.9937	2.6383	1.919	-0.6257
15	4.9885	4.6034	2.8459	2.3223	-0.6299
16	4.795	4.5757	2.9824	2.5704	-0.5134
17	4.9654	4.7553	2.7354	2.8849	-0.2609
18	4.9627	4.3054	2.9717	2.8497	0.2308
19	4.9167	4.774	2.2574	2.1228	0.7963
20	4.7373	4.6856	2.8315	2.9606	-0.6249
21	4.9526	4.8529	2.8159	2.1956	-1.3868
22	4.9985	4.6579	2.7469	2.1934	-2.1666
23	4.8887	4.4929	2.0285	2.328	-3.9179
24	4.8532	4.8041	2.5814	1.643	-4.4316

6.5. Operation and Emission Costs

Table 6 represents the operation and emission treatment costs. Negative values indicate a profit. It can be observed that some gains are made in the emission treatment costs to offset some of the operational costs, leading to an overall lower CEED cost.

Table 6: Hourly operation and emission treatment costs.

T	Operation Cost (\$/h)	Emission Treatment Cost(\$/h)
1	300.25	46.9
2	256	83.62
3	262.68	76.33
4	314.5	39.61
5	338.86	7.9
6	371.82	-31.92
7	380.69	-29.6
8	398.91	9.21
9	366.19	72.61
10	449	37.73
11	443.16	16.75
12	389.93	-39.23
13	524.8	-0.73
14	574.99	14.84
15	605.29	15.57
16	623.73	18.08
17	651.91	23.64
18	701.89	32.47
19	678.45	41.27
20	627.2	16.52
21	537.85	1.36
22	480.67	-13.81
23	356.76	-48.54
24	338.58	-57.98

7. CONCLUSION

A modified Mayfly Algorithm has been developed, known as the Modified Individual Experience Mayfly Algorithm (MIE-MA). This algorithm incorporates changes to the individual experience of each mayfly and enhances the balance between exploitation and exploration capabilities. The MIE-MA has been applied to a CEED problem with four constraints. The considered grid-connected microgrid consists of four dispatchable DGs and two renewable DGs. The MIE-MA achieved the best optimum cost of 11306.6 \$, compared to DA, IMA and PSO. This is an indication of the MIE-MA’s superior exploitation-exploration balance. The MIE-MA also achieved the best average optimum cost and standard deviation over 20 runs of 12163.48 \$ and 235.85\$ compared to DA, IMA, and PSO. This is an indication of the superior accuracy and consistency of the MIE-MA over the other algorithms. The hourly cost curve also showed that the MIE-MA almost always outperformed the other algorithms in each hour. The MIE-MA can still be improved further by further enhancing its exploitation abilities and so future works should look at enhancing the MIE-MA or applying the MIE-MA to other engineering problems.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Nicholas Kwesi Prah II: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Software, Visualization, Roles/Writing - original draft, Writing - review & editing. **Elvis Twumasi:** Conceptualization, Investigation, Methodology, Supervision, Validation, Visualization, Roles/Writing - original draft, Writing - review & editing. **Emmanuel Asuming Frimpong:** Investigation, Methodology, Supervision, Validation, Visualization, Roles/Writing - original draft, Writing - review & editing.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy has been completely observed by the authors.

APPENDIX

Table 1: Operational parameters of the DGs

	DG1	DG2	DG3	DG4
α (\$/MWh)	0.03	0.025	0.02	0.02
b (\$/MWh)	27.7	39.1	61.3	65.6
c(\$)	0	0	0	0
Minimum Power (MW)	1	1	0.8	0.8
Maximum Power (MW)	5	5	3	3
Ramp-up limits (MW)	2.5	2.5	3	3
Ramp-down limits (MW)	2.5	2.5	3	3

Table 2: Day-ahead demand, renewable output and grid price.

Hour	Demand (MW)	Renewable #1(MW)	Renewable #2(MW)	Grid price (\$/MWh)
1	8.73	0	0	15.3
2	8.54	0	0	10.97
3	8.47	0	0	13.51
4	9.03	0	0	15.36
5	8.79	0.63	0	18.51
6	8.81	0.8	0	21.8
7	10.12	0.62	0	37.06
8	10.93	0.71	0	22.83
9	11.19	0.68	0	21.84
10	11.78	0.35	0	27.09
11	12.08	0.62	0	37.06
12	12.13	0.36	0.75	68.95

13	13.92	0.4	0.81	65.79
14	15.27	0.37	1.20	66.57
15	15.36	0	1.23	65.44
16	15.69	0	1.28	79.79
17	16.13	0.05	1	115.45
18	16.14	0.04	0.78	110.28
19	15.56	0	0.71	96.05
20	15.51	0	0.92	90.53
21	14.00	0.57	0	77.38
22	13.03	0.60	0	70.95
23	9.82	0	0	59.42
24	9.45	0	0	56.68

REFERENCES

- [1] A. I. Omar, Z. M. Ali, M. Al-gabalawy, and S. H. E. A. Aleem, "Multi-objective environmental economic dispatch of an electricity system considering integrated natural gas units and variable renewable energy sources," *Mathematics*, vol. 8, no. 7, article 1100, 2020.
- [2] P. Kasinathan, K. Nagarajan, and V. Sangoden, "Chaotic self-adaptive interior search algorithm to solve combined economic emission dispatch problems with security constraints," *Computational Intelligence and Neuroscience*, vol. 29, no. 8, article 12026, 2019.
- [3] N. Karthik, A. K. Parvathy, and R. Arul, "Non-convex economic load dispatch using cuckoo search algorithm," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 5, no. 1, pp. 48–57, 2017.
- [4] J. K. Pattanaik, M. Basu, and D. P. Dash, "Review on application and comparison of metaheuristic techniques to multi-area economic dispatch problem," *Prot. Control Mod. Power Syst.*, vol. 2, article 17, 2017.
- [5] C. Wang and W. Song, "A modified particle swarm optimization algorithm based on velocity updating mechanism," *Ain Shams Eng. J.*, vol. 10, no. 4, pp. 847–866, 2019.
- [6] Z. Beheshti, and N. Branch, "A Review of population-based meta-heuristic algorithm," *Soft Computing Models in Industrial and Environmental Applications*, 2013.
- [7] A. R. Jordehi, "Dynamic environmental-economic load dispatch in grid-connected microgrids with demand response programs considering the uncertainties of demand, renewable generation and market price," Wiley no. July, pp. 1–17, 2020.
- [8] H. Bibi, A. Ahmad, F. Aadil, M. Kim, and K. Muhammad, "A Solution to Combined Economic Emission Dispatch (CEED) problem using Grasshopper Optimization Algorithm (GOA)," in *Proc. 2020 Int. Conf. Comput. Sci. Comput. Intell.*, pp. 712–718, 2020.
- [9] P. Chittur Ramaswamy and G. Deconinck, "Smart grid reconfiguration using simple genetic algorithm and NSGA-II," in *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, Berlin, Germany, 2012, pp. 1-8.
- [10] A. F. Sheta, "Solving the economic load dispatch problem using crow search algorithm," in *8th*

International Multi-Conference on Complexity, Informatics and Cybernetics (IMCIC 2017), 2017.

- [11] K. Nagarajan, A. Rajagopalan, S. Angalaeswari, L. Natrayan, and W. D. Mammo, "Combined economic emission dispatch of microgrid with the incorporation of renewable energy sources using improved mayfly optimization algorithm," *Computational Intelligence and Neuroscience*, article 6461690, 2022.
- [12] K. Zervoudakis and S. Tsafarakis, "A mayfly optimization algorithm," *Comput. Ind. Eng.*, vol. 145, article 106559, 2020.
- [13] S. Mo, Q. Ye, K. Jiang, X. Mo, and G. Shen, "An improved MPPT method for photovoltaic systems based on mayfly optimization algorithm," *Energy Reports*, vol. 8, pp. 141–150, 2022.
- [14] X. Wang, J. Pan, Q. Yang, L. Kong, V. Snášel, and S. Chu, "Modified mayfly algorithm for UAV path planning," *Drones*, vol. 6, no. 5, article 134, 2022.
- [15] V. K. Bhaskar, S. Ramesh, P. Chandrasekar, K. Karunanithi, and A. Raja, "An improved mayfly algorithm based optimal power flow solution for regulated electric power network," *International Journal of Advanced Technology and Engineering Exploration*, vol. 9, no. 92, pp. 979–995, 2022.
- [16] N. I. Nwulu and X. Xia, "Optimal dispatch for a microgrid incorporating renewables and demand response," *Renew. Energy*, vol. 101, pp. 16–28, 2017.
- [17] Y. Zhang, N. Gatsis and G. B. Giannakis, "Robust Energy Management for Microgrids With High-Penetration Renewables," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 4, pp. 944-953, 2013.
- [18] A. Kumar, M. Z. U. Khan, and D. M. A. Hussain, "Microgrids technology: a review paper," *Gyancity Journal of Electronics and Computer Science*, vol. 3, no. 1, pp. 11-23, 2018.
- [19] A. Khodaei, "Microgrid optimal scheduling with multi-period islanding constraints," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1383–1392, 2014.
- [20] M. Nemati, M. Braun, S. Tenbohlen "Optimization of unit commitment and economic dispatch in microgrids based on genetic algorithm and mixed integer linear programming," *Applied Energy*, vol. 210, pp. 944-963, 2018.

Copyrights

© 2023 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).





Iranian Association of
Electrical and Electronics
Engineers

Journal of Applied Research in Electrical Engineering

E-ISSN: 2783-2864

P-ISSN: 2717-414X

Homepage: <https://jaree.scu.ac.ir/>



Research Article

Improving the Diagnosis of Sudden Cardiac Death by Using Non-Linear Features of the ECG Signal and Hybrid RBF

Afshin Koliji¹ , Sara MihaDoost^{1,*} , Nematollah Ezzati¹ , and Ehsan Mostafapour² 

¹ Department of Electrical Engineering, Faculty of Industrial Technologies, Urmia University of Technology, Urmia, Iran

² Department of Electrical Engineering, Faculty of Electrical and Computer Engineering, Urmia University, Urmia, Iran

* Corresponding Author: a.koliji79@gmail.com

Abstract: Sudden Cardiac Death (SCD) leads to the killing of millions of people worldwide every year. In this article, sudden cardiac death is predicted by utilizing electrocardiogram signal processing. For this purpose, after extracting the signal of heart rate variations from the electrocardiogram signal, temporal and non-linear features have been extracted. In the next step, by applying LDA to the combined feature vector, the feature dimensions are reduced and finally, healthy people and high-risk people are classified through Hybrid-RBF classifiers. The obtained results show that there are features in the signal of heart rate variations related to risk-taking individuals near the occurrence of sudden cardiac death, that completely distinguish them from healthy persons. It has also been shown that from 6 minutes before the occurrence of cardiac death, this increase in the probability of risk is quite evident, so that as we get closer to the occurrence of the accident, the probability of its occurrence also increases, and this is enough time to adopt strategies to prevent it. The simulation results achieved by the data available in the MIT-BIH database prove the ability of the presented methods to achieve accurate diagnosis.

Keywords: Variations in heart rate, sudden cardiac death, linear features, nonlinear features.

Article history

Received 07 March 2023; Revised 29 July 2023; Accepted 13 August 2023; Published online 25 November 2023.

© 2023 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)

How to cite this article

A. Koliji, S. MihaDoost, N. Ezzati, and E. Mostafapour, "Improving the diagnosis of sudden cardiac death by using non-linear features of the ECG signal and hybrid RBF," *J. Appl. Res. Electr. Eng.*, vol. 2, no. 2, pp. 120-126, 2023. DOI: [10.22055/jaree.2023.43231.1068](https://doi.org/10.22055/jaree.2023.43231.1068)



1. INTRODUCTION

Cardiovascular diseases are one of the most common diseases of this century. Among them, one of the biggest reasons of death is cardiac arrhythmias in today's societies. Nowadays, in medicine, various tools and methods have been invented to investigate performance of the heart work such as the analysis of the heart's behaviour using the electrocardiogram signal. Of course, there are other methods to measure the behaviour of the heart, such as angiography, MRI, etc., in all these methods, the goal is to obtain different types of structural and functional information from the heart so that the specialist doctor can diagnose both conditions: heart disease and prediction of its occurrence.

Sudden cardiac death is an unexpected event that occurs due to known and unknown factors in a short period of time (less than an hour) [1, 2]. It is estimated that the death caused by Sudden Cardiac Death (SCD) alone is ten times higher than the death caused by accidents in Europe and the United

States, so that 300,000 people just die annually in the United States due to Sudden Cardiac Death [3, 4]. Usually, sudden cardiac death begins with a severe arrhythmia such as ventricular tachycardia, ventricular flutter or ventricular fibrillation [5, 6]. Fig. 1 shows the ECG signals of a patient with SCD a few seconds before SCD occurs. Statistics show that SCD increases with age. So that the incidence of SCD in young people (less than 30 years old) is 100 times lower than middle-aged or elderly people [7]. Relatively, women are more immune to SCD than men, although genetic factors also play an important role in the occurrence of sudden cardiac death [7].

Genetic mutations increase the risk factors associated with coronary artery diseases. The results of some studies show that African Americans have a higher risk of SCD than whites. In men, there are many factors that increase the risk of SCD, including various cancers (such as prostate cancer, lung cancer), accidents, chronic respiratory diseases, diabetes, and cerebrovascular diseases. In women, some

factors such as lung cancer and cerebrovascular diseases play a more important role than other factors in increasing the risk of SCD [8]. Research shows that nearly 50% of people diagnosed with SCD have not had any heart complications. In this type of people, risk classification is a challenging issue. In addition, 40% of SCD attacks occurred in people who had a history of heart disease with a left ventricular ejection fraction greater than 40%, the remaining 10% of people at risk for SCD are those with structural heart disease and left ventricular ejection fraction. fraction (LVEF) was less than 40%. Therefore, today, risk stratification and treatment tools such as internal cardiac defibrillators have little effect on the overall problem of SCD. It should be noted that a cardiac defibrillator is a battery-operated device that is only slightly larger than a pacemaker. Defibrillators are implanted under the skin, monitor the heart rhythm and correct it if necessary. Calculations based on genetic problems include only 2% of all SCD patients, so considering the high prevalence of SCD in people without a history of heart disease, most methods are focused on identifying and correcting common risk factors. Risk factors for SCD are similar to those for other heart diseases, including smoking, high blood pressure, and diabetes. For example, in Framingham, a US city with a population of 70,000, smoking increased the risk of SCD 2-3 times more than any other factor.

1.1. Previous Works

Studies represent that the QT waveform interval in the ECG signal and heart rate variations are two prominent indicators for assessing the risk of SCD [16, 17]. Despite this measurement, the QT interval is complicated and the results of the review demonstrate that the ability to predict using this method is weak [18]. Studies reveal that analysis based on HRV signal is a strong predictor [19]. Van Hougenhuys et al. [20] have been calculated the standard deviation, the mean value of the RR intervals, as well as the mean value of the standard deviation of the HRV signal. This study proves that the obtained values of the HRV signal in sick people are lower than in young normal people. Although linear methods have been used in HRV analysis, the capability of such methods is not effective in predicting SCD [21, 22]. Nowadays, researchers use non-linear methods to analyze the HRV signal and predict SCD from the HRV signal because these methods provide more information than linear methods [23].

Authors in [24] have been expressed the nature of the ECG signal with the short-term fractional scaling factor and showed that this factor is an independent and suitable factor for predicting SCD in 446 heart attack survivors with left ventricular function assessment less than 35%. In this research, it is claimed that the fluctuation analysis (Detrended DFA α) has the ability to determine the worst case (cardiac death with and without arrhythmia) during 1200 days of follow-up compared to the standard deviation of RR intervals and ventricular problem. In the research carried out in reference [25], it was shown that the short-term fractal scaling properties of the heart rate increase randomly in the elderly. Authors in [26] presented a new nonlinear algorithm using time-independent dimensions of point correlation extracted from the HRV signal. The accuracy of this method is 100% and its sensitivity is 85% for detecting VF/VT. Authors in [22] extracted time and frequency domain parameters from 26 cardiac patients after heart attack (low risk group and high-risk group) and used non-linear and normal entropy methods.

They reported 100% resolution, which was obtained by combining features in all fields. Authors in [21] used time and frequency domain parameters of HRV to predict SCD. They have affirmed that short-term low-frequency power during controlled breathing is a suitable tool for predicting SCD in patients with chronic heart problems. Reference [6] presented a home and personal protection system to predict SCD from ECG signals. They reported an accuracy of 92%. In addition, they used short-term HRV analysis to predict events preceding SCD, using four HRV features including HF, LF, and VHF parameters.

Finally, based on SVM and KNN classification methods in [33], the gained results were about 93% and 95% for SCD detection up to 6 minutes in [26]. In this article, we are trying to apply the Hybrid-RBF classification network to improve the percentages. The rest of this paper is designed as follows: in Section 2 we describe the SCD disease and processing the ECG signal to detect it. Section 3, is for describing the proposed feature extraction and the block diagram of the SCD classification method. Also, the proposed Hybrid RBF classifier is described in Section 3. Section 4, represents the results of SCD detection and Section 5 concludes the results.

2. SCD DIAGNOSING METHOD

In this section, we present the suggested method for diagnosing SCD.

2.1. Generalities

At this step, before starting of cardiac arrest, we divide the ECG signal into one-minute segments in patients who suffer SCD. It should be noted that the onset of cardiac arrest begins with the onset of a ventricular tachycardia. In this article, we can detect SCD before the onset of an attack, according to the assumption used in most articles. The one-minute intervals before SCD are divided into the first one-minute interval before SCD to the sixth one-minute interval before SCD; These intervals are independent of each other. Fig. 1 shows two minutes before the onset of SCD. For normal people, one-minute intervals were selected independently and at different times.

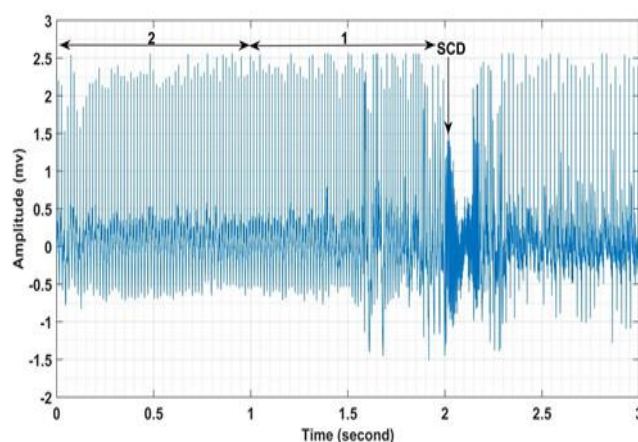


Fig. 1: ECG signal of a patient obtained 2 minutes before the onset of SCD and a few moments after it.

2.1. Noise Elimination

At first, the noise in the ECG signal must be removed. This action is necessary in order to correctly detect the QRS complex. Below are some sources of noise:

2.1.1. City electricity noise

This noise includes the frequency of 51 Hz of city electricity and its harmonics, which can be modelled by the sum of sinusoidal signals. The output of a research explains that the 51 Hz noise in the ECG signal and its harmonics have an amplitude equal to 51% of the peak-to-peak amplitude of the ECG signal [22]. To eliminate the city electricity noise, we use a median filter. Then, after passing the ECG signal through the median filter, we enter it into a band stop-pass filter. Due to the nature of the ECG signal, the coefficients of the band-stop filter are very significant. If coefficients are not selected correctly, it may cause transient states in the ECG signal, which is well explained in reference [26]. Fig. 2 illustrates the city electricity noise and its elimination from ECG signal.

3. FEATURE EXTRACTION AND REDUCTION METHODS

This section as the main part of this article discuss the extraction of features that increase the power of the classification in separating healthy people from risky people. Various features have been used in this article, which are mentioned below.

3.1. Non-Linear Features of the Time-Domain

The features of the time- domain have simple relationships, and due to the non-linear nature of the HRV signal, these types of features are not very useful in distinguishing healthy from unhealthy people, but the combination of these types of features with other non-linear features will increase the resolution of the classifier. Some important features of this area are as following

-Standard deviation of all NN (SDNN) distances:

$$DNN = \sqrt{\frac{1}{N} \sum_{i=1}^N (RR_i - Mean)^2} \quad (1)$$

- The standard deviation of the average NN (SDANN) intervals in all the minutes that exist in are same time interval:

$$SDANN = \sqrt{\frac{1}{N} \sum_{i=1}^N (Mean_i - Mean_{all})^2} \quad (2)$$

In the above equation, N is the total number of one-minute parts of NN intervals in a given part, and $Mean_i$ is the mean of NN intervals in one-minute segment. $Mean_{all}$ also indicates the average of all one-minute segments.

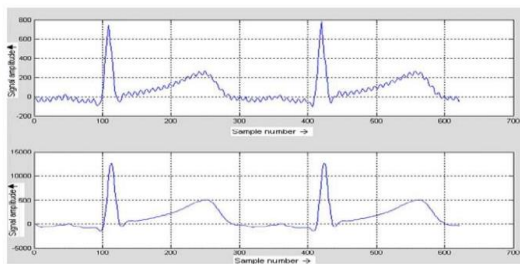


Fig. 2: The consequence of the eliminating city electricity noise from the ECG signal.

- The number of NN intervals whose difference is more than 51 ms:

$$NN50 = \sum_{i=1}^N \{ |RR_{i+1} - RR_i| \} > 50ms \quad (3)$$

- The value of NN50 is divided by the total number of NN distances in a specific segment:

$$SPNN50 = \frac{NN50}{N} \times 100 \quad (4)$$

- root mean square difference of NN intervals:

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (RR_{i+1} - RR_i)^2} \quad (5)$$

- The mean standard deviation of all NN distances for all one-minute segments in a specific part:

$$SDNN_{ind} = \frac{1}{N} \sum_{i=1}^N SDNN_i \quad (6)$$

- Heart rate standard deviation:

$$SDHR = \sqrt{\frac{1}{N} \sum (\text{Mean}_{(i)} - \text{Mean HR}_{ah})^2} \quad (7)$$

where Mean HR is the average heart rate.

After extracting these features, we reduce them using the linear discriminant analysis (LDA) feature selection method which itself improves our results dramatically.

Fig. 3 shows the block diagram which indicate the performance of the offered general method.

3.2. Classifier

In [32], Hybrid-RBF network was presented and presented to outperform the support vector machine (SVM) classifier. Fig. 4 shows the Hybrid RBF classifier. This network uses the k-means algorithm to train its hidden layer, which includes radial functions with decision capabilities. After finding the appropriate centres and variances for each radial function in the hidden layer, it is time to find the weights of the output layer, which we do through the adaptive algorithm. Although, as mentioned, there are other methods for this purpose. It should be noted that since the relationship between the data of the hidden layer and the output is linear, adaptive algorithms can be used to update the weights.

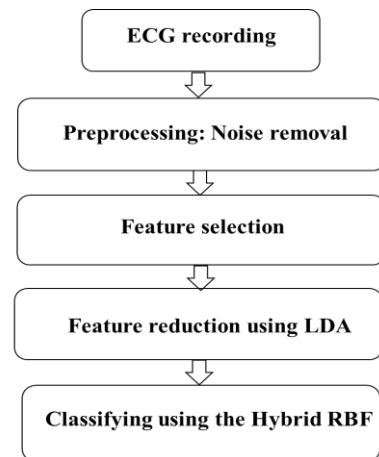


Fig. 3: Block diagram of the general method used in this article to predict sudden cardiac death.

The algorithm of this article is LMS, whose equations are as follows:

$$e(n) = d(n) - W^T(n)\Phi(x_i) \quad (8)$$

$$W(n+1) = W(n) + 2\mu e(n)\Phi(x_i) \quad (9)$$

where $d(n)$ is the desired output value known in the training phase and we estimate it in the testing phase. $W(n)$ is the general vector of weights w_i and $\Phi(x_i)$ defines the general vector of $\phi(x_i)$. $e(n)$ is the error value and μ is the step size, which is assumed constant here. This algorithm adjusts the weights of the values of the radial functions according to the

error between the estimated and the actual values. Therefore, we have the following two steps up to this section:

1) Using the unsupervised algorithm (k-means) to obtain centers and variances of radial functions.

2) Using the LMS algorithm to update the coefficients of the output layer.

As mentioned in [32] the classification procedure of the Hybrid RBF classifier can be at least 30 percent less than the SVM that is used in many SCD classification papers including [21].

Table 1: A comparison of the methods between previous researchers and the time domain and bi-spectrum methods four minutes before SCD and earlier

Features and reference number	Number of features	classifier	Percent correctness of 4 to 6 minutes before SCD
Combined-linear, nonlinear, and TF methods [27]	Linear time domain (5), Linear time domain (4), TF (11) and nonlinear (4)	KNN, MLP	4 min before = 83.96%
nonlinear and TF methods [28]	Nonlinear features (18), Sudden Cardiac Death Index	DT,SVM	4 min before = 92.11% (SVM)
Combined-linear, nonlinear and TF methods [17]	Bispectrum features of HRV signal (6) time-domain features (2)	SVM, KNN With LDA	4 min before= 94.46% (SVM) 5 min before = 91.07% (KNN) 6 min before = 92.77% (KNN)
Non-linear features [26]	Non-linear RQA (13), IncEn (2)	KNN, DT-SVM. Naïve Bayes	6 min before = 95% (KNN, DT-SVM. Naïve Bayes)
[18]	Linear (time, frequency domain), TF domain and, nonlinear methods (PP, DFA)	MLP	95.23% (4 min before) 83.88% (12 min before)
[19]	Linear (time, frequency domain), TF domain, and nonlinear methods (PP, DFA)	MLP, SVM, KNN	95.24% (4 min before) 84.28% (13 min before)
[20]	EEMD, linear (time, frequency domain), TF domain, and nonlinear methods (Ren, FuEn, dispersion entropy, Renyi distribution entropy (RdisEn), and improved multiscale permutation entropy	KNN	94.7% (4 min before) 96.1% (14 min before)
[21]	DWT, Nonlinear methods (Fractal Dimension (FD), DFA Hurst's exponent (H), Sample Entropy, Approximate Entropy, and Correlation Dimension (CD))	KNN, SVM, DT	92.11% (4 min before)
This work	Bispectrum features of HRV signal (6) time-domain features (2)	Hybrid-RBF	4 min before= 96.36% (Hybrid-RBF) 5 min before = 96% (Hybrid-RBF) 6 min before = 96% (Hybrid-RBF)

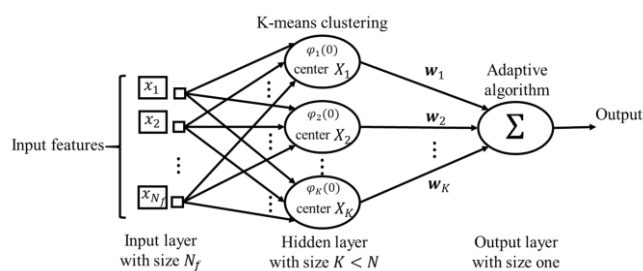


Fig. 4: Order of performance for K-means and adaptive algorithms in RBF network.

4. RESULTS

In this section, we will review and compare the results presented in this article and the previous work. To this end, we use a common database with the rest of the research.

4.1. Article Database

In this article, the database available on the reliable site Physionet [15] and the MIT-BIH database have been utilized. This database is used by many authoritative articles and is very useful for comparison with the work of other researchers in the field of sudden cardiac death. Table 1 shows the comparison of the SCD detection results for various papers and our proposed method.

As we can see, the results are based on two factors, first the number of minutes before SCD and the next one is accuracy. Our experience showed that as we get closer to the happening of the SCD, the detection of it becomes easier and the accuracy rises. For example, the detection of SCD, 4 minutes before its happening is more accurate than detecting it in 6 minutes before SCD. Also, our proposed method works superior than most of the given results in the literature in all the three cases of 4, 5 and 6 minutes detection before SCD.

5. CONCLUSION

In this paper, we introduce two methods to predict sudden cardiac death (SCD) with linear and nonlinear features extracted from heart rate variability (HRV) signal. It is a challenge to predict five and six minutes before the onset of SCD. This method includes four steps: pre-processing, feature extraction, feature reduction and classification. In the first step, QRS complexes are identified from the electrocardiogram (ECG) signal and then the HRV signal is extracted. In the second step, the linear and non-linear features of the HRV signal are obtained. In the next step, these features are reduced to one feature by the linear discriminant analysis (LDA) technique. Finally, the Hybrid-RBF classifier is used to classify the HRV signals. We utilized two databases, the MIT/BIH Sudden Cardiac Death (SCD) database and the Physiobank Normal Sinus Rhythm (NSR) database. In this work, we achieved the prediction of the occurrence of SCD for five and six minutes before SCD with an accuracy of more than 96%.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Afshin Koliji: Methodology, Software, Writing - original draft, Writing - review & editing. **Sara MihanDoost:** Supervision, Validation. **Ehsan Mostafapour:** Supervision, Validation. **Ehsan Mostafapour:** Conceptualization, Roles/Writing - original draft, Writing - review & editing.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy has been completely observed by the authors.

REFERENCES

- [1] J. G. Jeffrey, M. A. Christine, and R. J. Myerburg, "Cardiac arrest and sudden cardiac death," in P. L. MD, *Braunwald's Heart Disease. A Textbook of Cardiovascular Medicine*, Elsevier, 2022, pp. 1349-1386.
- [2] S. S. Chugh, "Sudden cardiac death with apparently normal heart: clinical implications of progress in pathophysiology," *Cardiac Electrophysiology Review*, vol. 5, pp. 394-402, 2001.
- [3] D. Lloyd-Jones et al., "Heart disease and stroke statistics—2010 update A report from the American Heart Association," *Circulation*, vol. 121, pp. e46-e215, 2010.
- [4] N. J. Pagidipati, and T. A. Gaziano, "Estimating deaths from cardiovascular disease: a review of global methodologies of mortality measurement," *Circulation*, vol. 127, pp. 749-756, 2013.
- [5] J. L. Jones, and O. H. Tovar, "The mechanism of defibrillation and cardioversion," in *Proceedings of the IEEE*, vol. 84, pp. 392-403, 1996.
- [6] T.-W. Shen, H.-P. Shen, C.-H. Lin, and Y.-L. Ou, "Detection and prediction of Sudden Cardiac Death (SCD) for personal healthcare," in *2007 29th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society*, 2007, pp. 2575-2578.
- [7] D. P. Zipes et al., "ACC/AHA/ESC 2006 guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: a report of the American College of Cardiology/American Heart Association Task Force and the European Society of Cardiology Committee for Practice Guidelines (Writing Committee to Develop Guidelines for Management of Patients With Ventricular Arrhythmias and the Prevention of Sudden Cardiac Death)," *Journal of the American College of Cardiology*, vol. 48, no. 5, pp. e247-e346, 2006.
- [8] E. C. Stecker et al., "Public health burden of sudden cardiac death in the United States," *Circulation: Arrhythmia and Electrophysiology*, vol. 7, no. 2, pp. 212-217, 2014.
- [9] H. J. Wellens et al., "Risk stratification for sudden cardiac death: current status and challenges for the future," *European heart journal*, vol. 35, pp. 1642-1651, 2014.
- [10] C. M. Albert, M. A. Mittleman, C. U. Chae, I.-M. Lee, C. H. Hennekens, and J. E. Manson, "Triggering of sudden death from cardiac causes by vigorous exertion,"

- New England Journal of Medicine*, vol. 343, pp. 1355-1361, 2000.
- [11] J. D. Kark, S. Goldman, and L. Epstein, "Iraqi missile attacks on Israel: The association of mortality with a life-threatening stressor," *Jama*, vol. 273, pp. 1208-1210, 1995.
- [12] C. M. Albert, J. E. Manson, N. R. Cook, U. A. Ajani, J. M. Gaziano, and C. H. Hennekens, "Moderate alcohol consumption and the risk of sudden cardiac death among US male physicians," *Circulation*, vol. 100, pp. 944-950, 1999.
- [13] C. M. Albert, H. Campos, M. J. Stampfer, P. M. Ridker, J. E. Manson, W. C. Willett, et al., "Blood levels of long-chain n-3 fatty acids and the risk of sudden death," *New England Journal of Medicine*, vol. 346, pp. 1113-1118, 2002.
- [14] C. M. Albert, J. Ma, N. Rifai, M. J. Stampfer, and P. M. Ridker, "Prospective study of C-reactive protein, homocysteine, and plasma lipid levels as predictors of sudden cardiac death," *Circulation*, vol. 105, pp. 2595-2599, 2002.
- [15] <http://physionet.org/>
- [16] V. Houshyarifar, M. C. Amirani, "An approach to predict sudden cardiac death (SCD) using time domain and bispectrum features from HRV signal," *Bio-medical materials and engineering*, vol. 27, pp. 275-285, 2016.
- [17] V. Houshyarifar, M. C. Amirani, "Early detection of sudden cardiac death using Poincaré plots and recurrence plotbased features from HRV signals," *Turkish Journal of Electrical Engineering & Computer Sciences*; vol. 25, no. 2, pp. 1541-53, 2017.
- [18] E. Ebrahimzadeh, M. S. Manuchehri, S. Amoozegar, "A time local subset feature selection for prediction of sudden cardiac death from ECG signal," *Med. Biol. Eng. Comput.*, vol. 56, pp. 1253-1270, 2018.
- [19] E. Ebrahimzadeh et al., "An optimal strategy for prediction of sudden cardiac death through a pioneering feature-selection approach from HRV signal," *Comput. Methods Programs Biomed.*, vol. 169, pp. 19-36, 2019.
- [20] M. Shi et al., "Early detection of sudden cardiac death by using ensemble empirical mode decomposition-based entropy and classical linear features from heart rate variability signals," *Frontiers in Physiology*, vol. 11, 2020.
- [21] U. R. Acharya et al., "An integrated index for detection of sudden cardiac death using discrete wavelet transform and nonlinear features," *Knowl.-Based Syst.*, vol. 83, pp. 149-158, 2015.
- [22] S. Mohd, and R. Jaafar, N. A. Nayan, N. H. Harun, "ECG-based detection and prediction models of sudden cardiac death: Current performances and new perspectives on signal processing techniques." *Journal of Electrocardiology*, vol. 15, no. 15, pp. 110-126, 2019.
- [23] J. R. Velázquez-González et al., "ECG-based identification of sudden cardiac death through sparse representations." *Sensors*, vol. 21, no. 22, article 7666, 2021.
- [24] M. Lewandowski, "A review of the commercially available ECG detection and transmission systems—the fuzzy logic approach in the prevention of sudden cardiac arrest," *Micromachines*, vol. 12, no. 12, article 1489, 2021.
- [25] P. Banerjee, S. Bhattacharjee, K. Dasgupta, and S. Sen. "Performance evaluation of machine learning classifiers for sudden cardiac arrest detection." *Journal of The Institution of Engineers (India): Series B*, pp. 1-7, 2022.
- [26] M. Khazaei, K. Raeesi, A. Goshvarpour, and M. Ahmadzadeh. "Early detection of sudden cardiac death using nonlinear analysis of heart rate variability." *Biocybernetics and Biomedical Engineering*, vol. 38, no. 4, pp.931-940, 2018.
- [27] D. Lai, Y. Zhang, X. Zhang, Y. Su, and M. B. B. Heyat. "An automated strategy for early risk identification of sudden cardiac death by using machine learning approach on measurable arrhythmic risk markers." *IEEE Access*, vol. 7, pp. 94701-94716, 2019.
- [28] U. R. Acharya et al., "An integrated index for detection of sudden cardiac death using discrete wavelet transform and nonlinear features," *Knowledge-Based Systems*, vol. 83, pp. 149158, 2015.
- [29] A. Martín-Yebra, L. Sörnmo, and P. Laguna. "QT interval adaptation to heart rate changes in atrial fibrillation as a predictor of sudden cardiac death," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 10, pp. 3109-3118, 2022.
- [30] A. Parsi, D. O'Loughlin, M. Glavin, and E. Jones, "Prediction of sudden cardiac death in implantable cardioverter defibrillators: A review and comparative study of heart rate variability features." *IEEE Reviews in Biomedical Engineering*, vol. 13, pp. 5-16, 2019.
- [31] D. Lai, Y. Zhang, X. Zhang, Y. Su, and M. B. B. Heyat. "An automated strategy for early risk identification of sudden cardiac death by using machine learning approach on measurable arrhythmic risk markers." *IEEE Access*, vol. 7, pp. 94701-94716, 2019.
- [32] S. Haykin, *Neural networks and learning machines*. Prentice Hall, pp. 230-263, 2008.
- [33] V. Houshyarifar, and M. C. Amirani. "An approach to predict Sudden Cardiac Death (SCD) using time domain and bispectrum features from HRV signal." *Bio-medical materials and engineering*, vol. 27, no. 2-3, pp. 275-285, 2016.

BIOGRAPHY



Afshin Koliji was born in Mohabad, Iran in 1980, he received his B. Sc. In 2001 and currently pursuing his M. Sc. In telecommunications engineering in Urmia University of technology, Urmia, Iran. His main field of research is Signal and image processing.



Sara Mihandoost received the B.S. degree in Electrical Engineering from Razi University of Kermanshah, Iran in 2008 and the M.S. and Ph.D. degrees in Communication Engineering from Iran University of Urmia in 2012 and 2017, respectively. In 2014, she joined the Urmia Graduate Institute. She currently is an assistant professor in Urmia University of technology. Her research interests include texture analysis, biomedical signal processing, pattern recognition and stochastic signal processing.



Nematollah Ezzati received his B.Sc. in Electrical and Electronic Engineering from Shiraz University and M.Sc. degree in Electrical and Electronic Engineering from Amirkabir University of Technology, and Ph.D. degree in Electrical and Electronic Engineering from Amirkabir university of Technology Tehran Iran in 2018. From 2018 he is an assistant professor in the

Department of Electrical Engineering of Urmia University of Technology, Urmia, Iran. His currently research interests are in distributed signal processing, cognitive radio network and radar (SAR & ISAR) design.

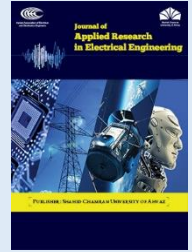


Ehsan Mostafapour was born in west Azarbayjan Province, Urmia, Iran, in 1988. He received his B.S. and the M.S. degrees, in 2010 and 2012, respectively, both in telecommunication engineering. He received his Ph.D. degree from the Department of Electrical Engineering, Urmia University in 2018. His research interests include stochastic and adaptive signal processing. He has published more than 30 papers in the field of adaptive sensor networks and reviewed more papers in the field of signal processing for various journals like ACES, JCE, WPC, IEEE TVT, IEEE Access and IET. Currently, Dr. Mostafapour works as a post-doc researcher for the Urmia University.

Copyrights

© 2023 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).





Research Article

Selection and Tuning Propagation Path Loss Model for Hawassa City, Ethiopia at 1800 MHz Frequency

Tamirat Yenealem^{1,*} , and Robel Getachew² 

¹ Electrical and Computer Engineering, Hawassa University, Hawassa, Ethiopia

² Ethio Telecom, Hawassa, Ethiopia

* Corresponding Author: tamiratyenealem@gmail.com

Abstract: Path loss models estimate the average path loss a signal experiences at a particular distance from a transmitter. However, each type of existing path loss propagation model is designed to predict path loss in a particular environment that may be inaccurate in other different; hence selecting the best path loss model and optimizing it will minimize that inaccuracy. This work presents a comparative analysis of five empirical path loss models, COST- 231, ECC-33, Hata, SUI, and Ericsson model, with respect to the measured data from the 14 selected sites in Hawassa city, Ethiopia at 1800 MHz frequency bands. A drive test methodology was adopted for data collection and Nemo Handy and Nemo Outdoor were used as measuring tools for the test. Error measuring tools such as root mean square error, mean absolute error, standard deviation, and mean absolute percentage error were used to select the terrain type of each site and the path loss model that best fits that site. The results show that not only Hawassa city consists of urban and sub-urban terrains but also ECC-33 and Hata are better estimators for Hawassa urban and sub-urban areas with RMSE of 4.18 and 7.86 respectively. The model tuning using the least square method reduced the RMSE of ECC-33 and Hata to 2.46 and 5.18 respectively. The reduction in RMSE shows that the tuned versions are close to the environment. Hence, using the tuned versions of the selected models will result in good cellular network design and enhance the service quality.

Keywords: Path loss, path loss model, urban, sub-urban, model tuning, least square method.

Article history

Received 26 March 2023; Revised 13 July 2023; Accepted 25 July 2023; Published online 9 December 2023.

© 2023 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)

How to cite this article

T. Yenealem, and R. Getachew, "Selection and tuning propagation path loss model for hawassa city, ethiopia at 1800 MHz frequency," *J. Appl. Res. Electr. Eng.*, vol. 2, no. 2, pp. 127-135, 2023.

DOI: [10.22055/jaree.2023.43085.1067](https://doi.org/10.22055/jaree.2023.43085.1067)



1. INTRODUCTION

Mainly all wireless communications involve the transmission of data in the form of electromagnetic (EM) waves. The transmitted information experiences a path loss, the reduction in power density (attenuation) of an electromagnetic wave or degradation in its signal strength, as it propagates between transmitter and receiver [1, 2]. Propagation of the radio waves especially in urban areas is quite complex because it consists of reflected, Scattering and diffracted waves produced by multipath propagation.

A propagation model which is a set of mathematical expressions, diagrams, and algorithms is used to represent the radio characteristics of a given environment and related signal strength degradations [3]. Radio propagation models are used to describe the relationship between the signal radiated by a

transmitter and the signal received by a receiver as a function of distance and other variables [4].

Generally, propagation models can be put into two categories: empirical propagation and deterministic propagation model [5]. Deterministic path loss models are those that use the laws governing the propagation of electromagnetic waves for the determination of the power of a received signal at a given distance from the transmitter by considering that specific environment; whereas empirical propagation models are mathematical formulations based on data measurement and observation of the given environment. Such propagation models are derived empirically from statistical analysis of large number of field measurements [6].

Direct deployment of the existing empirical models may give rise to high prediction errors due to two main reasons. Firstly, the environment they were initially built for is different from the one where the network is being deployed.

Secondly, some terrains are ambiguous and are difficult to categorize them as open, urban or sub-urban. Thus, this work is aimed at knowing the environment of Hawassa city based on results of the existing path loss models and the actual measured data, select the best path loss models from the existing ones and finally tuning the best models to make it as fit as possible at 1800 MHz frequency band. The fineness of the path loss models and terrain grouping is done based on the results from error measuring tools such as RMSE, MAE, SD and MAPE. The result of this work can be used by any telecom operator to make any decision in Hawassa city and in other cities and towns with approximately similar terrain.

2. EMPIRICAL MODELS

Empirical models are models based on the practical measurement data. They are useful to study the principle behavior of system-level concepts or to enable a rough estimation of the number of required sites in a large area for example in green field planning during a license auction. These models do not require site-specific terrain information. Instead input parameters are e. g. path loss decay exponents, effective antenna heights or average clutter loss factors characterizing the average propagation environment [7]. COST-231, ECC-33, SUI, Hata, and Ericsson models are the concerns of this work and hence discussed as follows.

2.1. COST-231 Model

The COST-231 Hata model is designed to be used in the frequency band from 500 MHz to 2000 MHz, BS antenna height from 30m to 200m and MS antenna height from 1m to 10m. The basic equation for path loss in dB [8, 9]:

$$PL = 46.3 + 33.9 \log_{10} f - 13.82 \log_{10} h_b - a_{hm} + (44.9 - 6.55 \log_{10} h_b) \log_{10} d + C_m \quad (1)$$

where, f is the frequency in MHz, d is the distance between base station and mobile station antennas in km, and h_b is the base station antenna height above ground level in meters. The parameter C_m is defined as 0 dB for suburban or open environments and 3 dB for urban environments. The parameter a_{hm} is defined for urban environments as:

$$a_{hm} = 3.2 (\log_{10} 11.75 h_r)^2 - 4.97 \quad \text{for } f > 400 \text{ MHz} \quad (2)$$

And for suburban or rural (flat) environments,

$$a_{hm} = (1.1 \log_{10} f - 0.7) h_r - (1.56 \log_{10} f - 0.8) \quad (3)$$

where, h_r is the mobile station antenna height above ground level.

2.2. ECC-33 Model

The ECC-33 path loss model is developed by the Electronic Communication (ECC). It is extrapolated from original measurements by Okumura and modified its assumptions [10]. The ECC-33 path loss model is empirical model composed from four terms and not to make double discussion it is going to be discussed in Section 7 [11].

2.3. Stanford University Interim (SUI) Model

The SUI model was developed under the Institute of Electrical and Electronics Engineers (IEEE) 802.16 working group for prediction of path loss in urban, suburban and rural environments. This model is defined for the Multipoint Microwave Distribution System (MMDS) frequency band

which is from 2.5 GHz to 2.7 GHz. In this model, terrains are grouped into terrain type A, B, and C. Type A is associated with maximum path loss and is appropriate for hilly terrain with moderate to heavy foliage densities. Type C is associated with minimum path loss and applies to flat terrain with light trees densities. Type B is associated characterized with either mostly flat terrains with moderate to heavy three density or hilly terrains with light tree densities [12]. The basic path loss equation with correction factors is presented in [10, 13]:

$$PL = A + 10\gamma \log_{10} \left(\frac{d}{d_o} \right) + X_f + X_h + s \quad \text{for } d > d_o \quad (4)$$

where $A = 20 \log_{10} \left(\frac{4\pi d_o}{\lambda} \right)$, $\gamma = a - b h_b + \frac{c}{h_b}$, and d is the distance between the Access Point (AP) and mobile station in meters, $d_o = 100m$ and s is a log normally distributed factor that is used to account for the shadow fading owing to tree and other cluster and has a valued between 8.2 dB and 10.6dB, and the parameter h_b is the base station height above ground in meters and should be between 10 m and 80 m. The constants used for a, b and c are given in Table 1.

The correction factors for the operating frequency and the mobile station antenna height for the model are:

$$X_f = 6 \log_{10} \left(\frac{f}{2000} \right) \quad (5)$$

and

$$X_h = -10.8 \log_{10} \left(\frac{h_r}{2000} \right) \quad \text{for terrain type A and B} \quad (6)$$

$$X_h = -20 \log_{10} \left(\frac{h_r}{2000} \right) \quad \text{for terrain type C} \quad (7)$$

where, f is the frequency in MHz and h_r is the mobile station antenna height above ground in meters.

2.4. Hata Model

The Hata model, also known as the Okumura-Hata model, is one of the most commonly used models for a macro cell environment to predict the radio signal attenuation. The model is valid for quasi-smooth terrain in an urban area. To avoid double discussion, the ranges of the used parameters for this model are given and discussed in under and the equations are explained in Section 7.2 [14].

2.5. Ericsson Model

To predict the path loss, the network planning engineers use software provided by Ericsson Company is called Ericsson model. The model also relies on the reformed Okumura-Hata model so that there is a room to change parameters according to the propagation environment. According to Ericsson model we can define the path loss as [10]:

$$L = a_0 + a_1 \log_{10} d + a_2 \log_{10} h_b + a_3 \log_{10} (h_b) \log_{10} (d) - 3.2 (\log_{10} (11.75 h_r)^2) + g(f) \quad (8)$$

where: $g(f) = 44.49 \log_{10} (f) - 4.78 (\log_{10} (f))^2$, f is the Frequency in (MHz), h_b is the transmission antenna height in (m), h_r is the Receiver antenna height in (m). The default values of these parameters (a_0, a_1, a_2 and a_3) for different terrain are given in Table 2.

Table 1: The parameters of SUI model in different environment [10].

Model parameter	Terrain A	Terrain B	Terrain C
A	4.6	4	3.6
b(m ⁻¹)	0.0075	0.0065	0.005
c(m)	12.6	17.1	20

Table 2: Values of parameters for Ericsson model [10].

Environment	a ₀	a ₁	a ₂	a ₃
Urban	36.2	30.2	12	0.1
Suburban	43.2*	68.93*	12	0.1
Rural	45.95*	100.6*	12	0.1

3. ERROR MEASUREMENT

Error measurement was done to evaluate the closeness of the predicted path loss to the measured path loss. In this work four types of error measurement techniques were used which are Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Standard Deviation (SD) and Mean Absolute Error (MAE). The values from these measurement tools are used to identify the approximate terrain type of a given cell and select the path loss model that nearly suits the cell.

3.1. RMSE

Root mean square error measures the average dispersion of a set of data predicted by a model from an observed data. It is the most used techniques in most papers that are done on performance comparison. In this paper also RMSE is mainly used to compare the calculated and measured path loss. It is calculated using the following formula [15]:

$$RMSE = \sqrt{\frac{1}{n} \sum_1^n (PL_m - PL_p)^2} \tag{9}$$

where, PL_m - is the measured path loss, PL_p - is the calculated path loss and n - is the number of measured data

3.2. MAPE

The mean absolute percentage error (MAPE) is a statistical measure of how accurate a prediction model is. It calculates this accuracy as a percentage, and can be calculated as the average absolute percent error for each time period minus measured values divided by measured values.

$$MAPE = \frac{1}{n} \sum_1^n \left| \frac{PL_m - PL_p}{PL_m} \right| \tag{10}$$

where: PL_m is the actual value and PL_p is the predicted value. MAPE is the most common measure used to predict error, and works best if there are no extremes to the data (and no zeros).

3.3. SD

Standard deviation (SD) measures how much a data is spread out around the mean or average [16]. It is calculated as follows:

$$SD = \sqrt{\frac{\sum_1^n ((PL_m - PL_p) - ME)^2}{n - 1}} \tag{11}$$

where, ME is mean error.

$$MAE = \frac{1}{n} \sum_1^n |PL_m - PL_p| \tag{12}$$

3.4. MAE

The Mean Absolute Error (MAE) is used to measure the closeness of the predicted values to the actually measured values. It is calculated as follows:

4. MODEL OPTIMIZATION ALGORITHM

The main purpose of the work is to select the best path loss models (among the mentioned ones) for Hawassa city and tune the selected model to make it more fit with the actual environment. From the available tuning algorithms, Linear Least Square method (LLSM) is used to tune the selected models.

LLSM is by far the most widely used modeling method. Not only it is the most widely used method but it is also plays a strong underlying role in many other modeling methods, including: nonlinear least squares regression, weighted least squares regression and LOESS. This method is a form of mathematical regression analysis which is used to determine the line of best fit for a set of data, providing a visual demonstration of the relationship between the data points. LLSM is an arithmetic tuning approach where all environmental influences are considered. The key idea of LLSM is to minimize the difference between the measured and predicted data in a way of mean square error function. Therefore, its simplicity and easy implementation was the main reason to choose LLSM for mode optimization process on this case study. The following equation shows the mathematical representation of LLSM:

$$E(a, b, c, \dots) = \sum_1^n [y_i - PR, i(x_i, a, b, c)]^2 \tag{13}$$

where, y_i = measured path loss values at distance x_i , $PR, i = (x_i, a, b, c)$ = model calculated path loss values at distance x_i and a, b and c are the model parameters, n represents number of experiment data.

The Path Loss model is considered optimized at the values of a and b where the equation bring minimum value of $E(a, b, c, \dots)$.

5. METHODOLOGY

Drive Test is used by many telecom industries as the best possible solution to collect signal strength, mobile network latency, voice call KPIs and optimization, it is also adopted in this work for data collection process. For the Drive test measurement process, measurements tools such as Nemo handy, Nemo outdoor, GPS, mobile phones, laptop, ACTIX software for data analysis and a vehicle have been used. There are a total of 44 sites which are installed and covered Hawassa city. At the beginning, planning was done to decide where the measurements should be taken. A Google earth assisted and physical observation planning was done on the entire city so that the best sites which can represent all the morphology types, open areas, suburban areas and urban areas of Hawassa city, can be selected. Doing so 14 sites which nearly represent all morphology types were selected. In this work, a sector is considered as a site because a sector facing to one side of an area and the other to another side of an area actually represents different area types. The selected sites and their absolute location are depicted in Fig. 1 and Table 3, respectively. A sample drive test displayed on google earth is put by Fig. 2.

Table 3: Selected sites.

No.	Site ID	Specific Location	Latitude	Longitude	Antenna Height
1	222078	Baleweld Church	7.0199	38.47249	29
2	222080	New Market	7.0227	38.5008	27
3	222082	Nigist Fura	7.0275	38.4848	35
4	222089	Liz Sefer	7.03561	38.48378	29
5	222096	Behind St. Gebriel church	7.0452	38.482	27
6	222097	Pay Station	7.045912	38.490708	33
7	222101	Chefe Genet Church	7.05038	38.51297	38
8	222102	Beza Collage	7.04801	38.47805	29
9	222105	Tele Branch	7.052475	38.470555	37
10	222110	Old Market	7.0597	38.4744	34
11	222112	Hawassa University (Techno)	7.05313	38.50423	30
12	222116	Debut Ez Condominium	7.07088	38.47975	34
13	222118	Industry Park	7.07464	38.50618	30
14	222120	Dato	7.08375	38.4898	35

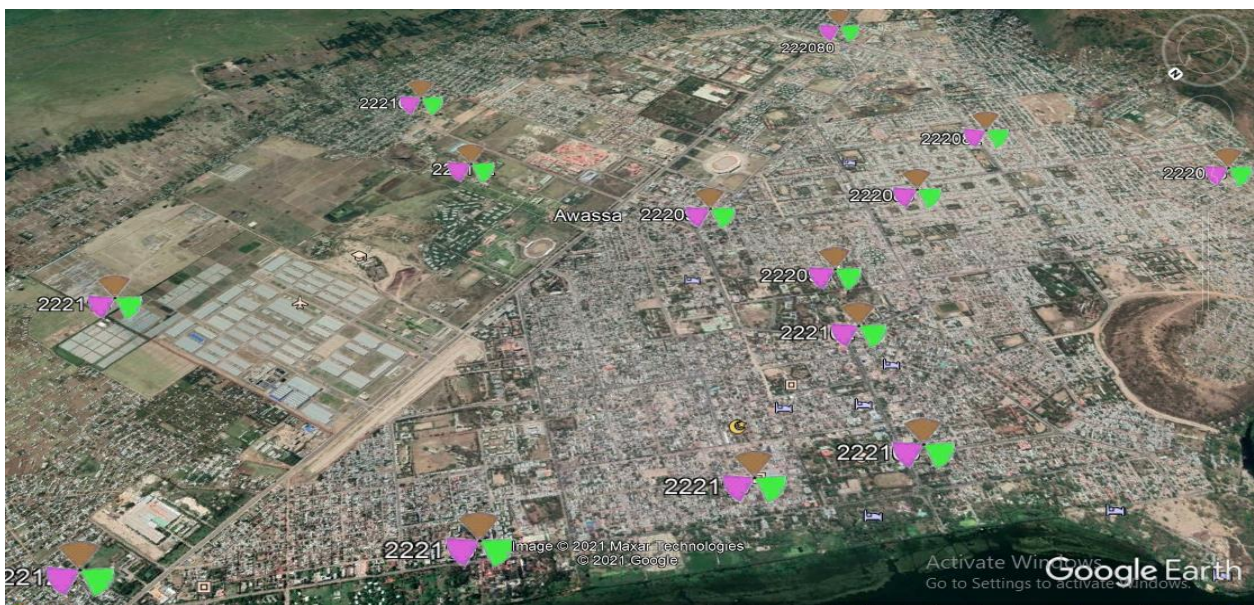


Fig. 1: Selected sites.



Fig. 2: Sample drive test of 1800 MHz displayed on Google earth.

For all sites, received signal strength was measured at a reference distant of 100m from the BS and at continuous intervals of 100m up to 1000m at every 15° interval. For every sector (site) 80 measurement value was taken. In this work, 240 measurement values have been taken for every base station. Based on the data from Ethio-telecom, all sites are configured with a transmission power of 43 dBm and have being used for path loss calculation at 1800 MHz frequency [17]. For 11 sites the measured path loss is depicted below in graph 3.

6. TERRAIN CATEGORIZATION

The terrain identification was done based on the values of RMSE, MAPE, MAE and SD of each site which are calculated from the measured path loss and the predicted path loss by each path loss model. The predicted value of each path loss model was calculated assuming every cell to be urban, sub-urban and rural. Doing so, the selected 14 sites nearest terrain is identified and presented in Table 4.

6.1. Path Loss Model Selection

Based on the presented terrain category, the performance of each path loss model was calculated using RMSE and presented in Table 5 and Table 6.

In the selection process, it should be noted that the smallest the error the better the model is. For urban sites, the one with smallest RMSE value is ECC-33 model and hence selected to be the one that fits best for the cells under this terrain category. For sub-urban sites, the one with smallest RMSE value is Hata model and hence selected to be the one that fits best for the cells under this terrain category. Though these models are relatively good, they can be made better fit for the actual Hawassa city terrain by tuning them as they are developed for a different terrain which is discussed in Section 7.

7. MODEL TUNING

In this section, the relatively best fit path loss models are tuned using LLSM so that they will better fit for Hawassa city at 1800 MHz frequency than their original version. In the model tuning process the average height of 1800 MHz base station antennas in urban areas and Sub Urban areas is rounded to 30m and 28m respectively, and the mobile station

height is 1.5m. Once optimized, it will be compared with the original one which is discussed in the upcoming sections.

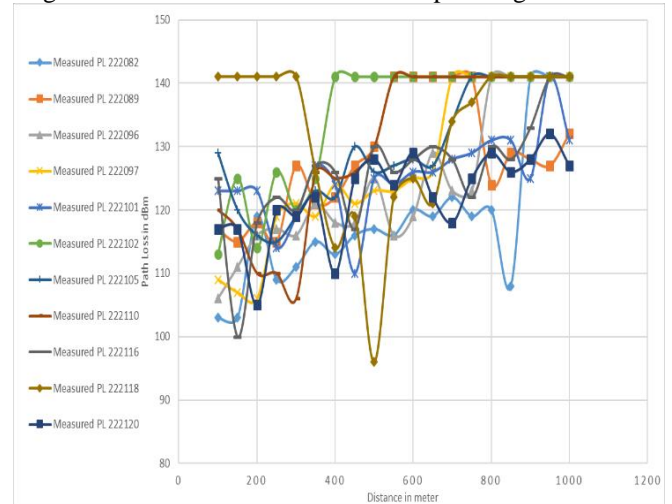


Fig. 3: Measured Path Loss of the eleven Sites at 1800 MHz.

Table 4: Urban area type Average measured path loss performance at 1800 MHz.

	RMSE				
Cost-231	SUI	Ericsson	ECC-33	Hata	
4.26	25.18	4.45	4.18	4.96	

Table 5: Sub-Urban area type Average measured path loss performance at 1800 MHz.

	RMSE				
Cost-231	SUI	Ericsson	ECC-33	Hata	
7.95	31.75	12.12	24.15	7.86	

7.1. Tuning ECC-33 for Urban Area Types

$$PL_{ECC-33} = A_{fS} + A_{bm} - G_b - G_r \tag{14}$$

where,

$$A_{fS} = 92.4 + 20 \log_{10} d + 20 \log_{10} f$$

$$A_{bm} = 20.41 + 9.83 \log_{10} d + 7.894 \log_{10} f + 9.56[\log_{10} f]^2$$

$$G_b = \log_{10} \left(\frac{h_b}{200} \right) \{ 13.958 + 5.8[\log_{10} d]^2 \}$$

$$G_r = 0.759h_r - 1.862$$

Table 6: 1800 MHz Sites and the area type they are categorized.

No.	Site ID	Latitude	Longitude	Antenna Height	Best Fit Model	Area Type
1	222082	7.0199	38.47249	35	ECC-33	Urban
2	222089	7.0227	38.5008	29	ECC-33	Urban
3	222096	7.0275	38.4848	27	ECC-33	Urban
4	222097	7.03561	38.48378	33	ECC-33	Urban
5	222101	7.0452	38.482	38	ECC-33	Urban
6	222102	7.045912	38.490708	29	Ericsson	Urban
7	222105	7.05038	38.51297	37	ECC-33	Urban
8	222110	7.04801	38.47805	34	ECC-33	Urban
9	222116	7.052475	38.470555	34	ECC-33	Urban
10	222118	7.0597	38.4744	30	ECC-33	Urban
11	222120	7.05313	38.50423	35	ECC-33	Urban
12	222078	7.07088	38.47975	29	Hata	Sub-Urban
13	222080	7.07464	38.50618	27	Cost-231	Sub-Urban
14	222112	7.08375	38.4898	30	Hata	Sub-Urban

where, f is the frequency in GHz, d is the distance between base station and mobile antenna in km, h_b is the base station antenna height in meters and h_r is the mobile antenna height in meters.

We can rewrite equation (1) as:

$$PL_{ECC-33} = c_1 + c_2 x_i + c_3 (x_i)^2 \quad (15)$$

where

$$x_i = \log_{10} d_i,$$

$$c_1 = 92.4 + 20 \log_{10} f + 20.41 + 7.894 \log_{10} f + 9.56[\log_{10} f]^2 - 13.958 \log_{10} \left(\frac{h_b}{200}\right) - 0.759 h_r + 1.862$$

$$c_2 = 9.83 + 20,$$

$$c_3 = -5.8 \log_{10} \left(\frac{h_b}{200}\right)$$

For this work A_{bm} is selected for the overall optimization/tuning process.

$$A_{bm} = K_1 + K_2 \log_{10} d + 7.894 \log_{10} f + 9.56[\log_{10} f]^2 \quad (16)$$

i.e., we should find K_1 and K_2 so that the optimized model, equation b fit better than the original one.

$$PL_{Tuned_ECC-33} = \widetilde{C}_1 + \widetilde{C}_2 2x_i + c_3 (x_i)^2 \quad (17)$$

Where

$$x_i = \log_{10} d_i$$

$$\widetilde{C}_1 = 94.262 + k1 + 27.894 \log_{10} f + 9.56[\log_{10} f]^2 - 13.958 \log_{10} \left(\frac{h_b}{200}\right) - 0.759 h_r$$

$$\widetilde{C}_2 = k2 + 20$$

where $k1$ is the tuned offset value of A_{bm} and $k2$ is the tuned slope value of the d -term in A_{bm} .

Then, the minimum error can be calculated by setting the following error function's partial derivative with respect to \widetilde{C}_1 and \widetilde{C}_2 to zero.

$$E(\widetilde{C}_1, \widetilde{C}_2) = \sum_1^n [y_i - PL_{ECC}(\widetilde{C}_1, \widetilde{C}_2)]^2 \quad (18)$$

$$\text{This means } \frac{\partial E}{\partial \widetilde{C}_1} = 0 \text{ and } \frac{\partial E}{\partial \widetilde{C}_2} = 0.$$

Solving them we get the following equations for \widetilde{C}_1 and \widetilde{C}_2 and values are given in [10].

$$\widetilde{C}_1 = \frac{n \sum_1^n y_i \sum_1^n (x_i)^2 - c_3 (\sum_1^n x_i)^2 \sum_1^n (x_i)^2 - n \sum_1^n x_i \sum_1^n y_i x_i + nc_3 \sum_1^n x_i \sum_1^n (x_i)^3 - nc_3 (\sum_1^n (x_i)^2)^2 + c_3 \sum_1^n (x_i)^2 (\sum_1^n x_i)^2}{n^2 \sum_1^n (x_i)^2 - n (\sum_1^n x_i)^2} \quad (19)$$

$$\widetilde{C}_2 = \frac{c_3 \sum_1^n x_i \sum_1^n (x_i)^2 + n \sum_1^n y_i x_i - nc_3 \sum_1^n (x_i)^3 - \sum_1^n y_i \sum_1^n x_i}{n \sum_1^n (x_i)^2 - (\sum_1^n x_i)^2} \quad (20)$$

Substituting the values into (19) and (20), the tuned coefficients of A_{bm} ($k1$ and $k2$) and the calibrated parameters

of the tuned ECC-33 model (\widetilde{C}_1 and \widetilde{C}_2) for urban terrains of hawassa city are found and presented in Table 7.

$$\widetilde{C}_1 = 117.469 + 27.894 \log_{10} f + 9.56[\log_{10} f]^2 - 13.958 \log_{10} \left(\frac{h_b}{200}\right) - 0.759 h_r \quad (21)$$

Then, the calibrated ECC-33 model equation for urban area types will be:

$$PL_{ECC-33} = A_{fs} + \widetilde{A}_{bm} - G_b - G_r \quad (22)$$

where,

$$A_{fs} = 92.4 + 20 \log_{10} d + 20 \log_{10} f$$

$$\widetilde{A}_{bm} = 23.207 + 8.187 \log_{10} d + 7.894 \log_{10} f + 9.56[\log_{10} f]^2$$

$$G_b = \log_{10} \left(\frac{h_b}{200}\right) \{13.958 + 5.8[\log_{10} d]^2\}$$

$$G_r = 0.759 h_r - 1.862$$

The performance of tuned version of ECC-33 model is presented in through RMSE, MAPE, MAE and SD.

The ECC-33 path loss model tuning reduces the RMSE value from 4.18 to 2.46 dB which is a visible improvement in terrain representation. Fig. 4 shows the measured average path loss, the predicted path loss by ECC-33 model and the predicted path loss by the tuned ECC-33 model. It can be clearly observed that the tuned ECC-33 model better fits measured average path loss.

7.2. Tuning Hata Model for suburban area types

The model is give below based on the assumption that Hawassa city is under medium city, and the terrain type under consideration is sub-urban.

$$PL_{Hata} = 40.9 + 33.9 \log_{10} f - 13.82 \log_{10} h_b - a_{hm} + (44.9 - 6.55 \log_{10} h_b) \log_{10} d - 2 [\log_{10} \left(\frac{f}{28}\right)]^2 \quad (23)$$

where,

$$a_{hm} = (1.1 \log_{10} f - 0.7) h_m - (1.56 \log_{10} f - 0.8)$$

The above equation can be rewritten as:

$$P_{LHata} = E_0 + E_s + B_s \log_{10} d \quad (24)$$

where, $E_0 = 40.9$, $E_s = 33.9 \log_{10} f - 13.82 \log_{10} h_b - a_{hm} - 2 [\log_{10} \left(\frac{f}{28}\right)]^2$, and $B_s = 44.9 - 6.55 (\log_{10} h_b)$

Table 7: Performance of Tuned ECC-33 after for urban area.

RMSE	MAPE	MAE	SD
2.46	1.54	2.04	1.53

Table 8: Tuned parameters of ECC-33 model for urban areas

\widetilde{C}_1	Tuned slope value of A_{bm} , \widetilde{C}_2	Tuned offset value of A_{bm} , $K1$	$K2$
Eq. (21)	28.187	23.207	8.187

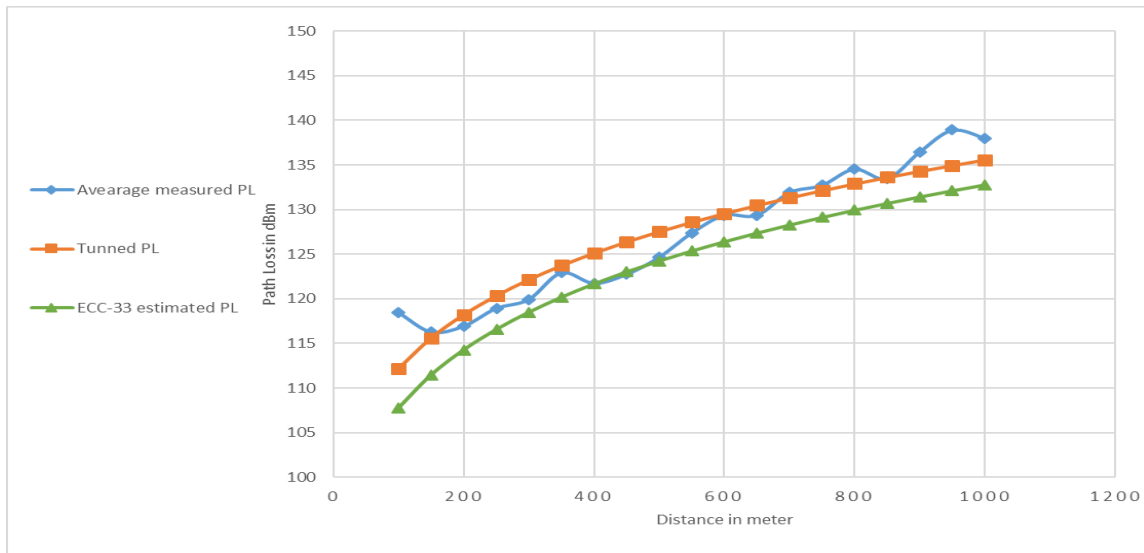


Fig. 4: Urban area Calibrated ECC-33 model Vs measured PL and original ECC-33 (1800).

where, E_0 is Initial offset parameter, E_s is Initial system design parameter, and B_s is the slope of the model curve. We can then write the total Hata path loss as:

$$P_{LHata} = a + bx \tag{25}$$

where $a = E_0 + E_s$, $b = B_s$ and $x = \log_{10} d$, a simplified logarithmic base.

Having that existing model, we need to make it more predictive by tuning the coefficients. That is, we need to have new coefficients \widetilde{E}_0 and \widetilde{B}_s for a new path loss equation that will better fit the actual measured data and be better representation of the actual Hawassa city suburban terrains.

$$P_{LTuned_Hata} = \widetilde{E}_0 + E_s + \widetilde{B}_s B_s \log_{10} d \tag{26}$$

The above equation can be written as:

$$P_{LTuned_Hata} = \hat{a} + \hat{b}x \tag{27}$$

Now let we have the values of \hat{a} and \hat{b} using Linear Least Square method, i.e. the minimum error can be calculated by setting the partial differentials of the following error function to zero.

$$E(\hat{a}, \hat{b}) = \sum_{i=1}^n [y_i - P_{LHata}(x_i, \hat{a}, \hat{b})]^2 \tag{28}$$

This means $\frac{\partial E}{\partial \hat{a}} = 0$ and $\frac{\partial E}{\partial \hat{b}} = 0$. Doing so we get:

$$\hat{a} = \frac{\sum_{i=1}^n (x_i)^2 \cdot \sum_{i=1}^n y_i - \sum_{i=1}^n y_i x_i \cdot \sum_{i=1}^n x_i}{n \sum_{i=1}^n (x_i)^2 - (\sum_{i=1}^n x_i)^2} \tag{29}$$

$$\hat{b} = \frac{n \sum_{i=1}^n y_i x_i - \sum_{i=1}^n y_i \cdot \sum_{i=1}^n x_i}{n \sum_{i=1}^n (x_i)^2 - (\sum_{i=1}^n x_i)^2} \tag{30}$$

Then we find the tuned Hata path loss model by substituting the tuned parameters \hat{a} and \hat{b} into the original Hata pathloss model that we begin the process with

$$\widetilde{E}_0 = \hat{a} - E_s \tag{31}$$

$$\widetilde{B}_s = \frac{\hat{b}}{B_s} = \frac{\hat{b}}{44.9 - 6.55 (\log_{10} h_b)} \tag{32}$$

where, a_{hm} is unchanged.

Evaluating equations (29)-(32), the tuned coefficients of $A_{bm}(\hat{a}$ and $\hat{b})$ and the calibrated parameters of the tuned Hata model (\widetilde{E}_0 and \widetilde{B}_s) for urban terrains of hawassa city are found and presented in Table 9. Finally, we get the tuned Hata path loss equation for the Hawassa city Sub-Urban area types will be:

$$P_{LHata} = 46.363 + 33.9 \log_{10} f - 13.82 \log_{10} h_b - a_{hm} + (43.06 - 6.28 \log_{10} h_b) \log_{10} d - 2 [\log_{10} (\frac{f}{28})]^2 \tag{33}$$

The performance of tuned version of Hata model is presented in Table 10 using RMSE, MAPE, MAE and SD. The Hata path loss model tuning resulted in reduction of the RMSE value from 7.86 to 5.18 dB which is a visible improvement in terrain representation. Fig. 5 shows the measured average path loss, the predicted path loss by Hata model and the predicted path loss by the tuned Hata model. It can be clearly observed that the tuned Hata model better fits measured average path loss and hence nearly represents the actual Hawassa city sub-urban terrain.

Table 9: Tuned parameters of Hata model for Sub-Urban areas.

Tuned parameter \hat{a}	Tuned parameter \hat{b}	E_s	B_s
130.135	33.969	83.772	35.421
E_0	\widetilde{E}_0	Slope Correction factor, \widetilde{B}_s	
40.9	46.363	0.959	

Table 10: Performance of tuned Hata model for Sub-Urban area.

RMSE	MAPE	MAE	SD
5.18	3.63	4.63	2.76

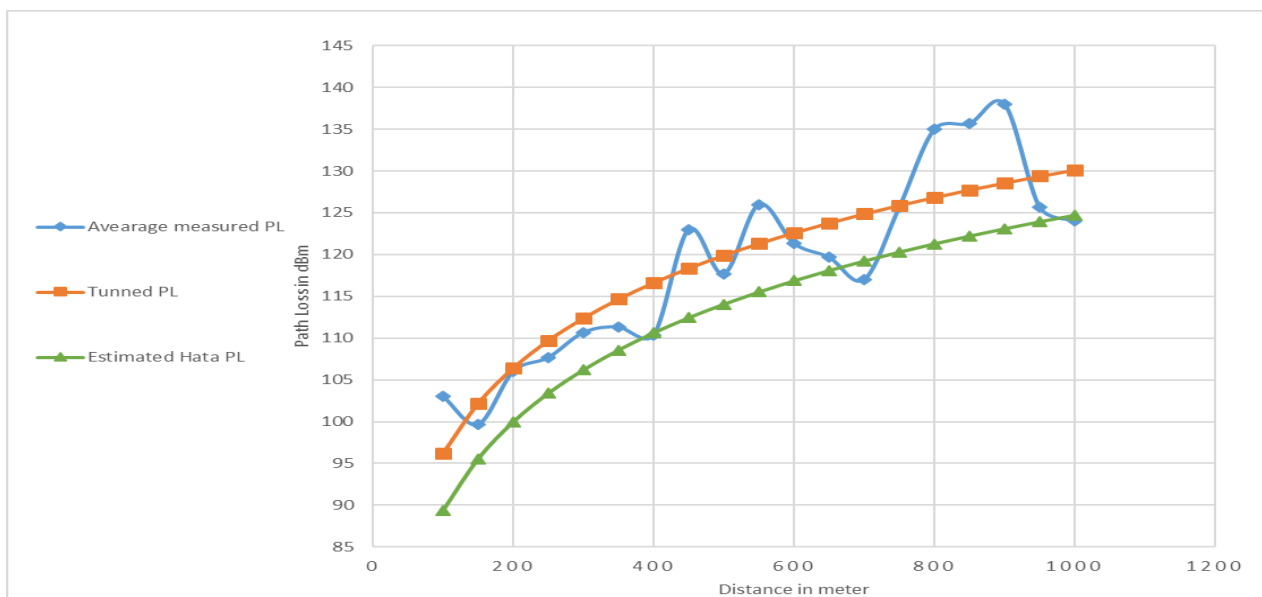


Fig. 5. Sub-Urban area calibrated Hata model vs. measured PL and original Hata (1800).

8. CONCLUSION

This work categorizes the terrain type of Hawassa city at 1800 MHz, selects the best path loss models among the five models, and optimizes the selected ones using Least Square Algorithm. Among the total 40 sites of Hawassa city, 14 cities that nearly represent the entire city terrain were selected. Among them, 11 sites were found to be urban for which ECC-33 with 4.18dB RMSE was the best. Tuning this model, the RMSE value was reduced to 2.46dB. The rest were suburban for which Hata model outperforms with 7.86dB RMSE. Tuning this model, the RMSE value was reduced to 5.18dB.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Tamirat Yenealem: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Roles/Writing - original draft, Writing - review & editing. **Robel Getachew:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Roles/Writing - original draft, Writing - review & editing.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy has been completely observed by the authors.

REFERENCES

- [1] J. Andrusenko, J. Burbank, and J. Ward, "Modeling and simulation for RF propagation," The Johns Hopkins University Design & Developers Fourm IEEE Globecom, 2009. [Online]. Available: https://www.stud.usv.ro/NACRC/NACRC/P1/propagat ionM_S.pdf.
- [2] O. N. Anthony, and O. O. Raphael, "Characterization of signal attenuation using pathloss exponent in South-South Nigeria," *International Journal of Emerging Trends & Technology in Computer Science*, vol. 3, no. 3, pp. 100-104, 2014.
- [3] A. Neskovic, N. Neskovic, and G. Paunovic, "Modern approaches in modeling of mobile radio systems propagation environment," *IEEE Communications Surveys*, vol. 3, no. 3, pp. 2-12, 2000.
- [4] R. Struzak, "Radio wave propagation basics," ICTP-ITU-URSI School on Wireless Networking for Development, 2006. [Online]. Available: http://wireless.ictp.it/school_2006/lectures/Struzak/RadioPropBasics-ebook.pdf
- [5] J. Wen, Y. Zhang, G. Yang, Z. He, and W. Zhang, "Path loss prediction based on machine learning methods for aircraft cabin environments," *IEEE Access*, vol. 7, pp. 159251-159261, 2019.
- [6] A. Hrovat, T. Javornik, S. Plevel, R. Novak, T. Celcer, and I. Ozimek, "Comparison of WiMAX field measurements and empirical path loss model in urban and suburban environment," in *WSEAS International Conference on Communications*, 2006.
- [7] S. R. Theodore, *Wireless communications: principles and practice*. 2nd ed.: Prentice Hall, 2001.
- [8] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, "Field strength and its variability in VHF and UHF land-mobile radio service," *Rev. Elec. Commun. Lab*, vol. 16, pp. 825-873, 1968.
- [9] I. Simi, I. Stani, and B. ZIRNI, "Minimax LS algorithm for automatic propagation model tuning," in *Proceeding of the 9th Telecommunications Forum (TELFOR 2001)*, Belgrade, 2001.
- [10] S. R. Saunders, and A. Aragón-Zavala. *Antennas and propagation for wireless communication systems*. John Wiley & Sons, 2007.

- [11] K. Pahlavan, and P. Krishnamurthy, *Principles of wireless networks*. Prentice Hall PTR, 2001.
- [12] E. Obot, M. Onuu, J. Arikpo, and C. Nwosu, "A comparative study of path loss models in the ultra-high frequency band in a vegetated environment," *FUNAI Journal of Science & Technology*, vol. 6, no. 1, pp. 91-104, 2023.
- [13] A. Zreikat, and M. Dordevic, "Performance analysis of path loss prediction models in wireless mobile networks in different propagation environments, in *3rd World Congress on Electrical Engineering and Computer Systems and Science*, 2017.
- [14] S. Islam, A. Mahmud, J. Uddin, and P. Podder, "An analytical analysis of path loss models for mobile cellular wireless communications," *International Journal on Cybernetics & Informatics*, vol. 9, 06/01 2020.
- [15] J. D. Gadze, K. Agyekum, S. J. Nuagah, and A. E. Ampoma, "Improved propagation models for LTE path loss prediction in urban & suburban Ghana," *ArXiv*, vol. abs/2001.05227, 2019.
- [16] T. T. Oladimeji, P. Kumar, and M. K. Elmezughi, "Performance analysis of improved path loss models for millimeter-wave wireless network channels at 28 GHz and 38 GHz," *PLOS ONE*, vol. 18, no. 3, 2023.
- [17] Ethio telecom, "Ericsson Ethiopia Physical data 10.38," 2015.

BIOGRAPHY



Tamirat Yenealem was born on Sep 1, 1983. He received the B.Sc. degree from Adama Science and Technology University in 2013, the M.Sc. degree from Addis Ababa University, Ethiopia in 2017, all in Electrical engineering (Communication Engineering). His research interests are digital communication system optimization, channel modeling, antenna design for specific application, Machine learning and artificial intelligence.



Robel Getachew was born on Nov 1, 1983. He received the B.Sc. degree from Addis Ababa University in 2013, the M.Sc. degree from the Hawassa University, Ethiopia in 2021, all in Electrical engineering. His research interests are digital communication system optimization, channel modeling and antenna design for specific application.

Copyrights

© 2023 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).





Iranian Association of
Electrical and Electronics
Engineers

Journal of Applied Research in Electrical Engineering

E-ISSN: 2783-2864

P-ISSN: 2717-414X

Homepage: <https://jaree.scu.ac.ir/>



Research Article

Different Types of Distributed Optimal Leader-Follower Consensus Protocol Design for a Class of High-Order Multi-Agent Systems

Farideh Azadmanesh , and Reza Ghasemi 

Department of Electrical Engineering, University of Qom, Qom, Iran

* Corresponding Author: Azadmanesh.2012@yahoo.com

Abstract: Different types of optimal leader-follower consensus of high-order multi-agent systems (MAS) under fixed, connected, and directed communication topology are presented in this paper. The dynamics of each agent including the followers and their corresponding leader is a linear high order system. First, the Linear Quadratic Regulator (LQR) problem is discussed to achieve the optimal consensus for high-order linear MAS with a guaranteed predefined phase and gain margin. Then stochastic leader-follower consensus problem of MAS in the presence of the Gaussian noise is designed. To tackle these problems, a set of fixed distributed control laws for each follower agent is designed, based on algebraic graph theory. Simulation results indicate the effectiveness of the proposed method and display the consensus in both cases via distributed control laws.

Keywords: LQR Controller, LQG Controller, multi-agent systems, leader-follower consensus, fixed topology.

Article history

Received 29 July 2022; Revised 27 August 2023; Accepted 03 October 2023; Published online 29 December 2023.

© 2023 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)

How to cite this article

F. Azadmanesh, and R. Ghasemi, "different types of distributed optimal leader-follower consensus protocol design for a class of high-order multi-agent systems," *J. Appl. Res. Electr. Eng.*, vol. 2, no. 2, pp. 136-144, 2023.

DOI: [10.22055/jaree.2023.41492.1064](https://doi.org/10.22055/jaree.2023.41492.1064)



1. INTRODUCTION

Simulation has become an indispensable tool for researchers to explore systems without having recourse to real experiments. Depending on the characteristics of the modeled system, the methods used to represent the system may vary. Multi-agent systems are, thus, often used to model and simulate complex systems. Control and optimization for multi-agent systems have drawn increasing interest in the past decade because they can be used to model practical examples in many fields, such as physical and information technology, economics, engineering, and social. In multi-agent systems, there are five types of problems: 1- Consensus 2- Containment 3- Tracking 4- Connectivity 5- Formation 6- Fault detection. In a consensus between agents, all agents try to achieve a constant value. This constant value is the average of the initial conditions of the agents. That is why it is called the average consensus. In such cases, the average consensus error is written and is tried to stably reduce this error to zero.

A. Bazaei et al. present a Linear Quadratic Gaussian (LQG) method to improve the tracking control performance of the constant-linear-velocity spiral reference [1]. This method is generally more robust compared to the inversion

control technique. In [2] an optimal control strategy based on LQR controller is proposed for a group of agents to maintain formations while moving towards the destination. X. Li et al. investigate the distributed suboptimal LQR controller problem for continuous-time multi-agent systems [3]. In this reference, the distributed controllers are designed based on the system's topological structure such that each subsystem can use the information of its state and its neighbors. A. Iftar et al. study present a Linear Quadratic Gaussian/Loop Transfer Recovery (LQG/LTR) design methodology for decentralized control systems [4]. In [5] upper bounds on the norm of the multiplicative error matrix are calculated. They account for the interactions between the subsystems. F. Zhang et al. study the decentralized optimal control of discrete-time system with input delay and propose decentralized and centralized optimal controllers by the optimal tracking control use the LQR problem with delay. In [6] a Leveraging nonlinear model predictive control and a linear quadratic regulator (LQR) is developed as new methods for the active control of the spatial distribution of several such buoyancy actuated agents. Thus, in demonstrated a novel application of feedback control theory with LQR to an emerging real-world application in multi-agent systems. In [7] Lebao Li et al. present some common control algorithms of quadrotor

unmanned aerial vehicle such as LQR, and analyse their merits and drawbacks. In addition, it is discussed that because of the limitations of single control algorithms, hybrid control schemes should be used to get the best performance. S. L. Nguyen et al. investigate the LQG game with a major player and a large number of minor players with mean field coupling [8].

In [9] A. M. Ferreira et al. present a self-tuning LQG/LTR method to design a controller for Thyristor Controlled Series Capacitors (TCSCs) with the objective of damping electromechanical power system oscillations. The optimal control problem in continuous space and time for collaborative multi-agent systems is investigated by W. Wiegerinck et al. [10]. In this method the agents in a stochastic environment have to distribute themselves over a number of targets. In [11] M. Nourian et al. study the large population leader-follower multi-agent systems where the agents have linear dynamics and are coupled via their quadratic cost functions. A mean field LQG stochastic control theory is presented for the large population dynamic game problem. Focusing on multi-robot navigation and collision avoidance applications, H. V. Henderson et al. propose a method to reduce the decentralized partially observed Markov decision process using the collection of decentralized LQG controllers for agents, thus, maximizing the joint performance of the team [12]. In [13] the problem of distributed linear-quadratic optimal control with agent-specified differential privacy requirements using the LQG problem has been studied. In [14] investigate the finite-time stability of a linear discrete time system with time-varying delays and its applications to the consensus problem of multi-agent system. In [15] using a control method similar to the LQR, the state-derivative and output-derivative feedbacks are derived for linear time-invariant systems. M. Rafiei sakhaei et al. provide a complete derivation for LQR optimal controllers and the optimal value function using basic principles from variational calculus [16]. As opposed to alternatives, the derivation does not rely on the Hamilton-Jacobi-Bellman equations, Pontryagin's maximum principle, or the Euler-Lagrange equations. It provides a different perspective of how and why key quantities such as the adjoint variable and Riccati equation show up in optimal control computations and their connection to the optimal value function. In [17] a linear quadratic optimal hierarchical control problem is studied in which by using semi-groups the Kronecker product is specified for large-scale dynamical systems. The systems are modeled by an interconnected system under multi-scale information exchange networks and the cost is minimized using an algebraic approach. In [18] both continuous and discrete time consensus problems are studied for multi-agent systems with linear time-invariant agent dynamics over randomly switching topologies. In addition, the effect of Markovian switching topologies and random link failures on consensus are revealed. In [19] a problem of H_∞ consensus for nonlinear multi-agent systems with time-delay is investigated. A dynamic output feedback protocol is designed such that the multi-agent system reaches consensus in mean square with a prescribed H_∞ performance level. In [20] a class of consensus protocols are presented for directed networks of agents with general linear dynamics and synchronous intermittent information. In addition, some analytical results on consensus tracking of multi-agent

systems with switching directed communication topologies are investigated. X. Wu et al study the distributed exponential consensus of delayed multi-agent systems with nonlinear dynamics under asynchronous switching [21]. L. D. Alvergue et al. investigate the output consensus control for continuous-time heterogeneous MASs, aimed at synchronizing all the agent's output to the desired trajectory generated by a reference model [22]. The controller synthesis is based on H_∞ loop shaping and LQG/LTR methods, thereby the local optimality and stability robustness is ensured. In [23] the consensus control problems are reviewed along with the recent progress on stochastic MASs and the latest results on consensus analysis and protocol design issues for MASs are presented. A new pre-tuning multivariable PID (Proportional Integral Derivative) controllers' method for quadrotors has been investigated by R. Guardeno et al. [24]. A procedure based on LQR theory is proposed for attitude and altitude control. In this procedure a considerable simplification is supposed for the design problem, since only one pre-tuning parameter is used. The results show that the design method proposed for multivariable PID controllers is robust in the face of uncertainties and external disturbances acting at plant input. In [25] the consensus protocol for linear multi-agent systems with communication noises are studied. Each agent is allowed to have its own time-varying gain to attenuate the effect of communication noises. It is proved that if all noise-attenuation gains are infinitesimal of the same order, then the mean square leader-following consensus can be reached. Furthermore, the convergence rate of the multi-agent system and the steady-state performance and the transient performance are investigated. In [26] has been investigated the control of continuous-time linear Gaussian systems over additive white noise wireless fading channels subject to capacity constraints. It is shown that a separation principle holds between the design of the communication and the control subsystems, implying that the controller that would be optimal in the absence of the communication channel is also optimal for the problem of controlling the system over the communication channel. In [27] the problem of reduced order LQG optimization is investigated in a finite horizon, linear time-varying system setting. The authors in this reference provide first-order necessary conditions for local optimality in the parameter space using four coupled matrix differential equations.

This paper is the extension of reference [33] that was published by these authors and was selected as a top paper in AREE2021 conference. In this paper a method is proposed in which it is assumed that the communication topology between the leader and its neighbors depends upon bounded and time-invariant functions. To the best of our knowledge, this idea has not been investigated so far especially for agents with high order dynamics. In this study, the LQR and LQG methodology based on the Riccati equation and algebraic graph theory is employed to show the convergence of followers to leader optimally, while the communication link between the leader and its neighbors are fixed over time.

The paper is organized as follows: In Section 2 some notations on graph theory and Kronecker product are phrased. Then, in Section 3 system dynamics and basic mathematical foundation and assumptions are discussed. In Section 4, the proposed LQR and LQG controllers are designed. Section 5 presents the simulation results and comparing controllers, and

in Section 6, the paper is concluded and the future research work is proposed.

2. GRAPH THEORY AND KRONECKER PRODUCT

A directed graph is denoted as $G = (V, E)$ where $V = \{1, 2, \dots, N\}$ is a finite and non-empty set of nodes (each node denotes the follower and there is N followers for $i = \{1, 2, \dots, N\}$ and also $E \subset V \times V$ is a set of edges, each edge denotes an ordered pair of nodes). An edge (v_i, v_j) in an undirected graph shows that the agent i can access the information of the agent j and it means that the agent j is the neighbor of agent i . An adjacency matrix is $A_a = [a_{ij}] \in R^{N \times N}$. Moreover, it is assumed that $a_{ii} = 0$ for $i = \{1, 2, \dots, N\}$. The set of neighbors of agent i is denoted by $N_i = \{j | (v_i, v_j) \in E\}$, Define the degree matrix as $D = \text{diag}\{d_1, \dots, d_N\}$ with $d_i = \sum_{j \in N_i} a_{ij}$. The symmetric Laplacian matrix corresponding to the directed graph G is defined as follows:

$$L = (D - A_a) \in R^{N \times N} \quad (1)$$

The leader agent is represented by vertex L and information is exchanged between the leader and the followers that are the neighbors of the leader [28]. A tool that is very useful in modeling and manipulating equations governing group motion is the Kronecker product \otimes [29]. Kronecker product is also known as tensor product or direct product. Suppose C is the field of mixed numbers and $C^{m \times n}$ is a set of matrices containing m rows and n columns with mixed elements. For any matrices, $A = [a_{ij}] \in C^{m \times n}$ and $B \in C^{p \times q}$ their Kronecker product denoted as $A \otimes B \in C^{mp \times nq}$, is defined by:

$$A \otimes B = [a_{ij}B] = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix} \quad (2)$$

For example, if $\dot{x} = Ax_i$ represents the dynamics of a single agent, the dynamics of N identical agents can be represented as $\dot{x} = (I_N \otimes A)$. Another important case is when A is an $N \times N$ order matrix representing the manipulation of scalar data from N agents, and that the manipulation needs to be applied to each value of a vector of length n . In that case, the manipulation can be represented by concatenating the N vectors of length n into a single vector of length Nn , and multiplying it by $A \otimes I_n$. The following property of the Kronecker product:

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD) \quad (3)$$

can be proved when all matrix operations are well-defined [30]. In particular one of the most important properties of the Kronecker product is as follows:

$$A \otimes B = (A \otimes I_p)(I_n \otimes B) = (I_m \otimes B)(A \otimes I_q) \quad (4)$$

3. PROBLEM STATEMENT

Consider a MAS consisting of N followers and a leader. The dynamics of followers are linear n^{th} order system is as follows:

$$\begin{cases} \dot{x}_i = Ax_i + Bu_i, & i = 1, \dots, N \\ y_i = Cx_i \end{cases} \quad (5)$$

where $x_i = [x_{i1}, \dots, x_{in}]^T \in R^n$ represents the states of agent i and also the term $u_i = [u_{i1}, \dots, u_{im}]^T \in R^m$ shows the control inputs of agent i . The matrices $A \in R^{n \times n}$ and $B \in R^{n \times m}$ are constant and represent the behavior of each follower [31]. The compact form of dynamics for N followers is as follows:

$$\begin{cases} \dot{X} = (I_N \otimes A)X + (I_N \otimes B)U \\ Y = (I_N \otimes C)X \end{cases} \quad (6)$$

where $X = [x_1, \dots, x_N]^T$, and $I_N = \text{eye}(N)$ and N is the number of agents.

The leader is the agent that is indexed by L , and the followers are the agents that are indexed by $i = 1, \dots, N$. The leader which is labelled with $i = L$ has linear n^{th} order dynamics is as follows:

$$\begin{cases} \dot{x}_L = Ax_L + Bu_L \\ y_L = Cx_L \end{cases} \quad (7)$$

where $x_L = [x_{L1}, \dots, x_{Ln}]^T \in R^n$ shows the states of the leader. Obviously, the dynamics of the leader is independent of the others, has no input, and is an autonomous system, and is not affected by any of the followers. In this paper, the system matrices are considered the same for all followers and as well as the leader, because this case has practical background such as birds group, fishes school, etc. The compact form of leader dynamics is as follows:

$$\begin{cases} \dot{X}_L = (I_N \otimes A)X_L + (I_N \otimes B)U_L \\ Y_L = (I_N \otimes C)X_L \end{cases} \quad (8)$$

where $X_L = Ix_L$, $I = 1 \otimes I_n$ and 1 is the N -vector of ones.

In this paper, the problem of designing distributed input law of U , which includes the inputs of all followers (i.e., $u_i, i = 1, \dots, N$) to achieve leader-follower consensus is investigated, which is discussed in the following definition.

Definition 1: The leader-follower consensus of system (5) and (7) is said to be achieved, if for each follower $i \in \{1, \dots, N\}$, there is a local state feedback controller u_i of $\{x_i | j \in N_i\}$ such that the closed-loop system satisfies the following equation [32]:

$$\lim_{t \rightarrow \infty} \|x_i(t) - x_L(t)\| = 0, \quad i = 1, \dots, N \quad (9)$$

for any initial condition $x_i(0), i = 1, \dots, N$.

Consider the neighbourhood tracking error of the follower as follows:

$$e_i = \sum_{j \in N_i} a_{ij} (x_j - x_i) + b_i (x_L - x_i) \quad (10)$$

If $x_i \in R^n$, then the overall error form is:

$$E = ((L + B_b) \otimes I_n)(X_L - X) \quad (11)$$

where $I_n = \text{eye}(n)$ and n is the number of states, and $B_b = b_i \otimes I_n$ and $E = [e_1(t), \dots, e_N(t)]^T \in R^N$ is the vector of overall error. As the main contribution of this article, it is assumed that the communication topology between the leader and its neighbors is bounded and also depends on functions. The time derivative of this error along the compact form of dynamics of (6) and (8) is as follows:

$$\dot{E} = ((L + B_b) \otimes I_n)(\dot{X}_L - \dot{X}) \quad (12)$$

Consider the following change of variable:

$$\dot{A} = I_N \otimes A \tag{13}$$

$$\dot{B} = I_N \otimes B \tag{14}$$

Putting the equations of \dot{A} and \dot{B} in the overall error dynamics result that, finally the total error dynamics is as follows:

$$\dot{E} = ((L + B_b) \otimes I_n) \dot{A} ((L + B_b) \otimes I_n)^{-1} E - ((L + B_b) \otimes I_n) \dot{B} U \tag{15}$$

Consider the following relationships:

$$A_n = ((L + B_b) \otimes I_n) \dot{A} ((L + B_b) \otimes I_n)^{-1} \tag{16}$$

$$B_n = ((L + B_b) \otimes I_n) \dot{B} \tag{17}$$

Replacing A_n and B_n from the above equations in (15), \dot{E} is obtained as shown in (18).

$$\dot{E} = A_n E - B_n U \tag{18}$$

4. DESIGNING LEADER-FOLLOWER CONSENSUS CONTROLLER

In this section, the time-invariant overall leader-follower consensus controller of U which consists of the inputs of all followers (i.e., $u_i, i = 1, \dots, N$) is designed. The LQG and LQR controllers are designed as follows.

The LQR cost function is as follows:

$$J = \int (E^T Q_i E + U^T R_i U) dt \tag{19}$$

Where matrix Q_i is positive semidefinite symmetric and matrix R_i is positive definite symmetric.

A hamiltonian function is derived as follows:

$$H = \frac{1}{2} E^T Q E + \frac{1}{2} U^T R U + p^T (A_n E - B_n U) \tag{20}$$

By calculating $\frac{\partial H}{\partial E}$ co-state equation, \dot{P} is obtained as follows:

$$\dot{P} = -Q E - A_n^T p \tag{21}$$

Using the sufficient condition $\frac{\partial H}{\partial U} = 0$ and after some mathematical manipulations, U is obtained as follows:

$$U = R^{-1} B_n^T p \tag{22}$$

Replacing U in (18), derives the following relation:

$$\dot{E} = A_n E - B_n R^{-1} B_n^T p \tag{23}$$

The following Riccati equation for LQR is obtained:

$$\dot{K} + K A_n + K A_n^T + Q - K B_n R^{-1} B_n^T K = 0 \tag{24}$$

Now the distributed overall control law can be considered as (25).

$$U = K E \tag{25}$$

The LQG cost function is as follows:

$$J = E(x_{(t_f)}^T H x_{(t_f)} + \int (E^T Q_i E + U^T R_i U) dt) \tag{26}$$

The following Riccati equation for LQG is obtained:

$$-\dot{S} = A^T S + S A - S B R^{-1} B S + Q \tag{27}$$

The Riccati equation was calculated in the final condition $S_{t_f} = H$ and the system interest was obtained as follows:

$$K = R^{-1} B^T S \tag{28}$$

Now consider the distributed overall control law as follows:

$$U = K E \tag{29}$$

5. SIMULATION RESULTS

In this section, the proposed methodology is applied to the following system. Suppose that a connected, fixed, and directed communication network is depicted in Fig. 1 which is shown by the Laplacian matrix (L). The adjacency matrix of the above MAS is as follows:

$$A_a = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \tag{30}$$

$$L = \begin{bmatrix} 1 & 0 & 0 & -1 \\ -1 & 2 & 0 & -1 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & -1 & 2 \end{bmatrix} \tag{31}$$

In this network, there are one leader and four followers. Bounded and functions for leader adjacency matrix are used, to keep the connectivity of the network. This means that if the network of agents is initially connected, it will stay connected as time passes.

The dynamics of each agent is a linear n^{th} time-invariant high order system, which is relevant to the control of the Robot kinematics system. The attention that there are 4 followers which are shown by $i = 1, \dots, 4$ and the leader is an agent which is shown by $i = L$:

$$\begin{cases} \dot{x}_i = A x_i + B u_i + \xi_i \\ y_i = C x_i + \eta_i \end{cases} \tag{32}$$

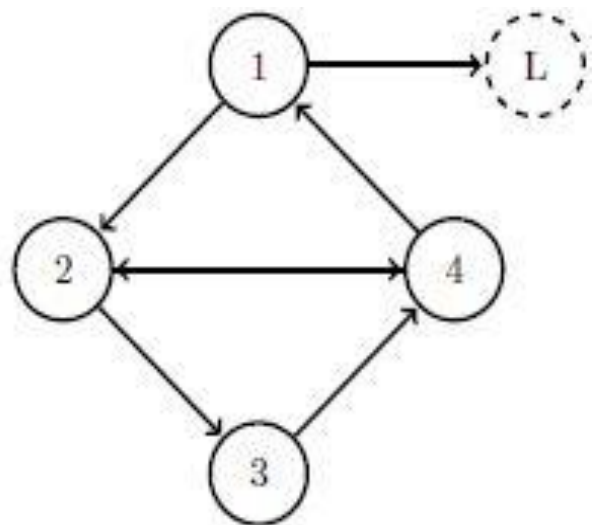


Fig. 1: The fixed and directed communication graph G for the MAS with one leader and four followers.

$$\begin{cases} \dot{x}_L = Ax_L + Bu_L \\ y_L = Cx_L \end{cases} \quad (33)$$

In which the ξ_i is state noise and the η_i is output noise. The variance and mean noises are considered as follows:

$$\xi_i \approx (1.1673 \times 10^{-6}, 1.0077 \times 10^{-8}) \quad (34)$$

$$\eta_i \approx (0.0012, 0.0101) \quad (35)$$

Consider the system matrices as follows:

$$A = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix} \quad (36)$$

$$B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (37)$$

$$C = [1 \ 0] \quad (38)$$

It is easy to investigate that A and B are stabilizable, as well as in here the matrix A is stable.

In the simulation the initial conditions for the agents are considered as follows:

$$x_1(0) = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}, x_2(0) = \begin{bmatrix} 0.3 \\ 0.4 \end{bmatrix}, x_3(0) = \begin{bmatrix} 0.5 \\ 0.6 \end{bmatrix}, x_4(0) = \begin{bmatrix} 0.7 \\ 0.8 \end{bmatrix}$$

$$, x_L(0) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

5.1. LQR Controller

The LQR controller has been implemented on the above system. The results and the diagrams are shown in the following.

Under the proposed compact form of control law (25), the states of each follower track the states of the leader starting from any initial conditions. Figs. 2 and 3 show the difference between the first states of all followers and the leader, and the difference between the second states of all followers and the leader, respectively. As seen in Fig. 2, the first state of the follower is converged to the first of the leader. Also, Fig. 3 reveals that the second state of the followers converges asymptotically to zero. The control input of each follower is $u_i \in R^1$. Figs. 4 and 5 demonstrate the LQR controller input and the LQR Lagrange coefficient of all followers, respectively. Lagrange coefficients are the effect of states on the Hamiltonian, which tend to zero.

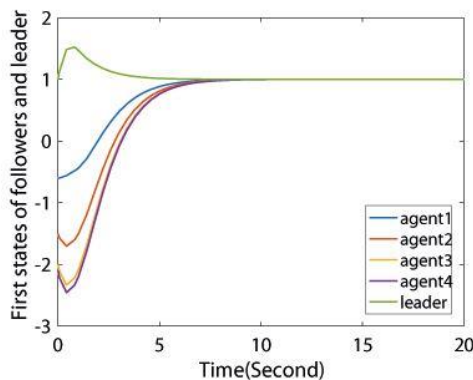


Fig. 2: The first states of followers and leader.

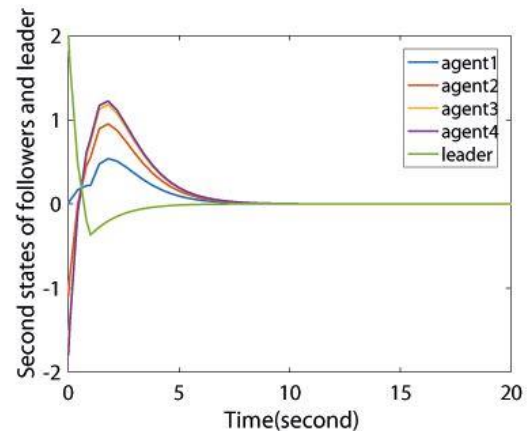


Fig. 3: The second states of the all followers and the leader.

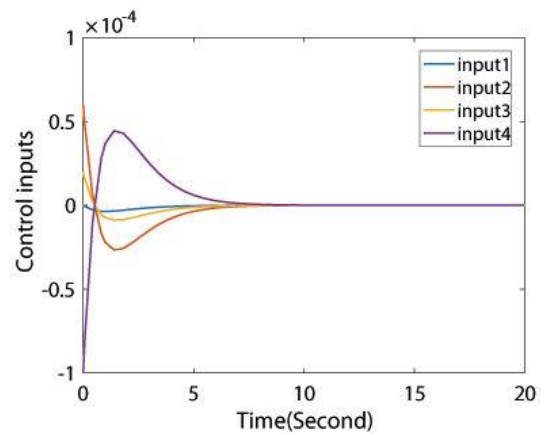


Fig. 4: The control inputs of the followers.

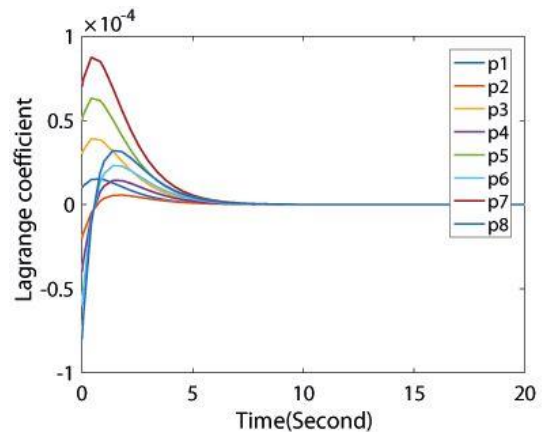


Fig. 5: The LQR Lagrange coefficient.

The consensus tracking errors of the first and the second states of the followers in the LQR controller design procedure are shown on Figs. 6 and 7, respectively. According to these figures, the consensus is well done. The results of the simulation show the hopeful efficiency of the proposed controller for MAS with fixed topology. The simulation results show the limitation of the control inputs, a convergence of the consensus error to zero, and the promising optimal performance with fixed topology.

5.2. LQG Controller

The LQG controller has been applied to the proposed system, the results and the diagrams are shown as follows. In Figs. 8 and 9 the first states of all followers and the leader and

the second states of all followers and the leader is shown, respectively.

According to the figure, the first states of the followers lead to the first states of the leader. According to the figure, the second states of the followers proceed to the second states of the leader too. The states of each follower track the states of the leader starting from any initial conditions. It has been shown that all states of followers track the states of leader according to the proposed methodology. In the following figure, the LQG controller input is depicted as. In the following figure, the LQG Lagrange coefficients of all followers are displayed.

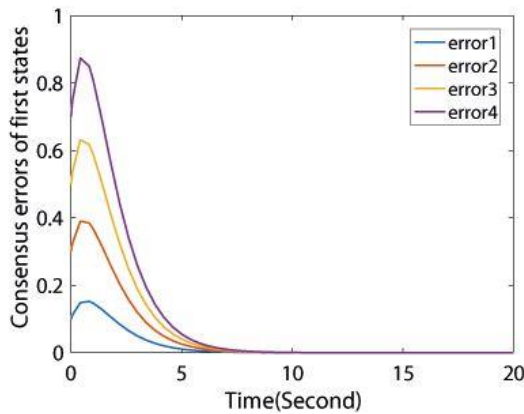


Fig. 6: The consensus errors of the first states of followers.

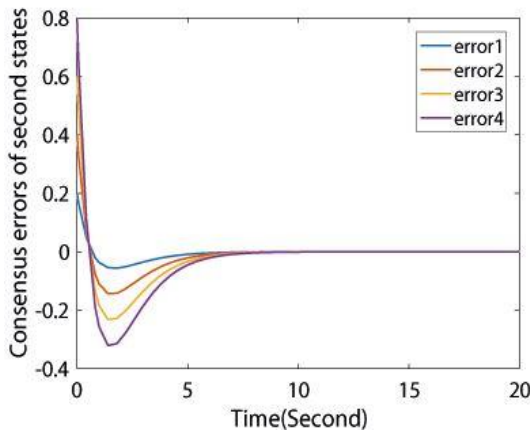


Fig. 7: The consensus errors of the second states of followers.

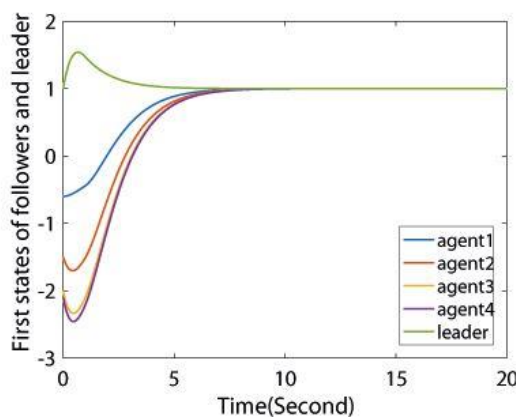


Fig. 8: The first states of followers and the leader.

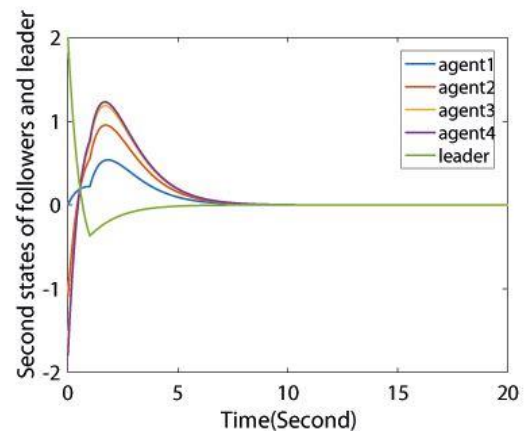


Fig. 9: The second states of the all followers and the leader.

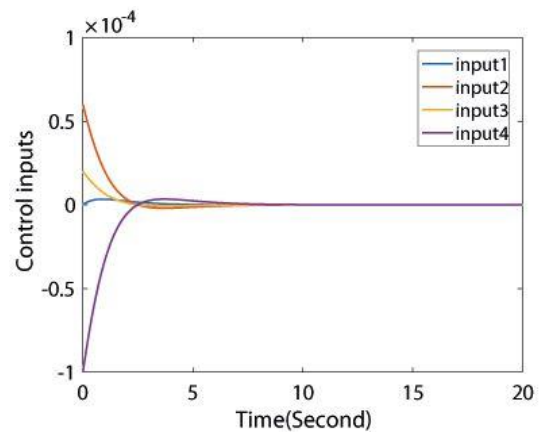


Fig. 10: The control inputs of the followers.

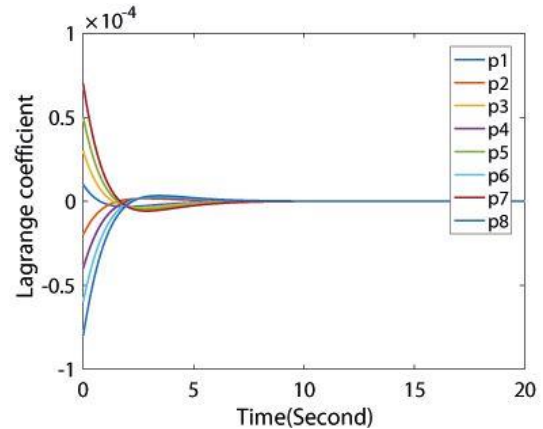


Fig. 11: The LQG Lagrange coefficient.

In the LQG controller design procedure, the consensus tracking errors of all states of the followers are shown as below. According to the figures, the consensus is well done. The results of the simulation show the hopeful performance of the proposed controller for the MAS with fixed topology and the limitation of the Control inputs, the convergence of the Consensus error to zero, and the optimal performance.

5.3. Comparison of LQR and LQG Tracking Error

The LQR and LQG tracking error is compared based on the Integral Square Error criterion in the following table. According to the table it is observed that convergence speed of the LQR methodology implemented is more than the LQG methodology.

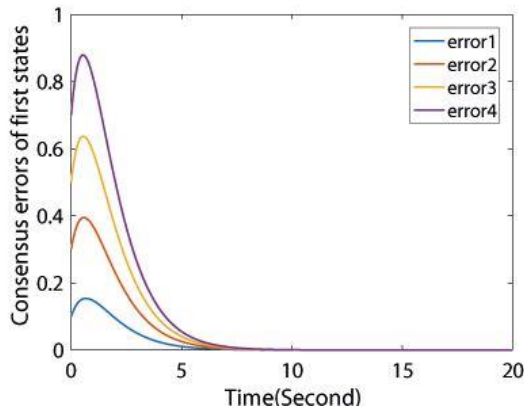


Fig. 12: The consensus errors of the first states of followers.

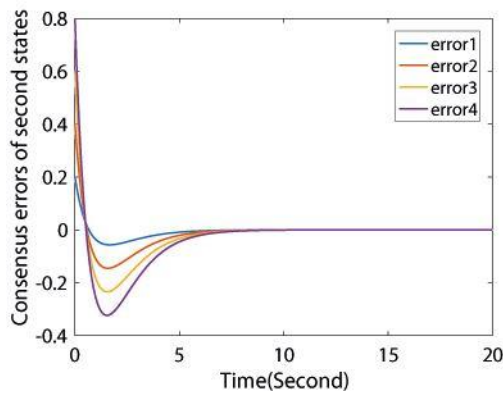


Fig. 13: The consensus errors of the second states of followers.

Table 1: ISE for LQR and LQG.

ISE	LQR	LQG
e1	0.1725	0.4808
e2	0.0682	0.1573
e3	1.1215	3.1002
e4	0.3194	0.7639
e5	2.9018	8.0021
e6	0.7634	1.8476
e7	5.5135	15.1869
e8	1.4003	3.4084

6. CONCLUSIONS

In this paper, the distributed leader-follower consensus of MAS is derived based on LQR and LQG under fixed, directed topology. The dynamics of each agent was a linear high order system. The distributed optimal controller is designated to reach a consensus for both the deterministic and stochastic MAS. The convergence of consensus error to zero in presence of the Gaussian noise is guaranteed. The prescribed gain and phase margin are achieved in our approach. The simulation results also indicated the promising efficiency of the presented methodology in stochastic and deterministic cases. Future research should focus on the formation of stochastic systems and the consensus of nonlinear MAS.

CREDiT AUTHORSHIP CONTRIBUTION STATEMENT

Farideh Azadmanesh: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Roles/Writing - original draft, Writing - review & editing. **Reza Ghasemi:** Formal analysis, Methodology, Project administration, Supervision, Validation,

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy has been completely observed by the authors.

REFERENCES

- [1] A. Bazaei, M. Maroufi, and S. R. Moheimani, "Tracking control of constant-linear-velocity spiral reference by LQG method," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 15568 - 15573, 2017.
- [2] H. Huang, C. Yu, and Q. Wu, "Distributed LQR design for multi-agent formations," in *49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 4535-4540.
- [3] X. Li, F. Zhang, and W. Wang, "The research of distributed suboptimal controller for continuous-time multi-agent systems," in *2017 Chinese Automation Congress (CAC)*, 2017, pp. 5648-5653.
- [4] A. Iftar, and U. Ozguner, "Local LQG/LTR controller design for decentralized systems," *IEEE Transactions on Automatic Control*, vol. 32, no.10, pp. 926-930, 1987.
- [5] F. Zhang, and W. Wang, "Decentralized control of discrete-time system with delay in mean field LQR problem," *Journal of Systems Science and Complexity*, vol. 28, no.4, pp. 755-772, 2015.
- [6] R. H. Krohn and T. R. Bewley, "Underactuated control and distribution of multi-agent systems in stratified flow environments," in *49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 2990-2995.
- [7] Lebao Li, Lingling Sun, and Jie Jin, "Survey of advances in control algorithms of quadrotor unmanned aerial vehicle," in *2015 IEEE 16th International Conference on Communication Technology (ICCT)*, 2015, pp. 107-111.
- [8] S. L. Nguyen and M. Huang, "Mean field LQG games with mass behavior responsive to a major player," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, 2012, pp. 5792-5797.
- [9] A. M. Ferreira, J. A. Barreiros, W. Barra, and J. R. B. de Souza, "A robust adaptive LQG/LTR tcsc controller applied to damp power system oscillations," *Electric Power Systems Research*, vol. 77, no. 8, pp. 956 - 964, 2007.

- [10] W. Wiegerinck, B. V. D Broek, and H. J. Kappen, "Stochastic optimal control in continuous space-time multi-agent systems," *CoRR*, arXiv preprint arXiv:1206.6866, 2012.
- [11] M. Nourian, P. E. Caines, R. P. Malhame, and M. Huang, "Mean field LQG control in leader-follower stochastic multi-agent systems: Likelihood ratio based adaptation," *IEEE Transactions on Automatic Control*, vol. 57, no. 11, pp. 2801–2816, 2012.
- [12] M. Rafieisakhaei, S. Chakravorty, and P. R. Kumar, "MT-LQG: Multi-agent planning in belief space via trajectory-optimized LQG," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 5583–5590.
- [13] M. Hale, A. Jones, and K. Leahy, "Privacy in feedback: The differentially private LQG," in *2018 Annual American Control Conference (ACC)*, 2018, pp. 3386–3391.
- [14] M. Hu, J. Cao, A. Hu, Y. Yang, and Y. Jin, "A novel finite-time stability criterion for linear discrete-time stochastic system with applications to consensus of multi-agent system," *Circuits, Systems, and Signal Processing*, vol. 34, pp. 41–59, 2015.
- [15] T. H. Abdelaziz, and M. Valášek, "State derivative feedback by LQR for linear time-invariant systems," *IFAC Proceedings Volumes*, vol. 38, no. 1, pp. 435 – 440, 2005.
- [16] A. Ansari and T. Murphey, "A variational derivation of LQR for piecewise time-varying systems," in *2015 American Control Conference (ACC)*, pp. 2260 – 2265, 2015.
- [17] D. Tsubakino, T. Yoshioka, and S. Hara, "An algebraic approach to hierarchical LQR synthesis for largescale dynamical systems," in *2013 9th Asian Control Conference (ASCC)*, 2013, pp. 1–6.
- [18] K. You, Z. Li, and L. Xie, "Consensus condition for linear multi-agent systems over randomly switching topologies," *Automatica*, vol. 49, no. 10, pp. 3125 – 3132, 2013.
- [19] J. Zhou, C. Sang, X. Li, M. Fang, and Z. Wang, " H_{∞} consensus for nonlinear stochastic multi-agent systems with time delay," *Applied Mathematics and Computation*, vol. 325, pp. 41 – 58, 2018.
- [20] G. Wen, Z. Duan, and G. Chen, "Distributed consensus of multi-agent systems with general linear node dynamics through intermittent communications," in *2012 24th Chinese Control and Decision Conference (CCDC)*, 2012, pp. 1–5.
- [21] X. Wu, Y. Tang, J. Cao, and W. Zhang, "Distributed consensus of stochastic delayed multiagent systems under asynchronous switching," *IEEE Transactions on Cybernetics*, vol. 46, no. 8, pp. 1817–1827, 2016.
- [22] L. D. Alvergue, A. Pandey, G. Gu, and X. Chen, "Consensus control for heterogeneous multiagent systems," *SIAM Journal on Control and Optimization*, vol. 54, no. 3, pp. 1719–1738, 2016. DOI:
- [23] L. Ma, Z. Wang, Q.-L. Han, and Y. Liu, "Consensus control of stochastic multi-agent systems: a survey," *Science China Information Sciences*, vol. 60, pp. 120–201, 2017.
- [24] R. Guardieño, M. J. López, and V. M. Sánchez, "Mimo pid controller tuning method for quadrotor based on LQR/LQG theory," *Robotics*, vol. 8, no. 2, 2019.
- [25] L. Cheng, Y. Wang, W. Ren, Z. Hou, and M. Tan, "On convergence rate of leader-following consensus of linear multi-agent systems with communication noises," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3586–3592, 2016.
- [26] C. D. Charalambous, A. Farhadi, and S. Z. Denic, "Control of continuous-time linear gaussian systems over additive gaussian wireless fading channels: A separation principle," *IEEE Transactions on Automatic Control*, vol. 53, no. 4, pp. 1013–1019, 2008.
- [27] W. M. Haddad and G. Tadmor, "Reduced order LQG controllers for linear time varying plants," *Systems and Control Letters*, vol. 20, no. 2, pp. 87 – 97, 1993.
- [28] H. Zhang, F.L. Lewis, and Z. Qu, "Lyapunov, adaptive, and optimal design techniques for cooperative systems on directed communication graphs," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 7, pp. 3026–3041, 2012.
- [29] H. V. Henderson, F. Pukelsheim, and S. R. Searle, "On the history of the kronecker product," *Linear and Multilinear Algebra*, vol. 14, no. 2, pp. 113–120, 1983.
- [30] H. Zhang and F. Ding, "On the kronecker products and their applications," *Journal of Applied Mathematics*, vol. 2013, 2013.
- [31] M. A. Ranjbar, R. Ghasemi, and A. Akramizadeh, "Time-varying leader-following consensus of high order multi-agent systems," *International Journal of Modelling, Identification and Control*, vol. 30, no. 4, pp. 333–341, 2018.
- [32] Z. Li, and Z. Duan, "*Cooperative Control of Multi-agent Systems: a Consensus Region Approach*," Taylor & Francis Group, 2014.
- [33] F. Azadmanesh, and R. Ghasemi, "Distributed Optimal Leader-Follower Consensus Protocol Design for a Class of High Order Multi-Agent System: Fixed-Time Topology Approach," in *The First National Conference on Applied Research in Electrical Engineering*, 2021.

BIOGRAPHY

Farideh Azadmanesh received his BSc in Electrical Engineering from the Sepahan University of Sciences and Technology of Isfahan and his MSc in Electrical Engineering from the University of Qom, Qom, Iran. His research interests include large-scale systems, adaptive control, robust control, microgrid, nonlinear control, intelligent systems and multi-agent systems.



Reza Ghasemi received his BSc in Electrical Engineering from the Semnan University in 2000, his MSc and PhD in Control Engineering from the Amirkabir University of Technology, Tehran, Iran, in 2004 and 2009, respectively. His research interests include large-scale systems, adaptive control, robust control,

nonlinear control, intelligent systems and multi-agent systems. He joined the Department of Electrical Engineering, University of Qom, Qom, Iran, where he is currently an Associate Professor of Electrical Engineering.

Copyrights

© 2023 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).





Iranian Association of
Electrical and Electronics
Engineers

Journal of Applied Research in Electrical Engineering

E-ISSN: 2783-2864

P-ISSN: 2717-414X

Homepage: <https://jaree.scu.ac.ir/>



Research Article

Morphological and Statistical Filter-Based Multi-Direction Building Detection for SAR Images

Fateme Amjadipour , Maryam Imani* , and Hassan Ghassemian

Image Processing and Information Analysis Lab, Tarbiat Modares University, Tehran, Iran

* Corresponding Author: maryam.imani@modares.ac.ir

Abstract: SAR images are used in many applications such as building detection. Extracting the building is very challenging due to the radar nature of the SAR images. However, due to the advantages of radar images such as day and night imaging, building extraction from SAR images is a hot topic. In this context, one of the main challenges is the effect of building orientation on the profile created in the SAR image. Also, the two geometric distortions of shadow and layover affect the SAR image. In most building extraction methods, shadow and double bounce are used as two main parameters in building detection. In this paper, different morphological profiles for detecting shadow index and double-bounce index (DMPSIDI) method have been developed using its combination with the method based on statistical features for building extraction. The DMPSIDI method is a morphological-based method that extracts buildings from SAR images independent of changing their profile. The proposed method is also robust to different data using weighting in the main parameters of shadow and double bounce.

Keywords: SAR, morphology index, building direction, statistical feature

Article history

Received 27 April 2022; Revised 04 December 2023; Accepted 31 December 2023; Published online 31 December 2023.

© 2023 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)

How to cite this article

F. Amjadipour, M. Imani, and H. Ghassemian, "Morphological and statistical filter based multi-direction building detection for SAR images," *J. Appl. Res. Electr. Eng.*, vol. 2, no. 2, pp. 145-151, 2023

DOI: [10.22055/jaree.2023.39566.1043](https://doi.org/10.22055/jaree.2023.39566.1043)



1. INTRODUCTION

Building extraction using satellite imagery is used to generate a spatial map of buildings. Synthetic Aperture Radar (SAR) radar images are very popular today due to the fact that they are capable of capturing in both day and night and can capture in any weather conditions. In the special application of building extraction, high resolution SAR images should be used. There are two categories of methods for extracting buildings from SAR images, which include feature-based and model-based methods. The statistical distribution model of SAR image is used in model-based methods to separate target from background. Markov random field (MRF) is a model-based method [1]. The MRF models can be constructed in Polarimetric SAR (PolSAR) images by using a combination of classical statistical distribution and Wishart mixture model [1]. The feature-based methods extract target by using building features in SAR image. These features are brightness, edge, shadow, and shape that are extracted from primary image. In this article, a feature-based method is used to extract double bounce and shadow area with an edge.

The amount of reflected energy from the scene depends on geometry and land cover in radar images. The scattering can be divided into four types: surface, diffusion, double bounce, and volume scattering [2]. Also, the slope in SAR images usually causes three important geometric distortions including layover, foreshortening, and radar shadows. Radar shadow is due to disability of radar beam in ground surface illumination. Relative to the direction of sending waves, these radar shadows are produced due to the positive inclination of objects which prevents the waves from reaching area behind them. So, shadows are seen with a very dark tone in radar images. Shadow in radar images is an important key to interpret the prominent features. So, this information about the scene such as the height of an object can be obtained through the shadow [3]. Grazing angle and slope are two important parameters in radar shadows that depend on size and intensity [4, 5]. So full shadow will be seen in image if slope of the back surfaces of hills and ridges of the ground are more than the grazing angle because they will not be illuminated by the radar pulse.

A building structure has a double bounce reflection against transmitting microwave signals due to its height and structure [6]. As a result, the building is seen as a bright spot in SAR images. The height of building also creates a shadow effect, and there is a dark area next to bright spots. Therefore, a significant percentage of buildings can be extracted using these two features. Another features of buildings in radar images is their rectangular shape and bright texture appropriate for feature extraction [7]. The morphology profile (MP) method is a way that extracts rectangular and line shapes of building in SAR images [8-10]. The size of building in urban areas is not the same. Also, the aspect angle that is the relationship between building orientation and flight line is different in one region [11]. So, the range of structural elements (SE) must be used [12]. On the other way, a combination of structural and statistical information can be used to accurately extract buildings. In [13], the bright and shadow areas are extracted by using the structural and statistical features of image. Then, the structures are detected separately by combining the results. The statistical properties are extracted using the power ratio (PR) method for shadow areas and the order statistics CFAR (OS-CFAR) method for bright areas. Structural features are also adjusted according to the shape and size of the building; and the Difference Morphological Profiles (DMP) method is used for this purpose. In [14], Using a combination of dictionary learning methods and statistical methods Constant False Alarm Ratio (CFAR) and PR, buildings were detected. On the other hand, one of controversial challenges in facing the distinguishing buildings from radar images is the change in radar cross section compared to the change in geometry [15, 16]. In [17] by using PolSAR images and eigenvalues of the coherency matrix, buildings are detected with different values of brightness level due to change of orientation. Subspace alignment has also been used to robust the method of data modification from different systems or the same as domain adaptation (DA). Sun et al [18] eliminated the effects of changing orientation of building using Geographic Information System (GIS) auxiliary data and used the Fully Convolutional Network (FCN) as a deep learning network. For this purpose, the Compound Gaussian Network (CG-Net) module was introduced, which has the ability to connection to other networks. Cho et al [19] improved effect of changing the shadow and bright area of building in different directions for target detection by fuzzy integration of double-bounce index (DI) and shadow index (SI). The DI index is for detecting bright areas and the SI is for detecting shadows in different sizes and directions.

This paper is an extension of our previous work presented in [20]. Our novelties and advantages in this article are as follow:

- Robustness of the proposed method with respect to changes in the orientation of buildings using fusion of two morphological indicators.
- Detection of buildings with flat and gables roofs.
- Reducing false alarm rate by weighting the indicators in different structural elements.
- Increasing detection ratio by using statistical information of the image

Fundamental of the proposed method are described by explaining Black top-hat (BTH) and White top-hat (WTH) method for DI and SI in section 2. Section 3 represents the data and experimental results. The conclusion and future work are followed in section 4.

2. PROPOSED APPROACH FOR BUILDING EXTRACTION

A useful feature for separating the target from the background in extraction of a building is its structural properties. The building in the image has a specific shape and size range in different urban areas. Due to this feature, a morphological method can be used. The difference morphological profiles for detect shadow index and double-bounce index (DMPSIDI), which is a morphological-based method, is presented in [21]. In this method, two indicators of SI and DI are used to detect shadow areas and double bounce, respectively. DMP is then used instead of MP to consider edge properties. Finally, to make the method more robust to different data, two areas of shadow and double bounce with different weights are considered. In this paper, to increase the accuracy of building detection, the results obtained from Weighted DMPSIDI (WDMPSIDI) are combined with the statistical features of the image and the results are improved. The flowchart of the proposed method is shown in Fig. 1.

2.1. Morphological Method

Spatial features of the image include shape, texture and edge. Using the morphological method, the shape and sometimes the edge can be extracted [22-24]. The determining element in this method is the structural element. The structural element can be linear or nonlinear. In a linear structural element unlike the nonlinear one, direction is important. Due to the rectangular shape of buildings, the structural element should be considered rectangular. There are several methods in morphology. In relation to building extraction, a method must be chosen that takes into account the two parameters of shadow and double bounce. The details of the proposed method are described below.

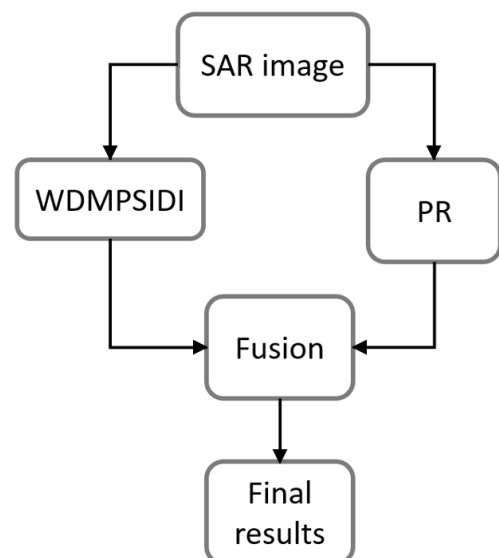


Fig. 1: The flowchart of the proposed method.

2.2. Top-hat Transform

To identify the building in the image, we need features that distinguish the building from the background. These features include double shadows and reflections that are dark and very light in the image and have a rectangular shape. The top-hat converter extracts small elements that are lighter or darker than their neighbours. BTH conversion is used to extract dark areas and WTH conversion is used for light areas [25]. The relationship between these two transformations is as follows:

$$BTH(f) = \phi(f) - f \tag{1}$$

$$WTH(f) = f - \gamma(f) \tag{2}$$

where f is the brightness of the image and γ is the opening operator and ϕ closing is the closing operator.

2.3. WDMPSIDI Method

This method is based on morphology and uses top-hat to extract two parameters of shadow and double bounce in different directions and sizes. Then, by merging the two obtained areas, the location map of the buildings is obtained. The flowchart of the WDMPSIDI method is shown in Fig. 2.

2.4. Double Bounce and Shadow Index

As shown in Fig. 3, the rectangular shape, the bright and shadow areas and the size of the building are the prominent features that distinguish it from the background. The shadow index (SI) is obtained using the BTH conversion and the double bounce index (DI) is obtained using the WTH conversion. n is the number of each step and MP is the image obtained from the morphology operator.

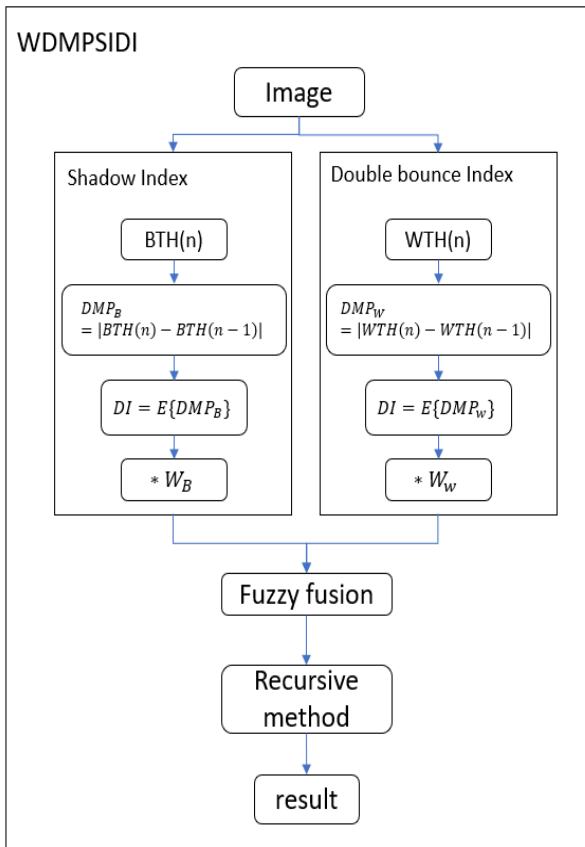


Fig. 2: The flowchart of the WDMPSIDI method.

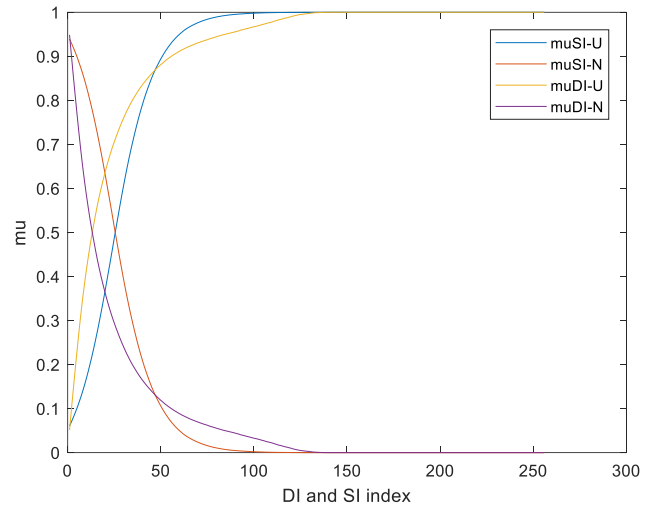


Fig. 3: Membership grade related to the two indicators SI and DI.

To identify different buildings, an image uses a rectangular structural element in different sizes and directions for opening and closing operators. Then, the conversion of BTH and WTH is applied to each, and finally, the two indices SI and DI are obtained by averaging.

$$SI = E \left\{ w_{ij_B} \left| MP_B(l_i, d_j) - MP_B(l_j, d_{j-1}) \right| \right\} \tag{3}$$

$$DI = E \left\{ w_{ij_w} \left| MP_w(l_i, d_j) - MP_w(l_j, d_{j-1}) \right| \right\} \tag{4}$$

$$i = 1, \dots, n \quad j = 2, \dots, m$$

where w_{ij} is a weight for SI and DI in element size l_i and direction of d_j . For the scale parameter l , since buildings have a limited size, the range l_{min} to l_{max} with step Δl is used. Multiscale and multidirectional shadow information is average of DMP.

2.5. Differential Morphological Profile

The DMP method has been used to consider feature of the building edge. In this method, the morphological profile obtained from each stage with the applied structural element is reduced differently from the previous stage. Relationships related to it are as follows:

$$DMP = MP(n) - MP(n - 1) \tag{5}$$

2.6. DMPSIDI Method

In this method, MP is obtained for each structural element related to BTH and WTH conversion. Then, to consider the edge property, each MP is subtracted from the MP of the previous step and the DMP of that step is obtained. Finally, by averaging the resulting DMPs, two indicators, DMPSI and DMPDI are obtained.

Now, to determine location of buildings, these two indicators must be combined. Fuzzy integration is used for this purpose. Fuzzy integration for crisp data has three steps. The first step is fuzzification of the morphological index. At this stage, using the cumulative distribution and its

complement, the degree of membership of each index is calculated. The membership grade chart obtained in this step is shown in the figure. The second stage is fusion of fuzzy memberships. To do this, the maximum function is used. The third step is defuzzification for decision making. At this stage, binary thresholding is performed.

2.7. Making Robust to Different Data

Depending on the urban area and geometry of the sensor, the number of shadows and brightness in different images is different. For this purpose, two indicators SI and DI with different weights are considered. An iterative loop is also used to estimate the most appropriate parameters required in this method. The loss function is considered in this MSE loop.

$$RMSE = \sqrt{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(x,y) - \hat{f}(x,y)]^2} \quad (6)$$

where $f(x,y)$ is the intensity of pixel in row x and column y , and M and N are numbers of rows and columns of image, respectively. The next challenge in some images is existence of the road, which is recognized as shadow of the building. To eliminate this false alarm, the differential recursive method has been used.

3. EXPERIMENTAL RESULTS

3.1. Used Data

The data was obtained from the Terra SAR-X satellite over Lille Strom, Norway on April 10, 2019 and the capture mode was starting spot light. There are various buildings in the study area. Geographically, it is 59.90 at longitude and 11.04 latitude west. Its polarization is HH and it is obtained in descending orbit. Fig. 4 shows the area under study.

3.2. Results of Implementation

In the first step, the necessary pre-processing approaches have been performed, including the removal of speckle noise and geometric correction of the data. After preparing the data for processing, the proposed method is applied to the data, and then, the appropriate parameters are estimated. The final result is obtained as a binary map. Fig. 5 shows Binary map of the building extracted using three methods. Methods are the proposed, SIDI and PR. In Fig. 6 optical image and GTM is shown.

3.3. Quantitative Evaluation

In most studies of this field, the evaluation of extracted buildings is usually done as an object-based problem. In this paper, a new method for evaluation is presented by using the GTM map prepared pixel-wise where the extracted buildings are evaluated pixel-wise. Table 1 provides a quantitative evaluation of the proposed method and two other methods that are SIDI and PR.

Table 1: Performance evaluation of building extraction.

Approach	FP	FN	TP	recall	precision	F1
Proposed	2920	4877	6148	55.76	67.80	0.6120
SIDI	3795	4196	6829	61.94	64.28	0.6309
PR	36626	1696	9329	84.62	20.30	0.3274

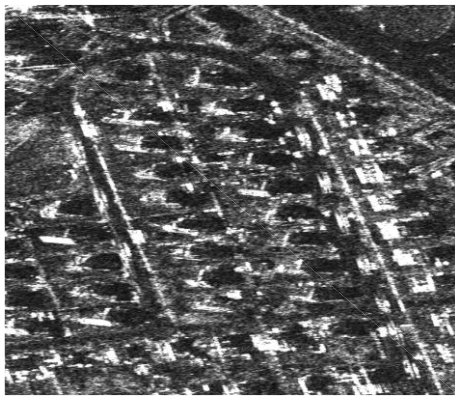


(a)

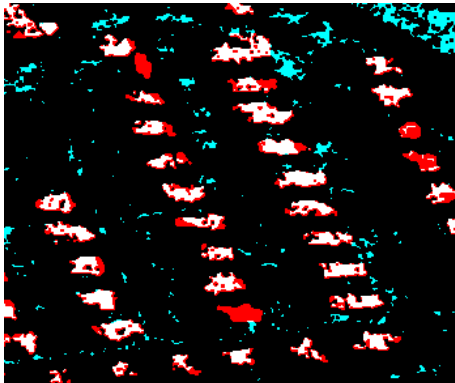


(b)

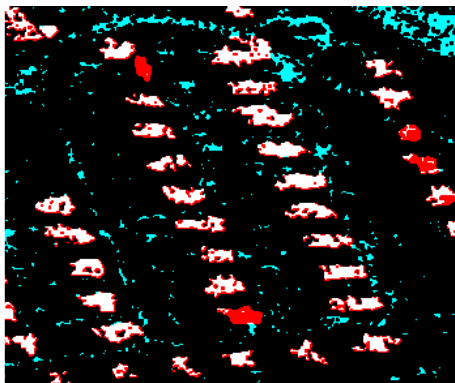
Fig. 4: Area overview, (a) optical image, (b) SAR image.



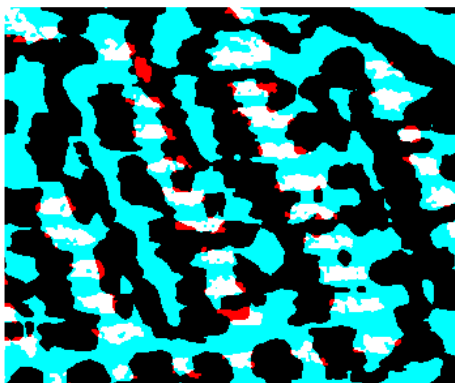
(a)



(b)



(c)



(d)

Fig. 5: Binary map of the building extracted using three methods, (a) SAR image, (b) false colour of binary map and ground truth of the proposed method and (c) result of the SIDI method, (d) result of the PR method.



(a)



(b)

Fig. 6: (a) google earth image, (b) ground truth.

4. CONCLUSION AND RECOMMENDATIONS

4.1. Conclusion

One of the major challenges in extracting a building is changing its brightness in SAR images. The morphological method is a suitable method to identify the building according to this challenge. Due to the phenomenon of shadow and double reflection, dark and bright areas are created in the image of the building. In the WDMPSIDI method, these regions are extracted using Top-Hat transform and finally merged using fuzzy fusion. To make the method resistant to data changes, two different coefficients w_1 and w_2 for shadow index and double bounce have been used. Another challenge is recognizing roads as false alarms. For this purpose, a recursive method is used in the proposed method. In this paper, by combining the results of WDMPSIDI method and using statistical features, the buildings have been extracted with good accuracy. A basic pixel method is also presented in the quantitative evaluation. This increases the accuracy of the assessment.

4.2. Future Work

One of the suggestions for future work is to develop fuzzy fusion to combine the detection results of shadow and

bright areas. Second, location of the building extracted from optical images can be used to increase the accuracy. Also in the evaluation section, a new method can be presented to evaluate the resulting images. The advantages of this method include recognition of tall buildings and high processing speed. However, this method has a lot of false positives in areas with small buildings, and also has low accuracy in certain imaging modes where shadow or bright area has a greater share than the other. To improve this problem, the wavelet transform can be used in future work and the resulting derivatives can be applied as input in this method.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Fateme Amjadipour: Conceptualization, Methodology, Software, Validation, Visualization, and Writing. **Maryam Imani:** Supervision, Review & Editing. **Hassan Ghassemian:** Investigation & Supervision.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy has been completely observed by the authors. This paper is an extension of the authors' previous work presented in 2022 International Conference on Machine Vision and Image Processing (MVIP) [20], which was selected as a top paper to be printed in this journal with some extensions.

REFERENCES

- [1] J. Yin, X., Liu, J. Yang, C. Y. Chu, and Y. L. Chang, "PolSAR image classification based on statistical distribution and MRF." *Remote Sensing*, vol. 12 no. 6, pp. 1027, 2020.
- [2] Y. Wang, W. Yu, X. Liu, C. Wang, A. Kuijper, and S. Guthe, "Demonstration and analysis of an extended adaptive general four-component decomposition," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 2573-2586, 2020.
- [3] O. Karabayır, U. Saynak, M. Z. Kartal, A. F. Coşkun, T. O. Gulum, and B. Bati, "Synthetic-range-profile-based training library construction for ship target recognition purposes of scanning radar systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 4, pp. 3231-3245, 2020.
- [4] S. Saha, F. Bovolo, and L. Bruzzone, "Building change detection in VHR SAR images via unsupervised deep transcoding," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 1917-1929, 2020.
- [5] O. Liu, O. Li, O. W. Yu, W. Hong, "Automatic building detection for multi-aspect SAR images based on the variation features," *Remote Sensing*, vol. 14, no. 6, 2022.
- [6] W. Sandhi, P. Sameczyński, and A. Gromek. "Data augmentation for building footprint segmentation in SAR images: an empirical study," *Remote Sensing*, vol. 14, no. 9, 2022.
- [7] J. Sejung, K. Lee, and W. H. Lee. "Object-Based High-Rise Building Detection Using Morphological Building Index and Digital Map." *Remote Sensing*, vol 14, no. 2, 2022.
- [8] M. Imani, and H. Ghassemian, "Spectral-spatial classification of high dimensional images using morphological filters and regression model," in *2016 6th International Conference on Intelligent and Advanced Systems (ICIAS)*, 2016, pp. 1-6.
- [9] M. Imani, and H. Ghassemian, "Edge patch image based morphological profiles for classification of multispectral and hyperspectral data", *IET Image Processing*, vol. 11, no. 3, pp. 164-172, 2017.
- [10] M. Imani, and H. Ghassemian, "Discriminant analysis in morphological feature space for high-dimensional image spatial-spectral classification", *J. Applied Remote Sensing*, vol. 12 no.1, pp. 0160241- 01602428, 2018.
- [11] S. Adelipour, and H. Ghassemian, "The fusion of morphological and contextual information for building detection from very high-resolution SAR images," *Electrical Engineering (ICEE), Iranian Conference on*, 2018, pp. 389-393.
- [12] M. Khodadadzadeh, O. Sarrafzade, and H. Ghassemian, "A fast traffic sign detection and classification system based on fusion of colour and morphological information," in *2010 6th Iranian Conference on Machine Vision and Image Processing*, 2010, pp. 1-5.
- [13] S. Adelipour, and H. Ghassemian, "Building extraction from very high-resolution synthetic aperture radar images based on statistical and structural information fusion," *Int. J. Remote Sens.*, vol. 40, no. 18, pp. 7113-7126, 2019.
- [14] S. Adelipour, and H. Ghassemian, "Building detection in very high resolution SAR images via sparse representation over learned dictionaries," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 11, no. 12, pp. 4808-4817, 2018.
- [15] Z. Guo, H. Liu, L. Pang, L. Fang, and W. Dou, "DBSCAN-based point cloud extraction for Tomographic synthetic aperture radar (TomoSAR) three-dimensional (3D) building reconstruction," *Int. J. Remote Sens.*, vol. 42, no. 6, pp. 2327-2349, 2021.
- [16] D. Pirrone, F. Bovolo, and L. Bruzzone, "An approach to unsupervised detection of fully and partially destroyed buildings in multitemporal VHR SAR images," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 5938-5953, 2020.
- [17] R. Gui, X. Xu, L. Wang, R. Yang, and F. Pu, "Eigenvalue statistical components-based pu-learning for polsar built-up areas extraction and cross-domain analysis," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 3192-3203, 2020.
- [18] Y. Sun, Y. Hua, L. Mou, and X. X. Zhu, "CG-Net: conditional GIS-aware network for individual building segmentation in VHR SAR images," *IEEE Trans. Geosci. Remote Sens.*, 2021.

- [19] K. Cho, S. E. Park, J. H. Cho, H. Moon, and S. H. Han, "Automatic urban area extraction from SAR image based on morphological operator," *IEEE Geosci. Remote Sens. Lett.*, 2020.
- [20] F. Amjadipour, H. Ghassemian, and M. Imani. "Building detection using very high resolution SAR images with multi-direction based on weighted-morphological indexes", in *2022 International Conference on Machine Vision and Image Processing (MVIP)*, pp. 1-6, 2022.
- [21] F. Amjadipour, H. Ghassemian, and M. Imani. "Estimation of free parameters of morphological profiles for building extraction using SAR images", in *2021 7th International Conference on Signal Processing and Intelligent Systems (ICSPIS)*, pp. 1-6, 2021.
- [22] A. Taghipour, H. Ghassemian, and F. Mirzapour, "Anomaly detection of hyperspectral imagery using differential morphological profile," in *2016 24th Iranian Conference on Electrical Engineering (ICEE)*, 2016, pp. 1219-1223.
- [23] M. Darvishnezhad, H. Ghassemian, M. Imani, "Spatial-Spectral morphological feature extraction for hyperspectral images classification", in *Proceedings of the 2019 International Conference on GeoSpatial, Agriculture, Natural Resources and Disaster Management*, pp. 1-6, October 2019.
- [24] M. Imani and H. Ghassemian, "spectral-spatial feature transformation with controlling contextual informations through smoothing filtering and morphological analysis", *International Journal of Information & Communication Technology Research*, vol. 10, no. 1, pp. 1-12, Winter 2018.
- [25] E. R. Dougherty, and R. A. Lotufo, "Hands-on morphological image processing," SPIE press, 2003.

BIOGRAPHY



Fateme Amjadipour received her B.S. degree in Electrical Engineering from Semnan University, Iran, in 2017, and his M.S. degree in Passive Defense (AFA) from Malek-Ashtar University of Technology, Iran, in 2020. Currently, she is pursuing a PhD at Tarbiat Modares University, focusing on the field of image processing and remote sensing.



Maryam Imani completed her Ph.D in Electrical Engineering, Communication, from Tarbiat Modares University, Tehran, Iran in 2015. She continued her research in Tarbiat Modares University as a postdoc. Since 2018, she has been with Tarbiat Modares University in Tehran, Iran,

where she is the Associate Professor of Computer and Electrical Engineering. Her research interests include statistical pattern recognition, machine learning, signal and image processing and remote sensing.

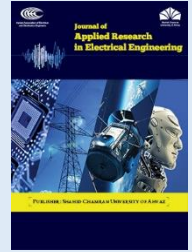


Hassan Ghassemian received the B.S.E.E. degree from Tehran College of Telecommunication in 1980 and the M.S.E.E. and Ph.D. degree from Purdue University, West Lafayette, USA in 1984 and 1988 respectively. Since 1988, he has been with Tarbiat Modares University in Tehran, Iran, where he is a Professor of Computer and Electrical Engineering. His research interests focus on Multi-Source Signal/Image Processing and Information Analysis and Remote Sensing.

Copyrights

© 2023 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).





Research Article

Design of a Tunable 4th Order OTA-C Band-Pass Filter for use in Front-End of ADC

Mohsen Ghaemmaghami , and Shahbaz Reyhani 

Department of Electrical Engineering, University of Guilan, Persian Gulf Highway, Rasht 41996-13776, Iran

* Corresponding Author: shahbaz@guilan.ac.ir

Abstract: This article presents a tunable fourth-order band-pass filter that is designed using operational trans-conductance amplifier (OTA), which can be used as an anti-aliasing filter (AAF) in the front-end of an analog to digital converter (ADC). It is necessary to use a suitable filter to prevent unwanted signals from entering the ADC. The bandwidth of the AAF should be designed according to the bandwidth of the ADC, therefore, matching these two bandwidths is one of the important challenges when using multi-bandwidth analog to digital converters in the digital communication applications. The proposed band-pass filter is designed and simulated in 180nm CMOS technology. The simulation results show that by changing the two bias voltages of the proposed filter, its bandwidth can be changed according to the frequency range of ADSL, ADSL2 and ADSL2+ communication standards and it effectively attenuates unwanted signals.

Keywords: Anti-aliasing filter, multi-bandwidth ADC, OTA-C filter, ADSL technology.

Article history

Received 04 January 2023; Revised 23 November 2023; Accepted 27 November 2023; Published online 29 January 2024.

© 2024 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)

How to cite this article

M. Ghaemmaghami, and S. Reyhani, "Design of a tunable 4th order OTA-C band-pass filter for use in front-end of ADC, *J. Appl. Res. Electr. Eng.*, vol. 2, no. 2, pp. 152-157, 2023. DOI: [10.22055/jaree.2023.42687.1065](https://doi.org/10.22055/jaree.2023.42687.1065)



1. INTRODUCTION

In order to use the continuous values in digital systems, analog to digital converters should be used. Analog to digital converters are circuits that convert continuous analog signals into discrete digital data [1, 2]. Digital cameras, biomedical applications and communication devices such as mobile phones and ADSL (Asymmetric Digital Subscriber Line) modems are devices that use an ADC [3-6]. Analog to digital converters have a structure similar to the block diagram in Fig. 1. In order to convert an analog signal to digital codes, first the input signal is passed through an anti-aliasing filter to prevent unwanted signals from entering the converter. Then, the output of this filter is sampled by the sampler circuit (for example: a sample and hold circuit) and sent to the next part, the quantizer, and according to the voltage level of each sample, a corresponding digital code is produced [2].

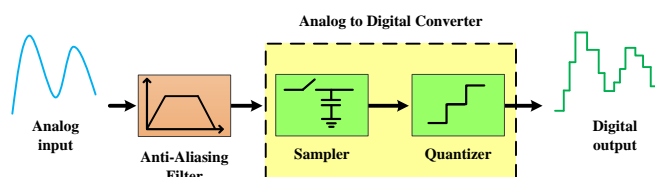


Fig. 1: Analog to digital converter block diagram.

Most ADCs need anti-aliasing filters, because according to the intended application, these filters can weaken some unwanted signals that are not in the required bandwidth of the converter. The passage of these signals reduces the performance of ADC [7]. On the other hand, new tunable ADCs that have the ability to change the bandwidth of their operation have been presented [8], which are used in various applications. In some digital communication systems, analog to digital converters should be able to change their bandwidth according to the need, in which a tunable anti-aliasing filter will be required. For example, ADCs used in communication devices that use ADSL technique must have the ability to convert in the bandwidths required by ADSL, ADSL2 and ADSL2+ standards, so it is necessary at the input of these types of converters and a proper tunable anti-aliasing filter should be placed in the front-end of the ADC [9, 10]. This paper is structured as follows: Section 2 describes the proposed band-pass filter architecture and its circuit details. Section 3 presents the simulation results, and finally Section 4 ends this paper with the conclusion.

2. STRUCTURE OF THE PROPOSED BAND-PASS FILTER

The block diagram of the proposed tunable fourth-order

band-pass filter is shown in Fig. 2. As shown in this figure, the structure of this filter consists of a low-pass and a high-pass filters, which their cut-off frequency can be adjusted by the external bias voltages of V_{b1} and V_{b2} .

2.1. Circuit Analysis of the Proposed Tunable OTA-C Filters

The OTAs which are actually known as differential voltage controlled current source (DVCCS). The value of "g" that shows their transconductance is calculated as follows (1):

$$g = \frac{I_{out}}{V_{in}} \quad (1)$$

Where I_{out} represents the output current and V_{in} is the value of the input voltage. The value of "g" can be mentioned as an essential parameter in the design such as active resistive capacitor (RC) circuits in which R is an important design parameter, and it is usually represented by an adjustable bias current (I_{bias}). The electronic adjustability of a circuit is one of the critical points that must be taken into account in a design to meet various design specifications [11, 12]. Due to the ability to adjust the transconductance in OTA, this type of amplifiers and capacitors are used in the OTA-C filter design. The OTA-C filters are categorized into two general structures, single output [13] or multiple outputs [14] with different types of designs. This research focuses on single output OTA-C filters. The proposed band-pass filter consists of two OTA-C fourth order low-pass and high-pass filters. To simplify the description of the operation of these filters, their second-order circuits are shown in Fig. 3.

The transfer function of these filters show the dependency of their cut-off frequency to the transconductance of OTAs. The transfer functions of the second-order high-pass and low-pass filters are shown in equations (2) and (3), respectively:

$$TF = \frac{S^2 C_1 C_2}{S^2 C_1 C_2 + S C_1 g_{m1} + g_{m1} g_{m2}} \quad (2)$$

$$TF = \frac{g_{m1} g_{m2}}{S^2 C_1 C_2 + S C_1 g_{m1} + g_{m1} g_{m2}} \quad (3)$$

Where g_{m1} and g_{m2} are the trans-conductance of the first and second amplifiers and C_1 and C_2 are the capacitors used in the proposed circuits. Assuming that the trans-conductance of OTAs and the capacitance of capacitors are equal, the cut-off frequency of the high-pass and low-pass filters (ω_0) is obtained from equation (4):

$$\omega_0 = \frac{g_m}{\sqrt{C_1 C_2}} \quad (4)$$

According to equation (4), the trans-conductance of OTAs can be changed by altering the bias voltage, thus the cut-off frequency of the filters can be tuned using the bias voltages.

The frequency response of the second-order high-pass and low-pass filters are shown in Figs. 4 and 5, respectively. As shown, by changing the bias voltage in the range of 0.3 -

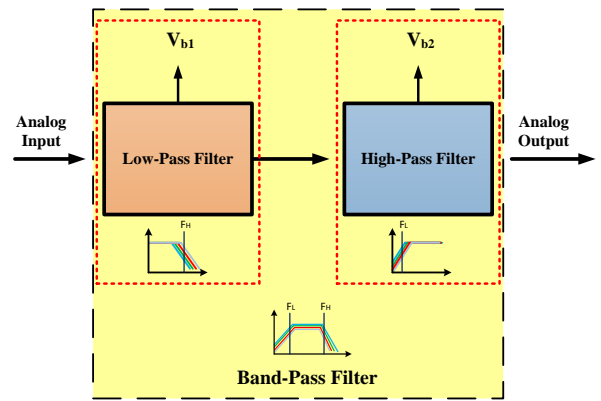


Fig. 2: Block diagram of the proposed band-pass filter.

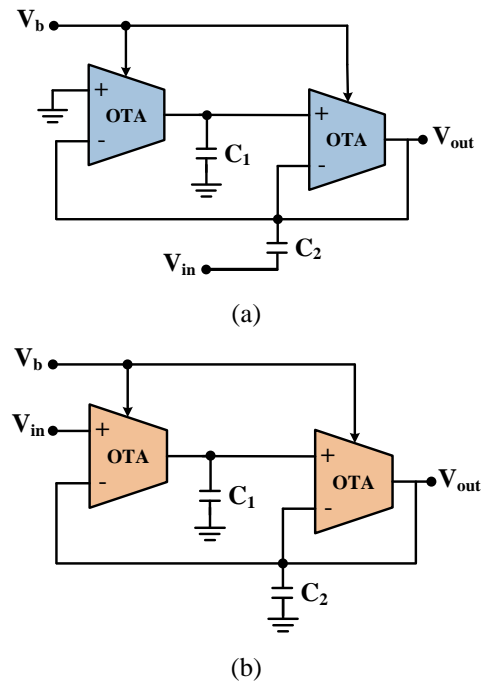


Fig. 3: Circuit schematic (a) high-pass (b) low-pass filter used in the proposed band-pass filter.

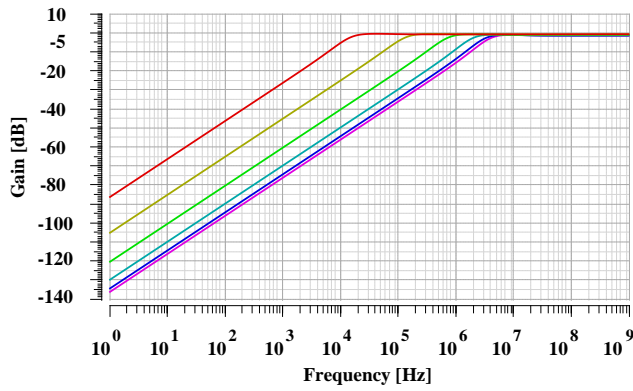
0.8 volts, the cut-off frequency in the high-pass filter is adjusted between 10 kHz - 5 MHz, and in the low-pass filter, by adjusting the voltage between 0.3 - 0.8 volts, the cut-off frequency is changed between 12 kHz - 4 MHz.

2.2. Proposed Transconductance Amplifier Circuit

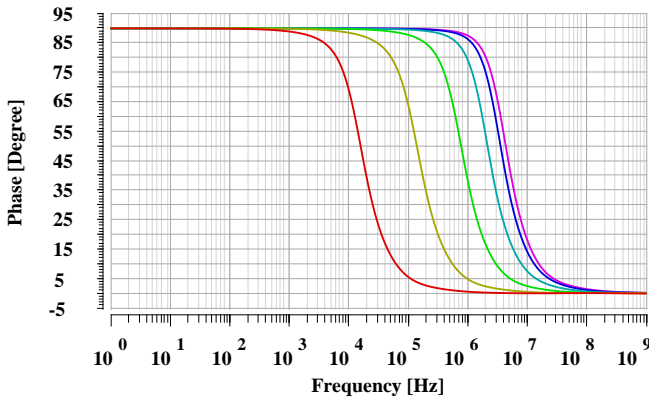
The proposed transconductance amplifier that is used in the high-pass and low-pass filters is shown in Fig. 6. In the proposed amplifier, the trans-conductance value can be adjusted with changing of the bias voltage (V_b). The proposed differential transconductance amplifier has an output gain of 91 dB, a bandwidth of 6.8 MHz, and a phase margin of 75 degrees, and a common-mode rejection ratio (CMRR) of 105 dB. Fig. 7 shows the gain and phase diagram of the transconductance amplifier. Table 1 shows information such as the sizing of the OTA-C used transistors, the drain currents and the calculated trans-conductance values for each transistors.

2.3. The Structure of the Proposed Band-Pass Filter

The OTA-C fourth-order band-pass filter circuit with adjustable feature using two high-pass and low-pass filters is

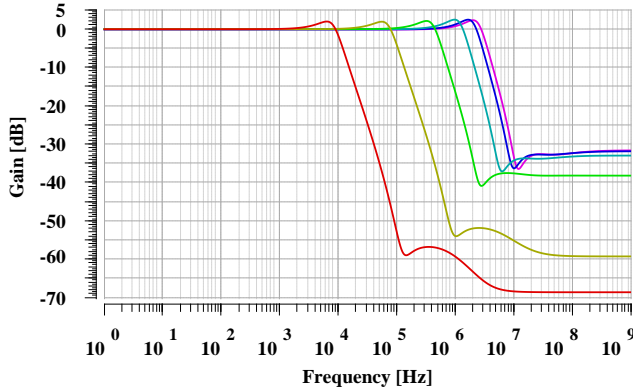


(a)

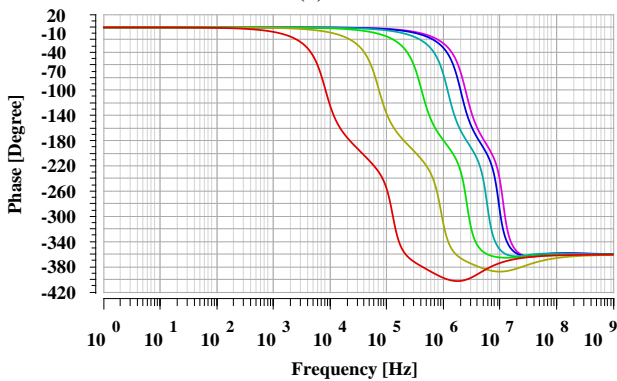


(b)

Fig. 4: Frequency response of the desired second-order high-pass filter including (a) gain and (b) phase.



(a)



(b)

Fig. 5: Frequency response of the desired second-order low-pass filter including (a) gain and (b) phase.

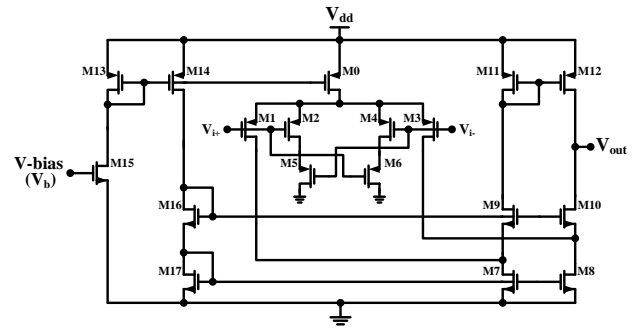
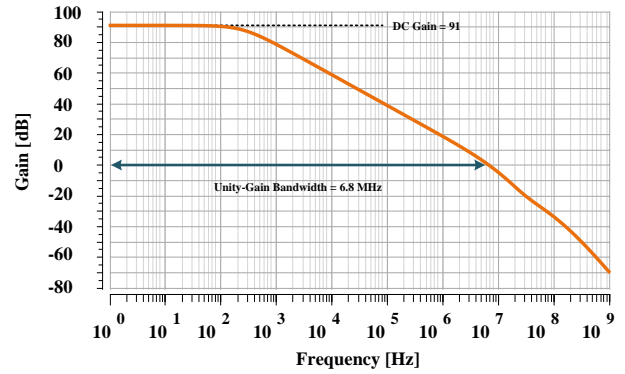
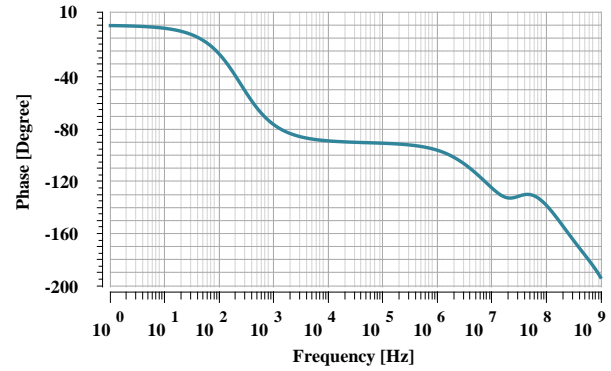


Fig. 6: Proposed trans-conductance amplifier circuit.



(a)



(b)

Fig. 7: The frequency response of the proposed operational trans-conductance amplifier for $V_b = 0.6V$ including (a) gain and (b) phase.

Table 1: The parameters of transistors in the proposed OTA

Transistor	W(μm)	L(μm)	M	$I_d(\mu A)$	$V_{dSat}(mv)$	$g_m(\mu S)$
M_0	4	2	7	6.15	210	83.5
$M_1 M_3$	2	2	6	1	106	31.5
$M_2 M_4$	2	2	6	2.1	106	11.1
$M_5 M_6$	2	2	6	2.1	81.2	18
$M_7 M_8$	2.5	2	3	4.74	106.5	77.5
M_9	2.5	2	1	2.7	132.3	37.8
M_{10}	2.5	2	1	2.7	131.4	38.2
$M_{11} M_{12}$	4	2	7	2.7	86.5	45.7
$M_{13} M_{14}$	4	2	7	6.2	119.3	84.5
M_{15}	2.25	2	2	6.2	141.7	83
$M_{16} M_{17}$	2	2	5	6.2	108	102

presented in Fig. 8. By changing the bias voltages of the trans-conductance amplifiers, the cut-off frequencies of the filter can be tuned. Table 2 shows the relation between the filter cut-off frequencies and the bias voltages. In the proposed band-pass filter circuit, the capacitance of the used capacitors is 5 pF.

3. SIMULATION RESULTS

The proposed band-pass filter was designed and simulated with CMOS 180nm technology at the transistor level. The proposed filter is tunable and its cut-off frequencies can be changed by adjusting the gm values of the OTAs using the bias voltages. There are two bias voltage settings of the proposed filter for working frequencies of special telecommunication standards, ADSL1 and ADSL2, are presented in Table 2. Also, the frequency responses and multiple simulation results of the proposed adjustable band-pass filter by setting the bias voltages are shown in Table 2 and Fig. 9, respectively. As shown in Fig. 9, the bandwidth of the proposed filter will be tuned with changing of the bias voltages. The layout of the proposed band-pass filter is shown in Fig. 10. The dimensions of the design are 573 um by 575 um, the area occupied by the proposed circuit is equal to 0.32 mm². Also, with specified bias voltages, the frequency responses related to the post-layout simulation of the proposed filter are shown in Fig. 11. By comparing the output results obtained from the post-layout and the pre-layout simulations of the proposed filter, it can be concluded that the frequency responses of the two are almost the same except for the attenuation in the upper stop band. The results of post-layout simulation are shown in Table 3.

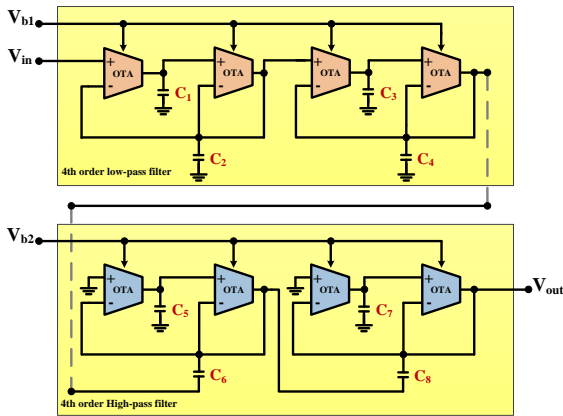


Fig. 8: Proposed fourth-order band-pass filter circuit.

Table 2: Frequency values required by changing the bias voltage for different applications in Pre-Layout simulation.

Curve	V _{BIAS1} (V)	V _{BIAS2} (V)	Low cut-off frequency (kHz)	High cut-off frequency (MHz)
ADSL 1	0.5	0.5	32	1.1
ADSL 2	0.5	0.6	32	2.2
Custom	0.28	0.52	11	1.3

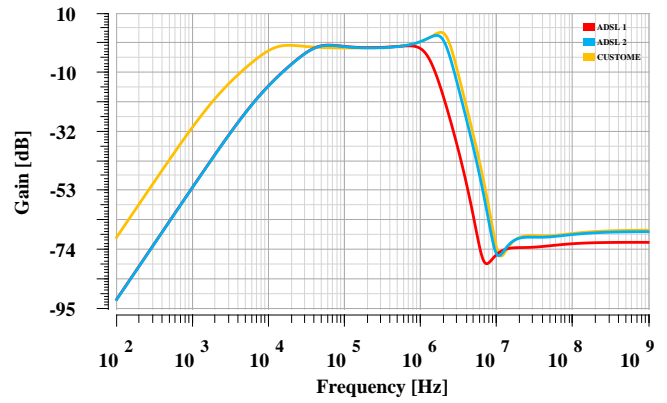


Fig. 9: The frequency response of the proposed filter for the bias voltage presented in Table 2.

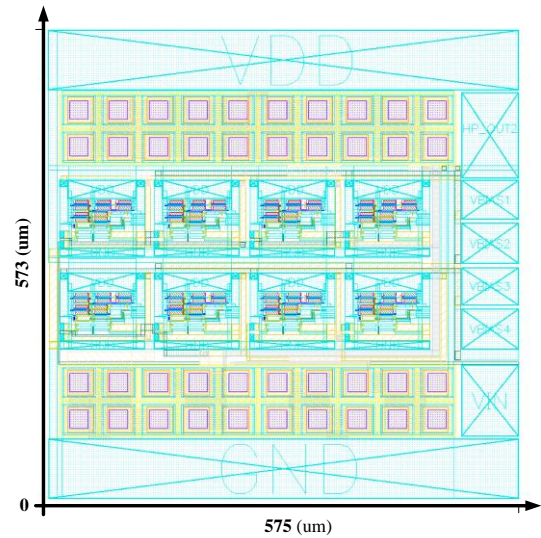


Fig. 10: Layout of the proposed band-pass filter.

Table 3: Frequency values required by changing the bias voltage for different applications in Post-Layout Simulation.

Curve	V _{BIAS1} (V)	V _{BIAS2} (V)	Low cut-off frequency (kHz)	High cut-off frequency (MHz)
ADSL 1	0.5	0.5	35	1.04
ADSL 2	0.5	0.6	33.2	2.15
Custom	0.28	0.52	18	1.19

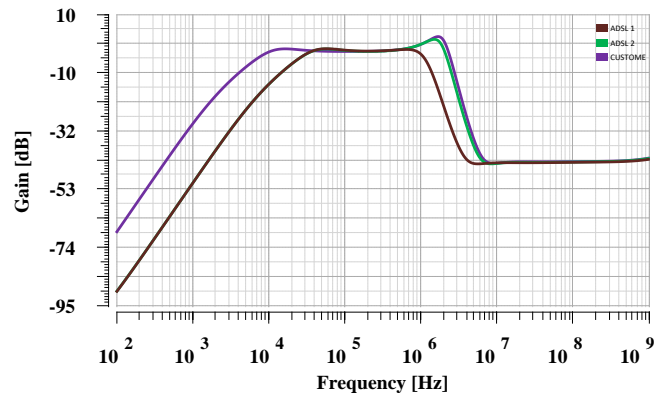


Fig. 11: The frequency response of the proposed filter for the bias voltage presented in Table 3.

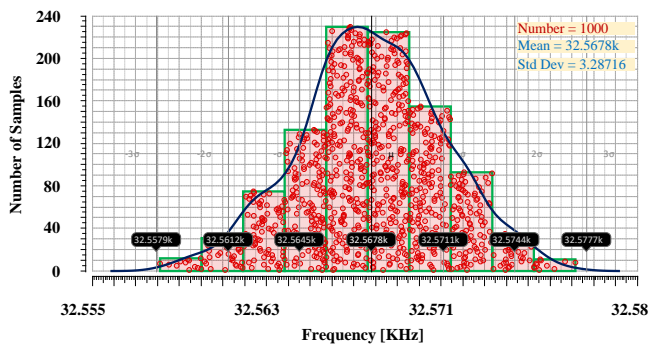
In order to investigate the effect of capacitor tolerance in the cutoff frequencies of the proposed filter, a Monte Carlo analysis is run. As shown in Fig. 12, the error created in the cut-off frequencies of the proposed filter is not significant for 2% deviation in the capacitance of the used capacitors. Table 4 provides the comparison of the proposed filter with previous reported band-pass filters. Among the reported band-pass filters, the proposed fourth-order band-pass filter exhibits a competitive feature despite the use of fewer amplifiers and lower capacitance.

4. CONCLUSIONS

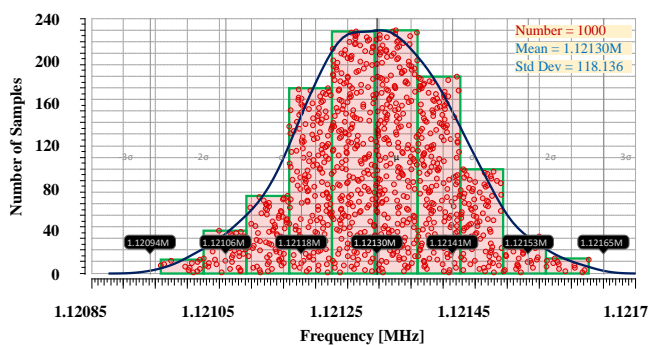
The OTA-C fourth-order band-pass filter, which has the ability to adjust the bandwidth, was designed with 180 nm CMOS technology and simulated using C.a.d.e.n.c.e software. Using the low-pass and high-pass filters, the proposed band-pass filter is designed that can be used as an anti-aliasing filter. In order to utilize the proposed anti-aliasing in the front-end of the ADC that is used in ADSL technologies, it is possible to adjust the lower and upper cut-off frequencies of the filter using two bias voltages. The simulation results showed that by adjusting the bias voltages, the bandwidth of the proposed filter can be tuned between 1.1 and 2.2 MHz. Also, the power consumption of the proposed filter is 327 μ W at supply voltage of 1.8V. The total capacitance and the number of trans-conductance amplifiers used in the proposed filter are 20 pF and 8, respectively.

Table 4: Comparison of proposed filter with the reported works.

	This work	[15]	[16]	[17]	[18]
Technology (μ m)	0.18	0.18	0.18	0.18	0.13
Filter type	Gm-C	Gm-C	Gm-C	Gm-C	Gm-C
Voltage (V)	1.8	1.8	1.2	1.8	0.6
Filter Order	4	8	2	5	1
Number of Amplifire	8	16	5	10	6
Central frequency (MHz)	1.12	0.5	0.13	1.17	2
Sum of Capacitance (pF)	20	88	24	45	30
Power consumption (μ W)	327	448	150	954	261



(a)



(b)

Fig. 12: Histograms of Monte Carlo analysis on the proposed filter (a) lower cut-off frequency (b) upper cut-off frequency.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Mohsen Ghaemmaghami: Formal analysis, Investigation, Investigation, Software, Roles/Writing - original draft. **Shahbaz Reyhani:** Methodology, Supervision, Validation, Writing - review & editing

DECLARATION OF COMPETING INTEREST

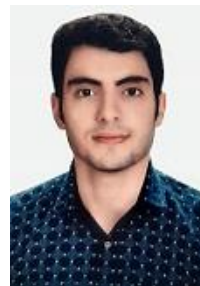
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy has been completely observed by the authors.

REFERENCES

- [1] Y. Zheng, Y. Zhao, N. Zhou, H. Wang, and D. Jiang, "A short review of some analog-to-digital converters resolution enhancement methods," *Measurement*, vol. 180, article 109554, 2021.
- [2] F. Maloberti, *Data converters*. Springer Science & Business Media, 2007.
- [3] Y. Sundarasaradula, T. G. Constandinou, and A. Thanachayanont, "A 6-bit, two-step, successive approximation logarithmic ADC for biomedical applications," in *2016 IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, 2016.

- [4] R. R. Thirrunavukkarasu, S. Kirthika, N. Nivetha, and N. Renuka, "Design of Sar-Adc based mixed signal architecture for low power application," in *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, 2019, pp.822-824.
- [5] T. Y. Lee, T. Butcher, T. Ishida, A. Panigada, and D. Meacham, "Digitally enhanced high speed ADC for low power wireless applications," in *2017 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*, 2017, pp.64-67.
- [6] M. Judy, A. Sodagar, R. Lotfi, and M. Sawan "Nonlinear signal-specific ADC for efficient neural recording in brain-machine interfaces" *IEEE Trans. Biomed. Circuits and Syst*, vol.8, no.3, pp.371-381, 2014.
- [7] B. C. Baker, "Anti-aliasing, analog filters for data acquisition systems," AN699, Microchip Technology Inc, 1999.
- [8] R. L. Geiger, and E. Sanchez-Sinencio, "Active filter design using operational transconductance amplifiers: A tutorial," *IEEE circuits and devices magazine*, vol, no.2, pp.20-32, 1985.
- [9] Q. Gu, *RF tunable devices and subsystems: Methods of modeling, analysis, and applications*. Springer International Publishing, 2015.
- [10] P. Sutthikarn, S. Chivapreecha, T. Jongsatoporn, and A. Trirat, "Tunable bandpass/bandstop digital filters based on 1 st-order allpass network instead of unit delay," in *5th International Conference on Engineering, Applied Sciences and Technology (ICEAST)*, 2019, pp.1-4.
- [11] C.-M. Chang, "New multifunction OTA-C biquads," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 46, no. 6, pp. 820-824, 1999.
- [12] A. Solaymanpour, and S. Reyhani, "Design of tunable low-power band-stop filter for elimination of 50 hz power-line noise," *Journal of Applied Research in Electrical Engineering*, vol. 1, no. 2, pp. 197-202, 2022.
- [13] C.-C. Hung, "Low voltage, low power CMOS analog circuit design techniques for mobile, portable VLSI applications," The Ohio State University, 1997.
- [14] P. Vervoort, and R. Wassenaar, "A CMOS VI converter with a constant transconductance for common and differential input voltages," in *International Conference on Electronics, Circuits & Systems 1995: Proceedings of the International Conference on Electronics, Circuits & Systems*, 1995, pp. 161-164.
- [15] G. B. Long, "A monolithic Gm-C filter based very low power, programmable, and multi-channel harmonic discrimination system using analog signal processing," M.Sc Thesis, University of Tennessee, 2019.
- [16] K. Garradhi, N. Hassen, T. Ettaghzouti, and K. Besbes, "A low voltage and low power OTA using bulk-driven technique and its application in Gm-c filter," in *2018 15th International Multi-Conference on Systems, Signals & Devices (SSD)*, 2018, pp.429-434.
- [17] B. Guthrie, J. Hughes, T. Sayers, and A. Spencer, "A CMOS gyrator low-IF filter for a dual-mode Bluetooth/ZigBee transceiver," *IEEE journal of solid-state circuits*, vol.40, no.9, pp.1872-1879, 2005.
- [18] G. Jha, M. A. Ibrahim, and M. Onabajo, "A low-power complex bandpass Gm-C filter with dynamic range expansion through adaptive biasing," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018, pp.1-5.

BIOGRAPHY



working with circuit design software.

Mohsen Ghaemmaghami was born in 1996. He received his B.Sc. degree in Electronic Engineering from the University of Ardakan, Yazd in Iran 2018. He is currently a graduate student in electronic engineering-integrated circuits at the University of Guilan, Rasht, Iran. His research interests include low-power integrated circuits, data converters, digital systems and



integrated circuits, data converters, digital systems and neural networks.

Shahbaz Reyhani was born in 1969. He received his B.Sc., M.Sc. and Ph.D. degree in electronic engineering from university of Guilan, Shiraz university and Shahid Beheshti university, respectively, in Iran. Dr. Reyhani is an assistant professor with department of electrical engineering, university of Guilan, Rasht, Iran. His research interests include low-power

Copyrights



© 2024 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).





Research Article

Nonlinear Observer-Based Control Design for a Three-Axis Inertial Stabilized Platform

Mohammad Mehdi Zohrei¹ , and Hamidreza Javanmardi^{2,*} 

¹ Department of Electrical and Electronic Engineering, Shiraz University of Technology, Shiraz 71557-13876, Iran

² School of Electrical and Computer Engineering, Shiraz University, Shiraz, Iran

* Corresponding Author: hamidjavanmardi@shirazu.ac.ir

Abstract: For decades in the aerospace and control sciences, the Inertial Stabilized Platform (ISP) system has been studied to improve the accuracy of recipient photos or target tracking. This paper presents a nonlinear observer-based control method for three Degrees Of Freedom (3-DOF) ISP systems. First, a new formula of the state space equation for the 3-DOF ISP system is proposed to make this model suitable for designing an observer-based control. Then, by measuring the angular positions as output feedback, the angular velocities are estimated by the nonlinear observer, and Lyapunov-based nonlinear control techniques are used to design the observer. Furthermore, the exponential stability and convergence of the observer system are proved. Finally, the auxiliary control signal is considered so that the dynamics of the designed observer become a simple linear form and are easily controlled by the state feedback controller. Simulation results illustrate the effectiveness and feasibility of the proposed control strategy.

Keywords: Exponential stability, inertial stable platform, lyapunov techniques, observer-based control, state feedback controller.

Article history

Received 01 February 2023; Revised 09 October 2023; Accepted 17 November 2023; Published online 29 January 2024.

© 2024 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)

How to cite this article

M. M. Zohrei, and H. Javanmardi, "Nonlinear observer-based control design for a three-axis inertial stabilized platform," *J. Appl. Res. Electr. Eng.*, vol. 2, no. 2, pp. 158-172, 2023. DOI: [10.22055/jaree.2023.42938.1066](https://doi.org/10.22055/jaree.2023.42938.1066)



1. INTRODUCTION

1.1. Motivation

Recently, receiving photos with high accuracy and quality from a moving target have received a lot of attention. Noise and disturbances from cameras and sensors mounted on moving vehicles such as airplanes, helicopters, boats, etc. have adverse effects and interfere with access to high-precision photos [1–3]. For this reason, the Inertial Stabilized Platform (ISP) is used to improve the accuracy of the recipient's photo or target tracking. The ISP plant is one of the most important parts of the inertial technology application system with broad applications in submarines, electronic telescopes, boats, satellites, helicopters, and spacecraft [4–7]. The ISP system including gimbals is used to isolate the movement of several cameras and sensors from changes in the angular position of the moving object [8]. Therefore, using this system, the accuracy of imaging and tracking the target is improved [9]. Moreover, various noises and disturbances exist, including internal disturbances such as mass imbalance torque, sensors measurement error and

friction force, and external disturbances generated by wind disturbances or moving items' motion and vibration [10]. In addition, the state space equation of the ISP system is very complex and has nonlinear dynamics. These statements show that achieving optimal performance control for the ISP system is challenging and has received much attention in recent decades [11].

1.2. Literature review

To improve the ISP plant performance, several control techniques have been suggested in the literature. In [12], a Proportional-Integral-Derivative (PID) controller is combined with a fuzzy technique to control the ISP system, although this method is not robust against disturbances and the controller performance is degraded against strong perturbations. Moreover, the performance of conventional PID controllers can be improved by combining them with the Neural Network approach [13, 14]. The backstepping technique is also suggested in some literature [15] but, it does not have satisfactory performance in the presence of disturbances and system uncertainties. To improve the control

performance against disturbances and uncertainties in the system, the controller must be robust. Therefore, several researchers have contributed to this field [16–18]. In [19–22], the H_∞ feedback controller is designed to be robust to various uncertainties and disturbances and has good performance. The nonlinear model predictive control is another efficient and robust controller that recently has been used for ISP systems [1, 23]. A sliding mode controller is also one of the most popular controller design techniques for the ISP system [24]. To improve the control performance, a combination of sliding mode techniques and a backstepping controller is widely used [25, 26]. Despite the advantages of this method, the sliding mode controller has a heavy switching gain, which causes chattering in the input signal control. The high switching in the control effort may damage the actuating motors to control the ISP system. In [8, 25], to decrease the chattering effect, the neural network approach has been used to estimate the disturbances imposed on the system. However, neural network training is usually time-consuming and can lead to weak transient performance in some complex systems. In addition, in all examined control techniques for ISP systems, it is assumed that all system variables are available and measurable. Whereas, in most practical systems, measuring all states of the system by direct observation is very costly and is usually accompanied by errors. An observer-based control design is one of the efficient approaches for accurately estimating the state of the ISP axis because it is simple and very applicable [27, 28]. Only a model of the desired system is required to design the observer. When an accurate numerical model of the plant is provided, the velocity observer can estimate an accurate steady state of the system. However, in dealing with an unknown real system model, various uncertainties and disturbances are always modeled, therefore accurate estimation of states is not so easy and also cannot be guaranteed [28, 29]. In the works of literature, various methods for estimating system states and then designing an observer-based controller are introduced. Such as extended state observer [30, 31], backstepping observer [32, 33], disturbance observer [34], and unknown input observer [35]. Also, in [36] the observer-based integral sliding mode control (SMC) for a class of stochastic linear flight systems subject to bounded disturbances is suggested. In [37], an adaptive observer design problem to estimate the parameters and state variables of Lipschitz systems has been investigated. The observer designed in [37] is suitable for systems that are non-linear in state variables and linear in unknown parameters. Also, the problem of designing a nonlinear adaptive observer for Lipschitz systems has been expanded in [38]. The gain for this nonlinear observer is systematically obtained using the linear matrix inequality technique.

1.3. Contributions

In this paper, an efficient simple-implementation observer-based controller method is proposed for a 3-DOF ISP system. To control the ISP system, in addition to angular positions, angular velocities with appropriate accuracy are required. However, high-precision gyroscopes are generally costly, and in addition, gyroscope rate drift often causes errors over time [39]. To reduce the above-mentioned problem of velocity gyros, the velocity observer is used to estimate the vector of velocity based on the Lyapunov theorem. For this purpose, first, using the kinetics and

kinematics of the ISP, the dynamic equations of each gimbal activated by the DC motor are obtained. Then the state variables of the ISP system are divided into two categories of angular position and angular velocity, and the state space equations are rewritten in a form that is suitable for designing a velocity observer. Next, with the help of Lyapunov-based nonlinear control techniques, a nonlinear observer is designed to improve the velocity estimation for the roll, pitch, and yaw gimbal of the ISP plant. By using the Lyapunov theorem, the overall ISP system stabilization and moving target tracking are guaranteed. Finally, the auxiliary control signal is considered in such a way that the dynamics of the designed observer become a simple linear form and is easily controlled by the state feedback controller. The suggested method is applied to the 3-DOF ISP case study, and the simulation results have shown the improvement of the designed controller.

Briefly, the presented paper has the following contributions and advantages over the current literature:

- i. Initially, a new formulation of the state-space equation is proposed for the 3-DOF ISP system so that this model is suitable for designing an observer-based control. For the designed controller to be robust against model uncertainties and external disturbances, the term of unknown bounded nonlinear disturbance in the ISP model is also considered in the driven model.
- ii. In almost all control strategies presented in the various literature for ISP systems, it is assumed that all system state variables with appropriate accuracy are available. However, one of the biggest challenges in the practical design of controllers for ISP systems, in addition to the issue of cost limitation and accuracy of measuring devices that always have errors, is the volume and weight added to the overall system. Therefore, in this paper, to improve the performance of the designed controller as well as reduce the implementation costs, it is assumed that the position vector θ is available and measurable but the angular velocity vector ω is not available which must be estimated. Therefore, it is no longer necessary to use expensive gyroscopes to directly measure the angular velocity, and an observer is used instead.
- iii. By using the state observer, one of the most important parts of the design is to prove the convergence of the estimated value to the actual value. In this paper, with the help of Lyapunov's theorem and by considering a new positive definite function, it is guaranteed that the state estimation error converges to zero asymptotically.
- iv. The suggested controller strategy is very simple and its implementation requires a low computational burden.

1.4. Paper Structure

The structure of the sections of this paper is as follows: in Section 2, the dynamic model of the 3-DOF ISP system is obtained and a suitable state space model for designing the velocity observer is derived. Section 3 presents the observer-based control method for the 3-DOF ISP plant and the exponential stability analysis. The simulation results are illustrated to show the effectiveness of the proposed method in Section 4. Finally, this paper is concluded in Section 5.

2. DYNAMIC OF 3-DOF ISP

Fig. 1 presents the configuration of the 3-DOF ISP plant which consists of the torque motor, outer loop (roll gimbal), middle loop (pitch gimbal), and inner loop (yaw gimbal). The imaging sensors and payloads are placed on the inner loop (yaw gimbal), and the position and orientation system (POS) is placed on the pitch gimbal to provide attitude-measured information for the rotation angles of the gimbal and ensure accurate pointing of the imaging sensors. In this section, first, the corresponding coordinate systems are defined and then a method for converting vectors from one coordinate system to another is introduced. The coordinate system of the 3-DOF ISP plant is shown in Fig. 1. By using Euler angles, the position of a fixed point can be defined as vector $\theta_A = [\theta_p \ \theta_r \ \theta_a]^T$, which (θ_p) is the position of the pitch angle, (θ_r) is a position of the roll angle, and (θ_a) is the position yaw angle. Moreover, roll-pitch-yaw velocities relating to the base coordinate are defined as vector $\dot{\theta}_A = [\dot{\theta}_p \ \dot{\theta}_r \ \dot{\theta}_a]^T$.

In Fig. 2 (x_r, y_r, z_r) , (x_a, y_a, z_a) , (x_p, y_p, z_p) and (x_b, y_b, z_b) are the roll coordinate, yaw coordinate, pitch coordinate, and base coordinate respectively.

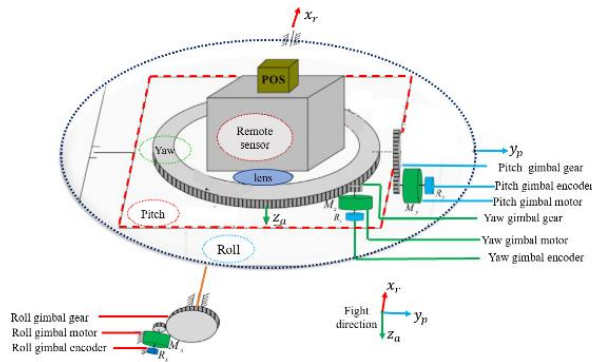


Fig. 1: The 3-DOF ISP configuration diagram [8].

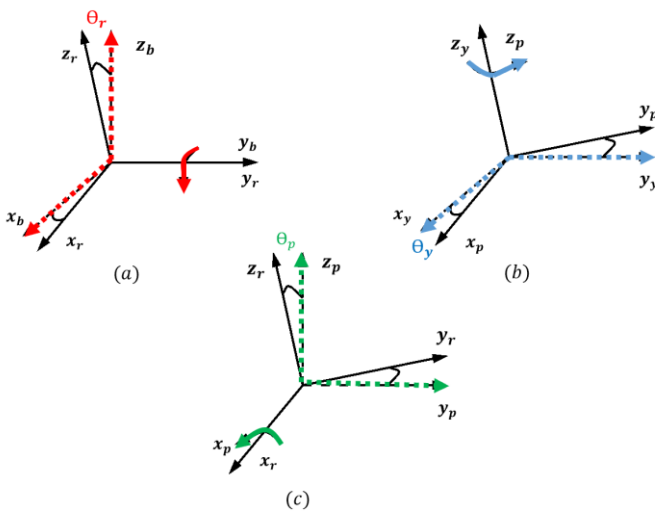


Fig. 2: coordinates of 3-DOF ISP plant: (a) roll gimbal, (b) yaw gimbal, (c) pitch gimbal.

2.1. Kinematics Model of the 3-DOF ISP Plant

According to Fig. 2, C_b^r , C_r^p , and C_p^a are transformation matrices from the base coordinate of the gimbal to the roll coordinate, the roll coordinate of the gimbal to the pitch coordinate, and the pitch of the gimbal coordinate to the yaw coordinate respectively defined as follows:

$$C_b^r = \begin{bmatrix} \cos(\theta_r) & 0 & -\sin(\theta_r) \\ 0 & 1 & 0 \\ \sin(\theta_r) & 0 & \cos(\theta_r) \end{bmatrix} \quad (1)$$

$$C_r^p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_p) & \sin(\theta_p) \\ 0 & -\sin(\theta_p) & \cos(\theta_p) \end{bmatrix} \quad (2)$$

$$C_p^a = \begin{bmatrix} \cos(\theta_a) & \sin(\theta_a) & 0 \\ -\sin(\theta_a) & \cos(\theta_a) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

According to the defined transformation matrices (1)-(3), the kinematics of the 3-DOF ISP plant in different coordinates are given as follows [8]:

$$\omega_{ir}^r = \begin{bmatrix} \omega_{irx}^r \\ \omega_{iry}^r \\ \omega_{irz}^r \end{bmatrix} = C_b^r \omega_{ib}^b + \begin{bmatrix} \dot{\theta}_r \\ 0 \\ 0 \end{bmatrix} \quad (4)$$

$$\omega_{ip}^p = \begin{bmatrix} \omega_{ipx}^p \\ \omega_{ipy}^p \\ \omega_{ipz}^p \end{bmatrix} = C_r^p \omega_{ir}^r + \begin{bmatrix} \dot{\theta}_p \\ 0 \\ 0 \end{bmatrix} \quad (5)$$

$$\omega_{ia}^a = \begin{bmatrix} \omega_{iax}^a \\ \omega_{ia y}^a \\ \omega_{iaz}^a \end{bmatrix} = C_p^a \omega_{ip}^p + \begin{bmatrix} 0 \\ 0 \\ \dot{\theta}_a \end{bmatrix} \quad (6)$$

where $\omega_{ir}^r = [\omega_{irx}^r \ \omega_{iry}^r \ \omega_{irz}^r]^T$, $\omega_{ip}^p = [\omega_{ipx}^p \ \omega_{ipy}^p \ \omega_{ipz}^p]^T$, $\omega_{ia}^a = [\omega_{iax}^a \ \omega_{ia y}^a \ \omega_{iaz}^a]^T$, and $\omega_{ib}^b = [\omega_{ibx}^b \ \omega_{iby}^b \ \omega_{ibz}^b]^T$ are the angular velocity vector in directions (x, y, z) of the roll-pitch-yaw and base, relative to the inertial coordinate, respectively.

2.2. The Dynamic Model of the 3-DOF ISP Plant

According to Newton–Euler theory, the dynamic model of the 3-DOF ISP plant can be written as follows [8]:

$$\dot{H}_k + \omega_k \times H_k = M_k, \quad k = a, p, r \quad (7)$$

In equation (7), $\omega_k = [\omega_{kx} \ \omega_{ky} \ \omega_{kz}]^T$ is defined as the vector of angular velocity and $H_k = [H_{kx} \ H_{ky} \ H_{kz}]^T$ is the vector of the angular momentum of inertia for the roll-pitch-yaw gimbal. In the following, the inertial angular momentum concerning the coordinates of inertia and the dynamic model of the roll, pitch, and yaw gimbal is defined.

2.2.1. The angular momentum of inertia and dynamic model of the pitch gimbal

The inertial angular momentum of the pitch gimbal is given as follows:

$$H_p = J_p \omega_{ip}^p + C_a^p H_a = \begin{bmatrix} (J_{px} + J_{ax}) \omega_{ipx}^p \\ (J_{py} + J_{ay}) \omega_{ipy}^p \\ (J_{pz} + J_{az}) \omega_{ipz}^p + J_{az} \dot{\theta}_a \end{bmatrix} \quad (8)$$

By substituting H_p and ω_{ip}^p to equation (7), the dynamic model of the pitch gimbal is defined by

$$\begin{aligned} & \left((J_{px} + J_{ax}) \omega_{ipx}^p \right)' \\ & + (J_{pz} + J_{az} - J_{py} - J_{ay}) \omega_{ipy}^p \omega_{ipz}^p + J_{az} \dot{\theta}_a \omega_{ipy}^p \\ & = M_{px} \end{aligned} \quad (9)$$

2.2.2. The angular momentum of inertia and dynamic model of the roll gimbal

The inertial angular momentum of the roll gimbal is given as follows:

$$H_r = J_r \omega_{ir}^r + C_p^r H_p \quad (10)$$

By substituting H_r and ω_{ir}^r to equation (7), the dynamic model of the roll gimbal is defined by

$$\begin{aligned} & [J_{ry} \omega_{iry}^r + (J_{py} + J_{ay}) \omega_{ipy}^p \cos \theta_p - J_{az} \dot{\theta}_a \sin \theta_p \\ & \quad - \sin \theta_p (J_{az} + J_{pz}) \omega_{ipz}^p] \\ & + (J_{rx} - J_{rz}) \omega_{irx}^r \omega_{irz}^r + (J_{ax} + J_{px}) \omega_{ipx}^p \omega_{irz}^r \\ & \quad - \sin \theta_p (J_{py} + J_{ay}) \omega_{ipy}^p \omega_{irx}^r \\ & - \cos \theta_p (J_{az} + J_{pz}) \omega_{ipz}^p \omega_{irx}^r \\ & \quad - \cos \theta_p J_{az} \dot{\theta}_a \omega_{irx}^r = M_{ry} \end{aligned} \quad (11)$$

2.2.3. The angular momentum of inertia and dynamic model of the yaw gimbal

The inertial angular momentum of the roll gimbal is given as follows:

$$H_a = J_a \omega_{ia}^a \quad (12)$$

By substituting H_a and ω_{ia}^a to equation (7), the dynamic model of the roll gimbal is defined by

$$J_{az} \dot{\omega}_{iaz}^a = M_{az} \quad (13)$$

where in the above equation, J_{mn} is the rotary momentums of the m gimbal around the n-axes.

Moreover, a DC motor with a gear-driven transmission is provided to generate the torque required to control the 3-DOF ISP plant. The dynamic equations of a gimbal gear-driven plant are considered as follows [8]:

$$\begin{cases} M_L = R_L F + T_{dL} \\ J_m \frac{d^2 \theta_{im}}{dt^2} = T_M - R_m F + T_{dm} \end{cases} \quad L = a, p, r \quad (14)$$

where, M_L is the torque applied on the gimbal, R_m is the armature resistance, T_{dL} is the total disturbance torque imposed on the gimbal, F is the interacting force between the two gears, J_m is the motor inertial, θ_{im} , θ_{il} and θ_{ir} are the attitudes of the motor, gimbal, and reference concerning the inertial space, respectively. Also T_M is the output torque of the motor is defined as follows:

$$T_M = K_t \frac{u - K_e (\dot{\theta}_{im} - \dot{\theta}_{ib})}{R_m} \quad (15)$$

In equation (15), K_t represents the motor torque constant, and K_e is known as the back Electromotive Force (EMF) constant. Moreover, the torque applied on each gimbal (M_L) is defined as follows:

$$M_L = NK_t \frac{u - K_e (\dot{\theta}_{il} - \dot{\theta}_{ib})}{R_m} + (NT_{dm} + T_{dL}) + N(N-1)J_m \dot{\theta}_{ib} - N^2 J_m \ddot{\theta}_{il} \quad (16)$$

By substituting (4)-(7) and (14) in (8)-(13), the dynamic model of the 3-DOF ISP plant with a DC motor and a gear-driven transmission for pitch, roll, and yaw is obtained as follows

$$\begin{aligned} & (J_{px} + J_{ax}) \omega_{ipx}^p \ddot{\theta}_p \\ & = -(J_{px} + J_{ax}) \dot{\omega}_{irx}^r \\ & \quad + J_{az} \dot{\theta}_a \omega_{ipy}^p \\ & \quad + (J_{pz} + J_{az} - J_{py} - J_{ay}) \omega_{ipy}^p \omega_{ipz}^p \\ & \quad + \frac{NK_t}{R_m} \left(u_p \right. \\ & \quad \left. - K_e N (\dot{\theta}_{ip}^p - \omega_{ibx}^a) \right) \\ & \quad + NT_{dm} + T_{dp} \\ & \quad + N(N-1)J_m \dot{\omega}_{ibx}^p \\ & \quad - N^2 J_m \ddot{\theta}_{ip}^p \end{aligned} \quad (17)$$

$$\begin{aligned} & (J_{ry} + \cos^2(\theta_p) (J_{ay} + J_{py})) \ddot{\theta}_r \\ & = \dot{\theta}_r \cos^2(\theta_p) \\ & \quad + (J_{ay} + J_{py}) \cos^2(\theta_p) \omega_{iby}^b \\ & \quad + (J_{ay} + J_{py}) \cos(\theta_p) \sin(\theta_p) \omega_{irz}^r \\ & \quad - (J_{az} + J_{pz}) \sin(\theta_p) \omega_{ipz}^p \\ & \quad - J_{az} \sin(\theta_p) \dot{\theta}_a \\ & \quad + (J_{rx} - J_{rz}) \omega_{irx}^r \omega_{irz}^r \\ & \quad + (J_{ax} + J_{px}) \omega_{ipx}^p \omega_{irz}^r \\ & \quad - (J_{ay} + J_{py}) \sin(\theta_p) \omega_{ipy}^p \omega_{irx}^r \\ & \quad - (J_{az} + J_{pz}) \cos(\theta_p) \omega_{ipz}^p \omega_{irx}^r \\ & \quad - J_{az} \cos(\theta_p) \dot{\theta}_a \omega_{irx}^r \\ & \quad + \frac{NK_t}{R_m} \left(u_r \right. \\ & \quad \left. - K_e N (\dot{\theta}_{ir}^r - \omega_{iby}^a) \right) \\ & \quad + NT_{dm} + T_{dr} \\ & \quad + N(N-1)J_m \dot{\omega}_{iby}^p \\ & \quad - N^2 J_m \ddot{\theta}_{ir}^r \end{aligned} \quad (18)$$

$$\begin{aligned}
 J_{az}\ddot{\theta}_a &= -J_{az}\dot{\omega}_{ipz}^p \\
 &+ \frac{NK_t}{R_m} \left(u_a \right. \\
 &\left. - K_e N (\dot{\theta}_{ia}^a - \omega_{ibz}^a) \right) \\
 &+ (NT_{dm} + T_{da}) \\
 &+ N(N-1)J_m\omega_{ibz}^a \\
 &- N^2J_m\ddot{\theta}_{ia}^a
 \end{aligned} \tag{19}$$

2.3. Deriving the State Space Equations of the 3-DOF ISP Plant

Remark 1: According to the state space obtained in [8], it can be seen that this state space equation is not affine with respect to the angular velocity variable, and therefore it is not suitable for designing a nonlinear observer. To solve this problem, in the following, a new formulation of the state-space equation for 3-DOF ISP plant is obtained, which is affine with respect to the angular velocity.

According to the equations presented in (17)-(19), the nonlinear state space model for a 3-DOF ISP plant with assumptions that $x_1 = \theta_p, x_2 = \dot{\theta}_p, x_3 = \theta_r, x_4 = \dot{\theta}_r, x_5 = \theta_a, x_6 = \dot{\theta}_a, y_1 = \theta_p, y_2 = \theta_r$ and $y_3 = \theta_a$ can be modeled as follows:

$$\begin{cases} \dot{\theta} = \omega \\ \dot{\omega} = h(\theta) + H(\theta)\omega + B(\theta)u \end{cases} \tag{20}$$

where $\theta = [\theta_p \ \theta_r \ \theta_a]^T$ is the angular position vector and $\omega = [\dot{\theta}_p \ \dot{\theta}_r \ \dot{\theta}_a]^T$ is the angular velocity vector of the gimbal. In equation (20), the matrix $B(\theta)$ is a 3×3 diagonal matrix as follows:

$$B(\theta) = \begin{bmatrix} \frac{NK_t}{(J_{px} + J_{ax})R_m} & 0 & 0 \\ 0 & \frac{NK_t}{(J_{ry} + (J_{ay} + J_{py})\cos^2(\theta_p))R_m} & 0 \\ 0 & 0 & \frac{NK_t}{J_{az}R_m} \end{bmatrix} \tag{21}$$

Also, matrix $H(\theta)$ is a functional matrix whose elements include only boundary functions such as $\sin(\theta)$ and $\cos(\theta)$ and is defined as follows:

$$H(\theta) = \begin{bmatrix} 0 & H_{12}(\theta) & H_{13}(\theta) \\ H_{21}(\theta) & H_{22}(\theta) & H_{23}(\theta) \\ H_{31}(\theta) & H_{32}(\theta) & 0 \end{bmatrix} \tag{22}$$

Finally, vector $h(\theta)$ is also expressed as follows:

$$h(\theta) = [h_1(\theta) \ h_2(\theta) \ h_3(\theta)] \tag{23}$$

The elements of matrix $H(\theta)$ and vector $h(\theta)$ are introduced in Appendix.

In the state space equations obtained in (20), to improve the performance of the designed controller as well as reduce the implementation costs, it is assumed that the position vector θ is available and measurable and the angular velocity vector ω is unmeasurable which must be estimated.

3. PROPOSED OBSERVER-BASED CONTROL

The suggested observer for dealing with the nonlinear model (20), in addition to the ability to estimate the angular velocity with very high accuracy, should be designed in such a way that it is possible to design a simple and efficient controller for the observed system. For this purpose, a nonlinear observer-based state feedback controller will be designed. In the following, some preliminary Lemmas for designing the proposed controller strategy are provided.

3.1. Preliminaries

In this subsection, some preliminary Lemmas are presented.

Lemma 1 (A Lyapunov exponential stability theorem) [40]: The nonlinear system $\dot{x} = f(x)$ is exponentially stable, if there exists a continuously differentiable function $V(x)$ and positive constants ρ_1, ρ_2, ρ_3 , and τ such that

$$\begin{aligned}
 \rho_1 \|x\|^\tau &\leq V(x) \leq \rho_2 \|x\|^\tau \\
 \frac{\partial V}{\partial x} \cdot f(x) &\leq -\rho_3 \|x\|^\tau
 \end{aligned} \tag{24}$$

where $\|x\| = \sqrt{x^T x}$ is defined as the Euclidean norm.

Proof: see [40].

Lemma 2: For any variables $\alpha, \beta \in \mathbb{R}^n$ the inequality (25) is always established:

$$\alpha^T \beta \leq \frac{1}{2} \|\alpha\|^2 + \frac{1}{2} \|\beta\|^2 \tag{25}$$

Proof:

$$\begin{aligned}
 \|\alpha - \beta\|^2 \geq 0 &\rightarrow \|\alpha\|^2 + \|\beta\|^2 - 2\alpha^T \beta \geq 0 \\
 &\rightarrow \alpha^T \beta \leq \frac{1}{2} \|\alpha\|^2 + \frac{1}{2} \|\beta\|^2
 \end{aligned} \tag{26}$$

Lemma 3: For any bounded matrix function $A(\theta) \in \mathbb{R}^{n \times n}$, the following inequalities always hold:

$$\|A(\theta)\| \leq \frac{1}{2} \gamma_1, \quad \gamma_1 > 1 \tag{27}$$

$$A(\theta) + A^T(\theta) \leq \frac{1}{2} \gamma_2 I, \quad \gamma_2 > 1 \tag{28}$$

where $\|A\| = \sqrt{\lambda_{\max}(A^T A)} = \sigma_{\max}(A)$ is defined as the spectral norm of a matrix A and $\sigma_{\max}(\cdot)$ represents the largest singular value of matrix A [41].

Proof:

a)

$$\begin{aligned}
 \|A(\theta)\| &= \sqrt{\lambda_{\max}(A^T(\theta)A(\theta))} = \sigma_{\max}(A(\theta)) \\
 &\leq \left(\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}} = \left(\sum_{k=1}^n \sigma_k \right)^{\frac{1}{2}}
 \end{aligned}$$

Therefore, $\gamma_1 = 2 * \sup(\sum_{k=1}^n \sigma_k)^{\frac{1}{2}}$, where σ_k represents the singular value of matrix A .

b)

$$A(\theta) + A^T(\theta) \leq |\lambda_{\max}\{A(\theta) + A^T(\theta)\}| I$$

Therefore, $\gamma_2 = 2 * \sup(|\lambda_{max}\{A(\theta) + A^T(\theta)\}|)$, where $\lambda_{max}(\cdot)$ represents the eigen value of matrix $A + A^T$.

3.2. Nonlinear Observer Design

In the first step, to design an observer for the 3-DIF ISP plant, we need to simplify equations (20). As can be seen, all elements on the main diagonal of the matrix $B(\theta)$ in equation (23) are always positive constants that depend on the parameters of the ISP system. This means that the $B(\theta)$ is always a non-singular and invertible matrix. Next, to eliminate the nonlinear term $h(\theta)$, the input vector u is considered as follows:

$$u = B^{-1}(\theta)(-h(\theta) + v) \tag{29}$$

In equation (29), the vector v is defined as the new input of the system, which will be determined in the following design steps. By substituting the control signal (29) in the system state space equation (20), one has:

$$\begin{cases} \dot{\theta} = \omega \\ \dot{\omega} = H(\theta)\omega + v \end{cases} \tag{30}$$

As mentioned, it is assumed that the position vector θ is available and measurable but the angular velocity vector ω is not available and must be estimated. Hence, there is a need for an observer who can estimate the internal states of the system. Assume that $\hat{\theta}$ and $\hat{\omega}$ are the estimated values of the internal states of the system. Then, the observer equations can be considered as follows:

$$\begin{cases} \dot{\hat{\theta}} = k_1\tilde{\theta} + \hat{\omega} \\ \dot{\hat{\omega}} = k_2\tilde{\theta} + H(\theta)\hat{\omega} + v \end{cases} \tag{31}$$

Note that the above estimation rules are applicable because only the available variables ($\hat{\omega}$, $\hat{\theta}$ and θ) are used and the non-measurable variable ω is not used. Also, the parameters k_1 and k_2 are constant coefficients that are determined according to the suggested Lyapunov function in the next steps. By definition $\tilde{\theta} = \theta - \hat{\theta}$ and $\tilde{\omega} = \omega - \hat{\omega}$ as the estimated system state error, one has:

$$\begin{aligned} \dot{\tilde{\theta}} &= \{\dot{\theta} - \dot{\hat{\theta}}\} = \{(\omega) - (k_1\tilde{\theta} + \hat{\omega})\} \\ &= -k_1\tilde{\theta} + (\omega - \hat{\omega}) \\ &= -k_1\tilde{\theta} + \tilde{\omega} \\ \dot{\tilde{\omega}} &= \{\dot{\omega} - \dot{\hat{\omega}}\} = (H(\theta)\omega + v) \\ &\quad - (k_2\tilde{\theta} + H(\theta)\hat{\omega} + v) \\ &= -k_2\tilde{\theta} + H(\theta)(\omega - \hat{\omega}) \\ &= -k_2\tilde{\theta} + H(\theta)\tilde{\omega} \end{aligned} \tag{32}$$

Consider the following Lyapunov candidate:

$$V_o = \begin{bmatrix} \tilde{\theta}^T & \tilde{\omega}^T \end{bmatrix} \begin{bmatrix} \gamma^2 & -\frac{1}{2}\gamma \\ -\frac{1}{2}\gamma & \frac{2}{3} \end{bmatrix} \begin{bmatrix} \tilde{\theta} \\ \tilde{\omega} \end{bmatrix} \tag{33}$$

$$= \gamma^2\tilde{\theta}^T\tilde{\theta} - \gamma\tilde{\theta}^T\tilde{\omega} + \frac{2}{3}\tilde{\omega}^T\tilde{\omega}, \quad \gamma = \max\{\gamma_1, \gamma_2\}$$

In matrix function (33), $\gamma > 1$, so $\begin{bmatrix} \gamma^2 & -\frac{1}{2}\gamma \\ -\frac{1}{2}\gamma & \frac{2}{3} \end{bmatrix} > 0$ and therefore, the Lyapunov function (33) is positive definite.

The time derivative of the Lyapunov function V_o is obtained by:

$$\dot{V}_o = 2\gamma^2\tilde{\theta}^T\dot{\tilde{\theta}} - \gamma\tilde{\theta}^T\dot{\tilde{\omega}} - \gamma\tilde{\omega}^T\dot{\tilde{\theta}} + \frac{4}{3}\tilde{\omega}^T\dot{\tilde{\omega}} \tag{34}$$

By considering the equation of the estimated system state error (32), the time derivative of the Lyapunov function (33) equals to:

$$\begin{aligned} \dot{V}_o &= 2\gamma^2\tilde{\theta}^T(-k_1\tilde{\theta} + \tilde{\omega}) \\ &\quad - \gamma\tilde{\theta}^T(-k_2\tilde{\theta} + H(\theta)\tilde{\omega}) \\ &\quad - \gamma\tilde{\omega}^T(-k_1\tilde{\theta} + \tilde{\omega}) \\ &\quad + \frac{4}{3}\tilde{\omega}^T(-k_2\tilde{\theta} + H(\theta)\tilde{\omega}) \\ &= (-2\gamma^2k_1 + \gamma k_2)\tilde{\theta}^T\tilde{\theta} \\ &\quad + \left(2\gamma^2 + k_1\gamma - \frac{4}{3}k_2\right)\tilde{\theta}^T\tilde{\omega} \\ &\quad - \gamma\tilde{\theta}^TH(\theta)\tilde{\omega} + \tilde{\omega}^T(-\gamma I \\ &\quad + \frac{4}{3}H(\theta))\tilde{\omega} \end{aligned} \tag{35}$$

To cancel the cross-product terms, we choose:

$$2\gamma^2 + k_1\gamma - \frac{4}{3}k_2 = 0 \rightarrow \begin{cases} k_1 = 2\gamma \\ k_2 = 3\gamma^2 \end{cases} \tag{36}$$

Therefore, substituting (36) in (35) leads

$$\begin{aligned} \dot{V}_o &= -\gamma^3\|\tilde{\theta}\|^2 - \gamma\tilde{\theta}^TH(\theta)\tilde{\omega} \\ &\quad + \tilde{\omega}^T(-\gamma I \\ &\quad + \frac{2}{3}(H(\theta) + H^T(\theta)))\tilde{\omega} \end{aligned} \tag{37}$$

According to Lemma 3:

$$\begin{aligned} |\gamma\tilde{\theta}^TH(\theta)\tilde{\omega}| &\leq \gamma\|\tilde{\theta}\|\|\tilde{\omega}\|\|H(\theta)\| \\ &\stackrel{\|H(\theta)\| \leq \frac{1}{2}\gamma}{\implies} |\gamma\tilde{\theta}^TH(\theta)\tilde{\omega}| \\ &\leq \frac{1}{2}\gamma^2\|\tilde{\theta}\|\|\tilde{\omega}\| \end{aligned} \tag{38}$$

Then with the help of Lemma 2:

$$\begin{aligned} \frac{1}{2}\gamma^2\|\tilde{\theta}\|\|\tilde{\omega}\| &= \frac{1}{2}\left(\gamma^{\frac{3}{2}}\|\tilde{\theta}\|\right)\left(\gamma^{\frac{1}{2}}\|\tilde{\omega}\|\right) \\ &\leq \frac{1}{4}\gamma^3\|\tilde{\theta}\|^2 + \frac{1}{4}\gamma\|\tilde{\omega}\|^2 \end{aligned} \tag{39}$$

substituting (28) and (39) in (37) yields:

$$\begin{aligned} \dot{V}_o &\leq -\frac{3}{4}\gamma^3\|\tilde{\theta}\|^2 - \frac{5}{12}\gamma\|\tilde{\omega}\|^2 \\ &\leq -\frac{\gamma}{3}(\|\tilde{\theta}\|^2 + \|\tilde{\omega}\|^2) \end{aligned} \tag{40}$$

The following theorem expresses the convergence of the designed observer.

Theorem 1: By designing the dynamics of the observer for system (30) in the form of equation (31), the angular

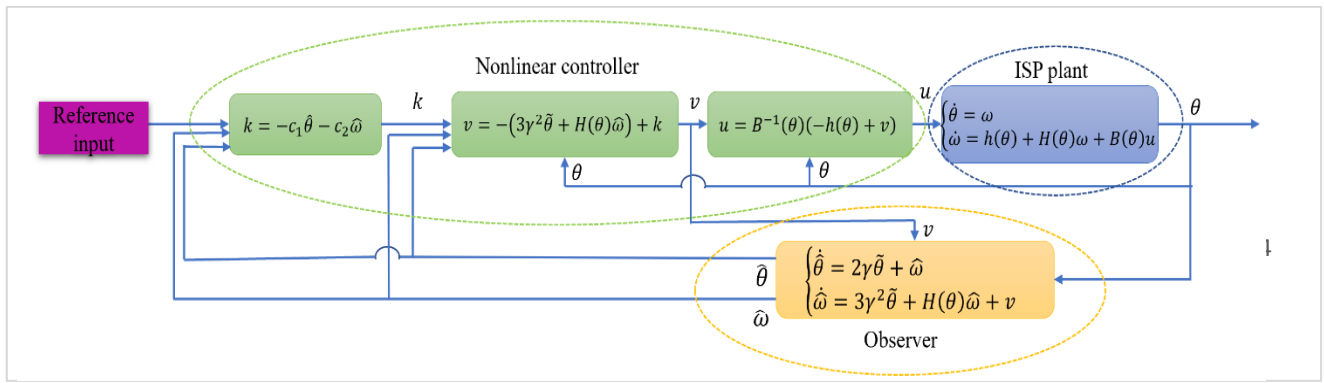


Fig. 3: Structure of the observer-based control strategy.

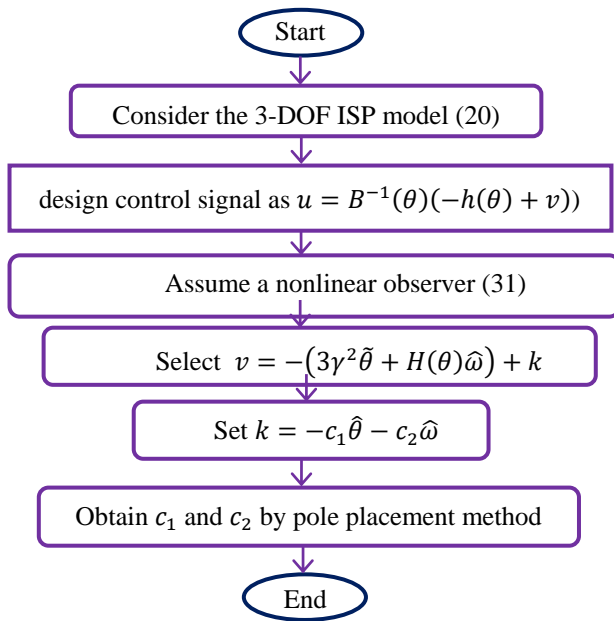


Fig. 4: The algorithm for designing observer-based control method for 3-DOF ISP system.

Table 1: Parameters of an ISP plant.

Parameter	Definition	Value
R_m	Armature resistance	7.56 Ω
K_t	Motor torque constant	0.127 Nm/Amp
K_e	Back EMF constant	0.128 V/rad/s
N	Drive ratio	104.7
J_L	Payload inertial	2.2 Kg m^2
J_m	Motor inertial	1.8006e-5 Kg m^2

$$T_{dm} = 0.06(\sin(t/5) + \sin(t/10)) \quad (53)$$

Also, to evaluate the robustness of the suggestion control strategy against the model uncertainties, 9 uncertain parameters have been added as shown in Table 2.

To evaluate the performance of the suggested observer-based control method, two simulation scenarios are considered. In the first scenario, only the ability to stabilize the proposed control is examined. In the second one, a different profile is considered for the roll, pitch, and yaw gimbal angle, and the controller's ability to track different outputs is evaluated.

Table 2: Variation Range of the moments of inertia [25].

J_p	J_a	J_r
$J_{px} = 0.769 + 0.307 * rand(t) - 0.$	$J_{ax} = 0.975 + 0.390 * rand(t) - 0.$	$J_{rx} = 1.048 + 0.419 * rand(t) - 0.$
$J_{py} = 0.925 + 0.381 * rand(t) - 0.$	$J_{ay} = 2.546 + 1.018 * rand(t) - 0.$	$J_{ry} = 2.763 + 1.105 * rand(t) - 0.$
$J_{pz} = 1.016 + 0.406 * rand(t) - 0.$	$J_{az} = 2.732 + 1.093 * rand(t) - 0.$	$J_{rz} = 2.984 + 1.194 * rand(t) - 0.$

4.1. Scenario 1 (Stabilization Results)

In this scenario, the steady-state equations of the 3-DOF ISP model (20), dynamics of the nonlinear observer (31), and the dynamics of the state feedback control system (49) have been simulated in MATLAB and the details of the stabilization control responses obtained by applying the proposed observer-based controller are provided in Figs. 5-8.

Figs. 5a and 5b show the pitch positions θ_p , and pitch velocities $\dot{\theta}_p$. As can be seen in Figs. 5a and 5b, the θ_p and $\dot{\theta}_p$ are stabilized by the designed controller and converge to zero in about 0.5 seconds. Figs. 6a and 6b shows that the performance of the proposed controller is in the acceptable range and it can stabilize roll positions θ_r and roll velocities $\dot{\theta}_r$, in less than 0.5 seconds. Figs. 7a and 7b illustrates that the stabilization time for regulated θ_a and $\dot{\theta}_a$ is about 0.45 seconds. Finally, Figs. 8a-8c shows the smooth control input signals u_p , u_r , and u_a , respectively. Based on these figures, it is clear that the designed observer has a very good performance and can accurately estimate the speed and position of each roll, pitch, and yaw gimbal angle in a short time.

4.2. Scenario 2 (Tracking Results)

To demonstrate the effectiveness and proper performance of the proposed controller strategy, the case of tracking desired references has been examined. In the following, we consider a different profile for each of the gimbal angles and examine the performance of the proposed controller.

4.2.1. Pitch position

It is assumed that the desired reference of the pitch angle is per with the following profile:

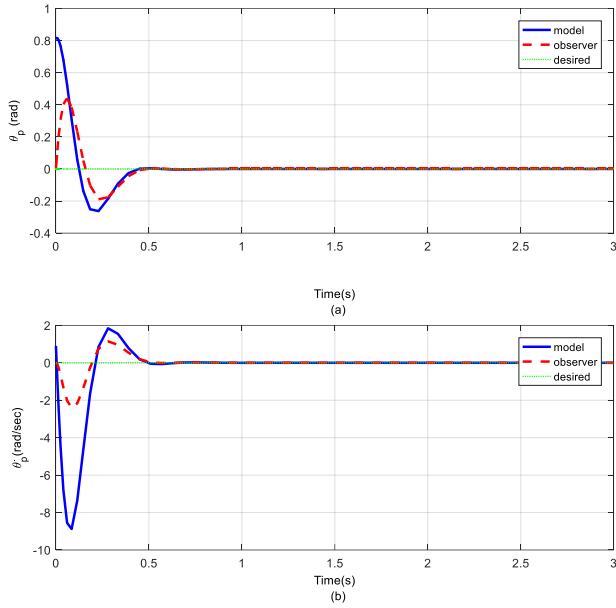


Fig. 5: State response of close-loop 3-DOF ISP system for scenario 1, (a) The pitch positions (θ_p), (b) The pitch velocities ($\dot{\theta}_p$).

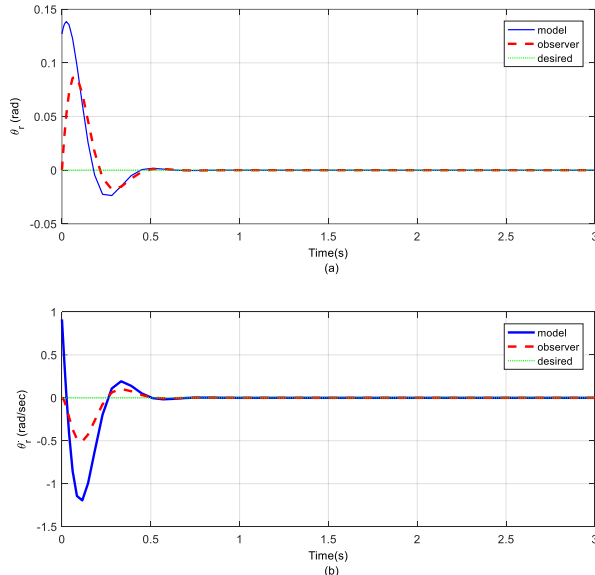


Fig. 6: State response of close-loop 3-DOF ISP system for scenario 1, (a) The roll positions (θ_r), (b) The roll velocities ($\dot{\theta}_r$).

$$\begin{cases} 1 & 1 \leq t < 5 \\ 1.5 & 5 \leq t < 10 \\ 1 & 10 \leq t < 15 \\ 0.1 * \sin(t) & 15 \leq t < 30 \end{cases} \quad (54)$$

By applying the proposed controller, the closed-loop pitch position and pitch velocities to track the reference input (54) for the observer and the model is shown in Fig. 9.

4.2.2. Roll position

The desired input to track roll angle is assumed to be as follows:

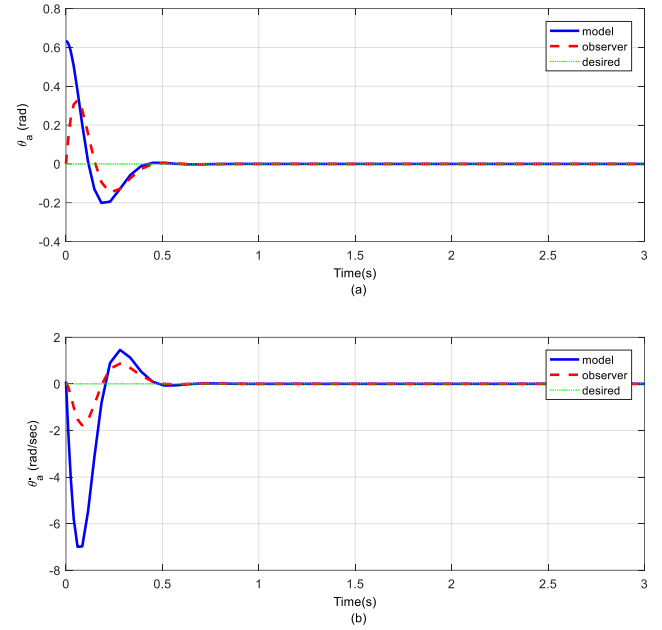


Fig. 7: State response of close-loop 3-DOF ISP system for scenario 1, (a): The yaw positions (θ_a), (b): The yaw velocities ($\dot{\theta}_a$).

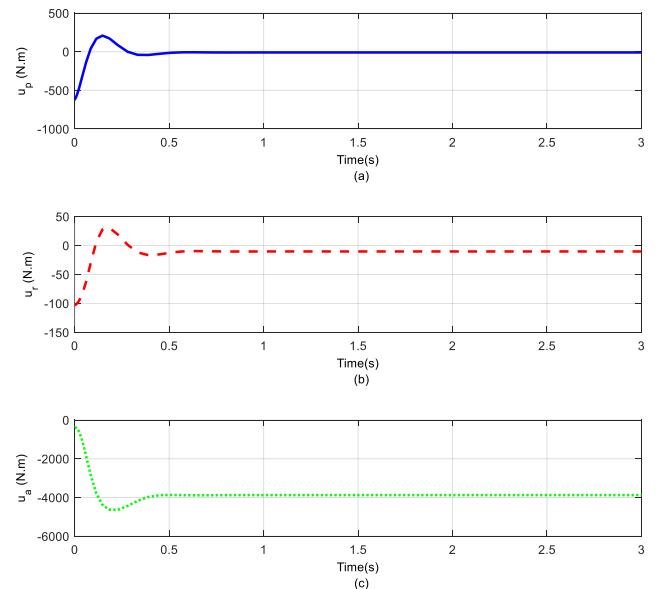


Fig. 8: Control signal of close-loop 3-DOF ISP system for scenario 1, (a) u_p , (b) u_r , (c) u_a .

$$\begin{cases} -2 & 1 \leq t < 5 \\ -1.5 & 5 \leq t < 10 \\ -1 & 10 \leq t < 15 \\ 0.1 * \sin(t) & 15 \leq t < 30 \end{cases} \quad (55)$$

How to track the profile (55) for the roll angle state is also illustrated in Fig. 10.

4.2.3. Yaw position

For the yaw angular variable, the desired input is considered as profile (56):

$$\begin{cases} 3 & 1 \leq t < 5 \\ 2.5 & 5 \leq t < 10 \\ 2 & 10 \leq t < 15 \\ 0.1 * \sin(t) & 15 \leq t < 30 \end{cases} \quad (56)$$

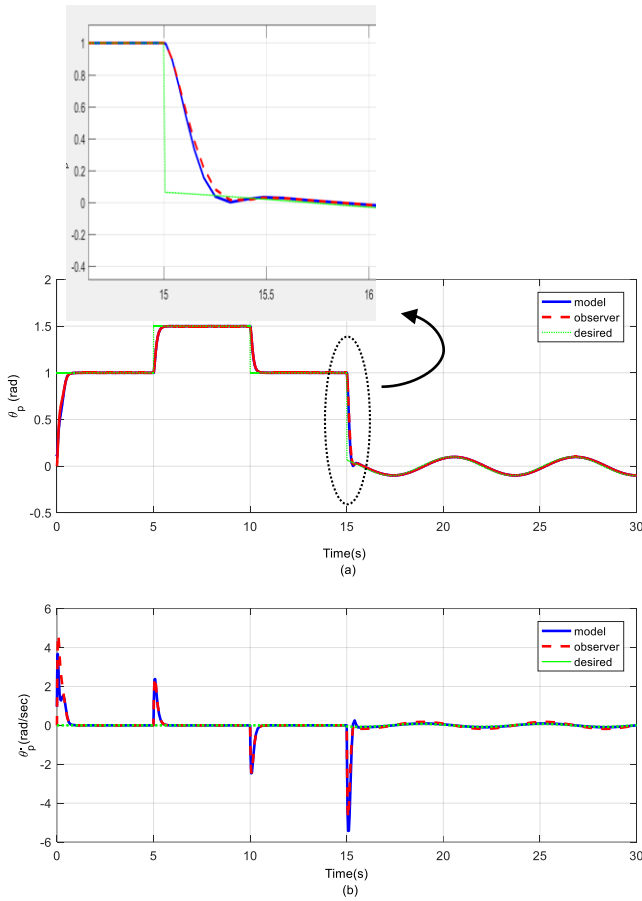


Fig. 9: State response of close-loop 3-DOF ISP system for scenario 2, (a) The pitch positions (θ_p), (b) The pitch velocities ($\dot{\theta}_p$).

The variations of the yaw angular variable to follow the above path are shown in Fig. 11.

Figs. 9-11 reveal that the angular positions of pitch, roll and yaw are well-tracked by the proposed control strategy. As can be seen, the closed-loop system is well able to follow the assumed paths for all three system states, and the amount of tracking error is acceptable. Finally, Fig. 12 shows the control efforts u_p , u_r , and u_a for Scenario 2.

Remark 3: It is important to mention that in most of the methods proposed to control the position of the ISP system in the current literature, it has been assumed that all required states with appropriate accuracy are available. In this paper, to reduce the number of measuring devices and also to increase the accuracy of the angular velocity value, instead of expensive gyroscopes, a velocity observer has been used. Also, in theorem 1, it was proved that the error of the designed nonlinear observer exponentially tends to zero and this allows us to estimate the angular velocity with very good accuracy. Given the above, comparing the observer-based control method proposed in this paper with previous methods is not fair because the performance of the output feedback controller at its best is the same performance as the state feedback controller when all system variables are available. To evaluate the performance and quality of the suggested observer algorithm, the following two performance indexes are utilized.

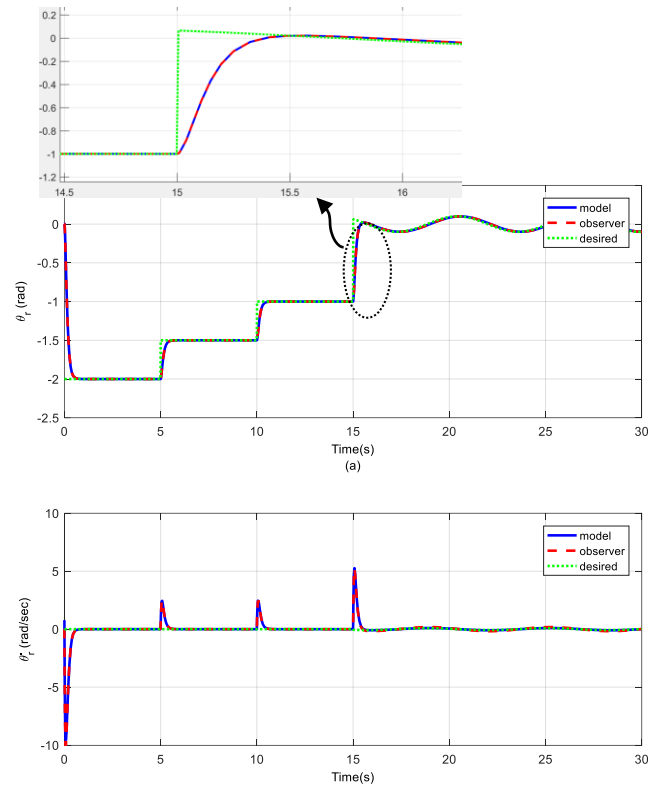


Fig. 10: State response of close-loop 3-DOF ISP system for scenario 2, (a) The roll positions (θ_r), (b) The roll velocities ($\dot{\theta}_r$).

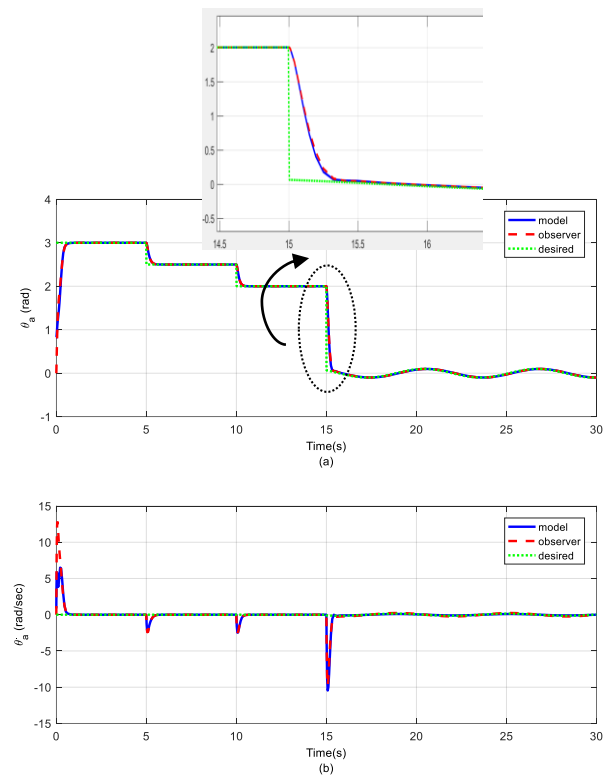


Fig. 11: State response of close-loop 3-DOF ISP system for scenario 2, (a) The yaw positions (θ_a), (b) The yaw velocities ($\dot{\theta}_a$).

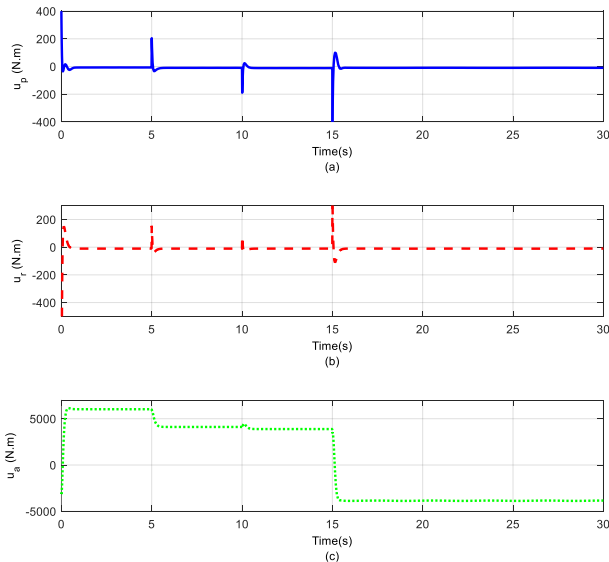


Fig. 12: Control signal of close-loop 3-DOF ISP system for scenario 2, (a) u_p , (b) u_r , (c) u_a .

4.2.4. The Integral Square Error (ISE) performance index

ISE is one of the indexes used to show the level of deviation of tracking errors in control systems and is defined as follows:

$$ISE = \int_0^{t_s} (y_{obs}(t) - y_{sys}(t))^2 dt \quad (57)$$

where $y_{obs}(t)$ and y_{sys} are defined as the observer output and system output respectively, and t_s represents the simulation time.

4.2.5. The Maximum Absolute Value (MAV) performance index

MAV criterion is used to evaluate the tracking accuracy, which is defined as formula (58):

$$MAV = \max_{t \in [0, t_s]} |y_{obs}(t) - y_{sys}(t)| \quad (58)$$

The browser performance for all variables in both scenarios is summarized in Table 3. Based on this table, it can be seen that the designed observer has a satisfactory performance and has been able to estimate the system variables with good accuracy.

Remark 4: Compared to other existing control strategies, the superiority of the strategy proposed in this paper is in cases such as:

- 1- Reducing the cost of construction and reducing the weight of the control system connected to the moving object through the removal of expensive gyroscopes
- 2- Easy implementation, and its efficiency.

3- CONCLUSION

This paper suggested the issue of observer-based control of the 3-DOF ISP system. To control ISP systems, in addition to measuring the angular position, angular velocity is also

Table 3: Parameter’s performance indexes for two scenarios.

Index	State variable	Value	
		Scenario 1	Scenario 2
ISE	pitch positions (θ_p)	0.0005	0.0018
	pitch velocities ($\dot{\theta}_p$)	0.0877	0.8458
	roll positions (θ_r)	0.0236	6.34e-5
	roll velocities ($\dot{\theta}_r$)	3.0731	0.0862
	yaw positions (θ_a)	0.0084	0.0210
	yaw velocities ($\dot{\theta}_a$)	1.1840	3.277
MAV	pitch positions (θ_p)	0.1190	0.1067
	pitch velocities ($\dot{\theta}_p$)	0.8483	1.4782
	roll positions (θ_r)	0.8436	0.0242
	roll velocities ($\dot{\theta}_r$)	6.2314	0.4230
	yaw positions (θ_a)	0.4053	0.7324
	yaw velocities ($\dot{\theta}_a$)	3.7367	6.304

required. Gyroscopes used to accurately measure angular velocity are expensive and also always have error measurement values due to gyroscope drift. On the other hand, there are limitations on the weight and volume of stabilization systems, therefore if angular velocity can be measured accurately and at the lowest possible cost, it is possible to achieve high performance and more economically in practical design. To reduce the expressed problems, the velocity observer is used to estimate the velocity vector based on the Lyapunov theorem. For this purpose, first, to design an efficient observer, the state space equation for the ISP system is rewritten to a new form. Then with the help of the Lyapunov-based nonlinear control technique, a nonlinear observer is designed to improve the velocity estimation of the ISP plant. Finally, through Theorem 1, the exponential stability and convergence of the observer system are proved. The simulation results showed that the proposed control method is efficient and can track different sources well. In future work, we can suggest some implementation issues that may come up in experimental settings, such as constraints on system state variables, state time delay, or input saturation for designing the high-performance ISP controller.

APPENDIX

Based on equations (17) - (19), the elements of the $H(\theta)$ and $h(\theta)$ are obtained as follows:

$$\begin{aligned}
 H_{12}(\theta) = & \omega_{ibz}^b * \cos(\theta_r) + \omega_{ibx}^b * \sin(\theta_r) \\
 & + (\sin(\theta_p) * (\sin(\theta_p) * (\omega_{ibz}^b \\
 & * \cos(\theta_r) - \omega_{ibx}^b * \sin(\theta_r)) \\
 & + \omega_{iby}^b * \cos(\theta_p)) * (J_{ay} - J_{az} \\
 & + J_{py} - J_{pz}) / (J_{ax} + J_{px}) \\
 & - (\cos(\theta_p) * (\cos(\theta_p) * (\omega_{ibz}^b \\
 & * \cos(\theta_r) - \omega_{ibx}^b * \sin(\theta_r)) \\
 & - \omega_{iby}^b * \sin(\theta_p)) * (J_{ay} - J_{az} \\
 & + J_{py} - J_{pz}) / (J_{ax} + J_{px})
 \end{aligned}$$

$$\begin{aligned}
 H_{13}(\theta) = & (J_{az} * (\sin(\theta_p) * (\omega_{ibz}^b * \cos(\theta_r) - \omega_{ibx}^b \\
 & * \sin(\theta_r)) + \omega_{iby}^b \\
 & * \cos(\theta_p)) / (J_{ax} + J_{px})
 \end{aligned}$$

$$\begin{aligned}
 H_{21}(\theta) = & -(J_{ax} * \omega_{ibx}^b * \sin(\theta_r) + J_{px} * \omega_{ibx}^b \\
 & * \sin(\theta_r) - J_{ax} * \omega_{ibz}^b * \cos(\theta_r) \\
 & - J_{px} * \omega_{ibz}^b * \cos(\theta_r)) / (J_{ry} + J_{ay} \\
 & * \cos^2(\theta_p) + J_{py} * \cos^2(\theta_p))
 \end{aligned}$$

$$\begin{aligned}
 H_{22}(\theta) = & (J_{ay} * \cos^2(\theta_p) + J_{py} * \cos^2(\theta_p) + J_{az} \\
 & * \sin^2(\theta_p) + J_{pz} * \sin^2(\theta_p) - J_{ay} \\
 & * \omega_{ibx}^b * \cos(\theta_p) * \cos(\theta_r) \\
 & * \sin(\theta_p) + J_{az} * \omega_{ibx}^b * \cos(\theta_p) \\
 & * \cos(\theta_r) * \sin(\theta_p) - J_{py} * \omega_{ibx}^b \\
 & * \cos(\theta_p) * \cos(\theta_r) * \sin(\theta_p) \\
 & + J_{pz} * \omega_{ibx}^b * \cos(\theta_p) * \cos(\theta_r) \\
 & * \sin(\theta_p) - J_{ay} * \omega_{ibz}^b * \cos(\theta_p) \\
 & * \sin(\theta_p) * \sin(\theta_r) + J_{az} * \omega_{ibz}^b \\
 & * \cos(\theta_p) * \sin(\theta_p) * \sin(\theta_r) \\
 & - J_{py} * \omega_{ibz}^b * \cos(\theta_p) * \sin(\theta_p) \\
 & * \sin(\theta_r) + J_{pz} * \omega_{ibz}^b * \cos(\theta_p) \\
 & * \sin(\theta_p) * \sin(\theta_r)) / (J_{ry} + J_{ay} \\
 & * \cos^2(\theta_p) + J_{py} * \cos^2(\theta_p))
 \end{aligned}$$

(59)

$$\begin{aligned}
 H_{23}(\theta) = & -(J_{az} * \sin(\theta_p) + J_{az} * \omega_{ibx}^b * \cos(\theta_p) \\
 & * \cos(\theta_r) + J_{az} * \omega_{ibz}^b * \cos(\theta_p) \\
 & * \sin(\theta_r)) / (J_{ry} + J_{ay} * \cos^2(\theta_p) \\
 & + J_{py} * \cos^2(\theta_p))
 \end{aligned}$$

$$\begin{aligned}
 H_{31}(\theta) = & \omega_{iby}^b * \cos(\theta_p) + \omega_{ibz}^b * \cos(\theta_r) * \sin(\theta_p) \\
 & - \omega_{ibx}^b * \sin(\theta_p) * \sin(\theta_r)
 \end{aligned}$$

$$\begin{aligned}
 H_{32}(\theta) = & \omega_{ibz}^b * \cos(\theta_p) * \sin(\theta_r) - \omega_{ibx}^b * \cos(\theta_p) \\
 & * \cos(\theta_r)
 \end{aligned}$$

And

$$\begin{aligned}
 h_1(\theta) = & (T_{dp} + N * T_{dm}) / (J_{ax} + J_{px}) - \dot{\omega}_{ibx}^b \\
 & * \cos(\theta_r) - \dot{\omega}_{ibz}^b * \sin(\theta_r) \\
 & - ((\sin(\theta_p) * (\omega_{ibz}^b * \cos(\theta_r) \\
 & - \omega_{ibx}^b * \sin(\theta_r)) + \omega_{iby}^b \\
 & * \cos(\theta_p)) * (\cos(\theta_p) * (\omega_{ibz}^b \\
 & * \cos(\theta_r) - \omega_{ibx}^b * \sin(\theta_r)) \\
 & - \omega_{iby}^b * \sin(\theta_p)) * (J_{ay} - J_{az} \\
 & + J_{py} - J_{pz}) / (J_{ax} + J_{px}) \\
 & + (J_m * N * (N - 1) * (\dot{\omega}_{ibx}^b \\
 & * \cos(\theta_r) - \dot{\omega}_{ibz}^b \\
 & * \sin(\theta_r))) / (J_{ax} + J_{px}) + (K_e \\
 & * K_t * N^2 * (\omega_{ibx}^b * \cos(\theta_r) \\
 & - \omega_{ibz}^b * \sin(\theta_r))) / (R_m * (J_{ax} \\
 & + J_{px}))
 \end{aligned}$$

$$\begin{aligned}
 h_3(\theta) = & \dot{\omega}_{iby}^b * \sin(\theta_p) - \cos(\theta_p) * (\dot{\omega}_{ibz}^b \\
 & * \cos(\theta_r) - \dot{\omega}_{ibx}^b * \sin(\theta_r)) \\
 & + (T_{da} + N * T_{dm}) / J_{az} + (J_m \\
 & * N * (N - 1) * (\omega_{ibz}^b * \cos(\theta_p) \\
 & * \cos(\theta_r) - \omega_{iby}^b * \sin(\theta_p) \\
 & + \omega_{ibx}^b * \cos(\theta_p) \\
 & * \sin(\theta_r))) / J_{az} + (K_e * K_t * N^2 \\
 & * (\omega_{ibz}^b * \cos(\theta_p) * \cos(\theta_r) \\
 & - \omega_{iby}^b * \sin(\theta_p) + \omega_{ibx}^b \\
 & * \cos(\theta_p) * \sin(\theta_r))) / (J_{az} * R_m)
 \end{aligned}$$

$$\begin{aligned}
 h_2(\theta) = & (\omega_{iby}^b * \cos^2(\theta_p) * (J_{ay} + J_{py}) + (J_{rx} \\
 & - J_{rz}) * (\omega_{ibx}^b * \cos(\theta_r) + \omega_{ibz}^b \\
 & * \sin(\theta_r)) * (\omega_{ibz}^b * \cos(\theta_r) \\
 & - \omega_{ibx}^b * \sin(\theta_r)) - \sin(\theta_p) \\
 & * (\cos(\theta_p) * (\omega_{ibz}^b * \cos(\theta_r) \\
 & - \omega_{ibx}^b * \sin(\theta_r)) - \omega_{iby}^b \\
 & * \sin(\theta_p)) * (J_{az} + J_{pz}) \\
 & + (\omega_{ibx}^b * \cos(\theta_r) + \omega_{ibz}^b \\
 & * \sin(\theta_r)) * (\omega_{ibz}^b * \cos(\theta_r) \\
 & - \omega_{ibx}^b * \sin(\theta_r)) * (J_{ax} + J_{px}) \\
 & - \cos(\theta_p) * (\cos(\theta_p) * (\omega_{ibz}^b \\
 & * \cos(\theta_r) - \omega_{ibx}^b * \sin(\theta_r)) \\
 & - \omega_{iby}^b * \sin(\theta_p)) * (\omega_{ibx}^b \\
 & * \cos(\theta_r) + \omega_{ibz}^b * \sin(\theta_r)) \\
 & * (J_{az} + J_{pz}) - \sin(\theta_p) \\
 & * (\sin(\theta_p) * (\omega_{ibz}^b * \cos(\theta_r) \\
 & - \omega_{ibx}^b * \sin(\theta_r))
 \end{aligned}$$

(60)

$$\begin{aligned}
& + \omega_{ib_y}^b * \cos(\theta_p) * (\omega_{ib_x}^b * \cos(\theta_r) + \omega_{ib_z}^b \\
& \quad * \sin(\theta_r)) * (J_{ay} + J_{py}) \\
& + \cos(\theta_p) * \sin(\theta_p) * (\omega_{ib_z}^b \\
& \quad * \cos(\theta_r) - \omega_{ib_x}^b * \sin(\theta_r)) \\
& \quad * (J_{ay} + J_{py}) / (J_{ay} + J_{py}) \\
& \quad * \cos^2(\theta_p) + J_{ry}) \\
& + (T_{dr} + N * T_{am}) / (J_{ay} + J_{py}) * \cos^2(\theta_p) \\
& \quad + J_{ry}) + (J_m * N * \dot{\omega}_{ib_y}^b * (N \\
& \quad - 1)) / (J_{ay} + J_{py}) * \cos^2(\theta_p) \\
& \quad + J_{ry}) + (K_e * K_t * N^2 \\
& \quad * \omega_{ib_y}^b) / (R_m * (J_{ay} + J_{py}) \\
& \quad * \cos^2(\theta_p) + J_{ry})
\end{aligned}$$

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Mohammad Mehdi Zohrei: Conceptualization, Funding acquisition, Investigation, Project administration, Roles/Writing - original draft, Writing - review & editing. **Hamidreza Javanmardi:** Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing - review & editing.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy has been completely observed by the authors.

REFERENCES

- [1] A. Altan, and R. Hacıoğlu, "Model predictive control of three-axis gimbal system mounted on UAV for real-time target tracking under external disturbances," *Mechanical Systems and Signal Processing*, vol. 138, p. 106548, 2020.
- [2] J. Mao, J. Yang, X. Liu, S. Li, and Q. Li, "Modeling and robust continuous TSM control for an inertially stabilized platform with couplings," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 6, pp. 2548–2555, 2020.
- [3] M. M. Zohrei, A. Roosta, and B. Safarinejadian, "Robust backstepping control based on neural network stochastic constrained for three axes inertial stable platform," *Journal of Aerospace Engineering*, vol. 35, no. 1, p. 04021117, 2022.
- [4] W. M. Abdel-Wahab et al., "Affordable large scale active-phased array antenna for Ka-band mobile SATCOM applications," in *2019 13th European Conference on Antennas and Propagation (EuCAP)*, 2019, pp. 1–4.
- [5] D.-H. Lee, D.-Q. Tran, Y.-B. Kim, and S. Chakir, "A robust double active control system design for disturbance suppression of a two-axis gimbal system," *Electronics*, vol. 9, no. 10, article 10, 2020.
- [6] C. D. Rodin, F. A. de Alcantara Andrade, A. R. Hovenburg, T. A. Johansen. "A survey of practical design considerations of optical imaging stabilization systems for small unmanned aerial systems," *Sensors*, vol. 19, no.21, p.4800, 2019.
- [7] V. Rosas-Cervantes, and S.-G. Lee, "3D localization of a mobile robot by using Monte Carlo algorithm and 2D features of 3D point cloud," *Int. J. Control Autom. Syst.*, vol. 18, no. 11, pp. 2955–2965, 2020.
- [8] Y. Zou, and X. Lei, "A compound control method based on the adaptive neural network and sliding mode control for inertial stable platform," *Neurocomputing*, vol. 155, pp. 286–294, 2015.
- [9] S. Cong, K. Deng, W. Shang, D. Kong, and H. Shen, "Isolation control for inertially stabilized platform based on nonlinear friction compensation," *Nonlinear Dyn*, vol. 84, no. 3, pp. 1123–1133, 2016.
- [10] S. Liu, H. Che, and L. Sun, "Research on stabilizing and tracking control system of tracking and sighting pod," *J. Control Theory Appl.*, vol. 10, no. 1, pp. 107–112, Feb. 2012.
- [11] M. Khayatian, and M. Arefi, "Adaptive dynamic surface control of a two axis gimbal system," *IET Science, Measurement & Technology*, vol. 10, 2016.
- [12] Y. Zhang et al., "Fuzzy-PID control for the position loop of aerial inertially stabilized platform," *Aerospace Science and Technology*, vol. 36, pp. 21–26, 2014,
- [13] F. Liu, H. Wang, Q. Shi, H. Wang, M. Zhang, and H. Zhao, "Comparison of an ANFIS and Fuzzy PID control model for performance in a two-axis inertial stabilized platform," *IEEE Access*, vol. 5, pp. 12951–12962, 2017.
- [14] M. Davanipour, H. Javanmardi, and N. Goodarzi, "Chaotic self-tuning PID controller based on Fuzzy wavelet neural network model," *Iran J Sci Technol Trans Electr Eng*, vol. 42, no. 3, pp. 357–366, 2018.
- [15] P. Setoodeh, A. Khayatian, and E. Farjah, "Backstepping-based control of a strapdown boatboard camera stabilizer," *International Journal of Control Automation and Systems*, vol. 5, pp. 15–23, 2007.
- [16] H. Javanmardi, M. Dehghani, M. Mohammadi, N. Vafamand, and T. Dragičević, "Optimal frequency regulation in AC mobile power grids exploiting bilinear matrix inequalities," *IEEE Transactions on Transportation Electrification*, pp. 1–1, 2021.
- [17] S. Pan, Y. Wu, J. Zhang, S. Zhou, and H. Zhu, "Modeling and control of a 2-degree-of-freedom gyro-stabilized platform driven by ultrasonic motors," *Journal of Intelligent Material Systems and Structures*, vol. 29, no. 11, pp. 2324–2332, 2018.
- [18] C. Yu, A. Li, S. Zhang, and X. Bai, "Robust finite time second-order sliding mode stabilization control for floated inertial platform," in *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, vol. 232, no. 9, pp. 1620–1627, 2018.
- [19] H. Javanmardi, M. Dehghani, M. Mohammadi, S. Siamak, and M. R. Hesamzadeh, "BMI-Based load frequency control in microgrids under false data injection attacks," *IEEE Systems Journal*, pp. 1–11, 2021.
- [20] W. A. Pradana, E. Joelianto, A. Budiyo, and W. Adiprawita, "Robust MIMO H ∞ integral-backstepping

- PID controller for hovering control of unmanned model helicopter," *Journal of Aerospace Engineering*, vol. 24, no. 4, pp. 454–462, 2011.
- [21] Z. Xiangyang, L. Xinyue, and Y. Chao, "Disturbances rejection based on robust H_∞ control for an aerial inertially stabilized platform," in *2019 14th IEEE International Conference on Electronic Measurement Instruments (ICEMI)*, 2019, pp. 1396–1401.
- [22] H. Javanmardi, M. Dehghani, M. Mohammadi, and M. R. Hesamzadeh, "A mesh-based partitioning algorithm for decreasing conservatism in solving bilinear matrix inequality problems," *Optimal Control Applications and Methods*, 2023.
- [23] H. R. Javanmardi, M. Dehghani, A. A. Safavi, and R. Abolpour, "Model predictive control of a class of uncertain nonlinear discrete time systems: The LMI approach," in *2016 24th Iranian Conference on Electrical Engineering (ICEE)*, 2016, pp. 323–328.
- [24] X. Zhou, Y. Shi, L. Li, R. Yu, and L. Zhao, "A high precision compound control scheme based on non-singular terminal sliding mode and extended state observer for an aerial inertially stabilized platform," *Int. J. Control Autom. Syst.*, vol. 18, no. 6, pp. 1498–1509, 2020.
- [25] F. Dong, X. Lei, and W. Chou, "A Dynamic model and control method for a two-axis inertially stabilized platform," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 1, pp. 432–439, 2017.
- [26] H. Ramirez-Rodriguez, V. Parra-Vega, A. Sanchez-Orta, and O. Garcia-Salazar, "Robust backstepping control based on integral sliding modes for tracking of quadrotors," *J Intell Robot Syst*, vol. 73, no. 1, pp. 51–66, 2014.
- [27] C. Afri, V. Andrieu, L. Bako, and P. Dufour, "State and parameter estimation: A nonlinear luenberger observer approach," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 973–980, 2017.
- [28] S. Lee and S. Jung, "An observer design technique for improving velocity estimation of a gimbal system," in *International Conference on Robot Intelligence Technology and Applications*, Springer, 2017, pp. 337–343.
- [29] N. Reichbach and A. Kuperman, "Recursive-Least-squares-based real-time estimation of supercapacitor parameters," *IEEE Transactions on Energy Conversion*, vol. 31, no. 2, pp. 810–812, 2016.
- [30] X. Lei, Y. Zou, and F. Dong, "A composite control method based on the adaptive RBFNN feedback control and the ESO for two-axis inertially stabilized platforms," *ISA Transactions*, vol. 59, pp. 424–433, 2015.
- [31] Z.-L. Zhao, and B.-Z. Guo, "A novel extended state observer for output tracking of MIMO systems with mismatched uncertainty," *IEEE Transactions on Automatic Control*, vol. 63, no. 1, pp. 211–218, 2018.
- [32] A. Grovlen and T. I. Fossen, "Nonlinear control of dynamic positioned ships using only position feedback: an observer backstepping approach," in *Proceedings of 35th IEEE Conference on Decision and Control*, Dec. 1996, vol.3. pp. 3388–3393.
- [33] I. Hosseini, M. Mirzaei, and M. H. Asemani, "Nonlinear output feedback for HL-20 flight control using back-stepping observer," *J Intell Robot Syst*, vol. 100, no. 3, pp. 1401–1416, 2020.
- [34] J. Valasek, and W. Chen, "Observer/Kalman filter identification for online system identification of aircraft," *Journal of Guidance, Control, and Dynamics*, vol. 26, no. 2, pp. 347–353, 2003.
- [35] A. F. Taha, A. Elmahdi, J. H. Panchal, and D. Sun, "Unknown input observer design and analysis for networked control systems," *International Journal of Control*, vol. 88, no. 5, pp. 920–934, 2015.
- [36] M. Liu, L. Zhang, P. Shi, and H. R. Karimi, "Robust control of stochastic systems against bounded disturbances with application to flight control," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 3, pp. 1504–1515, 2014.
- [37] M. Pourgholi, and V. J. Majd, "A new non-fragile H_∞ proportional—integral filtered-error adaptive observer for a class of non-linear systems and its application to synchronous generators," in *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 225, no. 1, pp. 99–112, 2011.
- [38] M. Pourgholi, and V. J. Majd, "A nonlinear adaptive resilient observer design for a class of Lipschitz systems using LMI," *Circuits Syst Signal Process*, vol. 30, no. 6, pp. 1401–1415, 2011.
- [39] M. Mirzaei, I. Hosseini, and V. Ghaffari, "MEMS gyroscope fault detection and elimination for an underwater robot using the combination of smooth switching and dynamic redundancy method," *Microelectronics Reliability*, vol. 109, p. 113677, 2020.
- [40] H. K. Khalil, *Nonlinear Systems*. Prentice Hall, 2002.
- [41] T. S. Shores, "Applied linear algebra and matrix analysis," in *Undergraduate Texts in Mathematics*. New York: Springer-Verlag, 2007.

BIOGRAPHY

Mohammad Mehdi Zohrei received his B. Sc. and M. Sc. degrees in Electrical Engineering from Shiraz University, Shiraz, Iran, in 1998 and 2003, respectively and his PhD from Shiraz University of Technology, Shiraz, Iran in 2022. He is interested include Inertially Stabilized Platform, Three degrees of freedom stabilizer, Nonlinear control approaches in ISP



Control.



robotic and artificial intelligence.

Copyrights

© 2024 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).





Research Article

Optimal Sizing and Placement of Series Capacitors in Distribution Networks Using Modified Elephant Herding Optimization Algorithm

Elvis Twumasi , Abdul-Fatawu Seini Yussif , and Emmanuel Asuming Frimpong

Department of Electrical and Electronic Engineering, College of Engineering, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

* Corresponding Author: etwumasi.coe@knust.edu.gh

Abstract: The optimal size and location of series capacitors is a critical challenge in a distribution network. In this paper, a novel approach for enhancing voltage stability in distribution networks through the optimal sizing and placement of series capacitors is proposed. The study introduces a technique to determine the optimal lines for connecting series capacitors based on line reactance and current. A Modified Elephant Herding Optimization (MEHO) algorithm was used to determine the reactance sizes of the series capacitors and the best lines to place them for optimum system performance. To evaluate the effectiveness of the proposed method, three series capacitors are placed and sized in the standard IEEE 33-bus radial distribution system for stability enhancement. A comparison is conducted between the proposed MEHO algorithm-based approach, the original Elephant Herding Optimization (EHO) algorithm, and the IGWO-TS-based methods reported in the literature. The evaluation is performed by analyzing the system voltage profile, total system losses, and system voltage deviation index under varying loading conditions of 30%, 100%, and 120% of the system nominal loading. Results demonstrate that the proposed MEHO algorithm-based approach outperforms the other two methods significantly in all the scenarios, highlighting its effectiveness in voltage stability enhancement in distribution networks.

Keywords: Elephant herding optimization algorithm, series capacitors, metaheuristic, voltage profile, distribution network.

Article history

Received 16 August 2023; Revised 28 November 2023; Accepted 01 December 2023; Published online 29 January 2024.

© 2024 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)

How to cite this article

E. Twumasi, A. F. S. Yussif, and E. A. Frimpong, "Optimal sizing and placement of series capacitors in distribution networks using modified elephant herding optimization algorithm," *J. Appl. Res. Electr. Eng.*, vol. 2, no. 2, pp. 173-181, 2023. DOI: [10.22055/jaree.2023.44430.1076](https://doi.org/10.22055/jaree.2023.44430.1076)



1. INTRODUCTION

The growing demand for reliable and high-quality electric power supply has necessitated the enhancement of voltage stability in distribution networks [1], as well as maximizing the power carrying capacity of the distribution lines [2]. Shunt reactive compensators such as shunt capacitor banks are known to be viable in providing reactive power compensation to minimize losses and improve voltage profile [3]. Voltage stability issues can lead to undesirable voltage deviations, system inefficiencies, and even cascading failures [4]. Series reactive compensation has been established as an effective method of maximizing power flows and reducing reactive power losses [5]. However, oversizing of series capacitors in distribution lines can lead to resonance as established in [5] and [6]. Therefore, the optimization of series capacitor placement and sizing has emerged as a promising approach to improve system voltage stability by

maximizing the power carrying capacity of distribution lines, and minimizing reactive power losses [7]. Unlike shunt capacitor banks that provide the reactive power at the demand end to reduce the reactive power flows on the distribution lines [8], series capacitors minimize the reactance of the distribution lines [9]. The placement and sizing of series capacitors in electrical power distribution networks have been widely investigated to alleviate voltage stability concerns. For instance, R. AL-Mula et al. [10] optimized the placement and sizing of series compensators to enhance voltage stability and reduce losses during increased power demand using PSO. Also, Mohammed Abd-EL-Hakeem et al proposed a method for optimal integration of series capacitors in distribution networks based on an Improved Grey Wolf Optimization (IGWO) algorithm and Tabu Search (TS) to improve the system voltage stability of a distribution network in Egypt [11]. Most of the proposed techniques used algorithms that tend to produce suboptimal solutions [12]. The challenge,

therefore, lies in determining better optimal locations and series capacitor ratings to achieve maximum voltage stability improvement while considering network constraints such as network topology and operational limitations [13].

To address this challenge, this research proposes a modified Elephant Herding Optimization (EHO) algorithm for the optimal sizing and placement of series capacitors in distribution networks [14]. The EHO algorithm is a population-based metaheuristic optimization technique inspired by the herding behavior of elephants in nature [15]. By adopting a modified version [16] of this algorithm to the problem of series capacitor optimization, this work aims to effectively determine the most suitable locations and sizes of series capacitors to enhance system voltage stability under different loading conditions.

Therefore, this paper develops a comprehensive and efficient methodology for the placement and sizing of series capacitors in distribution networks. The proposed approach integrates power system analysis, optimization techniques, and simulation studies to identify the optimal line and optimal sizes of series capacitors. The effectiveness of the proposed methodology is validated through extensive simulations and comparisons with existing methods in the literature.

The remainder of this paper is organized as follows: Section 2 presents the methodology, including the modified EHO algorithm and the formulation of the optimization problem. Section 3 discusses the simulation studies and the test system. Section 4 discusses the simulation results. Finally, Section 5 concludes the paper and outlines potential avenues for future research.

2. METHODOLOGY

2.1. Problem Formulation

To formulate the problem, a distribution line with a series capacitor connected is presented in Fig. 1 to illustrate the capacitor modeling. In Fig. 1, the amount of power flow from bus i to bus j without the capacitor can be expressed according to the power equation presented in (1) [17, 18]. Under heavily loaded conditions, the amount of power flow can be increased without affecting system voltage quality by reducing the line reactance value, X_c [19].

$$P_{flow} = \frac{V_i V_j}{X_T} \sin \delta \quad (1)$$

where; δ is the angle difference between V_i and V_j , and X_T is the total reactance of the line. The presence of the series capacitor results in a new total reactance, X_{Tnew} , expressed in (2) below to achieve an effective reduction in the overall line reactance as expressed in [20].

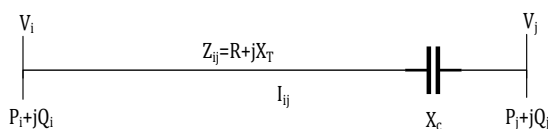


Fig. 1: Distribution Line with Series Capacitor.

$$X_{Tnew} = X_T - X_C \quad (2)$$

The X_{Tnew} is less than the X_T as a result of the presence of the capacitive reactance, X_c , of the series capacitor connected to the line. This increases the amount of maximum power the line can transmit safely within desired voltages [21].

However, within a given complex distribution network, the line to connect (location) and the amount of reactance sizing of the series capacitor plays a crucial role in its impact on the quality of power delivery [22]. Also, the oversizing of series capacitors can result in resonance [23]. To maximize the benefits and ensure stability, a two-step approach consisting of location determination and sizing is proposed as follows.

The proposed technique for determining the optimal lines to connect the series capacitors focuses on the loading level of the lines, and the reactance of the lines. This is on the basis that heavily loaded lines need series compensation the most, and series capacitors provide compensation by reducing the line reactance [24]. Hence the higher the line reactance, the more compensation a series capacitor can provide. The two conditions of the lines are combined to form an indicator of a priority list of distribution systems lines for series capacitor placement as presented in (3).

$$f_{ij} = |I_{ij} \times X_{ij}| \quad (3)$$

where I_{ij} represents the line current and X_{ij} represents the total line reactance. The value of f for all the lines in the distribution system is calculated from data from a backward-forward sweep load flow algorithm. The system lines are sorted in descending order according to the corresponding f values. This order of the system lines is considered the order for optimal placement of series capacitors in the distribution system for effective voltage profile enhancement. To enhance the system's voltage stability, the proposed objective function is developed based on the total system voltage deviation [25]. The system voltage profile is a good indicator of the system stability and the system losses [26]. The proposed MEHO algorithm [16] is then adopted to determine the optimal reactance values of the series capacitors to minimize the objective function expressed in (4). Just like in other areas of application where various metaheuristic algorithms have been adopted. For instance, the gravitation search algorithm in [26] has been used in optimal power flow for economic dispatch and emission.

$$\min(f_{obj}) = \sum_{i=1}^N |V_{ref} - V_i| \quad (4)$$

where the reference voltage, V_{ref} , is set to 1 p.u and V_i is the bus voltage.

2.2. Modified Elephant Herding Optimization (MEHO) Algorithm

The MEHO algorithm [16] is an improved version of the original EHO algorithm [15] that emulates the collective behavior of elephants in their natural habitat, specifically their herding and foraging patterns. It consists of two main phases: the clan updating operator and the separating updating operator. MEHO algorithm operates on the following assumptions:

1. The elephant population is composed of smaller groups known as clans, with a fixed number of members.
2. Male elephants tend to separate themselves from their clans and establish independent lives in each generation.
3. Within each clan, elephants coexist and are led by the strongest female member, known as the matriarch.

2.2.1. Initialization

To begin, the initial positions of N elephants in the population are generated randomly using (5). The population is then divided into smaller groups, referred to as clans, with an equal number of members.

$$x_j = x_{\min} + \alpha(x_{\max} - x_{\min} + 1) \times rand \quad (5)$$

where x_{\min} and x_{\max} are lower and upper bound of positions in the elephant population, and $rand \in [0,1]$ is a stochastic distribution.

2.2.2. Clan updating operator

In a given clan, C_i , the members coexist and are led by the matriarch. The matriarch's influence on the next generation determines the new position of each member, j , based on (6).

$$x_{ci,j}^{t+1} = x_{ci,j}^t + \alpha(x_{best,ci}^t - x_{ci,j}^t) \times r \quad (6)$$

where $x_{new,ci,j}$ and $x_{ci,j}$ represent the new and old positions of clan members respectively, $x_{best,ci}$ represents the position of the matriarch, $\alpha \in [0,1]$ is a factor that determines the level of the matriarch's influence on the clan members' new position, $r \in [0,1]$ and is a factor randomly generated using a uniform distribution.

The matriarch's new position is updated using equation (7).

$$x_{best,ci}^{t+1} = x_{best,ci}^t + \beta \times (x_{center,ci}^t - x_{best,ci}^t) \quad (7)$$

where $\beta \in [0,1]$ is a factor for determining the influence of $x_{center,ci}$ the matriarch's new position, while $x_{center,ci}$ represents the clan center calculated using (8).

$$x_{center,ci}^{t+1} = \frac{1}{n_{ci}} \times \sum_{j=1}^{n_{ci}} x_{ci,j,d}^t \quad (8)$$

where n_{ci} represents the number of elephants in the clan, and d is the dimension of the problem being solved in an interval ($1 \leq d \leq D$).

2.2.3. Separating updating operator

The male elephant leaves the clan when it attains puberty. To ensure the elimination of the worst clan member, it is assumed that the weakest clan member implements the

separation operator. The weakest member in each generation is replaced by a newly borne calf which is usually positioned close to the two strongest elephants in the clan. The weakest clan member is updated according to (9).

$$x_{worst,ci}^{t+1} = \frac{1}{2} \times (x_{best,ci}^t + x_{sec-best,ci}^t) \times R \quad (9)$$

where $x_{best,ci}$ and $x_{sec-best,ci}$ are the best two female members of the clan at the respective generation. R is a factor that gives the calf the freedom to roam within a certain range of the assigned position. It is randomly generated within the range $[0.5,1.01]$. The following pseudo codes can be used to implement the operators. With the detailed description provided in the clan updating operator and the separating updating operator, Fig. 4 serves as a flowchart guide for implementing the Modified Elephant Herding Optimization (MEHO) algorithm.

```

for  $Ci = 1$  to  $nClan$  (for all clans in the population) do
  for  $j=1$  to  $nCi$  (for all elephants in  $Ci$ ) do
    Update  $x_{ci,j}$  and generate  $x_{new,ci,j}$  using (6)
    If  $x_{ci,j} = x_{best,ci}$  then
      Update  $x_{ci,j}$  and generate  $x_{new,ci,j}$  using (7)
    end if
  end for j
end for Ci

```

Fig. 2: Pseudocode of clan updating operator.

```

for  $Ci = 1$  to  $nClan$  (all the clans in the elephant population) do
  Replace the worst elephant in clan  $Ci$  using (9).
end for Ci

```

Fig. 3: Pseudocode of separating updating operator.

The steps below outline the implementation of the Modified Elephant Herding Optimization (MEHO) algorithm.

Step 1: Initialization.

Set generation counter $t=1$, Initialize population, and maximum generation ($MaxGen$).

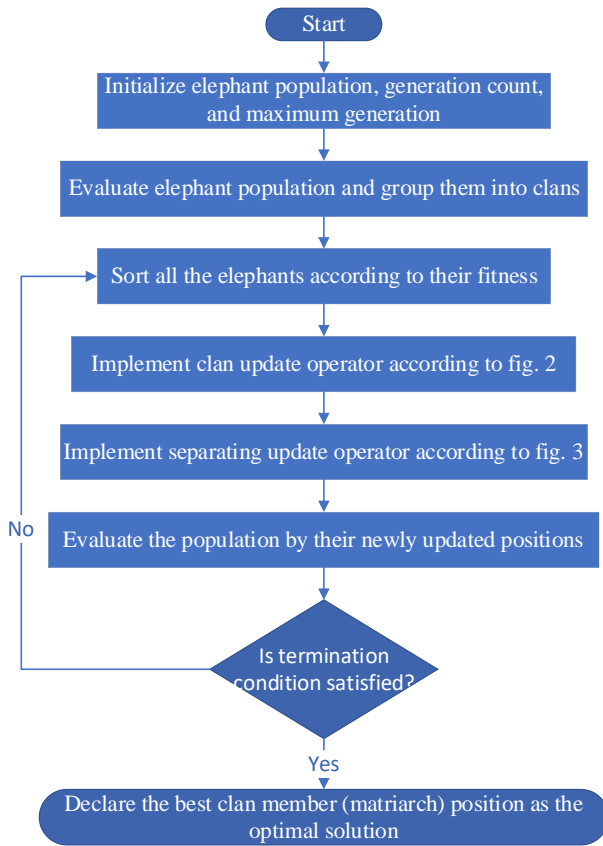


Fig. 4: Implementation of MEHO.

Step 2: While $t < MaxGen$, Do

Sort all the elephants according to their fitness.
 Implement the clan updating operator according to Fig. 2.
 Implement the separating operator according to Fig. 3.
 Evaluate the population by the newly updated positions.
 $t = t + 1$

Step 3: end while

2.3. Implementation

The following steps elaborate on the procedure to follow to implement the proposed method on a given electrical power distribution network.

1. Run a backward-forward sweep load flow algorithm (or any appropriate load flow method for distribution systems) on the system to determine the system data.
2. Calculate the values of f for all system lines using (8), sort the lines in descending order according to the values of f , and select the optimal lines to place capacitors starting from the top downwards till the number of capacitors wishes to be placed.
3. Integrate the capacitors at the various selected lines according to (7) to be sized by the MEHO algorithm.

4. Set the number of elephant clans to be equal to the number of series capacitors to be sized. The matriarch value of each clan will represent the size of a series capacitor.
5. Set all initial parameters of the MEHO algorithm, and run the algorithm to determine the sizes of the series capacitors.
6. Save the sizes of the matriarch elephants of all clans as the optimal sizes produced by the proposed MEHO algorithm-based method.

3. TEST SIMULATION

To establish the effectiveness of the performance of the proposed method, it was implemented in a MATLAB environment (MATLAB R2019a) and tested on the standard IEEE 33-bus radial distribution test system under varying load conditions. The simulation was carried out using an HP Pavilion laptop with a 64-bit AMD processor, 4.0 GB RAM, and 2.0GHz clock speed. The parameter settings used in simulating the Modified Elephant Herding Optimization (MEHO) algorithm to determine the series capacitor sizes are presented in Table 1.

3.1. IEEE 33-bus Radial Distribution Test System

The IEEE 33-bus radial distribution test system is one of the standard power distribution network systems developed for simulation tests in research works [27]. It has 33 buses and 32 lines, a total active and reactive power loading of 3.72MW and 2.3MVAR respectively, and total active power losses and reactive power losses of 202.67KW and 135.14KVAR respectively.

Table 1: Parameter settings for MEHO simulation.

Parameter	Value
Population size	30
Number of clans	3
Maximum Iterations	100
A	0.1
B	0.5

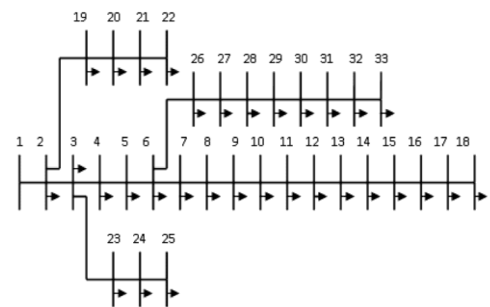


Fig. 5: Single line diagram of IEEE 33-bus system.

The proposed MEHO algorithm-based approach was used to place and size three (3) series capacitors in the IEEE 33-bus system to assess the performance in improving the system voltage stability. The outcome of this test is presented in the next section.

4. RESULTS AND DISCUSSION

The simulation results of the proposed method to integrate three series capacitors in the standard IEEE 33-bus radial distribution system are presented in this section. To establish the effectiveness of the proposed method, other methods reported in the literature and the original EHO algorithm were implemented as well to serve as benchmarks for comparison. The optimal locations and sizes produced by the methods are presented in Table 2. The impact of the various outcomes on the IEEE 33-bus system for the various methods is presented under three (3) cases of different system loading conditions. It is important to mention that series capacitors for distribution networks are available in varying sizes produced by manufacturers such as ABB, Schneider, Siemens, and so on [29]. Special requests can be made to meet certain requirements.

4.1. Case 1: System Performance Under Lightly Loaded Condition (30% of Nominal Load)

The performance of the system under lightly loaded conditions of 30% of the nominal loading condition is presented in Table 3. It covered the base system without compensation and with the compensation by all the methods considered.

Table 2: Simulation results on IEEE 33-bus system.

Method	Location (Size)
EHO	Line 5 (0.30p.u), Line 2 (0.19p.u), Line 27 (0.75p.u)
IGWO-TS	Line 2 (0.21p.u), Line 5 (0.50p.u), Line 3 (0.16p.u)
MEHO	Line 5 (0.70p.u), Line 2 (0.25p.u), Line 27 (0.90p.u)

Table 3: System Performance of IEEE 33-bus (30% loading).

IEEE 33-bus	Base Case	EHO	IGWO-TS [11]	MEHO
P Losses (KW)	16.49	16.43	16.41	16.38
% P Loss Reduction	-	0.36%	0.48%	0.67%
Q Losses (KVar)	10.98	7.53	6.54	5.38
% Q Loss Reduction	-	31.42%	40.44%	51.00%
Min. Voltage	0.97533	0.97697	0.97811	0.97836

The original EHO successfully reduced the total system active power losses to 16.43KW representing 0.36%, and that of the system total reactive power losses to 7.53KVar representing 31.42%. The IGWO-TS approach reported in [11] effectively reduced the total system active power losses to 16.41KW representing a 0.48% reduction, and the system total reactive power losses were reduced to 6.54KVar representing a 40.44% reduction. Finally, the proposed MEHO approach effectively reduced the total active power losses from 16.49KW to 16.38KW, representing the most marginal reduction of 0.67%, and the total reactive power losses to 5.38KVar representing 51.00%. Also, the MEHO produced the best voltage profile improvement with a minimum voltage of 0.97836p.u as shown in Fig. 6. The system voltage deviation index of the test system for various compensation conditions is presented in Fig 7. The base case scenario has a voltage deviation index of 0.014681, while the IGWO-TS [11], EHO, and MEHO methods produced voltage deviation indexes of 0.012734, 0.013268, and 0.012297 respectively. The proposed MEHO produced the best minimum index value of 0.012297.

4.2. Case 2: System Performance Under Nominal Load Condition (100% of Nominal Load)

Under nominal loading conditions, the performance of the system with the compensations for the various methods is presented in Table 4.

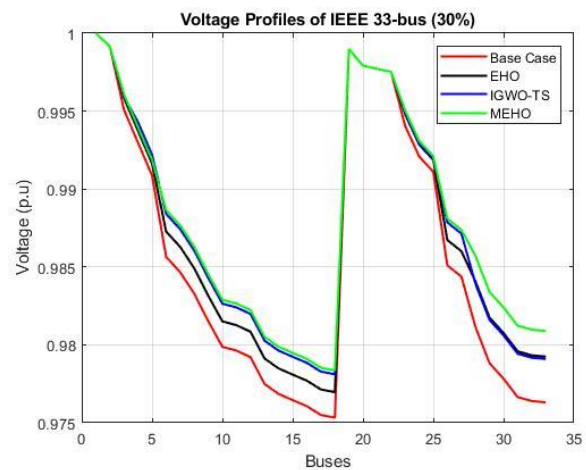


Fig 6: Voltage profiles of IEEE 33-bus (30% loading).

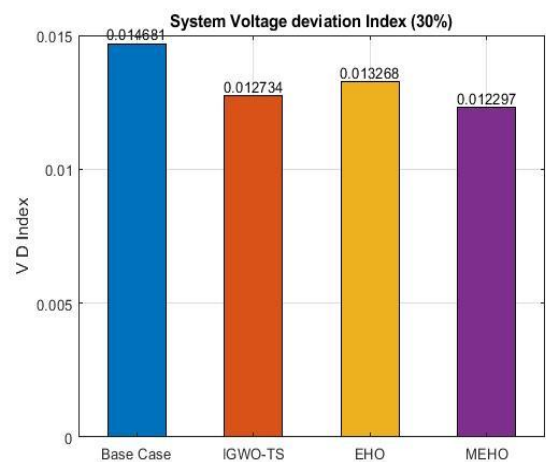


Fig. 7: System voltage deviation index (30% loading).

From Table 4, the original EHO successfully reduced the total system active power losses to 199.61KW representing 1.51%, and that of the system total reactive power losses to 91.75KVar representing 32.11%. The IGWO-TS approach reported in [11] effectively reduced the total system active power losses to 198.87 representing a 1.88% reduction, and the system total reactive power losses were reduced to 79.48 representing a 41.18% reduction. Finally, the proposed MEHO approach effectively reduced the total active power losses from 202.68KW to 197.68KW, representing the most marginal reduction of 51.92%. Also, the MEHO produced the best voltage profile improvement with a minimum voltage of 0.92432p.u as shown in Fig. 8.

Table 4: System Performance of IEEE 33-bus (100% loading).

IEEE 33-bus	Base Case	EHO	IGWO-TS [11]	MEHO
P Losses (KW)	202.68	199.61	198.87	197.68
% P Loss Reduction	-	1.51%	1.88%	2.47%
Q Losses (KVar)	135.14	91.75	79.48	64.97
% Q Loss Reduction	-	32.11%	41.18%	51.92%
Min. Voltage	0.91309	0.91924	0.9234	0.92432

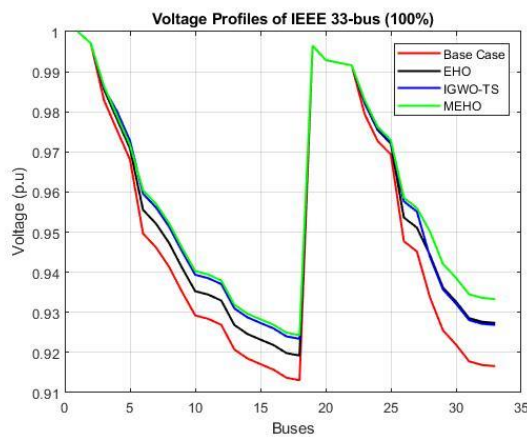


Fig 8: Voltage profiles of IEEE 33-bus (100% loading).

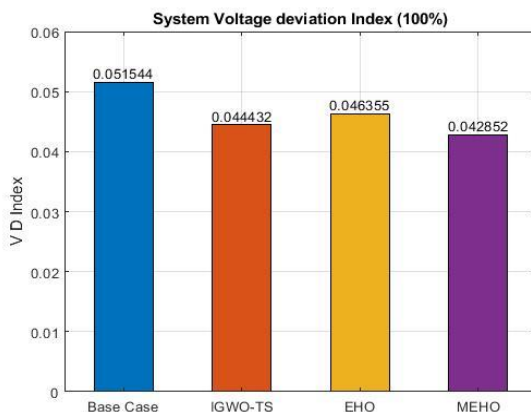


Fig 9: System voltage deviation index (100% lading).

To further assess the impact of the capacitors on the system voltage stability, the system voltage deviation index is presented in Fig 9. The base case scenario has a voltage deviation index of 0.051544, while the IGWO-TS, EHO, and MEHO methods produced voltage deviation indexes of 0.044432, 0.046355, and 0.042852 respectively. The proposed MEHO produced the best minimum index value of 0.042852.

4.3. Case 3: System Performance Under Heavily Loaded Condition (120% of Nominal Load)

The performance has also been assessed under a heavily loaded condition of 120% and the performance is presented in Table 5 below. Under a heavy loading condition of 120%, the total system active power losses and reactive power losses are 301.45KW and 201.10KVar respectively. The EHO technique successfully reduced them to 295.70KW (1.91%) and 136.00KVar (32.37%) respectively. Also, the IGWO-TS method reduced the total active power losses to 294.31KW representing 2.37%, and the total reactive power losses to 117.71KVar representing 41.47%. Again, the proposed MEHO approach produced the best reduction of the total active power losses to 292.11KW and total reactive power losses to 96.02KVar representing 3.10% and 52.25% reductions respectively. It also achieved the best voltage profile improvement as shown in Fig. 10 with a minimum voltage of 0.90781p.u.

Table 5: System performance of IEEE 33-bus (120% loading).

IEEE 33-bus	Base Case	EHO	IGWO-TS [11]	MEHO
P Losses (KW)	301.45	295.70	294.31	292.11
% P Loss Reduction	-	1.91%	2.37%	3.10%
Q Losses (KVar)	201.10	136.00	117.71	96.02
% Q Loss Reduction	-	32.37%	41.47%	52.25%
Min. Voltage	0.89384	0.90152	0.90666	0.90781

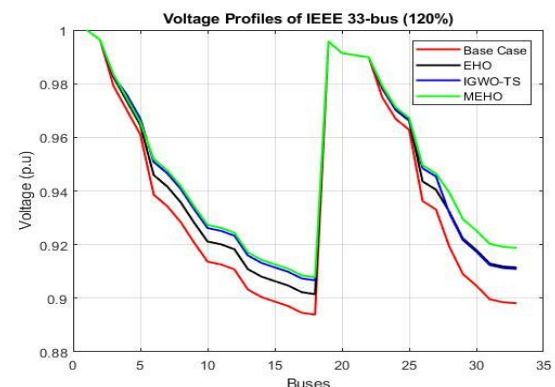


Fig 10: Voltage profiles of IEEE 33-bus (120% loading).

Fig. 11 presents the system voltage stability level of the test system with an illustration of the system voltage deviation index for the various placement methods. The lower the index value shows the more stable the system voltage. From Fig. 11, the base case situation has a system voltage deviation index value of 0.062893, the IGWO-TS produced 0.054095, EHO produced 0.056462, and the proposed MEHO produced the lowest index value of 0.052146. The proposed MEHO algorithm-based method produced the most stable system. From the results, the proposed method outperformed the other methods proposed in the literature by producing the lowest active power losses and the lowest reactive power losses. Also, the MEHO algorithm method produced the lowest system voltage deviation index, indicating the most stable system relative to the other methods. Again, the MEHO algorithm method exhibited a relatively better voltage profile. These performances are reflected in all the loading scenarios considered.

5. CONCLUSION AND RECOMMENDATION

A new approach has been proposed in this work to optimally place and size series capacitors in electrical power distribution networks to improve system voltage stability based on a Modified Elephant Herding Optimization (MEHO) algorithm. It has been tested on the standard IEEE 33-bus radial distribution test system by optimally placing and sizing three series capacitors, and the system performance under different loading conditions has been presented. The proposed method effectively enhanced the system voltage stability with minimum voltages of 0.97836p.u, 0.92432p.u, and 0.90781p.u under light loading, nominal loading, and heavy loading conditions respectively. It also resulted in a system voltage deviation index of 0.012297, 0.042852, and 0.052146 respectively under the three loading conditions. There were significant reductions in total reactive power losses of the system and slight reductions in the total active power losses in the various scenarios considered. These performances were superior to those of the other methods in the literature considered. Based on the performance of this work, the MEHO algorithm is recommended for application in determining the placement and sizing of series capacitors in distribution networks. Also, the proposed method is suitable for optimizing TSCS location and sizing. Future works should focus on exploring the potential of metaheuristic algorithms such as the MEHO algorithm for optimal joint integration of series and shunt capacitors for efficient power delivery.

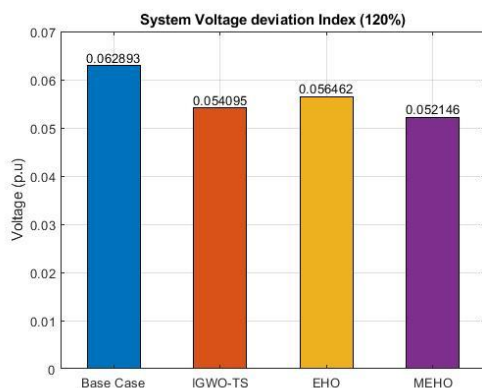


Fig 11: System voltage deviation index (120% loading).

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Elvis Twumasi: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Roles/Writing - original draft, Writing review & editing. **Abdul-Fatawu Seini Yussif:** Data curation, Investigation, Methodology, Software, Validation, Visualization, Roles/Writing - original draft, Writing - review & editing. **Emmanuel Asuming Frimpong:** Investigation, Methodology, Project administration, Supervision, Visualization, Writing - review & editing.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy has been completely observed by the authors.

REFERENCES

- [1] D. McCarrel, R. Bahry, A. Folkesson, and P. Bérubé, "Distribution series capacitor application for improved motor start and flicker mitigation," in *Proc. IEEE Power Eng. Soc. Transm. Distrib. Conf.*, 2006, no. 2, pp. 443–449.
- [2] S. Kannan, P. Renuga, and A. R. G. Monica, "Optimal capacitor placement and sizing using combined fuzzy-HPSO method," *Int. J. Eng. Sci. Technol.*, vol. 2, no. 6, pp. 75–84, 2011.
- [3] M. Y. Rofandy, A. Hasibuan, and R. Rosdiana, "Analysis of the effect of bank capacitor placement as voltage drop increase in distribution network," *Andalasian Int. J. Appl. Sci. Eng. Technol.*, vol. 2, no. 1, pp. 11–24, 2022.
- [4] G. Xu, S. Xiao, X. Xiao, and L. Li, "Research on the application of series capacitor compensation for typical medium voltage distribution network." in *2015 International Symposium on Material, Energy and Environment Engineering, 2015*.
- [5] I. Bucatariu, and F. Dan Surianu, "Optimal placement of fixed series capacitor in distribution networks," in *Proc. 9th WSEAS/IASME Int. Conf. Electr. Power Syst. High Voltages, Electr. Mach. POWER '09, 2009*, pp. 65–70.
- [6] A. Z. Abass, D. A. Pavlyuchenko, and Z. S. Hussain, "Methods comparison for optimal capacitor placement in distribution system," in *2020 Int. Multi-Conference Ind. Eng. Mod. Technol, 2020*.
- [7] Kathim, Al-Zohuari, and Majeed Zaidan, "Use series compensation in distribution networks 33 KV," *Int. J. Sci. Res.*, vol. 6, no. 6, pp. 2579–2583, 2017.
- [8] S. M. G. Mostafa, J. G. Shingh, and H. M. E. Haque, "An extensive literature review and new proposal on optimal capacitor placement in

- distribution systems," *J. Eng. Adv.*, vol. 01, no. 04, pp. 150–169, 2020.
- [9] Y. P. Soni, and E. Fernandez, "An artificial neural network approach to performance investigation of radial distribution system with series capacitor installation," *Electr. Power Components Syst.*, vol. 0, no. 0, pp. 1–11, 2023.
- [10] R. AL-MulaHumadi, "Optimal placement of series and compound compensators for power loss minimization based on WPSO algorithm," *Int. J. Comput. Appl.*, vol. 97, no. 12, pp. 1–7, 2014.
- [11] M. A. E. H. Mohamed, A. Y. Abdelaziz, M. M. F. Darwish, M. Lehtonen, and K. Mahmoud, "Optimum estimation of series capacitors for enhancing distribution system performance via an improved hybrid optimization algorithm," *Energy Sci. Eng.*, vol. 11, no. 6, pp. 1906–1924, 2023.
- [12] S. M. Parashar, G. Leena, M. Pande, Zaheeruddin, and J. Singh, "Dynamic capacitor placement to mitigate disaster in distribution system: a fuzzy approach," in *2019 Int. Conf. Power Electron. Control Autom. ICPECA 2019 - Proc.*, 2019, vol. 2019-November, no. 2, pp. 19–23.
- [13] S. Mokred, Q. Lijun, G. Kamara, and T. Khan, "Comparison of the effect of series and shunt capacitor application in 25kv radial power distribution network," in *2020 IEEE/IAS Ind. Commer. Power Syst. Asia*, 2020, pp. 822–830.
- [14] O. Baimakhanov, H. Şenyüz, A. Saukhimov, and O. Ceylan, "Heuristic optimization approaches for capacitor sizing and placement: A case study in kazakhstan," *Energies*, vol. 15, no. 9, pp. 1–16, 2022.
- [15] G. G. Wang, S. Deb, and L. D. S. Coelho, "Elephant herding optimization," in *Proc. 2015 3rd Int. Symp. Comput. Bus. Intell. (ISCBI)*, 2016, pp. 1–5.
- [16] A.-F. Seini Yussif, E. Twumasi, and E. A. Frimpong, "Performance enhancement of elephant herding optimization algorithm using modified update operators," *J. Nas. Tek. Elektro*, vol. 2, no. 30, 2023.
- [17] V. Golov, A. Kalutskov, and D. Kormilitsyn, "Controlled series compensation of high voltage lines to increase transmission capacity," in *Proc. 2020 Int. Ural Conf. Electr. Power Eng. Ural.*, 2020, pp. 377–382.
- [18] A. K. Hamza and M. F. Bonneya, "Step voltage regulator and capacitor placement to improve the performance of rural electrical distribution systems by CYME program," in *IOP Conf. Ser. Mater. Sci. Eng.*, 2019, vol. 518, no. 4.
- [19] S. Mokred, Q. Lijun, G. Kamara, and T. Khan, "Comparison of the effect of series and shunt capacitor application in 25kv radial power distribution network," in *2020 IEEE/IAS Ind. Commer. Power Syst. Asia*, 2020, no. September 2022, pp. 822–830.
- [20] E. Bakker, V. Deljou, and J. Rahmani, "Optimal placement of capacitor bank in reorganized distribution networks using genetic algorithm," *Int. J. Comput. Appl. Technol. Res.*, vol. 8, no. 4, pp. 89–87, 2019.
- [21] R. S. Basharin, M. D. Shipitsina, and A. N. Belyaev, "Improving controllability of flexible transmission lines with series compensation," in *Proc. 2021 IEEE Conf. Russ. Young Res. Electr. Electron. Eng. (ElConRus)*, 2021, pp. 1386–1390.
- [22] M. T. Hoq, J. Wang, and N. Taylor, "Impact of high levels of series compensation on line distance protection," *IET Conf. Publ.*, 2020, vol. 2020, no. CP771, pp. 1–6.
- [23] E. M. Carlini et al., "Transmission capacity and static stability improvement by series compensation: case study on South-North corridor of Italy," in *2022 AEIT Int. Annu. Conf.*, 2022, pp. 1–6.
- [24] P. R. Khade, and M. P. Thakre, "Advance approach to prevent the effect of mutual coupling series capacitor for series compensated parallel transmission lines," in *IEEE 2nd Int. Conf. Power, Energy, Control Transm. Syst. Proc. (ICPECTS)*, 2020, pp. 1–6.
- [25] A. Naderipour et al., "Spotted hyena optimizer algorithm for capacitor allocation in radial distribution system with distributed generation and microgrid operation considering different load types," *Sci. Rep.*, vol. 11, no. 1, pp. 1–15, 2021.
- [26] S. Aminzadeh, M. T. Hagh, and H. Seyedi, "Reactive power coordination between solid oxide fuel cell and battery for microgrid frequency control," *Journal of Applied Research in Electrical Engineering*, vol. 1, no. 2, pp. 121–130, 2022.
- [27] H. F. Kadom, A. N. Hussain, and W. K. S. Al-Jubori, "Optimal dual design based on capacitor placement and reconfiguration techniques for loss reduction and voltage enhancement," in *IOP Conf. Ser. Mater. Sci. Eng.*, 2020, vol. 745, no. 1, 2020.
- [28] B. Swapna et al., "Reactive power control in distribution network by optimal location and sizing of capacitor using Fuzzy and SFLA," *International Journal of Advances in Engineering and Management (IJAEM)*, vol. 5, no. 1, pp. 256–262, 2023.
- [29] Necessa Technologies, (January 28, 2024). *LV Capacitors QCap-L series capacitors for power factor correction ABB and power quality*. [Online]. Available: <https://necessatechnologies.com/lv-capacitors-qcap-l-series-capacitors/>.

BIBLIOGRAPHY

Elvis Twumasi is a lecturer at the Department of Electrical and Electronic Engineering at Kwame Nkrumah University of Science and Technology (KNUST), Kumasi Ghana. He has a BSc., MPhil, and Ph.D. in Electrical and Electronic Engineering from the same University.



Abdul-Fatawu Seini Yussif obtained his bachelor's degree in Electrical/Electronic Engineering from the Kwame Nkrumah University of Science Technology-Kumasi (KNUST). He also has an MPhil degree in Power Systems Engineering at the

same university.



Emmanuel Asuming Frimpong is a senior lecturer at the Department of Electrical and Electronic Engineering at Kwame Nkrumah University of Science and Technology (KNUST), Kumasi Ghana. He holds a BSc. in Electrical / Electronic Engineering from KNUST, MSc in Electrical Power

Systems from the University of Bath in the United Kingdom, and a Ph.D. in Electrical Engineering from KNUST.

Copyrights

© 2024 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).





Iranian Association of
Electrical and Electronics
Engineers

Journal of Applied Research in Electrical Engineering

E-ISSN: 2783-2864

P-ISSN: 2717-414X

Homepage: <https://jaree.scu.ac.ir/>



Research Article

Re-examination of the Doubly Fed Induction Machine (DFIM) Model Taking Into Account the Rotor Reactive Power

Mohammad Naser Hashemnia *

Department of Electrical Engineering, Mashhad Branch, Islamic Azad University, Mashhad, Iran

* Corresponding Author: hashemnia@mshdiau.ac.ir

Abstract: Modeling electric machines is crucial for analyzing their behavior and designing controllers. It is of the utmost importance to make use of a consistent equivalent circuit of the Doubly Fed Induction Machine (DFIM) that is applicable to a variety of operating modes. This is because it helps in the calculation of the machine's steady-state performance, converter ratings, and controller set-points. Traditional models of doubly fed induction machines employ the steady-state equivalent circuit of a wound-rotor induction machine with all rotor parameters referred to the stator through a frequency conversion. The present study investigates the validity of the traditional steady-state circuit model by taking into account the sequence change in rotor voltages and currents at super-synchronous speeds. The validity of phasor diagrams constructed using the traditional circuit is assessed, with a particular focus on super-synchronous operation in both motoring and generating modes. It has been demonstrated that the existing model is applicable to all rotor speeds (whether sub-synchronous or super-synchronous). However, caution should be exercised when utilizing expressions of rotor reactive power that involve dynamic dq and steady-state phasor models. Therefore, modified expressions are developed for rotor reactive power that are applicable regardless of the operating speed. The accuracy of the proposed method for different operating modes is confirmed by comprehensive simulation results developed with Matlab® Simulink. An investigation is also conducted into the sensitivity of rotor reactive power direction to parameter changes, and it is shown that machine parameter changes have a negligible effect on rotor reactive power direction.

Keywords: Doubly Fed Induction Machine (DFIM), dynamic model, phasor diagram, reactive power, super-synchronous speed.

Article history

Received 16 August 2023; Revised 06 December 2023; Accepted 24 December 2023; Published online 29 January 2024.

© 2024 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)

How to cite this article

M. N. Hashemnia, "Re-examination of the Doubly Fed Induction Machine (DFIM) model taking into account the rotor reactive power," *J. Appl. Res. Electr. Eng.*, vol. 2, no. 2, pp. 182-193, 2023.

DOI: [10.22055/jaree.2023.44572.1078](https://doi.org/10.22055/jaree.2023.44572.1078)



1. INTRODUCTION

1.1. Research Motivation

Wind energy conversion systems utilizing a Doubly Fed Induction Machine (DFIM) are by far the most prevalent topology worldwide [1]. This is due to the numerous benefits of DFIMs, such as decoupled control of active and reactive powers, the use of a partial scale converter resulting in reduced costs and losses, and so forth [2]. Significant research has been conducted on this topology for grid-connected and stand-alone applications, and the mathematical relationships of the machine are well-established [3–8]. Besides the generation mode of DFIM operation, a number of applications have emerged in which the aforementioned benefits make DFIM an excellent candidate for the motoring

mode of operation, including traction systems, marine propulsion and hybrid electric aircrafts [9-12]. A study of the machine's steady-state operation is essential for determining its steady-state characteristics and establishing the initial conditions for simulating its dynamic performance. In addition, sizing the converters necessitates a study of the active and reactive power ratings, which can be obtained most easily with the aid of equivalent circuit model equations [13]. Calculating the reference values of variables used as set-points in control loops requires valid and accurate model equations. Finally, load flow analysis of power systems with high penetration of wind turbines equipped with Doubly Fed Induction Generators (DFIGs) necessitates revisiting the traditional algorithms by taking the stator and rotor side

active/reactive powers into account [14-18]. Consequently, a consistent equivalent circuit that is applicable to various operating modes must be developed. Doubly fed induction generators (DFIGs) have been used predominantly in sub-synchronous mode, with super-synchronous operation restricted to wind gust conditions [19]. As demonstrated in [19], however, operation above synchronous speed results in higher efficiency, greater total output power and a simpler, less expensive power converter. Consequently, it is of the utmost significance to investigate DFIM in super-synchronous mode of operation. The super-synchronous study of DFIMs presents a challenge due to the appearance of negative values in the frequency of rotor circuit variables. The analysis of this issue can be traced back to 1983 [20], when a proposition was made to avoid referring the rotor circuit to the stator side and instead directly manipulate the rotor quantities at the rotor side. Nevertheless, the widely accepted representation of DFIMs is the stator-referred equivalent circuit. This choice is motivated by the desire to work with a circuit that encompasses a single frequency, namely the stator frequency, as opposed to a circuit that involves two different frequencies, namely the stator and rotor frequencies.

1.2. Literature Review

It has been claimed in [21, 22] that the traditional steady-state equivalent circuit exhibits deficiencies when employed in the super-synchronous mode of operation. A revised version of the model (referred to as the "accurate equivalent circuit" in references [21, 22]) is subsequently suggested for utilization in scenarios involving super-synchronous speeds, as an alternative to the widely accepted DFIM model. In effect, these references emphasize that the equivalent circuit structure depends on the rotor slip. This claim is supported by the change in the phase sequence of rotor variables at super-synchronous speeds, which influences the reactive power exchanged on the rotor side. The same authors put forward an iterative approach for integrating a DFIG into load flow analysis by utilizing its equivalent circuit in [23]. In [24], the same equivalent circuit was used to derive a framework for incorporating DFIG-based wind power plants into load flow analysis of distribution systems. Several operational limits, including stator and rotor currents, were considered when determining the DFIG's active and reactive power (PQ) limits. Consequently, it is essential to calculate the rotor current for any operating speed using an appropriate equivalent circuit. The forward–reverse sweep method was used to determine the load flow solution. A non-iterative method for calculating steady-state operating points of DFIG-based wind energy systems is employed in [25]. This calculation is a crucial component of the stability analysis. Since it is assumed that the DFIG operates at a particular power factor, the bus to which it is connected must be treated as a PQ bus. Both sub-synchronous and super-synchronous modes of operation are analysed using distinct equivalent circuits. In [26], a direct or non-iterative method for calculating the steady-state operating points of a DFIG is proposed. This direct initialization method is applied to both sub-synchronous and super-synchronous modes using distinct models. This method has the benefit of preventing non-convergence issues and decreasing the calculation time for DFIG internal variables.

The study presented in [27] examines a hybrid renewable energy system that incorporates a DFIG in conjunction with a direct current (DC) motor. Two distinct equivalent circuits are employed to represent the DFIG, each corresponding to a specific operating speed. Initial values of a DFIM are computed using two distinct equivalent circuits dependent on the rotor speed in [28]. These values are utilized for the machine's dynamic simulation. To demonstrate the efficacy of the proposed equivalent circuits, time-domain simulations and eigenvalue analysis are presented. In accordance with the assertion put forth by [21, 22] that the equivalent circuit changes considerably depending on the operating speed, a new unbalanced steady-state model for DFIG operating at various speeds is presented in [29].

1.3. Necessity of the Research Based on Challenges of the Literature

The proposed circuit in [21, 22] for super-synchronous speeds differs from the well-established DFIM circuit derived in any course of electric machinery analysis for wound rotor induction machines [30, 31] and the well-established context on DFIM modelling and control [3]. The additional modelling complication introduced by using two different circuit models based on the operating slip is the second reason why using two different circuit models based on the operating slip is questionable, since an equivalent circuit may not apply for some values of slip. Moreover, the authors in [21, 22] have yet to present corrective measures to be taken when dealing with the dynamic dq model of the machine; in fact, their study is confined to steady-state phasor analysis of the machine.

1.4. Novelty and Main Contributions of the Paper

In this paper, it is demonstrated that while it is necessary to modify some relationships when the machine operates at super-synchronous speeds, it is neither necessary nor recommended that the well-established model be modified based on the operating speed: the same stator-referred equivalent circuit as for sub-synchronous speed can be used. The phasor diagram is then constructed based on the equivalent circuit, with negative slip taken into account. For the actual voltages and currents of the rotor side, it suffices to consider that when there is a negative slip, resulting in a negative rotor angular frequency ($s < 0 \Rightarrow \omega_r < 0$), the direction of rotation of rotor phasors is reversed relative to the direction of rotation of stator phasors (taking the normal direction to be counter-clockwise, the direction of rotation of rotor phasors for super-synchronous speeds will be clockwise). Therefore, only rotor reactive power expressions based on steady-state and dynamic equations should consider the rotor slip sign.

The primary contributions of the paper can be enumerated as follows:

- Providing a comprehensive understanding of the negative sequence phenomenon in the super-synchronous mode of operation.
- Proposal of a novel approach to address the change in phase sequence of rotor variables at super-synchronous speeds, which has implications for a number of established relationships.

➤ Criticism of changing the DFIM equivalent circuit based on rotor speed conditions (as done in [21, 22]). Alternatively, a unified modelling framework that can accommodate all operating speeds is proposed.

➤ Modifying some steady-state and dynamic relations to take super-synchronous speed into account, while resorting to the widely accepted equivalent circuit.

1.5. Organization and Structure of the Paper

This paper is organized as follows. In Section 2, a super-synchronous operation is investigated for the very special case of pure capacitive load at rotor terminals to shed some light on the problem. The general load case is then investigated in Section 3, which gives an insight into rotor reactive power expression based on the operational speed. Afterwards, in Section 4, phasor diagrams for the four-quadrant operation of a DFIM with unity stator power factor operation are illustrated and justified. Rotor reactive power under dynamic conditions is examined in Section 5. Section 6 contains simulation results and discussions, followed by conclusions.

2. VOLTAGE-CURRENT RELATION FOR A ROTOR TERMINATED WITH A CAPACITIVE LOAD

To better illustrate the concept, consider a pure capacitive load connected to the rotor terminals of a DFIM rotating at super-synchronous speed as a particular case. In the given loading conditions, the relationship between the actual rotor phase voltage and current can be expressed by the following equation:

$$\bar{V}_r = \frac{-j}{C\omega_r} \bar{I}_r \quad (1)$$

where ω_r represents the electrical angular frequency of rotor variables. The over-barred variables (\bar{V}_r and \bar{I}_r) denote phasors. Considering the negative sign of ω_r due to negative slip, the phasor diagram can be drawn, as shown in Fig. 1. The speed of rotation of the phasors corresponds to the rotor frequency. At first glance, it might look like the rotor voltage phasor leads the rotor current phasor in the figure. This cannot be correct because the load is pure capacitive. However, taking the actual direction of rotation of phasors, as expected for a capacitive load, the rotor current phasor is still leading the rotor voltage phasor, both rotating clockwise. The situation is best depicted in Fig. 2.

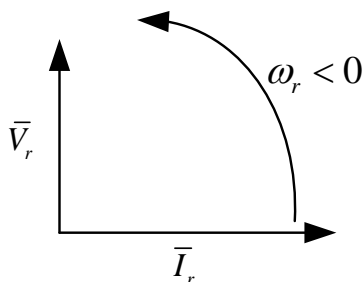


Fig. 1: Pure capacitive load phasor diagram for super-synchronous speed.

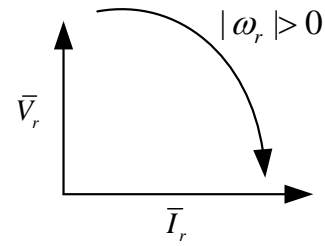


Fig. 2: Phasor diagram taking the actual direction of the phasors rotation.

The time variations of the voltage and current of rotor phase-a are expressed for better clarification. Assuming that the phase-a voltage at the rotor terminals is:

$$v_r(t) = V_{mr} \cos(\omega_r t) = V_{mr} \cos(s\omega_s t) \quad (2)$$

Rotor phase-a current is determined to be:

$$i_r(t) = C \frac{dv_r(t)}{dt} = cs\omega_s V_{mr} \cos(s\omega_s t + \frac{\pi}{2}) \quad (3)$$

where ω_s represents the electrical angular frequency of the stator. As the slip has been taken negative, the rotor voltage and current are expressed in an alternative way, resulting in the introduction of a positive frequency and the conversion of the negative multiplier of the cosine term in the rotor ($cs\omega_s < 0$) to a positive value:

$$v_r(t) = V_{mr} \cos(|s| \omega_s t), i_r(t) = c |s| \omega_s V_{mr} \cos(|s| \omega_s t + \frac{\pi}{2}) \quad (4)$$

This verifies the fact that the rotor current leads the rotor voltage.

It is now instructive to refer the rotor circuit to the stator (through frequency and turns-ratio transformations). The result is:

$$\frac{\bar{V}'_r}{s} = -\frac{j}{C's^2\omega_s} \bar{I}'_r \quad (5)$$

The above relation guarantees that the current of a pure capacitive load leads its voltage in the stator-referred equivalent circuit for both sub-synchronous and super-synchronous rotor speeds. A similar reasoning can be used to demonstrate that a pure inductive load connected to the rotor terminals will always have a current lagging its voltage, which is also supported in [32]. In fact, a capacitive (inductive) impedance is invariably capacitive (inductive), regardless of the operating speed. Instead of considering the rotor frequency to be negative when dealing with negative slips, it is preferable to consider the phase sequence to be reversed. The interested reader is referred to [33] for a more comprehensive examination of the negative frequency concept.

It is worth mentioning that the reactive power magnitude is increased on the stator side in comparison to the reactive power on the rotor side. This amplification of reactive power can be attributed to the division of frequency by slip when transforming from the rotor to the stator side.

In the rotor circuit, the reactive power produced by the capacitive load is:

$$Q_r = 3 |\bar{V}_r \parallel \bar{I}_r| \tag{6}$$

The stator-referred load also produces reactive power, and its value is determined by:

$$Q'_r = 3 \left| \frac{\bar{V}'_r}{s} \parallel \bar{I}'_r \right| = 3 \frac{|\bar{V}'_r| |\bar{I}'_r|}{|s|} = 3 \frac{|\bar{V}_r| |\bar{I}_r|}{|s|} = \frac{Q_r}{|s|} \tag{7}$$

The last expression shows that the reactive power produced by the capacitive reactance is amplified when the reactance is referred from its actual position (i.e., the rotor) to the stator side. This is supported by the expression of capacitive reactive power:

$$\begin{cases} Q_r = \frac{I_r^2}{C_r \omega_r} \\ Q'_r = \frac{I_r'^2}{C'_r \omega_s} = \frac{I_r^2}{C_r |s| \omega_r} \end{cases} \Rightarrow Q'_r = \frac{Q_r}{|s|} \tag{8}$$

3. VOLTAGE-CURRENT RELATION FOR THE GENERAL LOAD CASE

In the previous derivation, it was assumed that a pure capacitive load was connected to the rotor terminals of a DFIM operating at super-synchronous speed. Now, a general case will be examined. The relationship between rotor voltage and rotor current can be visualized more clearly with the aid of phasor diagrams, which depict the phasors of rotor voltage and current when they are on the rotor side and when they are referred to the stator side. Take the slip to be negative and assume that the rotor voltage and current phasors have a phase relationship when expressed in the actual rotor side, as shown in Fig. 3. As the slip is negative, the phasors rotate in a clockwise direction. Fig. 4 depicts the relative position of the stator-referred phasors appearing in the equivalent circuit. The direction of rotation of the phasors has now been reversed, and it is in the normal counter-clockwise direction. The referred rotor voltage phasor (\bar{V}'_r/s) is in the opposite

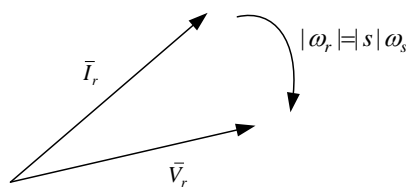


Fig. 3: Rotor phasors at super-synchronous speed.

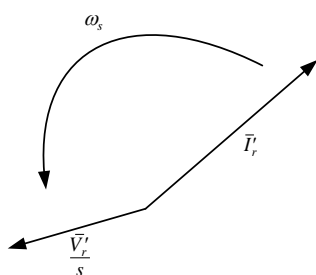


Fig. 4: Stator-referred rotor phasors.

direction with respect to the rotor voltage phasor (\bar{V}_r) as a result of a negative slip.

It is evident that in case the rotor voltage phasor leads the rotor current phasor on the rotor side, the referred rotor voltage phasor in the equivalent circuit (denoted as: \bar{V}'_r/s) also leads the referred rotor current phasor (\bar{I}'_r), regardless of whether the operation is in sub-synchronous or super-synchronous mode. It is observed that the relative position of the two phasors does not change if the rotational direction of the respective phasors is considered. This explanation confirms the intuitive expectation that the sign of reactive power from the rotor-side converter (RSC) to the rotor is identical to the sign of reactive power from the stator-referred rotor voltage source (\bar{V}'_r/s) to the stator-referred equivalent circuit, regardless of the operating mode, motoring or generating, sub-synchronous or super-synchronous speeds. This is shown schematically in Figs. 5a and 5b, in which the two reactive powers always have the same sign and differ only in magnitude ($Q'_r = Q_r/|s|$). The aforementioned expectation is consistent with [32], but the reactive power flow diagram in [34] ignores this fact, resulting in incompatible outcomes for super-synchronous speeds.

4. PHASOR DIAGRAMS FOR THE FOUR-QUADRANT OPERATION OF DFIM

In this section, phasor diagrams are drawn for a DFIM operating at unity stator power factor (zero stator reactive power). The diagrams are drawn for all four quadrants pertaining to the direction of stator active power and the slip sign. Therefore, the machine can operate in motoring/generating mode with sub/super-synchronous rotor speeds.

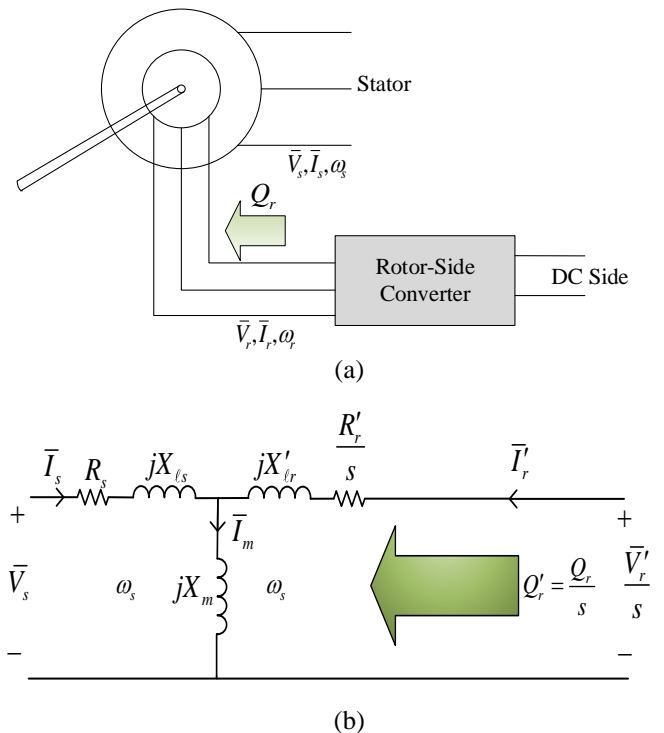


Fig. 5: (a) actual reactive power supplied into the rotor windings from the RSC, (b) reactive power supplied into the stator-referred equivalent circuit.

It is noteworthy in Fig. 6 that for sub-synchronous modes ((a) and (c)), the rotor angular frequency takes a positive value ($\omega_r > 0$), indicating that the phasors of actual (not stator-referred) rotor variables ($\bar{V}_r, \bar{I}_r, \bar{\psi}_r$) rotate in the same direction as the stator variables. In contrast, for super-synchronous modes ((b) and (d)), the slip and the rotor angular frequency become negative ($\omega_r < 0$); consequently, the direction of rotation of rotor phasors is reversed. All phasors associated with stator variables and rotor variables referred to the stator rotate at an angular frequency of ω_s .

The governing phasor equations in Fig. 5b are as follows:

$$\bar{V}_s = R_s \bar{I}_s + j\omega_s \bar{\psi}_s \tag{9}$$

$$\bar{\psi}_s = L_s \bar{I}_s + L_m \bar{I}'_r \tag{10}$$

$$\bar{V}'_r = R_r \bar{I}'_r + j\omega_r \bar{\psi}'_r \tag{11}$$

$$\bar{\psi}'_r = L_r \bar{I}'_r + L_m \bar{I}_s \tag{12}$$

In the following, a number of noteworthy findings in relation to the four quadrants illustrated in Fig. 6 are presented.

➤ Parts (a) and (b) represent motoring mode with a unity power factor; consequently, motoring convention dictates that the stator current phasor and stator voltage phasor are in phase. In contrast, the stator current phasor is 180 degrees out of phase with the stator voltage phasor in parts (c) and (d) corresponding to the generating mode.

➤ In motoring mode, the stator flux linkage phasor leads the rotor flux linkage phasor (referred to the stator). The opposite is true for generating mode.

➤ The angle between stator-referred rotor voltage and current phasors in parts (b) and (c), i.e., super-synchronous motoring and sub-synchronous generating modes, is less than 90 degrees. These two modes of operation are actually incompatible with a conventional induction machine with a shorted rotor winding. Consequently, these two new modes of DFIM operation require injection of active power into the rotor windings through an external circuit (the RSC).

➤ Assuming the stator's reactive power is zero and considering the reactive power requirement for core magnetization, one would expect that reactive power should be supplied to the rotor winding by the RSC in all four quadrants. From the viewpoint of the phasor diagram, the rotor current phasor should lag the rotor voltage phasor. Taking into account the positive direction of rotor phasors rotation in parts (a) and (c), it is evident that the rotor current

phasor lags the rotor voltage phasor. Also, in parts (b) and (d), when the reverse rotational direction of the rotor phasors is considered (due to negative ω_r), it is again interestingly observed that the rotor current phasor lags the rotor voltage phasor. This signifies the injection of reactive power into the rotor.

The validity of the phasor diagram results is supported by the consistency with physical considerations. It is worth noting that the phasor diagrams for all four quadrants in this study have been drawn using the same equivalent circuit, as shown in Fig. 5b. On the contrary, an alternative equivalent circuit has been utilised in [21, 22] to represent super-synchronous operating modes. This in turn increases the complexity of the problem and, if care is not taken in selecting the correct equivalent circuit, could lead to incorrect results.

5. EXPRESSION OF REACTIVE POWER BASED ON THE DYNAMIC MODEL

Up to now, the focus has been on the steady-state expression of the reactive power, employing phasor equations and the steady-state equivalent circuit. Expression of the equations in the synchronous dq reference frame is useful for gaining insight into the conditions that prevail under transient conditions. The dq transformation is implemented to investigate the relationship between the actual rotor side and stator-referred reactive powers. The particular steady-state case is taken into account as the primary objective is to determine the relative signs of reactive powers. Here, the same power-invariant abc to dq transformation is used as in [21, 22] in order to facilitate a comparison between the method presented in this paper and what has been presented in [21, 22].

Assuming that the rotor voltages are a balanced positive sequence set, they are referred to the synchronous reference frame (see (13)) where θ_s is the angle of the synchronous frame, θ_m is the rotor angle with respect to the stationary phase-a axis, and θ_0 is an initial phase angle, all expressed in electrical radians.

The slip angle can now be expressed as given by (14), where ω_r is the slip speed (which is also the angular frequency of the rotor voltages) in electrical radians/s.

$$\theta_r = \theta_s - \theta_m = \omega_r t \tag{14}$$

Recall that the rotor angular frequency was assumed to be positive ($\omega_r > 0$) in order to have positive sequence rotor voltages.

$$\begin{bmatrix} v_{dr} \\ v_{qr} \\ v_{0r} \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} \cos(\theta_s - \theta_m) & \cos(\theta_s - \theta_m - \frac{2\pi}{3}) & \cos(\theta_s - \theta_m - \frac{4\pi}{3}) \\ -\sin(\theta_s - \theta_m) & -\sin(\theta_s - \theta_m - \frac{2\pi}{3}) & -\sin(\theta_s - \theta_m - \frac{4\pi}{3}) \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} V_m \cos(\omega_r t + \theta_0) \\ V_m \cos(\omega_r t + \theta_0 - \frac{2\pi}{3}) \\ V_m \cos(\omega_r t + \theta_0 - \frac{4\pi}{3}) \end{bmatrix} \tag{13}$$

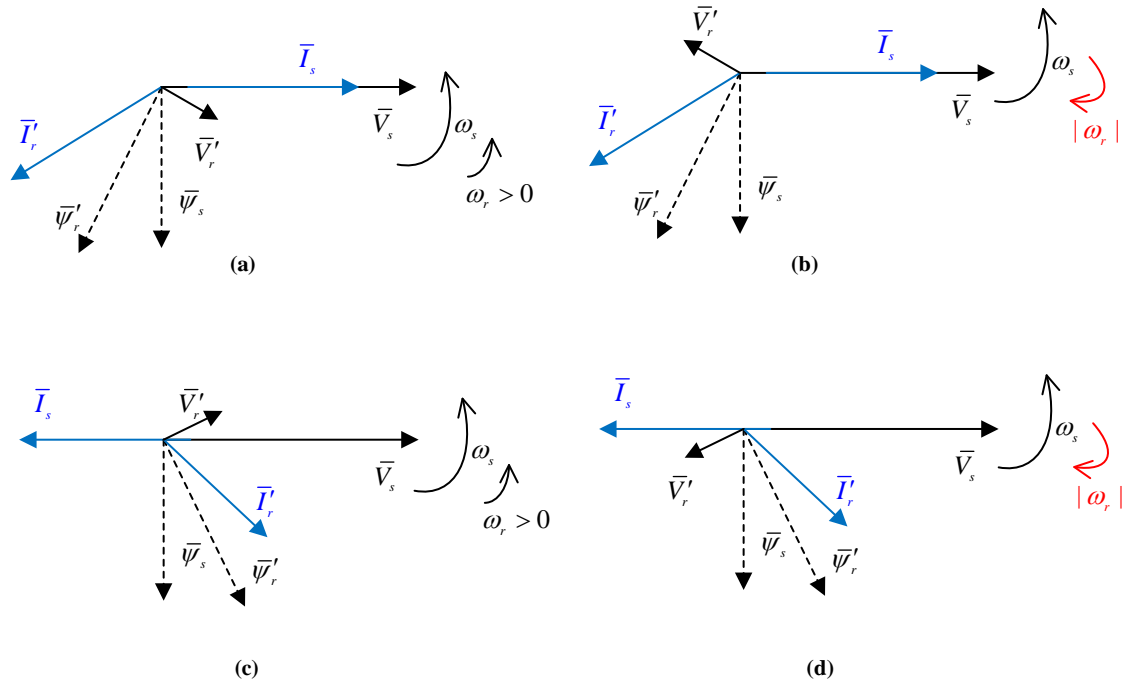


Fig. 6: Phasor diagram of DFIM for four quadrants (with unity stator power factor). (a) sub-synchronous motoring, (b) super-synchronous motoring, (c) sub-synchronous generating, (d) super-synchronous generating.

Using (13), the dq components of the rotor voltage can be easily determined. The result is:

$$v_{dr} = \sqrt{\frac{3}{2}}V_{rm} \cos(\theta_0), v_{qr} = \sqrt{\frac{3}{2}}V_{rm} \sin(\theta_0) \quad (15)$$

Because the voltages are assumed to be balanced, there will be no zero-sequence component ($v_{0r} = 0$).

The rotor voltage phasor can now be expressed using the d-axis and q-axis components in the synchronous reference frame, resulting in:

$$\bar{V}_r = v_{dr} + jv_{qr} = \sqrt{\frac{3}{2}}V_{rm} e^{j\theta_0} \quad (16)$$

It is observed that for ($\omega_r > 0$), a three-phase positive sequence rotor voltage set is converted to DC values in the synchronous reference frame.

Assuming that each rotor phase current lags its corresponding phase voltage by an angle of (φ), the three-phase rotor currents are represented as follows:

$$\begin{cases} i_{ar} = I_{rm} \cos(\omega_r t + \theta_0 - \varphi) = I_{rm} \cos(\omega_r t + \gamma) \\ i_{br} = I_{rm} \cos(\omega_r t + \gamma - \frac{2\pi}{3}) \\ i_{cr} = I_{rm} \cos(\omega_r t + \gamma - \frac{4\pi}{3}) \end{cases} \quad (17)$$

where γ has been defined to be: $\gamma = \theta_0 - \varphi$.

The dq components of the rotor current are derived as follows:

$$i_{dr} = \sqrt{\frac{3}{2}}I_{rm} \cos(\gamma), i_{qr} = \sqrt{\frac{3}{2}}I_{rm} \sin(\gamma) \quad (18)$$

Using (18), the following expression represents the rotor current phasor:

$$\bar{I}_r = i_{dr} + ji_{qr} = \sqrt{\frac{3}{2}}I_{rm} e^{j(\theta_0 - \varphi)} \quad (19)$$

Since the transformation is power-invariant, the active and reactive powers supplied into the rotor are expressed as follows:

$$P_r = \text{real}(\bar{V}_r \bar{I}_r^*) = (\frac{3}{2})V_{rm} I_{rm} \cos(\varphi) = v_{dr} i_{dr} + v_{qr} i_{qr} \quad (20)$$

$$Q_r = \text{imag}(\bar{V}_r \bar{I}_r^*) = (\frac{3}{2})V_{rm} I_{rm} \sin(\varphi) = v_{qr} i_{dr} - v_{dr} i_{qr}$$

In the preceding derivations ((13)-(20)), a sub-synchronous mode of operation corresponding to ($\omega_r > 0$) was assumed. As a next step, the super-synchronous mode of operation is studied. In this mode corresponding to ($\omega_r < 0$), the rotor voltages with the same expression will appear as a negative sequence set:

$$\begin{cases} v_{ar} = V_{rm} \cos(\omega_r t + \theta_0) = V_{rm} \cos(|\omega_r| t - \theta_0) \\ v_{br} = V_{rm} \cos(\omega_r t + \theta_0 - \frac{2\pi}{3}) = V_{rm} \cos(|\omega_r| t - \theta_0 + \frac{2\pi}{3}) \\ v_{cr} = V_{rm} \cos(\omega_r t + \theta_0 - \frac{4\pi}{3}) = V_{rm} \cos(|\omega_r| t - \theta_0 + \frac{4\pi}{3}) \end{cases} \quad (21)$$

The exact same d-axis and q-axis voltage components and rotor voltage phasor are obtained for negative sequence

voltages when they are transformed to the synchronous reference frame ((15) and (16)).

Now, the rotor currents at super-synchronous speed are expressed in their abc reference frame and then transformed to the synchronous reference frame. Again, it is supposed that the rotor phase current lags the phase voltage by an angle of (φ).

$$\begin{aligned} v_{ar} &= V_m \cos(\omega_r t + \theta_0) = V_m \cos(|\omega_r| t - \theta_0) \Rightarrow \\ i_{ar} &= I_m \cos(|\omega_r| t - \theta_0 - \varphi) = I_m \cos(\omega_r t + \theta_0 + \varphi) \end{aligned} \quad (22)$$

Similarly:

$$\begin{aligned} v_{br} &= V_m \cos(\omega_r t + \theta_0 - \frac{2\pi}{3}) = V_m \cos(|\omega_r| t - \theta_0 + \frac{2\pi}{3}) \\ \Rightarrow i_{br} &= I_m \cos(|\omega_r| t - \theta_0 + \frac{2\pi}{3} - \varphi) = \dots \\ &I_m \cos(\omega_r t + \theta_0 + \varphi - \frac{2\pi}{3}) \end{aligned} \quad (23)$$

and:

$$i_{cr} = I_m \cos(\omega_r t + \theta_0 + \varphi - \frac{4\pi}{3}) \quad (24)$$

Defining ($\gamma = \theta_0 + \varphi$), the rotor currents are transformed to the synchronous reference frame, resulting in:

$$i_{dr} = \sqrt{\frac{3}{2}} I_m \cos(\gamma), \quad i_{qr} = \sqrt{\frac{3}{2}} V_m \sin(\gamma) \quad (25)$$

Finally, the rotor current phasor is derived as follows:

$$\bar{I}_r = i_{dr} + j i_{qr} = \sqrt{\frac{3}{2}} I_m e^{j(\theta_0 + \varphi)} \quad (26)$$

Considering (16) and (26), it can be seen that, for super-synchronous operation, if the rotor phase currents lag (lead) their respective phase voltages by an angle of φ , the dq rotor currents will lead (lag) their respective dq rotor voltages by the same angle.

Since the relative angle between the rotor phase voltage and current has not changed (it can be assumed that a load is connected to the rotor terminals to fix the relative angles), it is expected that the rotor active and reactive powers remain the same as they were when the rotor was rotating at sub-synchronous speed. Manipulation of the dq voltage and current expressions ((15), (16), (25), and (26)) yields:

$$\begin{aligned} P_r &= (\frac{3}{2}) V_m I_m \cos(\varphi) = \text{real}(\bar{V}_r \bar{I}_r^*) = v_{dr} i_{dr} + v_{qr} i_{qr} \\ Q_r &= (\frac{3}{2}) V_m I_m \sin(\varphi) = -\text{imag}(\bar{V}_r \bar{I}_r^*) = -(v_{qr} i_{dr} - v_{dr} i_{qr}) \end{aligned} \quad (27)$$

Equation (27) demonstrates that while the expression for rotor active power is independent of whether the rotor is operating in sub-synchronous or super-synchronous mode, for the reactive power supplied to the rotor, it is important to consider whether the speed is sub-synchronous (where (20) is valid) or super-synchronous (where (27) is valid).

It is advisable not to modify the equivalent circuit referred to the stator and the phasor diagram based on this circuit in accordance with the operating speed of the

machine. Rather, while other relationships remain unchanged for super-synchronous speed, the expression for reactive power supplied into the rotor of the DFIM from the RSC should be modified. The following general expression for rotor reactive power is proposed for use at all operating speeds:

$$Q_r = \text{sign}(s) * \text{imag}(\bar{V}_r \bar{I}_r^*) = \text{sign}(s) * (v_{qr} i_{dr} - v_{dr} i_{qr}) \quad (28)$$

The coefficient ($\text{sign}(s)$) in (28) accounts for sub-synchronous (positive slip, therefore $\text{sign}(s) = 1$) and super-synchronous (negative slip, therefore $\text{sign}(s) = -1$) speeds.

6. RESULTS AND DISCUSSION

To demonstrate the effectiveness of the proposed rotor reactive power relation given by (28), a DFIM with the specifications listed in Table 1 is simulated in Matlab® Simulink under vector control using reference values of rotor speed and stator reactive power. The mechanical torque is taken as the input to the system. This could be load torque in the case of a motor or turbine torque in the case of a generator. As per the adopted motoring convention, positive values of mechanical torque and stator active power indicate motoring operation, whereas negative values of these variables indicate generation. In accordance with [21], the reference stator reactive power is kept at zero in order for the machine to be magnetized through the rotor; in other words, under this control scenario, positive reactive power must be supplied to the rotor in various operating modes. For the first 15 seconds, the reference speed is maintained at 0.7 pu (i.e., sub-synchronous speed), and then it changes to 1.3 pu (i.e., super-synchronous speed). Fig. 7 shows the active and reactive powers supplied into the stator winding for motoring operation. To simulate this mode of operation, a positive value of input torque (load torque) was used. Given the objective of this paper, which is not centered around control of the system's performance, transients occurring in the waveforms are not deemed important and emphasis is made on their steady-state values. Fig. 8 depicts the variation of rotor abc currents when in motoring mode at two distinct operating speeds. On the left side of the figure, corresponding to the sub-synchronous speed, the currents indicate a positive sequence. Conversely, on the right side of the figure, corresponding to the super-synchronous speed, the currents indicate a negative sequence. The same phase reversal phenomenon is observed for rotor voltages as well (Fig. 9). Fig. 10 illustrates the relative time displacement of the rotor phase-a voltage and current waveforms for motoring operation at sub-synchronous (left) and super-synchronous (right) speeds. It is interesting to note that the rotor current always lags the rotor voltage, regardless of the operating speed. This was expected because, with the stator's reactive power maintained at zero, the DFIM should be magnetized

Table 1: Specifications of the simulated DFIM.

Parameter	Unit	Value
Stator resistance	pu	0.05
Rotor resistance	pu	0.02
Stator leakage inductance	pu	0.1
Rotor leakage inductance	pu	0.1
Magnetizing inductance	pu	5
Base frequency	Hz	60

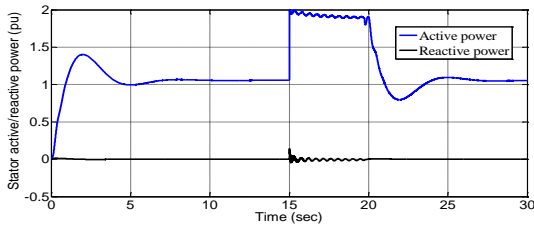


Fig. 7: Stator active and reactive powers for positive input torque (motoring operation).

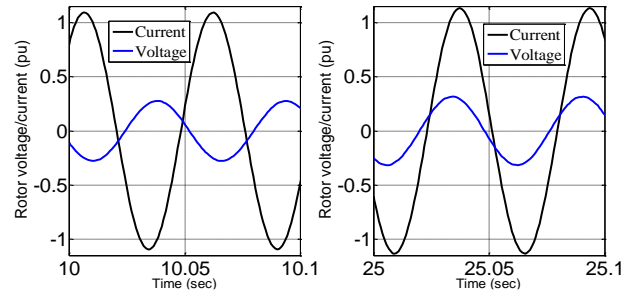


Fig. 10: Relative time displacement of rotor voltage and current for motoring mode (left: sub-synchronous, right: super-synchronous).

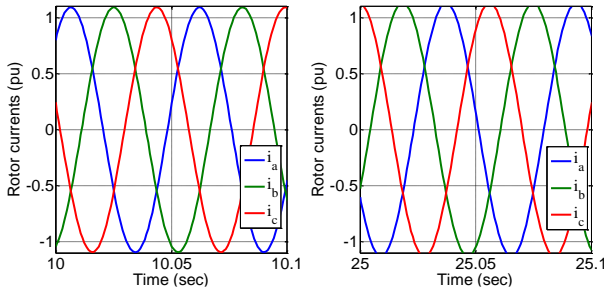


Fig. 8: Rotor phase currents for motoring mode (left: sub-synchronous, right: super-synchronous).

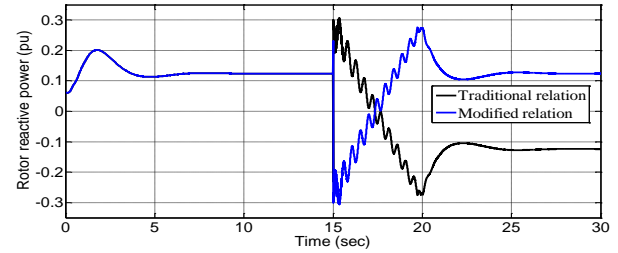


Fig. 11: Rotor reactive power using traditional and modified relations (motoring operation).

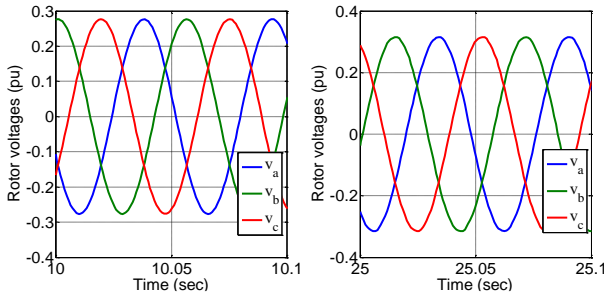


Fig. 9: Rotor phase voltages for motoring mode (left: sub-synchronous, right: super-synchronous).

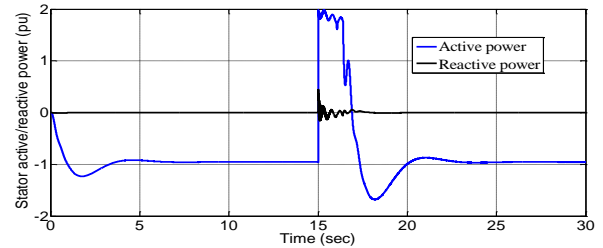


Fig. 12: Stator active and reactive powers for negative input torque (generating operation).

through the rotor; therefore, from the standpoint of the RSC, the rotor is inductive. For sub-synchronous speed (left), the rotor current lags the rotor voltage by more than 90 degrees, indicating that the rotor active power is negative (power supplied from the rotor to the RSC). In contrast, for super-synchronous speed (right), the rotor current lags the rotor voltage by an angle less than 90 degrees, indicating that the rotor active power is positive (power supplied into the rotor from the RSC). This is consistent with what has been reported in the literature [3]. Fig. 11 shows a comparison between the reactive power of the rotor calculated with the traditional equation (20) and the modified equation (28). Clearly, the rotor's steady-state reactive power must be positive for the machine to be magnetized. While both expressions produce similar results for sub-synchronous speeds, only the modified relation derived in this paper yields the correct rotor reactive power for super-synchronous speeds.

Fig. 12 shows the variation of stator active and reactive powers during generating operation. This mode of operation is achieved by keeping the input torque (turbine torque) negative. Fig. 13 depicts the variation of rotor *abc* currents in generating mode, where the phase sequence of rotor currents changes as the speed changes from sub-synchronous (left) to super-synchronous (right). This behavior was also observed for motoring operation (Fig. 8). As seen in Fig. 14, when the

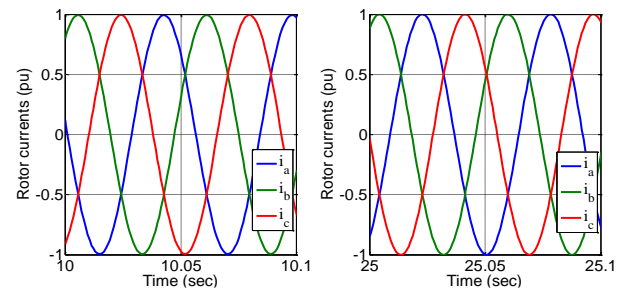


Fig. 13: Rotor phase currents for generating mode (left: sub-synchronous, right: super-synchronous).

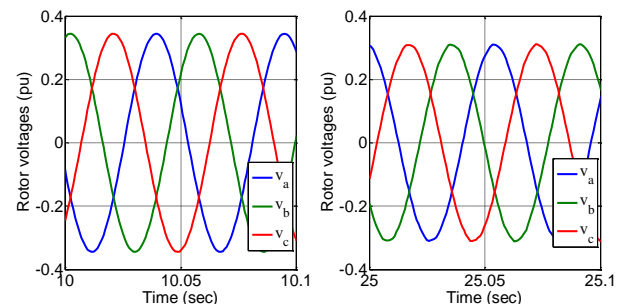


Fig. 14: Rotor phase voltages for generating mode (left: sub-synchronous, right: super-synchronous).

rotor speed goes from sub-synchronous to super-synchronous, the sequence of rotor voltages also changes. Variation of rotor phase-a voltage and current for generating operation is illustrated in Fig. 15 for both sub-synchronous (left) and super-synchronous (right) speeds. Similar to Fig. 10, rotor current lags rotor voltage under both conditions for the same reason.

For sub-synchronous generating operation, active power must be supplied to the rotor windings, necessitating a phase angle difference of less than 90 degrees between rotor voltage and current (left); for super-synchronous generating operation, active power must be drawn from the rotor to the RSC, necessitating a phase angle difference of greater than 90 degrees (right). This is again consistent with [3]. As was done for the motoring mode, the traditional and modified rotor reactive power expressions are compared for the generating mode in Fig. 16. Again, it is evident that both relations produce identical results for sub-synchronous speeds. Nonetheless, only the modified relation provides the correct sign of reactive power for super-synchronous speeds. To validate the proposed modified rotor reactive power relation (28) further, the reactive power requirement of the rotor for various stator reactive powers is obtained under super-synchronous rotor speeds for both motoring and generating modes, and indicated in Fig. 17. The predictions from the proposed method and the existing reactive power relation are both drawn for the sake of comparison. The flowchart of computations using the proposed method is depicted in Fig. 18. For a DFIM with parameters given in Table 1, super-synchronous operating mode with ($s = -0.25$) is investigated using the proposed modified relation and the traditional relation. The stator voltage is assumed to be $1\angle 0^\circ pu$. For this scenario, the stator side active power is taken to be $\pm 0.9 pu$. (positive for motoring and negative for generating operation). Stator reactive power is allowed to vary in the range $Q_{min} \leq Q_s \leq Q_{max}$ and the variation of rotor reactive power is investigated. Taking the equivalent circuit of DFIM into account, as the inductive elements in the circuit absorb reactive power, there is need of reactive power from the rotor side ($Q_r > 0$) to magnetize the machine when the stator feeds reactive power to the grid ($Q_s < 0$). This is in accordance with the predictions from the proposed formulation, but in contrast to the predictions of the traditional formulation, confirming the validity of the proposed method.

6.1. Sensitivity Analysis

For further analysis, the sensitivity of rotor reactive power direction to parameter changes (which can occur due to temperature rise or saturation effects) is investigated by varying the magnetizing inductance (L_m) and computing Q_r with respect to Q_s for motoring and generating modes, as illustrated in Fig. 19 and Fig. 20, respectively. The DFIM is again assumed to be at super-synchronous mode of operation ($s = -0.25$) with a stator voltage of $1\angle 0^\circ pu$. It is clear that the impact of machine parameter changes on rotor reactive power direction is minimal.

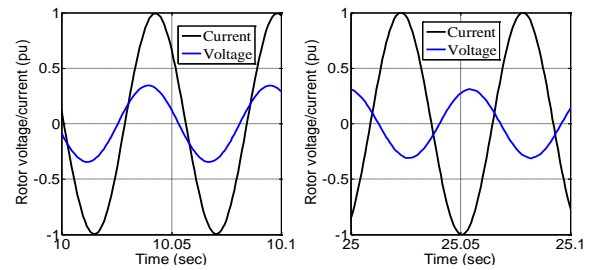


Fig. 15: Relative time displacement of rotor voltage and current for generating mode (left: sub-synchronous, right: super-synchronous).

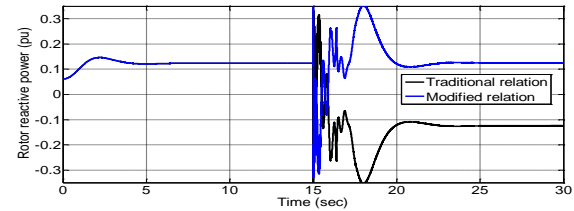


Fig. 16: Rotor reactive power using traditional and modified relations (generating operation).

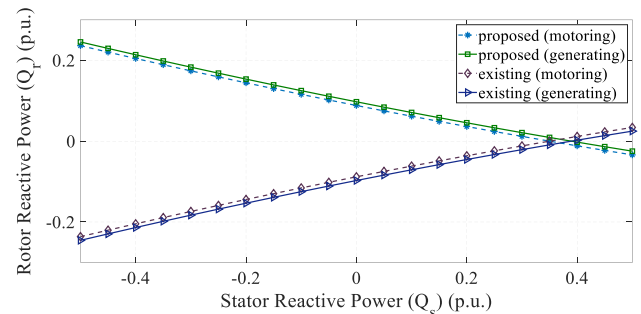


Fig. 17: Rotor reactive power versus stator reactive power for super-synchronous operation ($P_s = 0.9 pu$ for motoring mode and $P_s = -0.9 pu$ for generating mode).

7. CONCLUSION

A comprehensive investigation into the impact of super-synchronous operation of DFIMs on rotor reactive power was provided in this paper. Under such operating conditions, the slip is negative, causing the rotor frequency to be negative. Clearly, a negative rotor frequency has no physical significance, but it does indicate a change of the phase sequence of rotor voltages and currents. It has been demonstrated that the well-established stator-referred equivalent circuit of DFIM is valid for both sub-synchronous and super-synchronous speeds; yet, care must be taken when dealing with phasor diagrams drawn for super-synchronous speeds, where the actual rotor variables rotate in the opposite direction to the stator-referred variables. Also, when calculating rotor reactive power using dq rotor variables, the traditional relation must be modified to account for the sign of slip. While the focus of this article was on conventional (brushed) doubly fed induction machines, the scope of the study can be extended to include other types of doubly fed machines, such as brushless doubly fed machines with

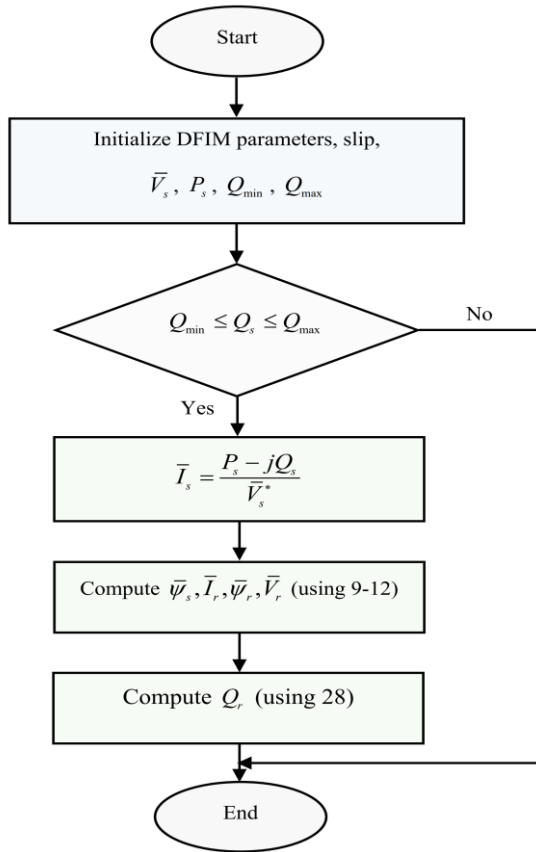


Fig. 18: Flowchart of Q_r versus Q_s calculation for $Q_{\min} \leq Q_s \leq Q_{\max}$ at a predefined slip and fixed P_s and \bar{V}_s .

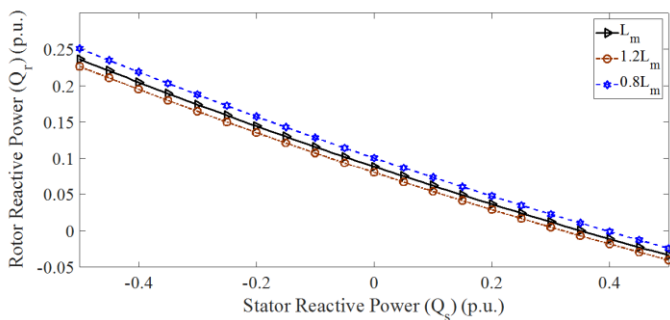


Fig. 19: Impact of variation of magnetizing inductance on rotor reactive power for motoring mode ($P_s = 0.9 pu$).

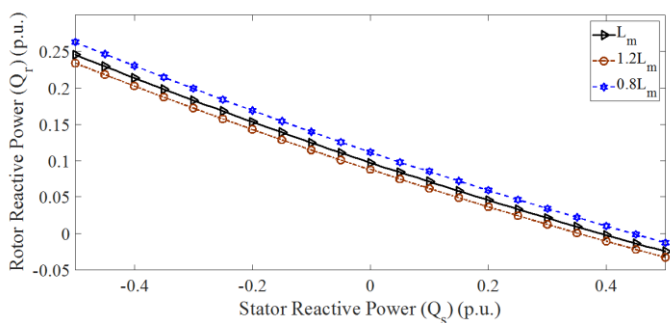


Fig. 20: Impact of variation of magnetizing inductance on rotor reactive power for generating mode ($P_s = -0.9 pu$).

induction or reluctance rotors. These electromechanical converters enjoy comparable advantages of a fractionally rated converter like the DFIM. Simultaneously, they eliminate the need for brushes and slip rings found in the DFIM, rendering them more maintenance-free. A comparable frequency conversion method is utilised for their modelling, where the parameters of the secondary stator winding are referred to the primary stator winding. Future research is going to focus on investigating the reactive power relation of the secondary winding, taking the potential reversal in phase sequence of its voltages and currents into account.

CREDiT AUTHORSHIP CONTRIBUTION STATEMENT

Mohammad Naser Hashemnia: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Roles/Writing - original draft, Writing - review & editing.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy has been completely observed by the authors.

REFERENCES

- [1] A. K. Abdulabbas, M. A. Alawan, and D. K. Shary, "Limits of reactive power compensation of a doubly fed induction generator based wind turbine system," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 5, pp. 2521–2534, Oct. 2023.
- [2] F. Z. Messaoud, T. Hamza, B. Ouamri, M. Abasi, and A. R. Zerek, "Decoupled SMC of DFIG Based Multi-Level Inverter," in *2021 IEEE 1st International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA)*, May 2021.
- [3] G. Abad, J. López, M. Rodríguez, L. Marroyo, and G. Iwanski, *Doubly fed induction machine: modeling and control for wind energy generation*. Oxford: Wiley-Blackwell, 2011.
- [4] L. Fan, and Z. Miao, *Modeling and Analysis of Doubly Fed Induction Generator Wind Energy Systems*. Academic Press, 2015.
- [5] E. N. Sanchez and Riemann Ruiz-Cruz, *Doubly Fed Induction Generators*. CRC Press, 2016.
- [6] D. Xu, F. Blaabjerg, W. Chen, and N. Zhu, *Advanced control of doubly fed induction generator for wind power systems*. Hoboken, New Jersey John Wiley Et Sons, 2018.
- [7] A. Abdelbaset, Y. S. Mohamed, A.-H. M. El-Sayed, A. E. Hussein, and A. Ahmed, *Wind Driven Doubly Fed Induction Generator*. Springer, 2017.

- [8] A. Sguarezi, *Model Predictive Control for Doubly-Fed Induction Generators and Three-Phase Power Converters*. Elsevier, 2022.
- [9] A. Dòria-Cerezo, M. A. Hossain, and M. Bodson, "Complex-Valued Sliding Mode Controllers for Doubly-Fed Induction Motors," *IEEE Transactions on Control Systems and Technology*, vol. 31, no. 3, pp. 1336–1344, May 2023.
- [10] F. Pourmirzaei-Deylami, and A. Darabi, "Steady-state performance analysis for optimal operation determination of doubly fed induction motors," *The Journal of Engineering*, Dec. 2022.
- [11] M. Zerzeri, A. Khedher, and F. Jallali, "Steady-state characteristics of DFIM : the potentialities of integration in electrical traction systems," *Research Square (Research Square)*, Feb. 2023.
- [12] X. Roboam, "A review of powertrain electrification for greener aircraft," *Energies*, vol. 16, no. 19, p. 6831, Jan. 2023.
- [13] S. Li, "Converters loading balance and stability verification for doubly-fed induction generator," *CSEE Journal of Power and Energy Systems*, 2022.
- [14] R. Gianto, "Steady-state model of DFIG-based wind power plant for load flow analysis," *IET Renewable Power Generation*, Mar. 2021.
- [15] R. Gianto, K. H. Khwee, H. Priyatman, and M. Rajagukguk, "Two-port network model of fixed-speed wind turbine generator for distribution system load flow analysis," *Telecommunication Computing Electronics and Control (TELKOMNIKA)*, vol. 17, no. 3, p. 1569, Jun. 2019.
- [16] R. Gianto, "Constant voltage model of DFIG-based variable speed wind turbine for load flow analysis," *Energies*, vol. 14, no. 24, p. 8549, Dec. 2021.
- [17] R. Gianto, "Constant power factor model of DFIG-Based wind turbine for steady state load flow studies," *Energies*, vol. 15, no. 16, pp. 6077–6077, Aug. 2022.
- [18] R. Gianto, "Integration of DFIG-Based variable speed wind turbine into load flow analysis," in *2021 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, 2021, pp. 63-66.
- [19] K. H. Shin, and T. A. Lipo, "A super-synchronous doubly fed induction generator option for wind turbine applications," in *2016 IEEE Energy Conversion Congress and Exposition (ECCE)*, 2016, pp. 1-7.
- [20] T. H. Ortmeier, "Negative frequency aspects of doubly fed machine analysis," *Proceedings of the IEEE*, vol. 71, no. 8, pp. 1017-1017, Aug. 1983.
- [21] V. S. S. Kumar, and D. Thukaram, "Accurate steady-state representation of a doubly fed induction machine," *IEEE Transactions on Power Electronics*, vol. 30, no. 10, pp. 5370-5375, Oct. 2015.
- [22] V. S. S. Kumar, and D. Thukaram, "Alternate proof for steady-state equivalent circuit of a doubly fed induction machine," *IEEE Transactions on Power Electronics*, vol. 31, no. 8, pp. 5378-5383, Aug. 2016.
- [23] V. S. S. Kumar, and D. Thukaram, "Accurate modeling of doubly fed induction generator based wind farms in load flow analysis," *Electric Power Systems Research*, vol. 155, pp. 363–371, Feb. 2018.
- [24] C.V.S. Anirudh, and V. S. S. Kumar, "Enhanced modelling of doubly fed induction generator in load flow analysis of distribution systems," *IET Renewable Power Generation*, vol. 15, no. 5, pp. 980–989, Jan. 2021.
- [25] D. R. Karthik, and S. M. Kotian, "Initialization of doubly-fed induction generator wind turbines using noniterative method," in *2019 8th International Conference on Power Systems (ICPS)*, 2019, pp. 1-6.
- [26] D. R. Karthik, S. M. Kotian, and N. S. Manjarekar, "A direct method for calculation of steady-state operating conditions of a doubly fed induction generator," in *2021 9th IEEE International Conference on Power Systems (ICPS)*, 2021, pp. 1-6.
- [27] Y. V. Pavan Kumar, and R. Bhimasingu, "Performance analysis of static versus rotary DC/AC power converters for hybrid renewable energy based microgrid applications," in *2016 IEEE Region 10 Conference (TENCON)*, 2016, pp. 1456-1461.
- [28] V. S. Sravan Kumar, "Computation of initial conditions for dynamic analysis of a doubly fed induction machine based on accurate equivalent circuit," in *2019 IEEE International Electric Machines & Drives Conference (IEMDC)*, 2019, pp. 307-313.
- [29] Y. Ju, F. Ge, W. Wu, Y. Lin, and J. Wang, "Three-phase steady-state model of doubly fed induction generator considering various rotor speeds," *IEEE Access*, vol. 4, pp. 9479-9488, 2016.
- [30] P. C. Krause, Oleg Wasynczuk, S. D. Sudhoff, S. Pekarek, and Institute Of Electrical And Electronics Engineers, *Analysis of electric machinery and drive systems*. Hoboken, New Jersey: Wiley, 2013.
- [31] P. C. Krause and T. R. Krause, *Introduction to modern analysis of electric machines and drives*. John Wiley & Sons, 2022.
- [32] T. J. E. Miller, "Theory of the doubly-fed induction machine in the steady state," in *XIX International Conference on Electrical Machines (ICEM 2010)*, 2010, pp. 1-6.
- [33] G. T. Heydt, "The meaning, analysis, and consequences of negative frequency in electric power systems," in *2022 North American Power Symposium (NAPS)*, 2022, pp. 1-6.
- [34] N. Bianchi, L. Alberti and S. Bolognani, "A design-oriented model of doubly-fed induction machine," in *2011 IEEE International Electric Machines & Drives Conference (IEMDC)*, 2011, pp. 557-562.

BIOGRAPHY

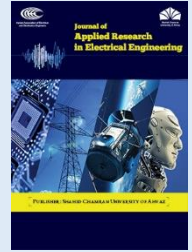
Mohammad Naser Hashemnia was born in Mashhad, Iran, in 1983. He received the B.S. degree in electrical power engineering from Ferdowsi University, Mashhad, Iran, in 2006 and the M.S. degree in electrical power engineering from the University of Tehran, Tehran, Iran, in 2008. He got his Ph.D. degree in electrical power

engineering from Sharif University of Technology, Tehran, Iran in 2013. Since 2015 he has been an assistant professor in Mashhad branch, Islamic Azad University, Mashhad, Iran. His main research interests include analysis, modelling and simulation of electric machines, advanced control methods for electrical drives and renewable energy systems.

Copyrights

© 2024 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).





Research Article

Full-Duplex Device-to-Device Relays in a Novel Hybrid and Adaptive Joint Relaying Network: Symbol Error Analysis and Optimum Power Allocation

Mehdi Basiri Abarghouei, and Reza Saadat *

Department of Electrical Engineering, Yazd University, Yazd, Iran

* Corresponding Author: rsaadat@yazd.ac.ir

Abstract: This paper proposes a new relaying protocol for transmitting from a cellular user to the base station with the joint cooperation of a Full-Duplex (FD)-enabled Device-to-Device (D2D) pair. In the proposed scheme, the receiver of the D2D acts as a relay, with the cooperation of its transmitter pair via D2D communication between them. The cooperation approach of the D2D receiver is chosen as Adaptive Decode-and-Forward (ADF), while the cooperation strategy of the D2D transmitter is chosen as either ADF, Amplify-and-Forward (AF), or Hybrid relaying protocol. These scenarios are named "Decode and Joint Cooperation," "Amplify and Joint Cooperation," and "Hybrid and Adaptive Joint Cooperation," respectively. The Average Symbol Error Probability (ASEP) of the system is studied over independent and identically distributed (i.i.d) complex Gaussian (Rayleigh envelope) channels, with perfect Channel State Information (CSI) in the presence of Residual Self-Interference (RSI) at the FD relays, as well as Co-Channel Interference (CCI). Moreover, closed-form and high Signal-to-Interference-plus-Noise Ratio (SINR) tight ASEP approximations are established. The optimum power allocation is formulated based on the approximate relations, and the optimal solutions and their characteristics are discussed in detail. Analytical comparisons and simulations confirm the theoretical results and demonstrate significant performance improvements.

Keywords: Cooperative communication, hybrid and adaptive DF/AF protocol, Device-to-Device (D2D) communication, full duplex communication, average symbol error probability, optimal power allocation.

Article history

Received 09 July 2023; Revised 23 August 2023; Accepted 17 November 2023; Published online 4 February 2024.

© 2024 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)

How to cite this article

M. Basiri Abarghouei, and R. Saadat, "Full-duplex device-to-device relays in a novel hybrid and adaptive joint relaying network: symbol Error analysis and optimum power allocation," *J. Appl. Res. Electr. Eng.*, vol. 2, no. 2, pp. 194-205, 2023. DOI: [10.22055/jaree.2023.44250.1074](https://doi.org/10.22055/jaree.2023.44250.1074)



1. INTRODUCTION

The current surge in demand for high-speed and reliable connectivity poses significant challenges to wireless networks [1]. In this context, cooperative communication has emerged as a promising technique in cellular networks. By leveraging the cooperation of one or more relays with cellular users, particularly those situated at the periphery of a cell, the performance parameters of the network can be enhanced [2-5]. The relaying protocols can be categorized into two groups, namely AF and DF [6]. Within the AF relaying protocol, the received signal from a source node is amplified and retransmitted. On the other hand, in the DF approach, a relay node demodulates and decodes the received source signal before transmission. To prevent error propagation and decoding errors, the Adaptive DF (ADF) technique can be employed. In this technique, the relay node switches its

operation mode based on certain criteria. Although AF requires low implementation complexity at the relay, it amplifies the noise power. When the quality of the source-relay channel is favourable and the signal can be accurately decoded at the relay node, the Adaptive DF (ADF) approach has been shown to outperform the AF approach [7, 8]. In order to leverage the benefits of both AF and ADF approaches, hybrid protocols, known as Hybrid Adaptive DF (HADDF), have been suggested in previous studies [9, 10]. These protocols involve the relay node selecting the most appropriate relaying approach (AF or DF) based on the

Channel State Information (CSI) between the users [11]. Device-to-device (D2D) communication is an emerging technology that presents numerous benefits, including wireless peer-to-peer services and higher spectral efficiency. It is used in various fields, including network traffic

offloading, public safety, social services, and applications like gaming and military applications [12]. This paper proposes a new hybrid and adaptive relaying protocol for transmitting data from a cellular user to the base station, utilizing the collaboration of a Full-Duplex (FD)-enabled Device-to-Device (D2D) pair in the context of next-generation wireless networks. The proposed approach involves the D2D receiver acting as a relay, employing an Adaptive-Decode-and-Forward (ADF) relaying approach. The D2D transmitter pairs with the receiver through existing D2D communication, using AF, ADF, or a Hybrid relaying protocol. The selection of the relaying approach is determined based on the D2D transmitter's ability to correctly detect the symbol sent by the cellular user. If successful, the DF approach is used; otherwise, the AF approach is employed. This strategy aims to achieve reliable and efficient information transmission while mitigating the performance degradation of Cellular Users (CUs) caused by D2D users reusing the cellular network's spectrum. Additionally, the use of Full-Duplex (FD) technology improves spectral efficiency, allowing nodes to transmit and receive signals simultaneously over the same frequency in the same time slot. However, it is important to consider the impact of self-interference (SI) on FD communication [15].

The remainder of this paper is organized as follows: Section 2 describes the considered system and presents the physical-layer mathematical framework. In Section 3, we analyze the symbol error probability in the presence of Residual Self-Interference (RSI) and Co-Channel Interference (CCI) in cases where the D2D transmitter cooperates with AF, ADF, or Hybrid relaying protocols, respectively. Due to the complexity of the closed-form formulation, we propose presenting tight approximations with a high Signal-to-Noise Ratio (SNR) to show the asymptotic performance of the systems. Based on this, in Section 4, we formulate the optimal power allocation problem and discuss its solution in cooperative approaches. Furthermore, we compare the efficiency of the hybrid relaying protocol with the case where the D2D transmitter cooperates with either AF or ADF approach. In Section 5, we provide numerical simulations to validate the theoretical findings. Section 6 concludes the paper.

2. SYSTEM MODEL AND MATHEMATICAL FRAMEWORK

Without loss of generality, in an uplink cellular network with both cellular and D2D communications, we focus on a simple model composed of a base station (D), a cellular user (S) and a pair of D2D users (T, R), as shown in Fig. 1. D2D users are operated in FD and equipped with a transmitting and a receiving antenna. The cellular user sends its information symbols to the base station with the help of D2D pair. We assume that the communication takes place in a Rayleigh flat fading environment, where the channel experiences random fluctuations. However, we assume perfect Channel State Information (CSI) at the receiver nodes, meaning that they have accurate knowledge of the channel conditions. This information helps in optimizing the power allocation and relaying protocols. In our system, each Full-Duplex (FD) node is subject to random Residual Self-Interference (RSI) between its transmit and receive antennas. We model the RSI channels as zero mean circularly symmetric complex Gaussian (ZM-CSCG) random variable [16].

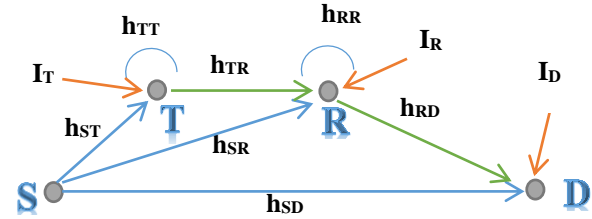


Fig. 1: A simplified model of the proposed hybrid and adaptive cooperative approach.

Moreover, it is assumed that each cellular user can share its uplink channel with a D2D communication in the network, which leads to interference between the D2D link and the cellular uplink. In addition, we fix the total transmitted power P and divide it between users as follows:

$$P_S = \delta_i P \tag{1}$$

$$P_T = \delta_i' P \tag{2}$$

$$P_R = (1 - \delta_i - \delta_i') P \tag{3}$$

in which the parameters δ_i and δ_i' are power allocation coefficients with $i \in \{AF, ADF, H\}$ which AF , ADF and H refers to the relaying strategy of the D2D transmitter. δ_i and δ_i' are in $(0,1)$ interval and satisfy the condition $\delta_i + \delta_i' < 1$.

The procedure can be delineated into two distinct phases: the transmission phase and the relaying phase. The destination effectively combines the signal received from the source during Phase 1 with the signal received from the relay during Phase 2. This combination is achieved through the utilization of maximum-ratio combining (MRC) techniques, allowing for the extraction of the transmitted symbols.

During the transmission phase, the cellular user S broadcasts its information to the destination node D with transmission power P_S and the information is also received by D2D receiver R and D2D transmitter T. The received signals at D, R and T are denoted by y_{SD} , y_{SR} and y_{ST} , respectively, and can be written as follows:

$$y_{SD} = \sqrt{P_S} h_{SD} x_S + n_{SD} + I_D \tag{4}$$

$$y_{SR} = \sqrt{P_S} h_{SR} x_S + n_{SR} + I_R + h_{RR} u_R \tag{5}$$

$$y_{ST} = \sqrt{P_S} h_{ST} x_S + n_{ST} + I_T + h_{TT} u_T \tag{6}$$

in which x_S is the transmitted information symbol with unit average energy and belongs to the MQAM constellation with $M = 2^K$. h_{SD} , h_{SR} and h_{ST} are the channel coefficients from the cellular user to the base station, D2D receiver and D2D transmitter and are modelled as ZM-CSCG random variable with variances σ_{SD}^2 , σ_{SR}^2 and σ_{ST}^2 , respectively. Also, n_{SD} , n_{SR} and n_{ST} are additive white Gaussian noises (AWGN) and are modeled as ZM-CSCG random variable with common variance N_0 . Further, I_D , I_R and I_T are CCI terms at D, R and T, respectively and are modelled as ZM-CSCG random variable with variances $\sigma_{I_D}^2$, $\sigma_{I_R}^2$ and $\sigma_{I_T}^2$. Moreover, h_{RR} and h_{TT} are the RSI channel coefficients at R and T with variances $\sigma_{h_{RR}}^2$ and $\sigma_{h_{TT}}^2$, respectively. Finally, u_R and u_T represent the signals transmitted by R and T to the D and R, respectively.

So we have $P_R = \mathbb{E}\{|u_R(t)|^2\}$ and $P_T = \mathbb{E}\{|u_T(t)|^2\}$. Therefore, the first terms in (4), (5) and (6) represent the desired signals and the others are noise terms because of AWGN, CCI and RSI. The sum of all noise terms are distributed as ZM-CSCG with effective noise power $\sigma_{n,SD}^2$, $\sigma_{n,SR}^2$ and $\sigma_{n,ST}^2$ respectively as follows:

$$\sigma_{n,SD}^2 = N_0 + \sigma_{I_D}^2 \quad (7)$$

$$\sigma_{n,SR}^2 = N_0 + \sigma_{I_R}^2 + P_R \sigma_{h_{RR}}^2 \quad (8)$$

$$\sigma_{n,ST}^2 = N_0 + \sigma_{I_T}^2 + P_T \sigma_{h_{TT}}^2 \quad (9)$$

The relaying phase of the proposed approach comprises two cooperation scenarios. In these scenarios, R acts as a relay and leverages the cooperation of its transmitter pair T through existing D2D communication. The selection of the DF or AF approach is based on the ability to accurately detect the symbol transmitted by the cellular user. It is assumed that ideal operation is maintained for analytical tractability. In practice, we may apply an SNR threshold at the relay nodes. If the received SINR is higher than the threshold, then the symbol is correctly decoded with a high probability. Details of threshold optimization at the relay can be found in [17]. In the relaying phase, cooperation occurs based on the algorithm shown in Fig. 2. Based on the cooperation strategy of the D2D transmitter node, three scenarios three scenarios can be imagined as follows:

- 1) In the first scenario D2D transmitter node cooperates with ADF relaying protocol. We call this scenario "Decode and Joint Cooperation".
- 2) In the second scenario D2D transmitter node cooperates with AF relaying protocol. We call this scenario "Amplify and Joint Cooperation".

- 3) In the last scenario D2D transmitter node cooperates with hybrid relaying protocol. We call this scenario "Hybrid and Adaptive Joint Cooperation".

Next, we focus on studying physical-layer mathematical framework of the relaying phase.

In the relaying phase of the "Decode and Joint Cooperation" scenario, the D2D transmitter T cooperates only if it decodes the transmitted symbol correctly. Therefore, u_T in (6) equals $\sqrt{\tilde{P}_T} \tilde{x}_S$, where \tilde{x}_S denotes the decoded symbol by user T, and \tilde{P}_T denotes the transmitted power. Clearly, \tilde{P}_T equals P_T in cooperation mode and is equal zero in non-cooperation mode. Then, the D2D receiver R receives the signal as follows:

$$y_{TR,ADF} = \sqrt{\tilde{P}_T} h_{TR} \tilde{x}_S + n_{TR} + I_R + h_{RR} u_R \quad (10)$$

in which h_{TR} is the channel coefficients from the relay R to the base station and is modelled as ZM-CSCG random variable with variances σ_{TR}^2 . Also, n_{TR} is AWGN and is modeled as ZM-CSCG random variable with common variance N_0 . Sum of all noise terms in (10) are distributed as ZM-CSCG with effective noise power $\sigma_{n,TR}^2$ as follows:

$$\sigma_{n,TR}^2 = N_0 + \sigma_{I_R}^2 + P_R \sigma_{h_{RR}}^2 \quad (11)$$

In the relaying phase of the "Amplify and Joint Cooperation" scenario, the D2D transmitter T simply amplifies the signal received from the cellular user and retransmits the resulting signal to the D2D receiver R.

Therefore, u_T in (6) equals $\sqrt{\frac{P_T}{\mathbb{E}\{|y_{ST}|^2\}}} y_{ST}$ and the D2D receiver R receives the signal as follows:

$$y_{TR,AF} = h_{TR} \beta_T y_{ST} + n_{TR} + I_R + h_{RR} u_R \quad (12)$$

Finally, in the relaying phase of the "Hybrid and Adaptive Joint Cooperation" scenario, the D2D transmitter T cooperates with the DF relaying protocol and switches to the AF approach in the case of decoding errors. So we have:

$$y_{TR,H} = \begin{cases} \sqrt{\frac{P_T}{\mathbb{E}\{|y_{ST}|^2\}}} h_{TR} y_{ST} + n_{TR} + I_R + h_{RR} u_R ; & AF \\ \sqrt{\tilde{P}_T} h_{TR} \tilde{x}_S + n_{TR} + I_R + h_{RR} u_R ; & DF \end{cases} \quad (13)$$

In the relaying phase of all of the mentioned scenarios, the D2D receiver R cooperates with the ADF relaying protocol. Therefore, we have:

$$y_{RD}^{J,scenario} = \sqrt{\tilde{P}_R} h_{RD} \tilde{x}_S + n_{RD} + I_D \quad (14)$$

in which h_{RD} is the channel coefficients from the relay R to the base station and is modelled as ZM-CSCG random variable with variances σ_{RD}^2 . Also, n_{RD} is AWGN and is modeled as ZM-CSCG random variable with common variance N_0 . Superscript "scenario" in $y_{RD}^{J,scenario}$ refers to the cooperation scenario of user T and, as a result, is selected from among the members of the set {ADF, AF, H}. \tilde{P}_R refers to the transmitted power of user R. As a result, \tilde{P}_R equals P_R in cooperation mode and is equal zero in non-cooperation mode. With knowledge of the channel coefficients h_{SD} , h_{ST} ,

Objective: To transmit data from a source node (cellular user) to a destination node (base station) with the cooperation of a D2D pair

Symbols used in the algorithm:

Cellular user: S
 Base station: D
 D2D transmitter: T
 D2D receiver: R

Begin

Step 1. User S transmits its data symbols to User D. Due to the broadcast phenomenon, users T and R also receive these data symbols.

Step 2. User T cooperates with User R using either the ADF, AF, or Hybrid relaying protocol based on the signal received in Step 1.

Step 3. User R cooperates with User S using the ADF relaying protocol based on the signals received in Steps 1 and 2.

Step 4. User D performs joint combining of the signal received from the user S in step 1 and that from the relay R in step 3, and extracts the transmitted symbols using maximum-ratio combining (MRC).

End of algorithm.

Fig. 2: Proposed relaying protocol algorithm.

h_{SR} and h_{RD} , the destination (base station) extracts the transmitted symbols by jointly combining the received signal y_{SD} in (4) and $y_{RD}^{J,scenario}$ in (14) using MRC detector. The instantaneous SINR of the MRC output is:

$$\gamma_D^{J,scenario} = \tilde{\gamma}_{RD}^{J,scenario} + \gamma_{SD} \quad (15)$$

where

$$\gamma_{SD} = \frac{P_S |h_{SD}|^2}{\sigma_{n,SD}^2} \quad (16)$$

$$\tilde{\gamma}_{RD}^{J,ADF} = \tilde{\gamma}_{RD}^{J,AF} = \tilde{\gamma}_{RD}^{J,H} = \tilde{\gamma}_{RD} \quad (17)$$

in which $\tilde{\gamma}_{RD} = \frac{\bar{P}_R |h_{RD}|^2}{\sigma_{n,RD}^2}$. Also, sum of all noise terms in (14) are distributed as ZM-CSCG with effective noise power $\sigma_{n,RD}^2$ as follows:

$$\sigma_{n,RD}^2 = N_0 + \sigma_{I_D}^2 \quad (18)$$

We also define $\gamma_{RD} = \frac{P_R |h_{RD}|^2}{\sigma_{n,RD}^2}$ which represents the SINR value of the signal received from the relay R if it decoded the transmitted symbol correctly.

3. SYMBOL ERROR ANALYSIS

In this section, we analyse the symbol error probability (SEP) performance. We also propose tight average SEP (ASEP) approximations at high SINRs to facilitate asymptotic performance analysis.

3.1. Closed Form Conditional SEP Analysis

We know from [18], if square MQAM signals with the utilization of instantaneous signal-to-noise ratios (SNRs) in the system, it becomes possible to calculate the conditional symbol error probability (SEP) in each transmission. This calculation is performed as follows:

$$SEP_{QAM} = \psi_{QAM}(\gamma) \quad (19)$$

where

$$\psi_{QAM}(\gamma) = 4K Q(\sqrt{b_{QAM}\gamma}) - 4K^2 Q^2(\sqrt{b_{QAM}\gamma}) \quad (20)$$

in which $K = 1 - \frac{1}{\sqrt{M}}$, $b_{QAM} = \frac{3}{M-1}$ and $Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty \exp(-t^2/2) dt$ is the Gaussian Q-function [19].

Substituting $\gamma_D^{J,scenario}$ (15) into (20), the conditional SEP can be obtained as $\psi_{QAM}(\gamma_D^{J,scenario})$. As stated earlier, relay R cooperates with ADF relaying protocol. Therefore, it operates with power P_R only if it decodes the transmitted symbol correctly. With knowledge of the channel coefficients h_{SR} , h_{ST} and h_{TR} , the D2D receiver R extracts the transmitted symbols by jointly combining the received signal y_{SR} in (5) and y_{TR} in (10), (12), or (13) depending on the cooperation approach of the D2D transmitter T using an MRC detector. The instantaneous SINR of the MRC output is:

$$\gamma_R^{J,scenario} = \gamma_{SR} + \gamma_{TR} \quad (21)$$

in which $\gamma_{SR} = \frac{P_S |h_{SR}|^2}{\sigma_{n,SR}^2}$. We also know that if user T cooperates with the AF approach, then $\gamma_{TR} = \gamma_{TR,AF} = \frac{P_S |h_{ST}|^2 P_T |h_{TR}|^2}{\sigma_{n,TR}^2 P_S |h_{ST}|^2 + \sigma_{n,ST}^2 P_T |h_{TR}|^2 + \sigma_{n,TR}^2 \sigma_{n,ST}^2}$. On the other hand in the

ADF or hybrid relaying protocol, it should be noted that user T tries to decode the transmitted symbol from the signal y_{ST} , and this is done successfully with a probability of $\psi_{QAM}(\gamma_{ST} = \frac{P_S |h_{ST}|^2}{\sigma_{n,ST}^2})$, which leads to $\gamma_{TR} = \frac{P_T |h_{TR}|^2}{\sigma_{n,TR}^2}$. Otherwise, $\gamma_{TR} = 0$ in the ADF approach, and $\gamma_{TR} = \gamma_{TR,AF}$ in the hybrid relaying protocol.

By Substituting $\gamma_R^{J,scenario}$ into (20), we can achieve probability of incorrect detection in the relay R as $\psi_{QAM}(\gamma_R^{J,scenario})$. Therefore, $\psi_{QAM}(\gamma_D^{J,scenario})$ equals to:

$$\begin{aligned} \psi_{QAM}(\gamma_D^{J,scenario}) &= \psi_{QAM}(\gamma_R^{J,scenario}) \psi_{QAM}(\gamma_{SD}) \\ &+ (1 - \psi_{QAM}(\gamma_R^{J,scenario})) \times \psi_{QAM}(\gamma_{SD} + \gamma_{RD}) \quad (22) \\ &\cong \psi_{QAM}(\gamma_R^{J,scenario}) \psi_{QAM}(\gamma_{SD}) + \psi_{QAM}(\gamma_{SD} + \gamma_{RD}) \end{aligned}$$

in which $\psi_{QAM}(\gamma_R^{J,scenario})$ could be ignored compared to 1.

3.2. ASEP Analysis and Tight Approximations

Averaging the conditional SEP in (20) over the instantaneous SINR γ , we obtain the average SEP (ASEP) as follows:

$$\Psi_{QAM}(\gamma) = \mathbb{E}\{\psi_{QAM}(\gamma)\} = F\left(1 + \frac{b_{QAM}}{2 \sin^2 \theta} \gamma\right) \quad (23)$$

where $\mathbb{E}\{\cdot\}$ is the expectation operator and

$$\begin{aligned} F(x(\theta)) &= \frac{4K}{\pi\sqrt{M}} \int_0^{\pi/2} \frac{1}{x(\theta)} d\theta \\ &+ \frac{4K^2}{\pi} \int_{\pi/4}^{\pi/2} \frac{1}{x(\theta)} d\theta \quad (24) \end{aligned}$$

in which $x(\theta)$ denotes a function with the variable θ . In order to derive the Average Symbol Error Probability (ASEP) formulation in equation (24), we utilized two special properties of the Gaussian Q-function. These properties are as follows for any $u \geq 0$ [18, 20]:

1. $Q(u) = \frac{1}{\pi} \int_0^{\pi/2} \exp\left(-\frac{u^2}{2 \sin^2 \theta}\right) d\theta$
2. $Q^2(u) = \frac{1}{\pi} \int_0^{\pi/4} \exp\left(-\frac{u^2}{2 \sin^2 \theta}\right) d\theta$

Even though the resulting ASEP formulation can be efficiently calculated numerically, it is very complex and it is hard to get insight into the system performance from these. To deal with this complexity, we present an approximation of (23) by replacing $\sin^2 \theta$ with 1 in the integrands in (24) as follows which is tight at high SNR values.

$$\Psi_{QAM}(\gamma) \approx \tilde{\Psi}_{QAM}(\gamma) = \frac{M-1}{2M(1+b\bar{\gamma})} \quad (25)$$

in which $b = \frac{b_{QAM}}{2}$ and $\bar{\gamma}$ is the expectation value of γ over the Rayleigh fading coefficient(s). Therefore, averaging the conditional SEPs $\psi_{QAM}(\gamma_{SD})$ and $\psi_{QAM}(\gamma_{SD} + \gamma_{RD})$ over the instantaneous SINRs γ_{SD} and γ_{RD}^J , we obtain:

$$\Psi_{QAM}(\gamma_{SD}) \approx \tilde{\Psi}_{QAM}(\gamma_{SD}) = \frac{M-1}{2M(1+b\bar{\gamma}_{SD})} \approx \frac{M-1}{2Mb\bar{\gamma}_{SD}} \quad (26)$$

$$\begin{aligned} \Psi_{QAM}(\gamma_{SD} + \gamma_{RD}) &\approx \tilde{\Psi}_{QAM}(\gamma_{SD} + \gamma_{RD}) \\ &= \frac{M-1}{2M(1+b\bar{\gamma}_{SD})(1+b\bar{\gamma}_{RD})} \\ &\approx \frac{M-1}{2Mb^2\bar{\gamma}_{SD}\bar{\gamma}_{RD}} \end{aligned} \quad (27)$$

where, $\bar{\gamma}_{SD}$ and $\bar{\gamma}_{RD}$ represent the expectation values of γ_{SD} and γ_{RD} over the Rayleigh fading coefficients. These values are equal to $\frac{P_S\sigma_{SD}^2}{\sigma_{n,SD}^2}$ and $\frac{P_R\sigma_{RD}^2}{\sigma_{n,RD}^2}$, respectively.

If the channel links h_{SR} , h_{ST} and h_{TR} are all available in the cooperative network involving users S, T, and R (i.e., $\sigma_{SR}^2 \neq 0$, $\sigma_{ST}^2 \neq 0$ and $\sigma_{TR}^2 \neq 0$), then it is possible to obtain highly accurate approximations of ASEP through the ADF and AF approaches as follows using a similar technique as used in [22] and provided that the SINR is sufficiently high.

$$\begin{aligned} \Psi_{QAM}(\gamma_R^{J,ADF}) &\approx \tilde{\Psi}_{QAM}(\gamma_R^{J,ADF}) \\ &= \frac{1}{b^2\bar{\gamma}_{SR}} \left(\frac{A^2}{\bar{\gamma}_{ST}} + \frac{B}{\bar{\gamma}_{TR}} \right) \end{aligned} \quad (28)$$

$$\begin{aligned} \Psi_{QAM}(\gamma_R^{J,AF}) &\approx \tilde{\Psi}_{QAM}(\gamma_R^{J,AF}) \\ &= \frac{B}{b^2\bar{\gamma}_{SR}} \left(\frac{1}{\bar{\gamma}_{ST}} + \frac{1}{\bar{\gamma}_{TR}} \right) \end{aligned} \quad (29)$$

The constants A and B, which are dependent on the constellation size M, can be calculated as follows:

$$A = \frac{4K}{\pi\sqrt{M}} \int_0^{\frac{\pi}{2}} \sin^2\theta \, d\theta + \frac{4K^2}{\pi} \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \sin^2\theta \, d\theta = \frac{M-1}{2M} + \frac{K^2}{\pi}$$

$$B = \frac{4K}{\pi\sqrt{M}} \int_0^{\frac{\pi}{2}} \sin^4\theta \, d\theta + \frac{4K^2}{\pi} \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \sin^4\theta \, d\theta = \frac{3(M-1)}{8M} + \frac{K^2}{\pi}$$

In addition, $\bar{\gamma}_{SR}$, $\bar{\gamma}_{ST}$ and $\bar{\gamma}_{TR}$ represent the expectation values of γ_{SR} , γ_{ST} and γ_{TR} over the Rayleigh fading coefficients. These values can be calculated as $\frac{P_S\sigma_{SR}^2}{\sigma_{n,SR}^2}$, $\frac{P_S\sigma_{ST}^2}{\sigma_{n,ST}^2}$ and $\frac{P_T\sigma_{TR}^2}{\sigma_{n,TR}^2}$, respectively.

The following theorems provide accurate approximations of ASEP for the proposed joint cooperation approach in various scenarios. These approximations are obtained by averaging the respective SEP formulations for each scenario using (22), and then applying (23) to (29).

Theorem 1: The ASEP of the cooperation system using "Decode and Joint Cooperation" scenario and MQAM modulation can be approximated as:

$$\begin{aligned} \Psi_{QAM}(\gamma_D^{J,ADF}) &= \mathbb{E}\{\psi_{QAM}(\gamma_D^{J,ADF})\} \\ &\cong \tilde{\Psi}_{QAM}(\gamma_D^{J,ADF}) \\ &= \frac{M-1}{2Mb^2\bar{\gamma}_{SD}} \left[\frac{1}{b\bar{\gamma}_{SR}} \left(\frac{A^2}{\bar{\gamma}_{ST}} + \frac{B}{\bar{\gamma}_{TR}} \right) + \frac{1}{\bar{\gamma}_{RD}} \right] \end{aligned} \quad (30)$$

Proof By averaging relation (22) with $\psi_{QAM}(\gamma_R^{J,scenario})$ for the "Decode and Joint Cooperation" scenario, and then applying (26), (27), and (28), we obtain (30).

Theorem 2: If all of the channel links h_{SD} , h_{SR} , h_{RD} , h_{ST} and h_{TR} are available (i.e., $\sigma_{SD}^2 \neq 0$, $\sigma_{SR}^2 \neq 0$, $\sigma_{RD}^2 \neq 0$, $\sigma_{ST}^2 \neq 0$ and $\sigma_{TR}^2 \neq 0$), then the ASEP of the cooperation system using "Amplify and Joint Cooperation" scenario and MQAM modulation can be tightly approximated as:

$$\begin{aligned} \Psi_{QAM}(\gamma_D^{J,AF}) &= \mathbb{E}\{\psi_{QAM}(\gamma_D^{J,AF})\} \cong \tilde{\Psi}_{QAM}(\gamma_D^{J,AF}) \\ &= \frac{M-1}{2Mb^2\bar{\gamma}_{SD}} \left[\frac{B}{b\bar{\gamma}_{SR}} \left(\frac{1}{\bar{\gamma}_{ST}} + \frac{1}{\bar{\gamma}_{TR}} \right) + \frac{1}{\bar{\gamma}_{RD}} \right] \end{aligned} \quad (31)$$

Proof By averaging relation (22) with $\psi_{QAM}(\gamma_R^{J,scenario})$ for the "Amplify and Joint Cooperation" scenario, and then applying (26), (27), and (29), we obtain (31).

Theorem 3: If all of the channel links h_{SD} , h_{SR} , h_{RD} , h_{ST} and h_{TR} are available (i.e., $\sigma_{SD}^2 \neq 0$, $\sigma_{SR}^2 \neq 0$, $\sigma_{RD}^2 \neq 0$, $\sigma_{ST}^2 \neq 0$ and $\sigma_{TR}^2 \neq 0$), then the ASEP of the cooperation system using "Hybrid and Adaptive Joint Cooperation" scenario and MQAM modulation can be tightly approximated as:

$$\begin{aligned} \Psi_{QAM}(\gamma_D^{J,H}) &= \mathbb{E}_\gamma\{\psi_{QAM}(\gamma_D^{J,H})\} \cong \tilde{\Psi}_{QAM}(\gamma_D^{J,H}) \\ &= \frac{M-1}{2Mb^2\bar{\gamma}_{SD}} \left\{ \frac{M-1}{2Mb\bar{\gamma}_{SR}} \left[\frac{B}{b\bar{\gamma}_{ST}} \left(\frac{1}{\bar{\gamma}_{ST}} + \frac{1}{\bar{\gamma}_{TR}} \right) + \frac{1}{\bar{\gamma}_{RD}} \right] + \frac{1}{\bar{\gamma}_{RD}} \right\} \end{aligned} \quad (32)$$

Proof By averaging (22) with $\psi_{QAM}(\gamma_R^{J,scenario})$ for the "Hybrid and Adaptive Joint Cooperation" scenario, and then applying (26), (27), (28) and (29), we obtain (32).

4. POWER OPTIMIZATION AND PERFORMANCE COMPARISON

Optimizing the transmitted power to achieve optimal performance levels with the minimum transmitted power is crucial, particularly for fifth-generation (5G) and beyond cellular networks [21]. The findings from the previous section highlight that the ASEP in cooperative approaches is dependent not only on the transmitted power but also on the power allocation method employed. To determine the optimal power allocation for each cooperative approach, we can solve the following optimization problems:

$$\begin{aligned} &(\delta_{scenario}^*, \delta'_{scenario}^*) \\ &= \arg \min_{\delta_{scenario}, \delta'_{scenario}} \Psi_{QAM}(\gamma_D^{scenario}) \end{aligned} \quad (33)$$

such that

$$0 < \delta_{scenario}, \delta'_{scenario} < 1$$

and

$$0 < \delta_{scenario} + \delta'_{scenario} < 1$$

The exact analytical expressions for ASEP in integral form are not practical for evaluating the proposed power allocation rules. Therefore, we use $\tilde{\Psi}_{QAM}(\gamma_D^{J,ADF})$, $\tilde{\Psi}_{QAM}(\gamma_D^{J,AF})$ and $\tilde{\Psi}_{QAM}(\gamma_D^{J,H})$ instead of $\Psi_{QAM}(\gamma_D^{J,ADF})$, $\Psi_{QAM}(\gamma_D^{J,AF})$ and $\Psi_{QAM}(\gamma_D^{J,H})$ in (33). We then employ (1), (2), and (3) and take the derivative with respect to the power allocation coefficients, setting the resulting derivatives to 0 to obtain the optimal power allocation solutions. So we have:

$$\begin{cases} \frac{\partial \tilde{\Psi}_{QAM}(\gamma_D^{J,ADF})}{\partial \delta_{J,ADF}} = 0 \\ \frac{\partial \tilde{\Psi}_{QAM}(\gamma_D^{J,ADF})}{\partial \delta'_{J,ADF}} = 0 \end{cases} \quad (34)$$

$$\begin{cases} \frac{\partial \tilde{\Psi}_{QAM}(\gamma_D^{J,AF})}{\partial \delta_{J,AF}} = 0 \\ \frac{\partial \tilde{\Psi}_{QAM}(\gamma_D^{J,AF})}{\partial \delta'_{J,AF}} = 0 \end{cases} \quad (35)$$

$$\begin{cases} \frac{\partial \tilde{\Psi}_{QAM}(\gamma_D^{J,H})}{\partial \delta_{J,H}} = 0 \\ \frac{\partial \tilde{\Psi}_{QAM}(\gamma_D^{J,H})}{\partial \delta'_{J,H}} = 0 \end{cases} \quad (36)$$

Since obtaining closed-form solutions for the aforementioned sets of equations is challenging, we rely on numerical methods to determine the optimal power allocation coefficients with high precision. Nonetheless, there exist common special cases that simplify the equation complexities, which are discussed in the following.

Case 1: If we assume a fixed power allocation for the D2D transmitter and neglect the effects of co-channel interference (CCI) and residual self-interference (RSI) components in comparison to the Gaussian noise term, the following results hold:

$$\delta_{J,ADF} = \frac{D_1^{J,ADF}}{\sqrt[3]{D_2^{J,ADF}}} + \sqrt[3]{D_2^{J,ADF}} + D_3^{J,ADF} \quad (37)$$

$$\delta_{J,AF} = \frac{D_1^{J,AF}}{\sqrt[3]{D_2^{J,AF}}} + \sqrt[3]{D_2^{J,AF}} + D_3^{J,AF} \quad (38)$$

$$D_4^{J,H} \delta_{J,H}^4 + D_3^{J,H} \delta_{J,H}^3 + D_2^{J,H} \delta_{J,H}^2 + D_1^{J,H} \delta_{J,H} + D_0^{J,H} = 0 \quad (39)$$

where $D_1^{J,ADF}$, $D_2^{J,ADF}$, $D_3^{J,ADF}$, $D_1^{J,AF}$, $D_2^{J,AF}$, $D_3^{J,AF}$, $D_0^{J,H}$, $D_1^{J,H}$, $D_2^{J,H}$, $D_3^{J,H}$ and $D_4^{J,H}$ are constants with the following definitions. We also define $\bar{\gamma} = \frac{P}{N_0}$.

$$D_1^{J,ADF} = \frac{\sigma_{RD}^2 \left(\frac{-3A^2}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} + (\delta'_{J,ADF} - 1) \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,ADF}} \right)}{6} + \left(\frac{\delta'_{J,ADF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,ADF}}}{6} \right)^2$$

$$D_2^{J,ADF} = \left(\left(\left(\frac{\delta'_{J,ADF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,ADF}}}{6} \right)^3 + \left(\frac{\sigma_{RD}^2 \left(\frac{-3A^2}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} + (\delta'_{J,ADF} - 1) \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,ADF}} \right)}{24} \right) \right)$$

$$\left(\frac{\delta'_{J,ADF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,ADF}}}{6} + \frac{\frac{-3A^2}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} \sigma_{RD}^2 (\delta'_{J,ADF} - 1)}{4} \right)^2 - \left(D_1^{J,ADF} \right)^3 - \left(\frac{\delta'_{J,ADF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,ADF}}}{6} \right)^3 - \left(\frac{\sigma_{RD}^2 \left(\frac{-3A^2}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} + (\delta'_{J,ADF} - 1) \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,ADF}} \right)}{24} \right) \times \left(\frac{\delta'_{J,ADF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,ADF}}}{6} + \frac{\frac{-3A^2}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} \sigma_{RD}^2 (\delta'_{J,ADF} - 1)}{4} \right)$$

$$D_3^{J,ADF} = - \frac{\delta'_{J,ADF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,ADF}}}{6}$$

$$D_1^{J,AF} = \frac{\sigma_{RD}^2 \left(\frac{-3B}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} + (\delta'_{J,AF} - 1) \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}} \right)}{6} + \left(\frac{\delta'_{J,AF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}}}{6} \right)^2$$

$$D_2^{J,AF} = \left(\left(\left(\frac{\delta'_{J,AF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}}}{6} \right)^3 + \left(\frac{\sigma_{RD}^2 \left(\frac{-3B}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} + (\delta'_{J,AF} - 1) \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}} \right)}{24} \right) \right)$$

$$\left(\frac{\delta'_{J,AF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}}}{6} + \frac{\frac{-3B}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} \sigma_{RD}^2 (\delta'_{J,AF} - 1)}{4} \right)^2 - \left(D_1^{J,AF} \right)^3 - \left(\frac{\delta'_{J,AF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}}}{6} \right)^3 - \left(\frac{\sigma_{RD}^2 \left(\frac{-3B}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} + (\delta'_{J,AF} - 1) \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}} \right)}{24} \right) \times$$

$$\left(\frac{\delta'_{J,AF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}}}{6} + \frac{\frac{-3B}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} \sigma_{RD}^2 (\delta'_{J,AF} - 1)}{4} \right)^2 -$$

$$\left(\frac{\delta'_{J,AF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}}}{6} + \frac{\frac{-3B}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} \sigma_{RD}^2 (\delta'_{J,AF} - 1)}{4} \right)^2 -$$

$$\left(D_1^{J,AF} \right)^3 - \left(\frac{\delta'_{J,AF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}}}{6} \right)^3 - \left(\frac{\sigma_{RD}^2 \left(\frac{-3B}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} + (\delta'_{J,AF} - 1) \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}} \right)}{24} \right) \times$$

$$\left(\frac{\delta'_{J,AF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}}}{6} + \frac{\frac{-3B}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} \sigma_{RD}^2 (\delta'_{J,AF} - 1)}{4} \right)^2 -$$

$$\left(\frac{\delta'_{J,AF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}}}{6} + \frac{\frac{-3B}{b\bar{\gamma}\sigma_{ST}^2\sigma_{SR}^2} \sigma_{RD}^2 (\delta'_{J,AF} - 1)}{4} \right)^2 -$$

$$D_3^{J,AF} = -\frac{\delta'_{J,AF} - 1 - \sigma_{RD}^2 \frac{-2B}{b\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2\delta'_{J,AF}}}{6}$$

$$D_0^{J,H} = \frac{2B(1-M)\sigma_{RD}^2(1-\delta'_{J,H})^2}{Mb^2(\bar{\gamma})^2\sigma_{SR}^2\sigma_{ST}^4}$$

$$D_1^{J,H} = \frac{4B(M-1)(1-\delta'_{J,H})\sigma_{RD}^2}{Mb^2\bar{\gamma}\sigma_{SR}^2\sigma_{ST}^4} - \frac{3B(M-1)(1-\delta'_{J,H})^2\sigma_{RD}^2}{2Mb^2(\bar{\gamma})^2\delta'_{J,H}\sigma_{SR}^2\sigma_{ST}^2\sigma_{TR}^2}$$

$$D_2^{J,H} = \frac{3B(M-1)(1-\delta'_{J,H})\sigma_{RD}^2}{Mb^2(\bar{\gamma})^2\delta'_{J,H}\sigma_{SR}^2\sigma_{ST}^2\sigma_{TR}^2} - \frac{2B(M-1)\sigma_{RD}^2}{Mb^2\bar{\gamma}\sigma_{SR}^2\sigma_{ST}^4} + \frac{(1-M)(1-\delta'_{J,H})^2\sigma_{RD}^2}{Mb\delta'_{J,H}\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2}$$

$$D_3^{J,H} = \frac{2(M-1)(1-\delta'_{J,H})\sigma_{RD}^2}{Mb\delta'_{J,H}\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2} - \frac{3B(M-1)\sigma_{RD}^2}{2Mb^2(\bar{\gamma})^2\delta'_{J,H}\sigma_{SR}^2\sigma_{ST}^2\sigma_{TR}^2} + \delta'_{J,H} - 1$$

$$D_4^{J,H} = \frac{(1-M)\sigma_{RD}^2}{Mb\delta'_{J,H}\bar{\gamma}\sigma_{SR}^2\sigma_{TR}^2} + 2$$

Case 2: If we assume a fixed power allocation for the cellular user and neglect the effects of co-channel interference (CCI) and residual self-interference (RSI) components in comparison to the Gaussian noise term, the following results hold:

$$\delta'_{J,ADF} = \frac{\sqrt{\frac{B\sigma_{RD}^2}{b\delta_{J,ADF}\bar{\gamma}\sigma_{TR}^2\sigma_{SR}^2}}}{1 + \sqrt{\frac{B\sigma_{RD}^2}{b\delta_{J,ADF}\bar{\gamma}\sigma_{TR}^2\sigma_{SR}^2}}}(1 - \delta_{J,ADF}) \quad (40)$$

$$\delta'_{J,AF} = \frac{\sqrt{\frac{B\sigma_{RD}^2}{b\delta_{J,AF}\bar{\gamma}\sigma_{TR}^2\sigma_{SR}^2}}}{1 + \sqrt{\frac{B\sigma_{RD}^2}{b\delta_{J,AF}\bar{\gamma}\sigma_{TR}^2\sigma_{SR}^2}}}(1 - \delta_{J,AF}) \quad (41)$$

$$\delta'_{J,H} = \frac{\sqrt{\frac{(M-1)\sigma_{RD}^2}{2Mb\bar{\gamma}\delta_{J,H}\sigma_{SR}^2\sigma_{TR}^2} \left[\frac{B}{b\bar{\gamma}\delta_{J,H}\sigma_{ST}^2} + 1 \right]}}{1 + \sqrt{\frac{(M-1)\sigma_{RD}^2}{2Mb\bar{\gamma}\delta_{J,H}\sigma_{SR}^2\sigma_{TR}^2} \left[\frac{B}{b\bar{\gamma}\delta_{J,H}\sigma_{ST}^2} + 1 \right]}}(1 - \delta_{J,H}) \quad (42)$$

We interpret the results of optimal power allocation as follows:

1. In AF cooperation systems, the power allocation does not depend on the modulation scheme used. This is different from ADF cooperation systems, where the optimal power allocation may vary based on the modulation scheme employed. This is

because in the AF approach, the relay amplifies and forwards the received signal to the destination, regardless of the type of signal received. On the other hand, in ADF cooperation systems, the relay forwards information to the destination only if it correctly decodes the received signal. The process of decoding at the relay necessitates the utilization of particular modulation information, leading to the development of a power allocation scheme that is contingent upon the modulation technique employed.

2. The optimal power allocation is independent of the direct link between the cellular user and the base station, and depends solely on the channel links associated with the relay.
3. The optimal power allocated to the cooperator user in each cooperation scenario is lower than the power allocated to the cellular user. If the link quality between the cellular user and cooperator user is significantly worse than that between the cooperator user and the base station, then the power allocated to the cellular user approaches the maximum power level P. When the link quality between the cellular user and cooperator user is poor, it becomes difficult for the cooperator user to correctly decode the transmitted symbol. Thus, its forwarding role is less important, and it makes sense to allocate more power to the source. Conversely, when the link quality between the cellular user and cooperator user is very good, the cooperator user can always decode the transmitted symbol accurately, and we can consider it as a copy of the cellular user and allocate almost equal power to them.
4. In the proposed hybrid and adaptive approach, the relay benefits from the cooperation of user T, resulting in a higher probability of correctly detecting the symbol sent by the cellular user. As a result, based on the explanation in the previous paragraph, solving the power allocation problem in this approach tends to allocate half of the total available power (total power minus the power allocated to user T), more than in the AF and ADF approaches. In the proposed hybrid and adaptive approach, the relay benefits from the cooperation of user T, resulting in a higher probability of correctly detecting the symbol sent by the cellular user. As a result, based on the explanation in the previous paragraph, solving the power allocation problem in this approach tends to allocate half of the total available power (total power minus the power allocated to user T), more than in the AF and ADF approaches.

5. It is worth noting that by increasing the power available to the cellular user, both the desired signal power and power of the RSI component increase simultaneously. Conversely, reducing the transmitted power causes both the desired signal power and power of the RSI component to decrease. Therefore, the best transmission quality from the cellular user to the relay is achieved with intermediate values of transmitted power (neither high nor low). In this condition, and based on the explanations of the previous paragraphs regarding optimal power allocation, the smallest proportion of available power will be assigned to the cellular user.

The superiority of the proposed relaying protocol's performance compared to the ADF and AF approaches is evident from the comparison of their average symbol error probability tight approximations in (28) to (32). We conduct a comparative study on the performance of the proposed relaying protocol, utilizing asymptotically tight ASEP approximations and optimal power allocation solutions. Initially, we compare the average symbol error probability of the "Decode and Joint Cooperation" scenario with that of the "Amplify and Joint Cooperation" scenario, using the parameter $\lambda_{J,AF}^{J,ADF}$, which is defined as follows:

$$\lambda_{J,AF}^{J,ADF} = \frac{\Psi_{QAM}(\gamma_D^{J,ADF})}{\Psi_{QAM}(\gamma_D^{J,AF})} \approx \frac{\tilde{\Psi}_{QAM}(\gamma_D^{J,ADF})}{\tilde{\Psi}_{QAM}(\gamma_D^{J,AF})} = \frac{\frac{1}{b} \left(\frac{A^2}{\bar{\gamma}_{ST}} + \frac{B}{\bar{\gamma}_{TR}} \right) \gamma_{RD,SR}^r + 1}{\frac{B}{b} \left(\frac{1}{\bar{\gamma}_{ST}} + \frac{1}{\bar{\gamma}_{TR}} \right) \gamma_{RD,SR}^r + 1} \quad (43)$$

in which $\gamma_{RD,SR}^r \triangleq \frac{\bar{\gamma}_{RD}}{\bar{\gamma}_{SR}}$.

Continuing on, we will compare the average symbol error probability of the "Hybrid and Adaptive Joint Cooperation" scenario with that of the "Decode and Joint Cooperation" scenario, using the parameter $\lambda_{J,ADF}^{J,H}$, which is defined as follows:

$$\lambda_{J,ADF}^{J,H} = \frac{\Psi_{QAM}(\gamma_D^{J,H})}{\Psi_{QAM}(\gamma_D^{J,ADF})} \approx \frac{\tilde{\Psi}_{QAM}(\gamma_D^{J,H})}{\tilde{\Psi}_{QAM}(\gamma_D^{J,ADF})} = \frac{\frac{M-1}{2M} \left[\frac{B}{b\bar{\gamma}_{ST}} \left(\frac{1}{\bar{\gamma}_{ST}} + \frac{1}{\bar{\gamma}_{TR}} \right) + \frac{1}{\bar{\gamma}_{TR}} \right] \gamma_{RD,SR}^r + 1}{\left(\frac{A^2}{\bar{\gamma}_{ST}} + \frac{B}{\bar{\gamma}_{TR}} \right) \gamma_{RD,SR}^r + 1} \quad (44)$$

We will now discuss the ratios $\lambda_{J,AF}^{J,ADF}$ and $\lambda_{J,ADF}^{J,H}$ for the following cases.

Case 1: If the channel link quality between the cellular user and the D2D receiver is much worse than that between the D2D receiver and the base station (i.e., $\gamma_{RD,SR}^r \gg 1$), then

$\lambda_{J,AF}^{J,ADF}$ tends towards the value $\frac{\frac{1}{b} \left(\frac{A^2}{\bar{\gamma}_{ST}} + \frac{B}{\bar{\gamma}_{TR}} \right)}{\frac{B}{b} \left(\frac{1}{\bar{\gamma}_{ST}} + \frac{1}{\bar{\gamma}_{TR}} \right)}$. In this case, if the

channel link quality between the cellular user and the D2D transmitter is much worse than that between the D2D transmitter and receiver (i.e., $\bar{\gamma}_{TR} \gg \bar{\gamma}_{ST}$), then $\lambda_{J,AF}^{J,ADF} \rightarrow \frac{A^2}{B}$, which depends on the modulation size M . For instance, when $M=4$, $\lambda_{J,AF}^{J,ADF} \rightarrow 1.3214$, while for sufficiently high values of M , $\lambda_{J,AF}^{J,ADF} \rightarrow 1.0175 > 1$. Consequently, in this case, the "Decode and Joint Cooperation" scenario outperforms the "Amplify and Joint Cooperation" scenario, and this superiority decreases with an increase in M . On the other hand, if the channel link quality between the cellular user and the D2D transmitter is much better than that between the D2D transmitter and receiver (i.e., $\bar{\gamma}_{TR} \ll \bar{\gamma}_{ST}$), then $\lambda_{J,AF}^{J,ADF} \rightarrow 1$, which implies that, in this case, the performance of the "Decode and Joint Cooperation" scenario is almost the same as that of the "Amplify and Joint Cooperation" scenario. As the ADF cooperation protocol necessitates a decoding process at the relay, we propose the use of the AF approach in this case to reduce the system complexity.

It should be noted that in this case, the parameter $\lambda_{J,ADF}^{J,H}$ tends towards the value $\frac{\frac{M-1}{2M} \left[\frac{B}{b\bar{\gamma}_{ST}} \left(\frac{1}{\bar{\gamma}_{ST}} + \frac{1}{\bar{\gamma}_{TR}} \right) + \frac{1}{\bar{\gamma}_{TR}} \right]}{\left(\frac{A^2}{\bar{\gamma}_{ST}} + \frac{B}{\bar{\gamma}_{TR}} \right)}$, which

demonstrates the superiority of the "Hybrid and Adaptive Joint Cooperation" scenario over the "Decode and Joint Cooperation" scenario.

Case 2: If the channel link quality between the cellular user and the D2D receiver is much better than that between the D2D receiver and the base station (i.e., $\gamma_{RD,SR}^r \rightarrow 0$), then $\lambda_{J,AF}^{J,ADF}$ and $\lambda_{J,ADF}^{J,H}$ tend towards 1. This implies that in this case, the performance of the aforementioned scenarios is nearly the same.

5. NUMERICAL SIMULATIONS AND DISCUSSION

To illustrate the aforementioned theoretical analysis, we conducted computer simulations. We utilized the network topology depicted in Fig. 1 for our study. In all simulations, we normalized the transmitted power and the variance of the channel coefficients to the noise power. Additionally, we set $\sigma_{SD}^2 = 1$. The default values of the other parameters are listed in Table 1.

Table 1: Default values of simulation parameters.

Parameter	Default value/ N_0
P	20 dB
σ_{SR}^2	10
σ_{RD}^2	10
σ_{ST}^2	20
σ_{TR}^2	100
$\sigma_{ID}^2, \sigma_{IR}^2, \sigma_{IT}^2$	2 dB
$\sigma_{hRR}^2, \sigma_{hTT}^2$	-10 dB
Maximum error of $\delta_{scenario}^*$, $\delta'_{scenario}^*$	0.01% of transmitted power

First, we perform simulations to compare the exact analytical ASEP curves with the asymptotically tight ASEP approximations of the proposed relaying protocol scenarios. In this simulation, we set the power of the cellular user and the D2D Transmitter User to 50% and 10% of the total transmitted power, respectively. The results shown in Fig. 3 confirm the close correspondence between the exact analytical and approximate ASEPs. Moreover, it can be observed that although these approximations are derived for high SNR values, they also provide good approximations for moderate SNR scenarios. Therefore, they can be used instead of analytical relationships to determine the optimal power allocation, particularly for middle and high values of SNR. Next, we compare the exact analytical ASEP results of the proposed relaying protocol in all scenarios with the non-cooperative approach using different power allocation schemes. The results in Fig. 4 demonstrate that non-optimal power allocation can lead to performance degradation in all scenarios, resulting in ASEP being worse than the non-cooperative approach. The following simulation is dedicated to the optimum power allocation solution graphs in the proposed relaying protocol scenarios. Based on the discussions presented in Section 4, it is claimed that the optimal power allocation, based on solving (33), follows a general rule: the power allocation criterion is the SINR value of the signal received at the relay (D2D receiver) compared to the SINR value of the signal received at the base station by the relay. The higher this criterion, the greater the percentage of the total transmitted power allocated to the relay, which will not exceed half of the total transmitted power. Conversely, the lower this criterion, the greater the percentage of the total transmitted power allocated to the cellular user. This crucial rule will be confirmed in the following simulation. In Fig. 5, the power allocation coefficients are plotted in terms of the total transmitted power. It can be observed that an increase in the total transmitted power results in an increase in the power of the Relay RSI component in addition to increasing the desired signal power in the relay. At the base station, due to the absence of RSI, only the desired signal power increases. Consequently, the quality of the signal received at the relay decreases compared to the quality of the signal received at the base station. Therefore, with an increase in the total transmitted power, a higher percentage of this power is allocated to the cellular user. A similar discussion can be used to justify the power allocation graphs of the D2D transmitter. However, in this case, the power allocation criterion is the SINR value of the signal received at the D2D transmitter compared to the SINR value of the signal received at the D2D receiver by the D2D transmitter.

The simulations presented in Fig. 6 have demonstrated the distinct capability of the "Hybrid and Adaptive Joint Cooperation" scenario to reduce the total transmitted power required at each ASEP level. The results obtained show a significant reduction in the total transmitted power required to guarantee the ASEP levels compared to the AF and ADF approaches. This is precisely the reason why (as illustrated in Fig. 7), among the proposed approach scenarios, the most significant power savings are assigned to "Hybrid and Adaptive Joint Cooperation", followed by "Decode and Joint Cooperation" scenario, and "Amplify and Joint Cooperation"

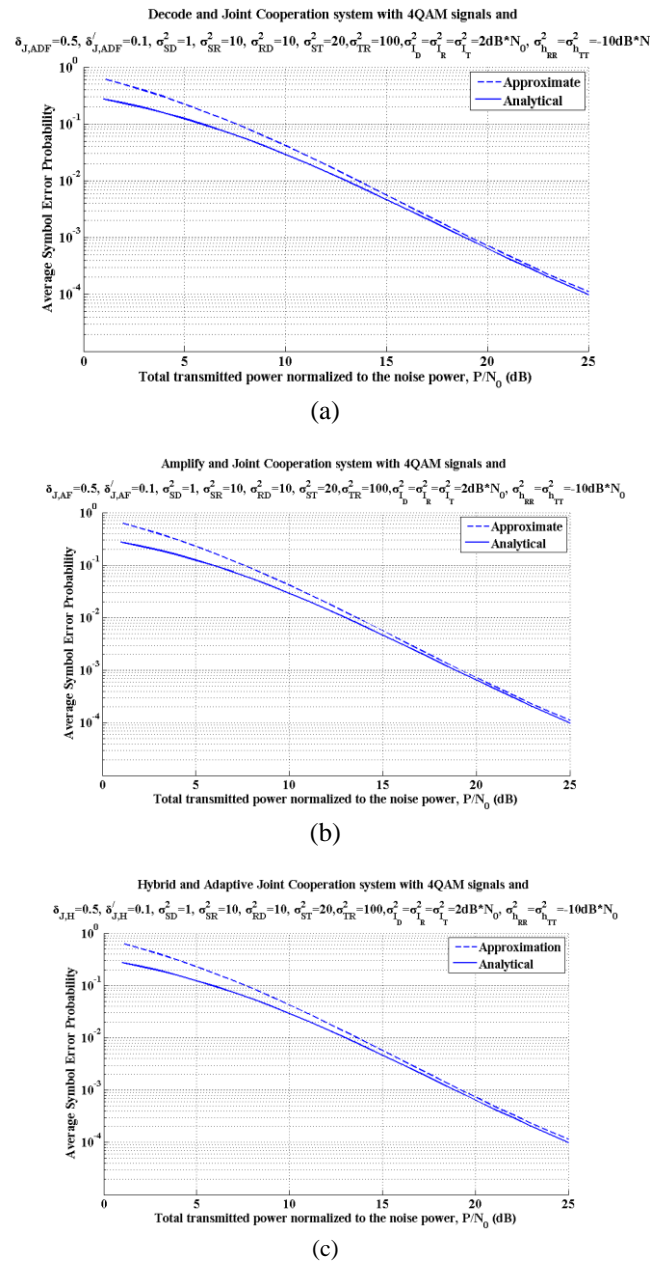


Fig. 3: Comparison of exact analytical and approximate ASEP's in the proposed relaying approach scenarios, (a) decode and Joint Cooperation, (b) amplify and Joint Cooperation, (c) hybrid and adaptive joint cooperation.

scenario, respectively. It should be noted that this improvement is more pronounced at smaller threshold levels. This is because guaranteeing low levels of error probability necessitates ensuring a high probability of correct detection in the relay, where the role of D2D transmitter cooperation is crucial.

6. CONCLUSION

This paper has presented a novel relaying protocol for transmitting from a cellular user to the base station with the joint cooperation of a Full-Duplex (FD)-enabled Device-to-Device (D2D) pair. In the suggested scheme, the recipient of the D2D pair functions as a relay, collaborating with its transmitting counterpart through D2D communication. The

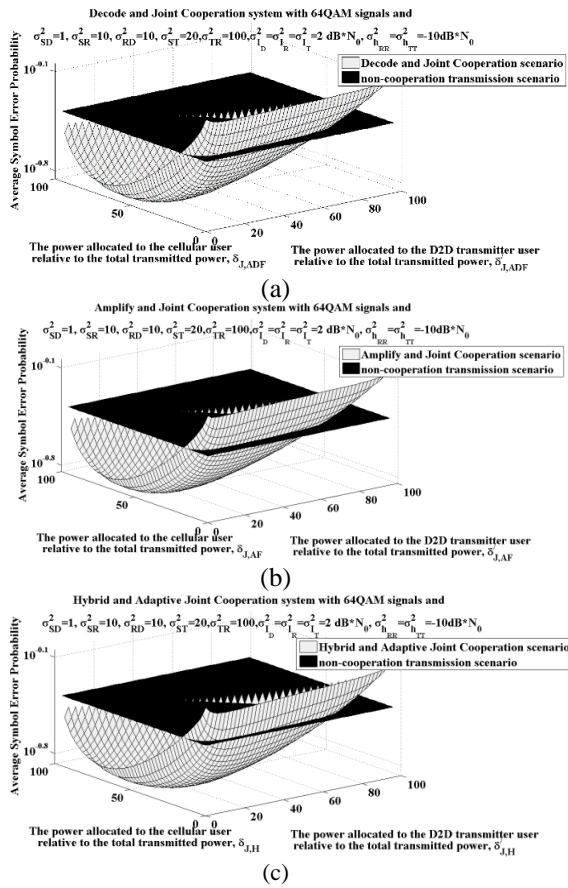


Fig. 4: Comparison of exact analytical ASEP's in the proposed relaying approach for different power allocation schemes with non-cooperation approach, (a) decode and joint cooperation, (b) amplify and joint cooperation, (c) hybrid and adaptive joint cooperation.

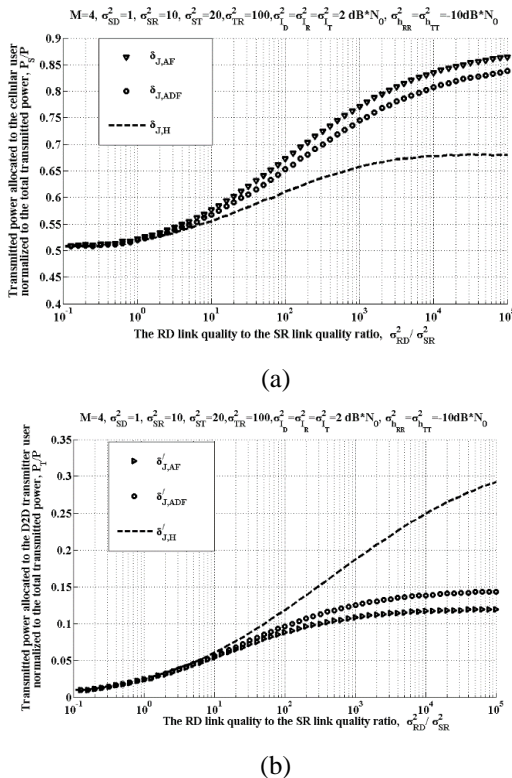


Fig. 5: Power allocation coefficients in the proposed relaying approach versus total transmitted power, (a) P_S/P , (b) P_T/P

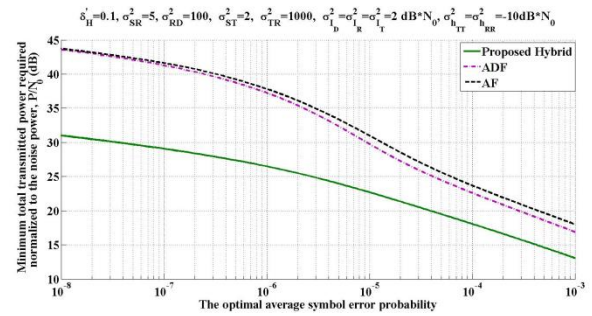


Fig. 6: Minimum total transmitted power required versus ASEP in the relaying approaches.

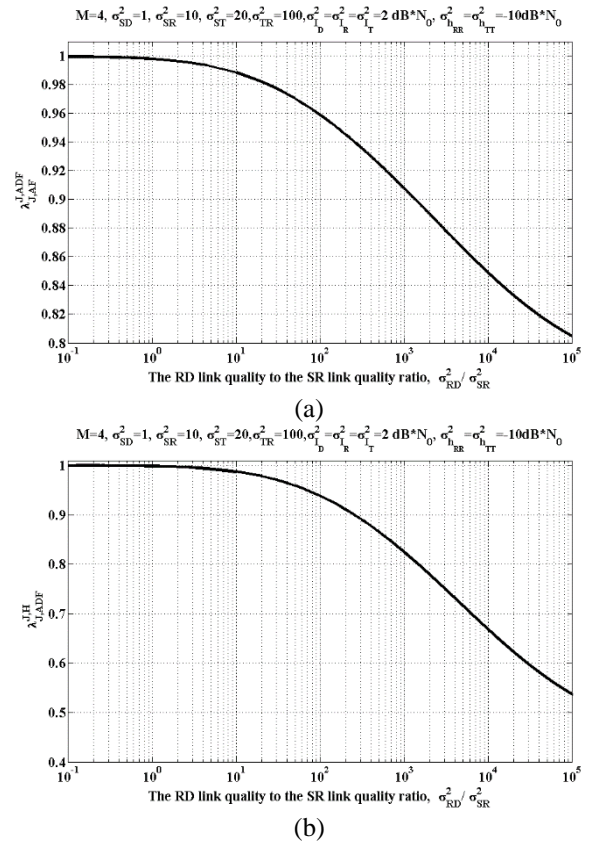


Fig. 7: ASEP comparison of various scenarios of the proposed relaying protocol, (a) $\lambda_{J,ADF}^J$, (b) $\lambda_{J,ADF}^H$.

cooperative strategy employed by the D2D recipient is referred to as Adaptive Decode-and-Forward (ADF), while the cooperative strategy adopted by the D2D transmitter can be either ADF, AF, or a Hybrid relaying protocol. These scenarios are denoted as "Decode and Joint Cooperation," "Amplify and Joint Cooperation," and "Hybrid and Adaptive Joint Cooperation," respectively. The Average Symbol Error Probability (ASEP) of the system has been studied over independent and identically distributed (i.i.d) complex Gaussian (Rayleigh envelope) channels, with perfect Channel State Information (CSI) in the presence of Residual Self-Interference (RSI) at the FD relays, as well as Co-Channel Interference (CCI). Moreover, closed-form and high Signal-to-Interference-plus-Noise Ratio (SINR) tight ASEP approximations have been established. The optimum power allocation has been formulated based on the approximate relations, and the optimal solutions and their characteristics have been discussed in detail. Analytical comparisons and simulations have confirmed the theoretical results and have demonstrated significant performance improvements.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Mehdi Basiri Abarghouei: Formal analysis, Investigation, Investigation, Software, Roles/Writing - original draft. **Reza Saadat:** Methodology, Supervision, Validation, Writing - review & editing

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy has been completely observed by the authors.

REFERENCES

- [1] S. Kurma, P. K. Sharma, S. Dhok, K. Singh, and C.-P. Li, "Adaptive AF/DF two-way relaying in FD multi-user URLLC system with user mobility," *IEEE Transactions on Wireless Communications*, vol. 21, no. 1, pp. 123-135, 2022.
- [2] R. Zhu, T. Li, J. Guo, and Y. Huang, "Relay selection scheme for AF system with partial CSI and optimal stopping theory," *Tsinghua Science and Technology*, vol. 25, no. 2, pp. 302-312, 2020.
- [3] M. C. Ju, and H. C. Yang, "Optimum design of energy harvesting relay for two-way decode-and-forward relay networks under max-min and max-sum criterions," *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 6682-6697, 2019.
- [4] G. Li, D. Mishra, Y. Hu, Y. Huang, and H. Jiang, "Adaptive relay selection strategies for cooperative NOMA networks with user and relay cooperation," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11728-11742, 2020.
- [5] Bahmani, and M. Biguesh, "Feasibility of joint power optimization of multiple source-destinations in an AF relay network," *IEEE Signal Processing Letters*, vol. 28, pp. 339-343, 2021.
- [6] J. N. Laneman, and G. W. Wornell, "Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2415-2525, 2003.
- [7] T. E. Hunter, and A. Nosratinia, "Diversity through coded cooperation," *IEEE Transactions on Wireless Communications*, vol. 5, no. 2, pp. 283-289, 2006.
- [8] M. Basiri Abarghouei, and A. M. Doost Hoseini, "Cooperative communication with imperfect channel information: Performance analysis and optimum power allocation," *Physical Communication*, vol. 4, no. 3, pp. 144-155, 2011.
- [9] Liu, C. Zhou, Q. Hu, and H. Zhao, "A novel efficient cooperative diversity protocol for wireless networks," in *Proceedings of the 2009 International Conference on Communications, Circuits and Systems*, Milpitas, CA, USA, 2009, pp. 456-460.
- [10] J. Zhang, J. Jiang, J. Bao, B. Jiang, and C. Liu, "Improved relay selection strategy for hybrid decode-amplify forward protocol," *J. Commun.*, vol. 11, pp. 297-304, 2016.
- [11] D. P. Setiawan, and H. Zhao, "Performance analysis of hybrid AF and DF protocol for relay networks," in *Proceedings of the 2017 International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC)*, 2017, pp. 123-135.
- [12] O. Maraqa, A.S. Rajasekaran, S. Al-Ahmadi, H. Yanikomeroglu, and S. M. Sait, "A survey of rate-optimal power domain NOMA with enabling technologies of future wireless networks," *IEEE Commun. Surveys Tutorials*, vol. 22, no. 4, pp. 2192-2235, 2020.
- [13] Z. Zhang, K. Long, A. V. Vasilakos, and L. Hanzo, "Full-duplex wireless communications: Challenges, solutions, and future research directions," *Proceedings of the IEEE*, vol. 104, no. 7, pp. 1369-1409, 2016.
- [14] P. Raut, and P. K. Sharma, "Full-duplex multi-user pair scheduling with time-selective fading and imperfect CSI," in *Proceedings of the National Conference on Communications (NCC)*, 2019, pp. 1-6.
- [15] Z. Zhang, Z. Ma, Z. Ding, M. Xiao, and G. K. Karagiannidis, "Full-duplex two-way and one-way relaying: Average rate, outage probability, and trade-offs," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 3920-3933, 2016.
- [16] G. Liu, F. R. Yu, H. Ji, V. C. M. Leung, and X. Li, "In-band full-duplex relaying: A survey, research issues and challenges," *IEEE Communications Surveys and Tutorials*, vol. 17, no. 2, pp. 500-524, 2nd Quarter, 2015.
- [17] W. P. Siritwongpairat, T. Himsoon, W. Su, and K. J. R. Liu, "Optimum threshold-selection relaying for decode-and-forward cooperation protocol," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, 2006, pp. 1015-1020.
- [18] M. K. Simon, and M.-S. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proceedings of the IEEE*, vol. 86, no. 9, pp. 1860-1877, 1998.
- [19] J. G. Proakis, *Digital Communications*. 4th ed., New York, NY: McGraw-Hill, 2001.
- [20] J. W. Craig, "A new, simple and exact result for calculating the probability of error for two-dimensional signal constellations," in *Proc. IEEE MILCOM*, 1991, pp. 25.5.1-25.5.5.
- [21] K. Singh, S. Biswas, M.-L. Ku, and M. F. Flanagan, "Transceiver design and power control for full-duplex ultra-reliable low-latency communication systems," *IEEE Trans. Wireless Commun.*, pp. 1-1, 2021.
- [22] W. Su, A. K. Sadek, and K. J. R. Liu, "Cooperative communication protocols in wireless networks: performance analysis and optimum power allocation,"

Wireless Personal Communication, vol. 44, no. 2, pp. 181-217, 2008.

BIBLIOGRAPHY



Mehdi Basiri Abarghouei received his B.S. degree in Electrical Engineering (first class honors) from the University of Isfahan, Iran, in 2007, and his M.S. degree in Electrical Engineering from Isfahan University of Technology, Iran, in 2010. Currently, he is pursuing a PhD at Yazd University, focusing on the field of wireless communication.



Reza Saadat was born on February 22, 1968 in Isfahan, Iran. He received his Bachelor of Science (B.Sc.) degree in electronics engineering from Isfahan University of Technology, Isfahan, Iran, in 1992. Continuing his academic journey, he pursued a Master of Science (M.Sc.) degree in communication systems engineering, which he successfully obtained in 1994. Driven by his passion for research and advancement, he went on to complete his Ph.D. degree in electrical engineering in 1999, further solidifying his expertise. He is currently Associate Professor in the Department of Electrical Engineering at Yazd University, Yazd, Iran. Throughout his distinguished career, he has focused his research efforts on various areas within the realm of electrical engineering. Notably, his research interests revolve around RSS based positioning, resource allocation in wireless systems, massive-MIMO technology, and wireless sensor networks.

Copyrights

© 2024 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).

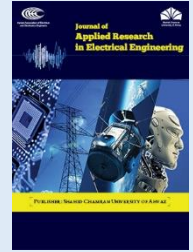


Shahid Chamran
University of AhvazIranian Association of
Electrical and Electronics
Engineers

Journal of Applied Research in Electrical Engineering

E-ISSN: 2783-2864

P-ISSN: 2717-414X

Homepage: <https://jaree.scu.ac.ir/>

Research Article

Performance Analysis of a Steep-Slope Bi-channel GaSb-GaAs Extended Source Tunnel Field Effect Transistor With Enhanced Band to Band Tunneling Current

Zahra Ahangari 

Department of Electronic, Yadegar- e- Imam Khomeini (RAH) Shahr-e-Rey Branch, Islamic Azad University, Tehran, Iran

* Corresponding Author: z.ahangari@gmail.com

Abstract: This paper presents a comprehensive investigation of the electrical properties of a heterojunction tunnel field effect transistor with enhanced electrical tunneling current. The proposed device structure incorporates an extended source region and two parallel channels positioned above and below the source region. This configuration effectively amplifies the tunneling area, leading to a significant improvement in the on-state current. Moreover, the inclusion of an embedded oxide region between the source and drain regions confers the device with a high resistance to short-channel effects. The combination of materials in both the source and channel region results in a staggered band alignment at the tunneling junction. This specific configuration leads to a lower threshold voltage for the initiation of tunneling. The impact of critical design parameters on the device performance has been thoroughly examined. A 2D variation matrix has been developed to compute the threshold voltage and on-state current variation based on the source doping density and gate workfunction, which serve as essential design parameters to optimize the electrical performance of the device. Furthermore, the device has achieved a unity current-ratio frequency of 300 GHz indicating its suitability for high-frequency applications. Additionally, the proposed structure provides an on-state current of 2.24×10^{-4} (A/ μm), an off-state current of 1.24×10^{-15} (A/ μm), an on/off current ratio of 1.81×10^{11} , and a subthreshold swing of 5 (mV/dec). These characteristics make the device viable for energy-efficient, high-speed digital circuits.

Keywords: Tunnel field effect transistor, band-to-band tunneling, workfunction, heterojunction.

Article history

Received 21 September 2023; Revised 08 January 2024; Accepted 28 January 2024; Published online 28 January 2024.

© 2024 Published by Shahid Chamran University of Ahvaz & Iranian Association of Electrical and Electronics Engineers (IAEEE)

How to cite this article

Z. Ahangari, "Performance analysis of a steep-slope bi-channel GaSb-GaAs extended source tunnel field effect transistor with enhanced band to band tunnelling current," *J. Appl. Res. Electr. Eng.*, vol. 2, no. 2, pp. 206-213, 2023.

DOI: [10.22055/jaree.2024.44844.1082](https://doi.org/10.22055/jaree.2024.44844.1082)



1. INTRODUCTION

The reduction in size of traditional Metal-Oxide-Semiconductor Field Effect Transistor (MOSFET) dimensions is the leading factor in the advancement of low-power high-speed integrated circuits. Nevertheless, when the MOSFET dimensions are scaled down to the nanoscale level, undesirable effects, commonly referred to as short channel effects, come into play, ultimately impairing device performance and elevating power consumption [1-6]. The conventional MOSFET is restricted by the thermionic electron emission-type transport and has a fundamental limit for the sub-threshold swing. However, Boltzmann limit dictates that it cannot be smaller than 60 (mV/dec) under room temperature conditions. In order to advance computational technology, a new class of devices is required. The Tunnel Field Effect Transistor (TFET), also known as the steep-slope device, is one of the most promising candidates.

It exploits interbond tunnelling, a quantum mechanical effect, as a transport mechanism. As a result, sub-thermionic transport can be achieved at room temperatures, along with low off current and a favourable on/off current ratio. The conventional TFET is a gated p-i-n diode that is reverse biased. The modulation of the tunneling barrier width at the interface of the source and channel region is achieved through the gate bias. When in the off-state, the tunneling barrier is sufficiently thick, and as a result, the carriers cannot tunnel through. However, by applying an adequate amount of bias to the gate, carriers of the opposite type with respect to the source region accumulate in the channel. This accumulation creates a steep p⁺-n⁺ tunneling junction with a thin barrier width that facilitates carrier band to band tunneling [7-10]. The initial TFET device suffers from low on-state current due to tunneling occurring only at a limited region at the interface of source and channel regions. However, recent studies have

shown that the band to band tunneling current of a TFET is dependent not only on the source doping density but also on the band gap of the material used in designing the device. Effective band alignment at the source-channel interface can amplify the tunneling rate. Homojunction TFETs employ similar materials at the tunneling junction, while heterojunction TFETs use different materials leading to different band alignments at the tunneling junction, including staggered gap and broken gap heterojunctions. A heterojunction formed by suitable semiconductor materials of different band gaps, mainly from group III-V, can be constructed and arranged to further boost the tunneling current [11-14]. Expanding the tunneling area is an additional approach, alongside material engineering, that can be employed to enhance the tunneling rate. Various structures such as L-shaped [15-17], F-shaped [18,19], and electron-hole bilayer TFET [20-22] and vertical TFET [23-25] exhibit an enlarged tunneling region. Nevertheless, the intricate nature of their fabrication process poses a challenge.

In this paper, a new heterojunction vertical tunnel field-effect transistor is presented with two parallel channels and an improved band-to-band tunneling region, leading to a significant enhancement in the on-state current of the device. The proposed device employs GaSb as the source material and high band gap GaAs in the channel and drain regions. Through the creation of a staggered heterostructure at the source/channel interface, a robust electric field is established, thereby facilitating band-to-band tunneling current with low transition voltage. In the proposed extended source TFET, band to band tunneling occurs across the thickness of the intrinsic channel and in the vertical direction, leading to a substantial increase in the tunneling area and a notable improvement in the on-state current. Furthermore, a comprehensive analysis is conducted to examine the influence of crucial design parameters on the device performance, and the corresponding electrical measures in the analog and digital domains are evaluated. Additionally, a 2D variation matrix of the threshold voltage and on-state current is generated to determine the ideal design parameters that promote optimal efficiency of the device. In prior investigations, the primary emphasis was placed on optimizing the electrical properties of a TFET device through the simultaneous attainment of the minimum possible off-state current, the lowest average subthreshold swing over a broad range of drain current, and the maximum achievable on-state tunneling current. In essence, in order to enhance the band-to-band tunneling current and establish a step-like transition from the off-state to the on-state, the transmission probability of the tunneling barrier must approach unity for a small deviation in the gate voltage around the threshold voltage. This necessitates a significant modulation of the energy bands through gate bias and an exceedingly thin tunneling barrier. Various approaches exist to enhance TFET performance, encompassing Doping Engineering [26-29], Material Engineering [30], as well as Geometry and Structure Engineering [31]. The TFET device in [23] examines the proposed vertical tunneling TFET based on silicon material. Essentially, owing to the relatively high effective mass of silicon, an alternative material combination becomes essential in order to amplify the tunneling rate. In [28], in contrast to the conventional TFETs with a p-i-n doping profile, both the source and channel regions possess a similar

doping profile, which could potentially facilitate the fabrication process of the device. Nevertheless, the optimization of the gate workfunction and channel doping density is imperative in achieving efficient performance. In [30], a p-type electron-hole bilayer TFET with GaSb is introduced as a low-power device with an enhanced tunneling window. In [31], a three-dimensional nanowire TFET is proposed to enhance the gate controllability and boost the tunneling rate. Although the fabrication process for this device is complex, it leads to significant improvement in electrical characteristics. The device presented in this paper integrates the benefits of previously documented tunnel field-effect transistors (TFETs) to create a device with enhanced electrical efficacy. The device utilizes a heterojunction approach founded on material engineering principles, and a widened source region is utilized to amplify the rate of tunneling. The incorporation of a dual parallel channel introduces an additional tunneling junction that significantly enhances the on-state current.

The present paper is structured in the following manner: Firstly, the proposed TFET schematic along with critical simulation models are introduced in the subsequent section. Subsequently, the device performance outcomes and discussions regarding the crucial physical and structural design parameters are presented. Ultimately, the conclusion section provides an overview of the paper.

2. DEVICE STRUCTURE AND SIMULATION SET UP

The illustration of Bi-channel Extended Source Tunnel Field Effect Transistor (BESTFET) is depicted in Fig. 1. Fig. 1a displays the preliminary two-dimensional diagram of the BESTFET, while Fig. 1b portrays its on-state functionality by delineating the tunneling region in the parallel channel. The utilization of GaSb as a source material and high band gap GaAs in the channel and drain regions results in a lattice matched staggered type heterojunction that promotes band to band tunneling rate, as depicted. Experimentally, the utilization of the metalorganic vapor phase epitaxy (MOVPE) technique allows for the successful incorporation of carbon as a dopant in order to produce p-type GaSb. Moreover, the fabrication of p-GaSb can also be achieved through the diffusion process by introducing Zinc (Zn) as a dopant.

The aforementioned material combinations allow for the attainment of a high on-state current without any occurrence of ambipolarity in the suggested heterojunction device. The proposed BESTFET comprises two parallel channels with a significant band to band tunneling area. In contrast, conventional TFET experiences tunneling in a limited region near the gate oxide area at the source-channel interface, leading to inadequate on-state current. However, the BESTFET offers two parallel channels with an enhanced band to band tunneling area, which efficiently enhances the on-state current. The fundamental physical and structural design parameters of the structures investigated in this study are presented in Table 1. The Numerical simulations are conducted using the ATLAS device simulator [32], and the device performance is evaluated using the following models:

- (1) Band to band tunneling; Basically, TFETs exploits the

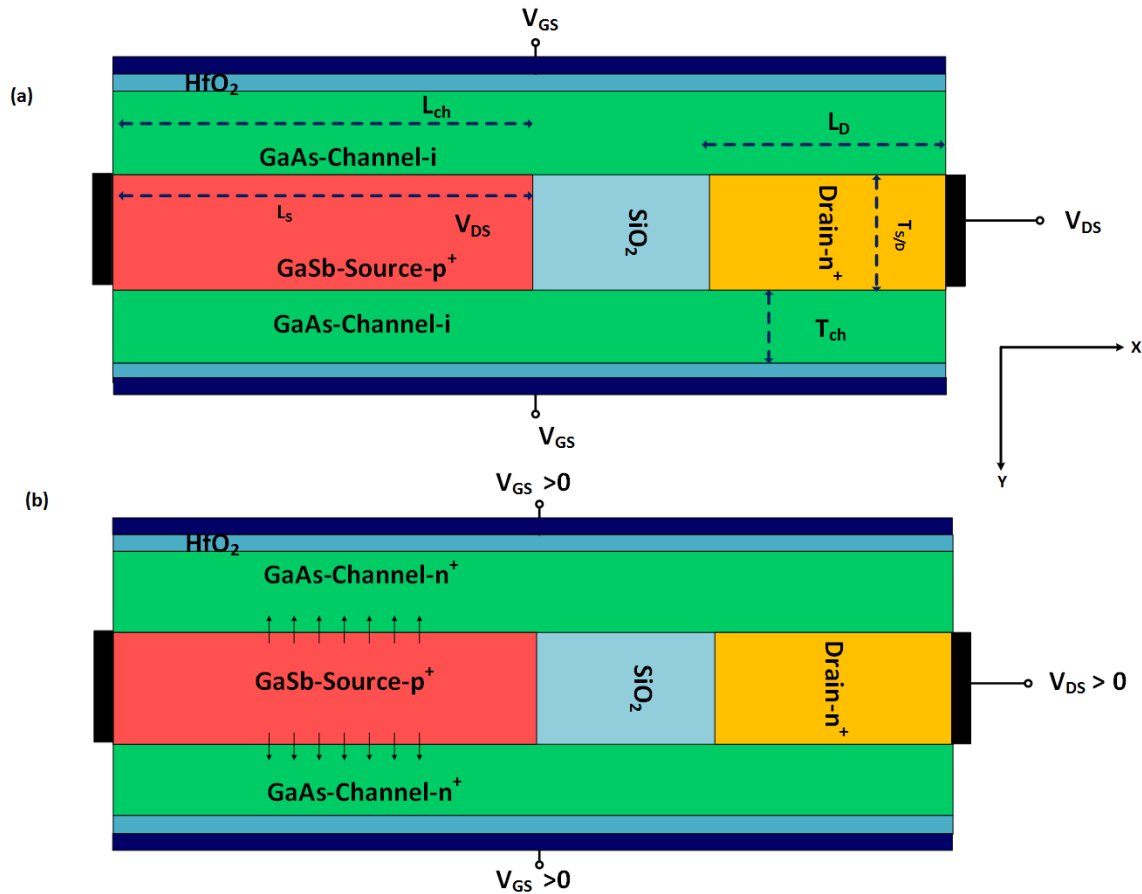


Fig. 1: (a) schematic of BESTFET with heterojunction material combination at the source and channel region, (b) on-state operation of BESTFET, in which tunneling occurs along the channel thickness in two parallel channels.

quantum mechanical phenomena better known as nonlocal band to band tunneling as the main current mechanism. In principle, band to band tunneling occurs when the channel conduction band aligns with the source valence band. Following the band alignment, energetic overlap emerge promptly which leads to a steep switching of the device from off-state to on-state. (2) Band gap narrowing; The energy band gap is a crucial factor that can significantly impact the rate of tunneling. Empirical evidence suggests that the contraction of the band gap takes place in regions with heavy doping because of the emergence of supplementary impurity bands caused by the overlapping of impurity states. (3) Mobility models; Mobility models take into account the effect of dopant density and lattice scattering on carrier mobility, as well as the influence of vertical and horizontal electric fields. In areas with high levels of doping, the mobility of carriers is reduced as dopant atoms serve as potent scatter centers, thereby diminishing the average time between collisions. Additionally, the carrier drift velocity and effective mobility can be significantly impacted by the electric field emanating from the gate and drain bias. (4) Trap assisted tunneling (or Shockley-Read-Hall recombination) models; It is evident that the process of band-to-band tunneling involves the movement of an electron from the valence band across the band gap to the conduction band. However, the occurrence of energy states within the band gap due to the presence of traps and defects can lead to tunneling before the onset of gate bias

assisted band to band tunneling. Consequently, a decline in subthreshold swing can be expected. (5) The proposed BESTFET contains a channel region that is confined between the gate insulator and the source region according to the quantum confinement model. As the thickness of the channel scales down below 5nm, the quantum capacitance is affected by the finite number of density of states in the channel, leading to variations in the gate capacitance. Thus, appropriate quantum confinement models must be utilized to ensure accurate results.

Table 1: Initial physical and structural design parameters for the BESTFET.

Parameters	Value
Channel Length (L_{ch})	60nm
Extended Source Length(L_s)	30nm
Channel Thickness (T_{ch})	5nm
Gate Insulator Thickness-HfO ₂	2nm
Source Doping Density	3×10^{19} (cm ⁻³)
Channel Doping Density	intrinsic
Drain Doping Density	1×10^{18} (cm ⁻³)
Gate Workfunction (WF_G)	4.0 eV
Source and Drain Thickness	5nm
Length of SiO ₂	10nm

3. RESULTS AND DISCUSSIONS

The energy band bending occurring at the interface between the source and channel regions can be roughly estimated as a barrier that resembles a triangle. Utilizing the Wentzel-Kramers-Brillouin approximation (WKB), the probability of tunneling in the BESTFET can be elucidated.

$$T_{WKB} \approx \exp\left(-\lambda \frac{4\sqrt{2m^*}E_G^{3/2}}{3q\hbar(E_G+\Delta\Phi)}\right) \quad (1)$$

where λ shows the tunneling length and shows the extension of the transition region at the interface of source and channel regions, E_G denotes the bandgap energy of the source region and $\Delta\Phi$ represents the tunneling window, where the BTBT is allowed to occur and m^* is the carrier effective mass. Essentially, the smaller value of λ demonstrates a more pronounced band bending at the tunneling region. From a theoretical standpoint, the increase in gate bias regulates the tunneling probability, T_{WKB} , through the reduction of λ and simultaneous increment of the $\Delta\Phi$ window. To enhance the operational effectiveness of TFET, it is imperative to meticulously engineer the tunneling junction in order to ensure a high BTBT rate and exceptional gate controllability across the tunneling barrier. In simpler terms, the tunneling barrier should be sufficiently thin to initiate tunneling. The energy band diagram for the BESTFET is presented in Fig. 2 for both the off-state ($V_{GS}=0V$ $V_{DS}=1V$) and on-state ($V_{GS}=1V$ $V_{DS}=1V$) conditions from the source to the channel. In the off-state, the electron density within the channel is inadequate to generate a narrow tunneling barrier, resulting in an absence of tunneling. However, when the gate voltage surpasses the threshold voltage value, the energy band within the channel is pulled down, and a significant reduction in the tunneling barrier at the source-channel interface is observed. Based on the energy band diagrams, the tunneling barrier width in the off-state is approximately 5nm, and by employing the gate bias in the on-state, this barrier reduces to 3nm. Due to this phenomenon, there is a significant surge in the drain current which happens suddenly, resulting in a subthreshold swing that is much lower compared to the conventional MOSFET Boltzmann limit. Threshold voltage, as defined, is the minimum gate voltage necessary to initiate tunneling, leading to a rapid increase in drain current from the off-state to the on-state. Fig. 3 illustrates the density of both electrons and holes in the BESTFET device, both in the off-state and on-state. Essentially, in the absence of the gate bias, the electron density in the channel is insufficient for the creation of a sharp tunneling junction. Nonetheless, through the application of a positive bias to the sidewall gates, a build-up of electrons within the channel enables the formation of a thin tunneling barrier to facilitate tunneling. Fig. 4 displays the impact of drain bias variation on the transfer characteristics of the BESTFET. It is observed that the threshold voltage and off-state current of the device remain unaffected by drain bias variation. Tunneling field-effect transistors suffer from a major issue of drain induced source tunneling (DIST), wherein the electric field lines arising from the drain electrode affect the tunneling barrier width at the source and channel region interface in the absence of gate bias. However, the proposed structure deals with this issue by separating the drain region from the tunneling junction using a thick oxide layer, which significantly reduces the DIST effect. It is evident that the carrier velocity and, correspondingly, the on-

state current increase as the drain bias is incremented. The subthreshold swing (SS) in the proposed device can be approximately defined as:

$$SS \approx \frac{\ln(10)}{q} W \quad (2)$$

in which q is the basic electric charge, W is defined as the tunneling barrier width. Essentially, as the gate bias is elevated towards adequately positive magnitudes, the width of the tunneling barrier is fundamentally diminished. This signifies the emergence of tunneling and a substantial reduction in the subthreshold swing. The subthreshold swing is precisely defined as the gate bias needed to achieve a six-fold change in the drain current, spanning from the minimum off-state current. The results demonstrate that for $V_{DS}=1V$, 2.24×10^{-4} (A/ μm), an off-state current of 1.24×10^{-15} (A/ μm), an on/off current ratio of 1.81×10^{11} , and a subthreshold swing of 5 (mV/dec) can be achieved, which makes this device feasible for high-speed low-power digital applications.

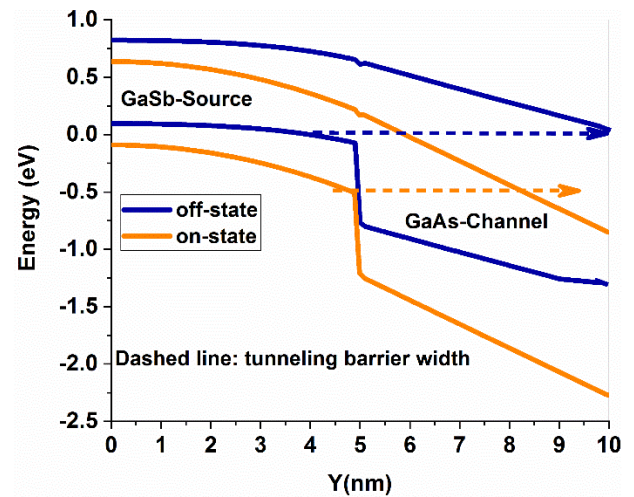


Fig. 2: Energy band diagram of the BESTFET in the off-state and on-state operation along the vertical direction from the source region towards lower channel region.

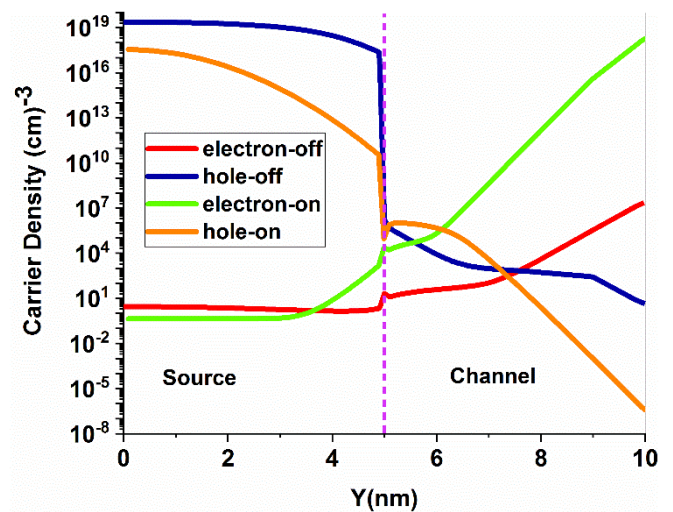


Fig. 3: Carrier density along the vertical direction from the source to lower channel region in the off-state and on-state.

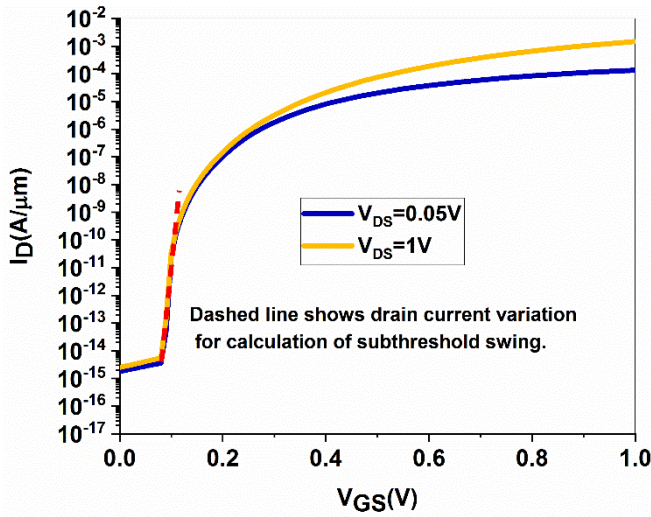


Fig. 4: Transfer characteristics of BESTFET as the drain bias is varied.

Fig. 5 depicts the transfer characteristics of the BESTFET with respect to the parametrization of the metal gate workfunction. Following materials are considered for the gate electrode: Hafnium (3.9eV), Zirconium (approximately 4eV), Indium (4.1 eV), Molybdenum (4.2 eV) and Titanium, Silver, Niobate (4.3 eV). It is evident that the distinction between the workfunction of the gate and the channel can indeed influence the carrier density within the channel. The findings indicate that an increase in the gate workfunction results in a reduced electron density in the channel, leading to a shift in the on-state to off-state transition voltage towards higher positive values. It is imperative to note that a narrow tunneling barrier necessitates a sufficient electron density. As such, a gate workfunction engineering framework is necessary for optimal device performance with a minimal threshold voltage value.

In order to initiate tunneling, a sharp p⁺-n⁺ tunneling junction is fundamental. The density of doping in the source is a crucial factor that can impact the tunneling rate. The transfer characteristics of BESTFET are depicted in Fig. 6, which demonstrates the effect of source doping density. The findings reveal that decreasing the source doping density leads to a significant increase in the threshold voltage and a decrease in the tunneling rate. The thickness of the space charge region at the interface of the source and channel regions is inversely related to the doping concentration of the source region. It is clear that a high concentration of dopants in the source region reduces the width of the tunneling barrier, thus increasing the electric field at the primitive gate electric field. Consequently, the increased electric field across the tunneling junction results in an elevation of the tunneling current. Therefore, to achieve a sharp tunneling junction and minimize resistance in the source region, a higher doping density should be utilized. The proposed device is characterized by an extended tunneling window that is situated at the interface of the source and channel region. The on-state and off-state variation with respect to the length of the extended source region is illustrated in Fig. 7. It is observed that the on-state current increases considerably as the L_s increases, which is indicative of an increment in the

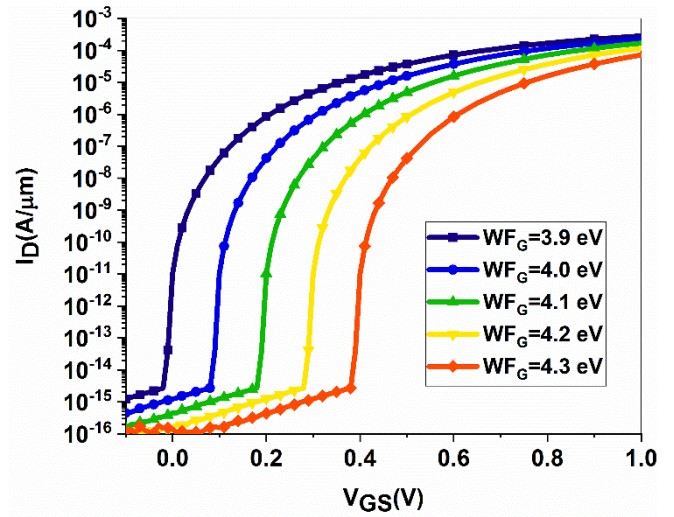


Fig. 5: Transfer characteristics of BESTFET as the gate workfunction is parametrized, V_{DS}=1V.

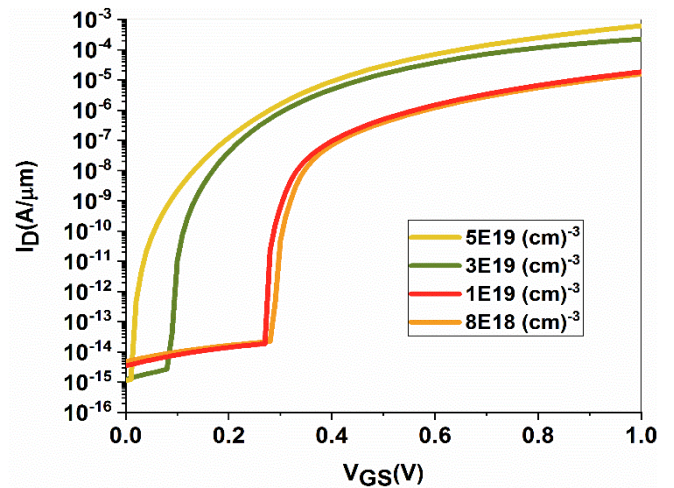


Fig. 6: I_D-V_{GS} curves of BESTFET as the source doping density is varied, V_{DS}=1V.

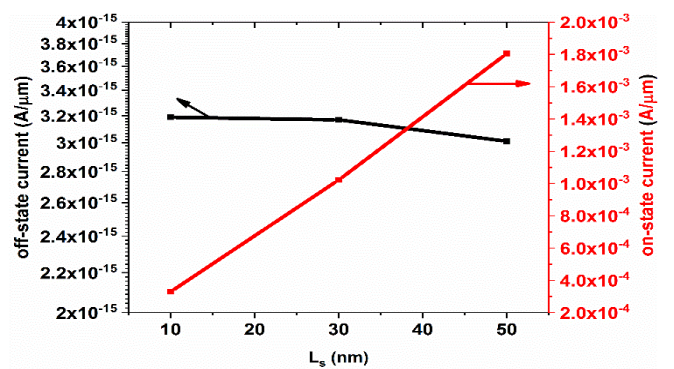


Fig. 7: Off-state and on-state current variation of the BESTFET versus source length.

band to band tunneling rate. Conversely, the off-state current exhibits negligible variation with respect to L_s. This is due to the embedded oxide region that lies between the channel and the drain side, as well as the wide band gap of the drain region, which limit the effect of the drain electric field on the tunneling rate.

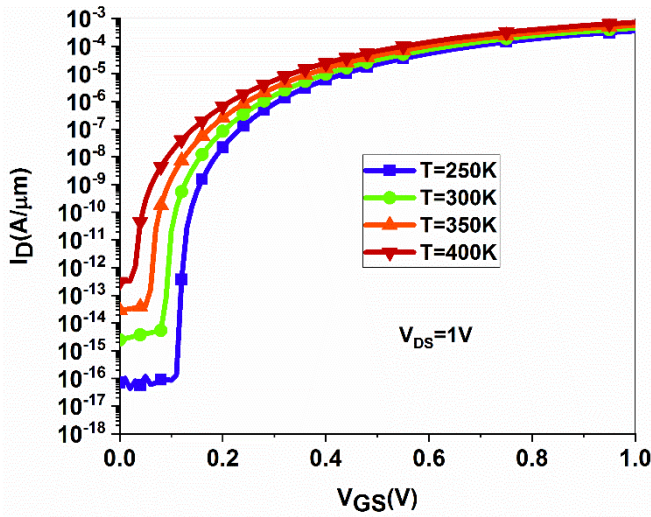


Fig. 8: Transfer characteristics of the BESTFET as the temperature is parametrized, $V_{DS}=1V$.

Temperature, a fundamental physical measure, has a profound impact on device performance as illustrated in Fig. 8 that showcases the effect of temperature on the transfer characteristics of BESTFET. It is evident that temperature has a notable influence on the off-state current and threshold voltage, while the on-state current shows a negligible variation. Specifically, an increase in temperature causes a variation in the density of minority carriers in the source region, leading to a contribution of minority electrons to the leakage current. In addition, scattering resulting from imperfections in the crystal structure, such as ionized impurities, assumes as the dominant mechanism when temperatures are low. Due to the decreased agitation of the atoms in the cooler lattice, scattering within the lattice becomes less significant. However, the motion of the carriers also slows down. Considering that a carrier moving at a slower pace is more likely to experience a stronger scattering effect when interacting with a charged ion compared to a carrier with higher momentum, events of impurity scattering lead to a reduction in mobility as the temperature decreases. However, in the on-state, the tunneling current predominantly consists of majority carriers and hence temperature does not have a significant effect on the drain current.

The findings indicate that the gate workfunction and source doping are crucial design parameters that significantly impact the p^+-n^+ tunneling junction, leading to a modification in the tunneling rate. In Fig. 9, the 2D variation matrix of the threshold voltage and on-state current is computed as a function of source doping density and gate workfunction. The point of maximum threshold voltage is situated in the uppermost left-hand corner of the variation matrix, indicating that a combination of a high gate workfunction value and low source density leads to an increase in the width of the depletion layer at the tunneling junction. To attain the lowest possible positive threshold voltage and high on-state current with steep subthreshold swing for n-type TFET, a sharp p^+-n^+ tunneling junction is essentially required. The variation matrix reveals that the proper combination of source doping density and gate workfunction in the bottom right side of the contour can be adjusted to achieve the lowest possible positive threshold voltage. Consequently, decreasing the

transition bias for the onset of tunneling can result in maximum drain current, which can be observed in the bottom right corner of the variation matrix. It is noteworthy to state that the implementation of a heavily doped region in the source region can serve as a means to decrease the parasitic resistance of the source. Additionally, the manipulation of the gate workfunction engineering framework can be utilized to adjust the threshold voltage and thus the on-state current.

The illustration of the frequency response of the apparatus is evident in Fig. 10. The results demonstrate the achievement of a unity current-ratio frequency of 300GHz, thereby emphasizing the feasibility of employing the device in high-frequency applications. The proposed device possesses two sidewall channels as well as an extended source region. The presence of a thin film channel region leads to the existence of capacitance between the gate and the source/drain electrode. This capacitance plays a crucial role in the degradation of the current gain at high frequency. At high frequency, the impedance of the gate and the source/drain capacitance deteriorates, resulting in a decline in the current gain. This decline occurs due to the inability of the gate bias to effectively modulate the width of the tunneling barrier.

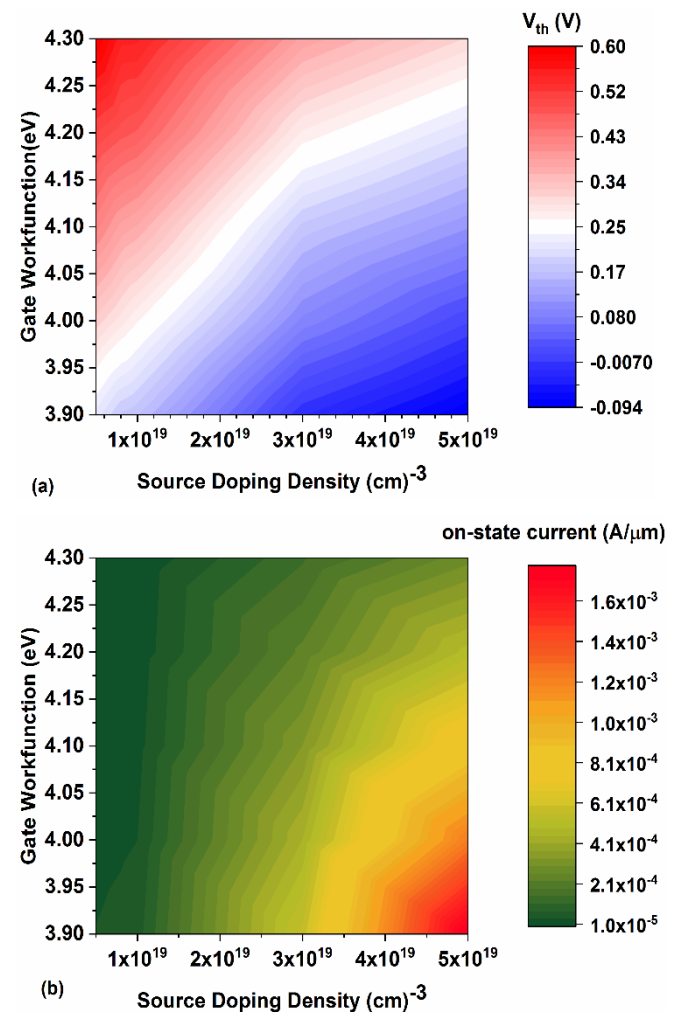


Fig. 9: (a) 2D variation contour of threshold voltage, (b) on-state current variation as a function of source doping density and gate workfunction value.

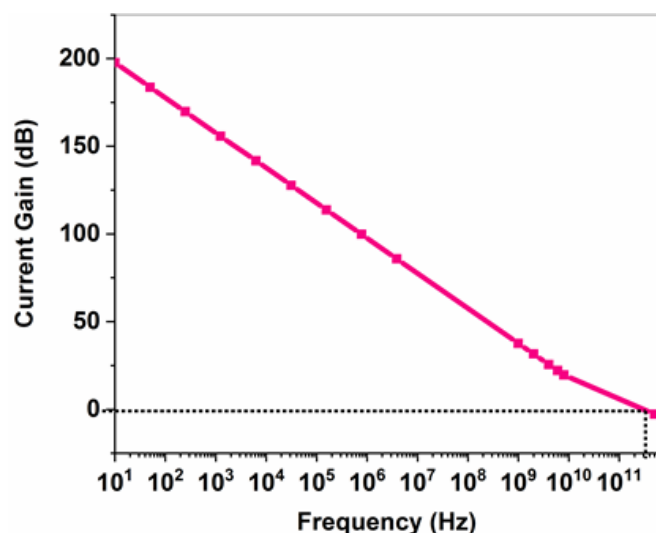


Fig. 10: Current gain versus frequency of the BESTFET for calculating the unity current-ratio frequency of the device.

4. CONCLUSION

In this study, a bi-channel TFET device with two sidewall parallel channels is subjected to a comprehensive analysis of its electrical characteristics. The study also evaluates the influence of crucial structural and design parameters on the device's performance. The findings showcase the device remarkable resistance to short channel effects, thereby highlighting its suitability for employment in the nanoscale regime. Furthermore, the gate workfunction and source doping density emerge as critical design parameters that significantly impact the device's performance. A 2D variation contour is computed for the threshold voltage and on-state current, taking into account the gate workfunction and source doping density variation. The results reveal that in order to attain a low-power high-speed TFET device, the source region should possess a high doping density. Additionally, the electrical performance of the device can be optimized through precise gate workfunction engineering framework.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Zahra Ahangari: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Roles/Writing - original draft, Writing - review & editing.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The ethical issues; including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy has been completely observed by the authors.

REFERENCES

[1] S. Jha, and S. K. Choudhary, "A comparative analysis of the short-channel effects of double-gate, tri-gate and

gate-all-around MOSFETs", *International Journal of Nanoparticles*, vol. 12, no. 1-2, pp. 112-121, 2020.

- [2] F. Á. Herrera et al., "Advanced short-channel-effect modeling with applicability to device optimization-Potentials and scaling", *IEEE Transactions on Electron Devices*, vol. 66, no. 9, pp. 3726-3733, 2019.
- [3] Q. Xie, J. Xu, and Y. Taur, "Review and critique of analytic models of MOSFET short-channel effects in subthreshold", *IEEE Transactions on Electron Devices*, vol. 59, no. 6, pp. 1569-1579, 2012.
- [4] H. Ghorbani, and J. L. R. Martinez, "A feedforward active gate voltage control method for SiC MOSFET driving," *Journal of Applied Research in Electrical Engineering*, vol. 2, no. 1, pp. 87-94, 2023.
- [5] M. Arab Nezhad, and A. Mahani, "Design of low-power approximate logarithmic multipliers with improved accuracy," *Journal of Applied Research in Electrical Engineering*, vol. 2, no. 1, pp. 95-102, 2023.
- [6] A. Davami, and M. H. Shahrokh Abadi, "Investigating the effects of phosphorene nanotubes (PNTs)-on-gold substrates on the enhancements of the sensitivity of SPR biosensors", *Journal of Applied Research in Electrical Engineering*, vol. 1, no. 2, pp. 169-174, 2021.
- [7] K. M. C. Babu, and E. Goel, "Evolution of tunnel field effect transistor for low power and high speed applications: a review," *Silicon*, vol. 14, no. 17, pp. 11051-11060, 2022.
- [8] S. Chander, S. K. Sinha, and R. Chaudhary, "Comprehensive review on electrical noise analysis of TFET structures," *Superlattices and Microstructures*, vol. 161, p. 107101, 2022.
- [9] R. Goswami, and B. Bhowmick, "Comparative analyses of circular gate TFET and heterojunction TFET for dielectric-modulated label-free biosensing," *IEEE Sensors Journal*, vol. 19, no. 21, pp. 9600-9609, 2019.
- [10] K. S. Singh, S. Kumar, and K. Nigam, "Impact of interface trap charges on analog/RF and linearity performances of dual-material gate-oxide-stack double-gate TFET", *IEEE Transactions on Device and Materials Reliability*, vol. 20, no. 2, pp. 404-412, 2020.
- [11] J. E. Jeyanthi, T. S. A. Samuel, A. S. Geege, and P. Vimala, "A detailed roadmap from single gate to heterojunction TFET for next generation devices," *Silicon*, vol. 14, no. 7, pp. 3185-3197, 2022.
- [12] C.-Y. Chen, H. Ilatikhameneh, J. Z. Huang, G. Klimeck, and M. Povolotskyi, "Impact of body thickness and scattering on III-V triple heterojunction TFET modeled with atomistic mode-space approximation," *IEEE Transactions on Electron Devices*, vol. 67, no. 8, pp. 3478-3485, 2020.
- [13] N. Oliva, J. Backman, L. Capua, M. Cavalieri, M. Luisier, and A. M. Ionescu, "WSe₂/SnSe₂ vdW heterojunction Tunnel FET with subthermionic characteristic and MOSFET co-integrated on same WSe₂ flake," *npj 2D Materials and Applications*, vol. 4, no.1, p. 5, 2020.

- [14] U. S. Shikha, B. Krishna, H. Harikumar, J. Jacob, A. Pradeep, and R. K. James, "OFF current reduction in negative capacitance heterojunction TFET," *Journal of Electronic Materials*, vol. 52, no. 4, pp. 2695-2707, 2023.
- [15] I. Chahardah Cherik, and S. Mohammadi, "Germanium-source L-shaped TFET with dual in-line tunneling junction," *Applied Physics A*, vol. 127, no. 7, p. 525, 2021.
- [16] S. Singh and B. Raj, "Analytical and compact modeling analysis of a SiGe hetero-material vertical L-shaped TFET," *Silicon*, pp. 1-11, 2021.
- [17] S. S. Shin, J. H. Kim, and S. Kim, "L-shaped tunnel FET with stacked gates to suppress the corner effect," *Japanese Journal of Applied Physics*, vol. 58, no. SD, p. SDDE10, 2019.
- [18] S. Yun *et al.*, "F-shaped tunnel field-effect transistor (tfet) for the low-power application," *Micromachines*, vol. 10, no. 11, p. 760, 2019.
- [19] P. Singh, and D. S. Yadav, "Design and investigation of f-shaped tunnel fet with enhanced analog/rf parameters," *Silicon*, pp. 1-16, 2021.
- [20] L. Lattanzio, L. De Michielis, and A. M. Ionescu, "The electron-hole bilayer tunnel FET," *Solid-State Electronics*, vol. 74, pp. 85-90, 2012.
- [21] Ashita, S. A. Loan, A. G. Alharbi, and M. Rafat, "Ambipolar leakage suppression in electron-hole bilayer TFET: Investigation and analysis," *Journal of Computational Electronics*, vol. 17, pp. 977-985, 2018.
- [22] J. L. Padilla, C. Alper, A. Godoy, F. Gamiz, and A. M. Ionescu, "Impact of asymmetric configurations on the heterogate germanium electron-hole bilayer tunnel FET including quantum confinement," *IEEE Transactions on Electron Devices*, vol. 62, no. 11, pp. 3560-3566, 2015.
- [23] Z. Ahangari, and S. Mahmodi, "Design and sensitivity analysis of steep-slope bi-channel vertical tunnel field effect transistor," *Silicon*, vol. 13, pp. 1917-1924, 2021.
- [24] V. Ambekar and M. Panchore, "Realization of high-speed logic functions using heterojunction vertical TFET," *Applied Physics A*, vol. 129, no. 3, p. 166, 2023.
- [25] J. Singh, S. Singh, and N. Paras, "Design and integration of vertical TFET and memristor for better realization of logical functions," *Silicon*, vol. 15, no. 2, pp. 783-792, 2023.
- [26] S. Meriga and B. Bhowmick, "Investigation of a dual gate pocket-doped drain engineered tunnel FET and its reliability issues", *Applied Physics A*, vol. 129, no. 2, p. 104, 2023.
- [27] R. Paul, "Simulation and performance evaluation of charge plasma based dual pocket biosensor using sige-heterojunction TFET design," *Silicon*, vol. 15, no. 5, pp. 2147-2162, 2023.
- [28] Z. Ahangari, "Performance optimization of a nanotube core-shell semi-junctionless p⁺p⁺n heterojunction tunnel field effect transistor," *Indian Journal of Physics*, vol. 95, pp. 1091-1099, 2021.
- [29] K. Kumar, A. Kumar, V. Mishra, and S. C. Sharma, "Implementation of band gap and gate oxide engineering to improve the electrical performance of SiGe/InAs Charged plasma-based junctionless-TFET", *Silicon*, vol. 15, no. 3, pp. 1303-1313, 2023.
- [30] A. Masoudi, Z. Ahangari, and M. Fathipour, "Performance optimization of a nanoscale GaSb P-channel electron-hole bilayer tunnel field effect transistor using metal gate workfunction engineering", *Materials Research Express*, vol. 6, no. 9, p. 096311, 2019.
- [31] U. Mushtaq, N. Kumar, S. Anand, and I. Amin, "Design and performance analysis of core-shell dual metal-dual gate cylindrical GAA silicon nanotube-TFET", *Silicon*, vol. 12, pp. 2355-2363, 2020.
- [32] ATLAS User Manual 2018 (Santa Clara, USA: Silvaco International).

BIOGRAPHY



Zahra Ahangari received her PhD degree in Electrical Engineering-Electronics from Islamic Azad University, Science and Research Branch, in the year 2013. From the year 2004 onwards, Zahra served as a research assistant in the Device Modelling and Simulation laboratory of University of Tehran under supervision of Professor Morteza Fathipour. Currently, Zahra is the faculty member of electrical engineering department at Islamic Azad University, Yadegar-e-Imam Khomeini (RAH) Shahr-e-Rey Branch. Zahra's main areas of research are modelling and simulation of nanoelectronic devices, nano sensors, and solar cells.

Copyrights

© 2024 Licensee Shahid Chamran University of Ahvaz, Ahvaz, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution –Non-Commercial 4.0 International (CC BY-NC 4.0) License (<http://creativecommons.org/licenses/by-nc/4.0/>).



Journal of Applied Research in Electrical Engineering

Vol. 2, No. 2, Summer and Autumn 2023

jaree.scu.ac.ir

New Approach in Minerals Cracking Using Electromechanical Breakdown Theory Asaad Shemshadi, and Mohamad Reza Khojaste	103
Combined Economic Emission Dispatch in a Grid Connected Microgrid Using an Improved Mayfly Algorithm Nicholas Kwesi Prah II, Elvis Twumasi, and Emmanuel Asuming Frimpong	112
Improving the Diagnosis of Sudden Cardiac Death by Using Non-Linear Features of the ECG Signal and Hybrid RBF Afshin Koliji, Sara MihanDoost, Nematollah Ezzati, and Ehsan Mostafapour	120
Selection and Tuning Propagation Path Loss Model for Hawassa City, Ethiopia at 1800 MHz Frequency Tamirat Yenealem, and Robel Getachew	127
Different Types of Distributed Optimal Leader-Follower Consensus Protocol Design for a Class of High-Order Multi-Agent Systems Farideh Azadmanesh, and Reza Ghasemi	136
Morphological and Statistical Filter-Based Multi-Direction Building Detection for SAR Images Fateme Amjadipour, Maryam Imani, and Hassan Ghassemian	145
Design of a Tunable 4th Order OTA-C Band-Pass Filter for use in Front-End of ADC Mohsen Ghaemmaghami, and Shahbaz Reyhani	152
Nonlinear Observer-Based Control Design for a Three-Axis Inertial Stabilized Platform Mohammad Mehdi Zohrei, and Hamidreza Javanmardi	158
Optimal Sizing and Placement of Series Capacitors in Distribution Networks Using Modified Elephant Herding Optimization Algorithm Elvis Twumasi, Abdul-Fatawu Seini Yussif, and Emmanuel Asuming Frimpong	173
Re-examination of the Doubly Fed Induction Machine (DFIM) Model Taking Into Account the Rotor Reactive Power Mohammad Naser Hashemnia	182
Full-Duplex Device-to-Device Relays in a Novel Hybrid and Adaptive Joint Relaying Network: Symbol Error Analysis and Optimum Power Allocation Mehdi Basiri Abarghouei, and Reza Saadat	194
Performance Analysis of a Steep-Slope Bi-channel GaSb-GaAs Extended Source Tunnel Field Effect Transistor with Enhanced Band-to-Band Tunneling Current Zahra Ahangari	206



Iranian Association of Electrical
and Electronics Engineers



Shahid Chamran
University of Ahvaz