

# **SCHOLARLY PUBLICATIONS**

*A CURRENT AWARENESS BULLETIN  
OF RESEARCH OUTPUT*

**@DTU**

(112<sup>th</sup> Edition)

**APRIL 2022**

**BY: CENTRAL LIBRARY**

**DELHI TECHNOLOGICAL UNIVERSITY**

(FORMERLY *DELHI COLLEGE OF ENGINEERING*)

GOVT. OF N.C.T. OF DELHI

SHAHBAD DAULATPUR, MAIN BAWANA ROAD

DELHI 110042

# **PREFACE**

This is the **One Hundred Twelfth** Issue of Current Awareness Bulletin started by Delhi Technological University, Central Library. The aim of the bulletin is to compile, preserve and disseminate information published by the faculty, students and alumni for mutual benefits. The bulletin also aims to propagate the intellectual contribution of Delhi Technological University (DTU) as a whole to the academia.

The bulletin contains information resources available in the internet in the form of articles, reports, presentations published in international journals, websites, etc. by the faculty and students of DTU. The publications of faculty and student which are not covered in this bulletin may be because of the reason that the full text either was not accessible or could not be searched by the search engine used by the library for this purpose.

The learned faculty and students are requested to provide their uncovered publications to the library either through email or in CD, etc. to make the bulletin more comprehensive.

This issue contains the information published during **April, 2022**. The arrangement of the contents is alphabetical. The full text of the article which is either subscribed by the university or available in the web is provided in this bulletin.

**Central Library**



# CONTENTS

1. A survey on IRS NOMA integrated communication networks, Sandeep Kumar, Poonam Yadav, **8.Manpreet Kaur** and Rajesh Kumar, CSE, DTU
2. An analytical insight of discussions and sentiments of Indians on Omicron-driven third wave of COVID-19 using twitter data, Deepika Vatsa and **6.1.Ashima Yadav**, CSE, DTU
3. An attention-based deep-learning system with fMRI functional connectivity optimized frequency EEG microstates classifies distinct temporal cortical communications of different cognitive tasks, Swati Agrawal, Vijayakumar Chinnadurai and **3.Rinku Sharma**, Applied Physics, DTU
4. Applicability of ML-IoT in Smart Healthcare Systems: Challenges, Solutions & Future Direction, **6.Nikhil Sharma** and **3.Prashant Giridhar Shambharkar**, CSE, DTU
5. Capacitance Characteristics Behavior of 0.5 Order FC Using CFOA Based FC Multiplier, **6.Rakesh VERMA**, **3.Neeta PANDEY** and **3.Rajeshwari PANDEY**, Electronics, DTU
6. Characterization of finished surface through thermal additive centrifugal abrasive flow machining for better surface integrity, **6.Parvesh Ali**, R S Walia, **3.Qasim Murtaza** and **3.Ranganath Muttanna Singari**, Mechanical, DTU
7. Comparative Analysis of Software Reliability Prediction Using Machine Learning and Deep Learning, **8.Akshat Jindal**, **8.Ashi Gupta** and **3.Rahul**, Software Engineering, DTU
8. Design of Low Cost Bio-impedance Measuring Instrument, **3.Rajesh Birok** and **3.Rajiv Kapoor**, Electronics, DTU

9. DETECTING TROPICAL CYCLONE FROM THE BASIC OVERVIEW OF LIFE OF EXTREMELY SEVERE CYCLONIC STORM, TAUKTAE, **7.Monu Yadav** and **3.Laxmi Narayan Das**, Applied Mathematics, DTU
10. Four-mode squeezed states in de Sitter space: A study with two field interacting quantum system, Sayantan Choudhury, Sudhakar Panda, **8.Nilesh Pandey** and Abhishek Roy, Applied Physics, DTU
11. Geopolymer Concrete: A Material for Sustainable Development in Indian Construction Industries, **6.Manvendra Verma**, **3.Nirendra Dev**, Ibadur Rahman, Mayank Nigam, Mohd. Ahmed and Javed Mallick, Civil, DTU
12. Global, regional, and national burden of colorectal cancer and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019, **3.Rajesh Sharma**, USME, DTU
13. Inverse Analog Filters: History, Progress and Unresolved Issues, Raj Senani, **3.Data Ram Bhaskar** and **7.Ajishkek Raj**, Electrical, DTU
14. K-means Embedded Deep Transform Learning for Hyperspectral Band Selection, **3.Anurag Goel** and Angshul Majumdar, IT, DTU
15. Multi objective study in powder mixed near dry electric discharge machining by using utility concept, **6.Sanjay Sundriyal**, **3.Vipin**, and **3.R S Walia**, Mechanical, DTU
16. On improving the performance of dynamic positive-feedback source-coupled logic (D-PFSCL) through inclusion of transmission gates, **6.Ranjana Sivaram**, Kirti Gupta and **3.Neeta Pandey**, Electronics, DTU
17. Performance Evaluation using Online Machine Learning Packages for Streaming Data, **6.Santosh Kumar Ray** and **3.Seba Susan**, IT, DTU

18. Real Time Analysis of Material Removal Rate and Surface Roughness for Turning of Al-6061 using ANN and GA, *8.Abhishek Jha, 8.Baibhav Kumar and 3.Ashok Kumar Madan*, Mechanical, DTU
19. S-DCNN: stacked deep convolutional neural networks formalware classification, *3.Anil Singh Parihar, 8.1.Shashank Kumar and 8.1.Savya Khosla*, CSE, DTU
20. Temperature Analysis of a Novel 5nm Stacked Oxide Top Bottom Gated Junctionless FinFET for improved switching and efficiency, *8.Lakshya Gangwani and 8.Sumanyu Hajela*, Applied Physics, DTU

1. *Vice Chancellor*

2. *Pro Vice Chancellor*

3. *Faculty*

4. *Teaching-cum-Research Fellow*

5. *Asst. Librarian*

6. *Research Scholar*

7. *PG Scholar*

8. *Undergraduate Student*

1.1. *Ex Vice chancellor*

2.1. *Ex Pro Vice Chancellor*

3.1. *Ex Faculty*

4.1. *Alumni*

5.1 *Others*

6.1. *Ex Research Scholar*

7.1. *Ex PG Scholar*

8.1. *Ex Undergraduate Student*



# A survey on IRS NOMA integrated communication networks

Sandeep Kumar<sup>1</sup> · Poonam Yadav<sup>2</sup> · Manpreet Kaur<sup>3</sup> · Rajesh Kumar<sup>2</sup>

Accepted: 16 March 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

Intelligent reflecting surfaces aided communication have been emerging as strong candidates to support the 6G wireless physical platforms. IRS has shown promising qualities in enhancing the spectral efficiency of wireless networks because of its capability to alter the conduct of interacting electromagnetic waves through intelligent handling of the reflections phase shifts. Also, NOMA proves itself to be superior among the other multiple access techniques as it supports a greater number of users using non-orthogonal resource allocation. This paper brings a survey over the IRS-assisted NOMA networks. The IRS and NOMA technologies, and their physical working principles are first introduced in the paper. The state-of-the-art of the IRS-assisted NOMA communication networks is next presented followed by a discussion of related performance parameters for analysis. Afterward, it discusses the resource allocation, and secrecy requirements in the IRS–NOMA networks. Furthermore, it presents the relevant work related to the optimization of energy efficiency, power efficiency and coverage. A comparison of IRS–NOMA network with MIMO–NOMA, and relay aided NOMA network is provided. Finally, a few exciting open challenges for IRS-assisted NOMA networks are identified including optimization problem using ML, identifying implementing scenarios of NOMA or OMA with IRS, PLS, and terahertz communication.

**Keywords** Intelligent reflecting surfaces · Multiple access · OMA · NOMA · Relay networks · IRS assisted NOMA networks

## 1 Introduction

The future wireless networks are expected to support highly efficient, secure, reliable, and flexible design for emerging applications of 6G and beyond [1]. In order to achieve this goal, rigorous efforts have been undertaken in the research and development of wireless communications. As the wireless environment is modelled as uncontrolled entity, to mitigate its impairments several approaches such as using

multiple antennas, employing complex encoding and decoding algorithms at the end-points of the communication link, and adding additional network infrastructure, e.g., relays is worked upon to make the transmission reliable. These solutions, however, may increase the network complexity, the network power consumption, and the network deployment cost which is comprehensively discussed in subsequent sections. Due to its minimal complexity, intelligence, and energy-efficient features, IRS is a very promising technology for the development of beyond 5G or 6G wireless communications in recent years [2–5]. By aptly deploying an IRS in the environment, an extra communication link that goes through the IRS can be built between the transmitter and receiver, and thus, better support diverse user requirements, such as extended coverage, increased data rate, minimized power consumption, and enhanced secure transmissions. Not only theoretically attractive, IRS also possesses various advantages in terms of practical implementation such as it is of low hardware and energy cost, and can be easily deployed on the environment objects, e.g., the facades of buildings. Figure 1 illustrates the application of IRS in 5G and beyond communication networks. In comparison with conventional active relaying technologies, IRS consumes much less energy

---

✉ Manpreet Kaur  
manpreettiet@gmail.com

Sandeep Kumar  
sann.kaushik@gmail.com

Poonam Yadav  
poonam.yadav.25@gmail.com

Rajesh Kumar  
rk@nerist.ac.in

<sup>1</sup> Central Research Laboratory, BEL, Ghaziabad, India

<sup>2</sup> Department of Electronics and Communication, NERIST, Nirjuli, AP, India

<sup>3</sup> Department of Computer Science and Engineeringg, Delhi Technological University, New Delhi, Delhi, India

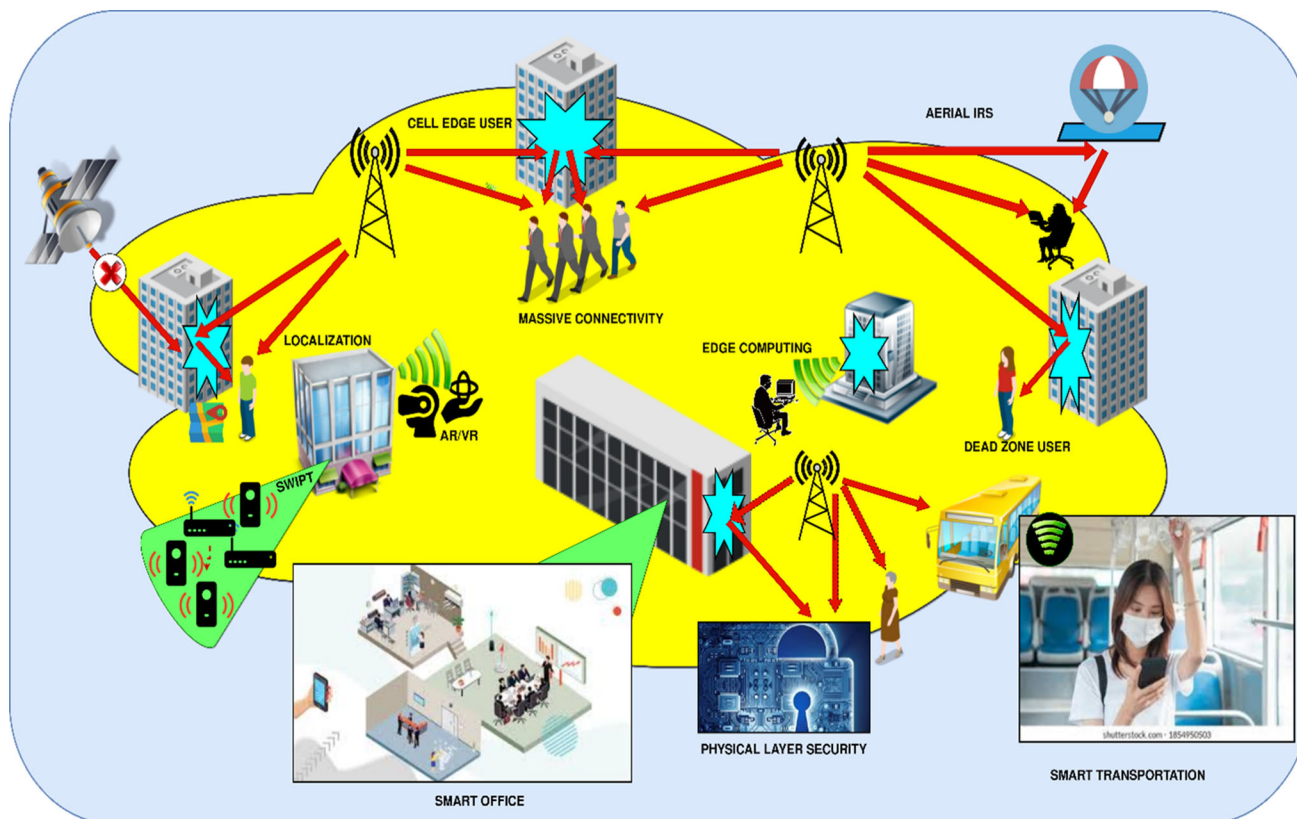


Fig. 1 IRS applications in 5G and beyond networks

and is less expensive since they do not employ RF chains or high power-consuming components. IRS can be seamlessly integrated into existing wireless communication networks to realize different objectives, including the improvement of the energy efficiency and the mitigation of inter-user interference [6]. To fully extract the benefits provided by IRS, it is necessary to investigate the integration of IRS with other transmission technologies for next-generation communication systems. Among the MA techniques NOMA has received considerable attention in the design of the next-generation networks. With NOMA, multiple UEs are co-scheduled and share the same radio resources in time, frequency, and/or code so the problem of spectrum scarcity can be solved up to some extent. In particular, many forms of NOMA have been developed to encourage dynamic spectrum sharing among users by opportunistically exploiting the user's heterogeneous channel conditions [7]. Conventionally, a user's channel conditions are viewed as a type of fixed and non-tunable phenomenon which is solely determined by the user's propagation environment. Therefore, if the user's channel conditions are not ideal for the application of NOMA, like if the users have similar channel conditions, the performance gain of NOMA can be quite limited. IRS enables a paradigm shift for the design of intelligent NOMA,

since the use of an IRS ensures that the propagation environment can be effectively and intelligently customized for the needs of NOMA. The application of IRS into NOMA is also promising, as IRS can be utilized to introduce desirable channel gain differences among users and lead to performance improvement of NOMA systems [8, 9]. By enabling multiple users to access the same time/frequency resources, NOMA can achieve higher spectral efficiency and energy efficiency, and better support massive connectivity when compared to OMA. Nevertheless, to obtain a decent performance gain of NOMA over OMA, the users are required to have a large channel gain disparity. In the conventional NOMA systems, it becomes difficult to achieve this since the user channels are determined by the highly stochastic propagation environments. A recent study over the IRS–NOMA communication technology introduces several capabilities of this integrated technology to improvise the spectral efficiency, user fairness, latency, and secrecy requirements. The characteristics of IRS to provide configurable channel gains via constructively or destructively adding the user signals can be used to suppress the inter-user interference, and thus, lead to improved throughput or fairness of NOMA systems.

## 1.1 Review of the existing literature

Several studies over the review of the IRS assisted network is performed in the literature. This has to be noted that several notions are given to IRS that are interchangeably used in the literature. Each of such terminologies emphasizes one particular feature of the smart surface [10]. Table 1 includes the list of acronyms used throughout the paper. To understand the basic characteristics of each notion, a comparison is provided in Table 2. A detailed tutorial on the IRS technology as a new paradigm enabling the future wireless network is provided in [11]. In [12], authors review IRS technology and key implementation technical challenges. A detailed review on the performance optimization of IRS-assisted wireless networks under several communication scenarios is performed in [13]. In [14] authors summarize RIS-empowered wireless networks in particular assuming RIS as an efficient transmitter. Likewise, [15] and [16] provide a complete overview of applications, merits, primary concept, and future research prospects as well as reflecting design issues on RIS, namely IRS channel estimation and resource allocation. Review of the principles of IRS and dispel three specific misconceptions regarding their functionality and performance gains is explained in [17]. In [18], comparison and contrast between RIS and relay are performed on the basis of their communication performance of measures. An overview of different implementations of RIS using meta-surface and reflect array, with appropriate channel modelling as well as problems and opportunities in RIS integrated wireless networks is performed in [19]. In [20] authors deal with the analysis of PLS over the RIS-assisted future wireless networks and provide a comprehensive review on articles related to PLS performance in IRS networks. The review of the related work is summarized in Table 3.

## 1.2 Motivation

Apart from these, several surveys on IRS have recently appeared in [26, 34, 33] but none of the prior works have overviewed and studied the integrated IRS–NOMA environment. IRS and NOMA are two different communication techniques that are envisioned to meet the requirements of 6G networks. These two important enabling technologies are complementary to each other, where the use of NOMA can improve the spectral efficiency and connectivity of IRS systems, and the use of IRSs ensures that the users' propagation environments can be effectively and intelligently customized for the implementation of NOMA. As a result, it is still completely not clear what are the expectations from the IRS assisted NOMA networks. Motivated by this research gap, this survey paper deals with the overview of the concept of the IRS-assisted NOMA networks. In view of this, the main

**Table 1** List of Acronyms

AF	Amplify and forward
ADC	Analog to digital converter
ADMM	Alternating direction method of multipliers
ASC	Average secrecy capacity
BER	Bit error rate
BS	Base station
BW	Bandwidth
CDMA	Code division multiple access
CSI	Channel state information
CLT	Central limit theorem
DAC	Digital to analog converter
DC	Direct current
DF	Delay and forward
DRL	Deep reinforcement learning
ED	Eavesdropper
EE	Energy efficiency
FDMA	Frequency division multiple access
FD	Full duplex
FET	Field effect transistor
IRS	Intelligent reflecting surfaces
HD	Half duplex
HARQ-CC	Hybrid automatic repeat request with chase combining
LIM	Large intelligent metasurfaces
LIS	Large intelligent surfaces
LOS	Line-of-sight
LS	Least squared
LTE	Long term evolution
MA	Multiple access
MEMS	Micro electro mechanical switch
MISO	Multiple-input–single-output
MIMO	Multiple-input-multiple-output
MRC	Maximal ratio combining
MSE	Mean squared error
NOMA	Non-orthogonal multiple access
OMA	Orthogonal multiple access
OFDMA	Orthogonal frequency division multiple access
OP	Outage probability
PIS	Passive intelligent surfaces
PLS	Physical layer security
RIS	Reconfigurable intelligent surfaces
SC	Selection combining
SDM	Software defined metasurfaces
SDR	Semi definite relaxation

**Table 1** (continued)

AF	Amplify and forward
SDS	Software defined surfaces
SIC	Successive interference cancellation
SINR	Signal-to-interference-noise-ratio
SOCP	Second-order cone programming
SNR	Signal-to-noise ratio
SRA	Smart reflect array
TDMA	Time division multiple access
UAV	Unmanned ariel vehicle
UE	User equipment
UL	Uplink

**Table 2** List of several notions of IRS and their characteristics

Notion	Characteristic	References
IRS	Emphasize more on reflecting property of the smart surface	[10, 11, 21]
LIS	Considers infinite surface length or massive antennas	[22–24]
RIS	Focuses more on “reconfigurability” of the smart surfaces means that the angle of reflection can be reconfigured (via software) regardless of the angle of incidence	[14, 6]
LIM	Considers a massive number of antennas for the asymptotic analysis of metasurfaces	[25, 26]
SRA	Emphasizes the surface’s reflection function	[27, 28]
PIS	Underlines the passive reflection without consuming transmit power	[29, 30]
SDS	Considers software defined interaction between surface and incoming wave	[31]
SDM	Considers both the metallic pattern through which the surface is engineered and software defined interaction of radios	[32]

input of this survey is to bring a thorough understanding of the functioning and challenges of IRS aided NOMA networks.

### 1.3 Major contributions

The aim of this article is to provide a comprehensive review of the implementation and application of this integrated environment of the IRS and NOMA technologies. The core contributions of this review are highlighted in the points mentioned below.

- A descriptive explanation of the theory of working of IRS system is provided which can enable the readers to understand the concept and importance of the IRS system in communication easily. It also includes the description of the hardware architecture of the IRS system which can be helpful for the core hardware design engineer to further enhance the hardware architecture of the IRS systems.
- The importance and need of MA techniques in the transmission are also highlighted with the main focus on the NOMA technique. A comparative study of OMA and NOMA techniques is also explained in the paper which proves the superiority of the NOMA to fulfill the demands of the 5G and beyond networks.
- The theory of IRS-aided NOMA networks and its application over the various fading channels to analyse the performances is comprehensively reviewed in the paper. To add on this performance optimization of the IRS aided NOMA networks, secrecy requirements, importance of CSI, and comparison with MIMO–NOMA and relay aided NOMA is also reviewed.
- In the end, future research directions and open exploration issues are identified which can empower the researchers to add on some productive outcomes in the mentioned areas.

### 1.4 Paper organization

This paper is organized as follows. Section 2 explains the principle of operation and the hardware architecture of the IRS system. Importance of MA techniques in wireless communication and the concept of NOMA transmission is discussed in Sect. 3. Section 4 includes state-of-art of IRS aided NOMA networks and a review of the performance parameters. It also reviews the optimization problems, PLS, and the importance of CSI in IRS NOMA networks. A comparison of IRS–NOMA with the current MIMO–NOMA and relay aided NOMA technology is provided in Sect. 5. Section 6 consists of a few research ideas followed by a conclusion in Sect. 7. The overall framework of this paper is shown in Fig. 2.

## 2 Theory of IRS-aided wireless communication

The wireless communication network systems are designed based on the postulate that the wireless transmission environment is uncontrollable that may impair the signal quality due to the unwanted reflections and refractions, and deleterious interference [35–37]. To meet the requirements of the next-generation networks which are envisioned to support high data rate with URLLC, there is a need for a technology that can transform the stochastic and uncontrollable wireless



**Table 3** Review of the existing literature and its contribution

References	Year	Main contribution	Comparison to this survey
[11]	2020	Provides a comprehensive tutorial over principle, state-of-the-art results to address the prime communication challenges, and potential directions for further investigation	Unlike the mentioned surveys, this paper brings a comprehensive theoretical understanding of the IRS along with the NOMA concept. This paper first discusses the working principle of the IRS. Then it highlights the importance of MA techniques in wireless communication with a main emphasis on the NOMA technique and provides a comparison between OMA and NOMA. Then a mathematical analysis to understand the IRS–NOMA integrated environment is provided. Then various performance metrics of IRS- assisted NOMA networks are provided. The secrecy requirements and the importance of CSI in the IRS–NOMA link are reviewed to improvise the study. Along with this, the performance optimization parameters such as energy efficiency, power efficiency, resource allocation, and coverage optimization are reviewed. This paper afterward surveys the comparison of IRS–NOMA with MIMO–NOMA and relay based NOMA technology. Lastly, various research scopes are highlighted on the integrated environment of IRS NOMA networks
[12]	2020	Examines the most common wireless communication applications and compares competitive benefits to existing technologies along with hardware architecture and signal model of the IRS	
[13]	2020	Provides overview over several performance measures and analytical methodologies for IRS-assisted wireless networks performance optimization	
[14]	2019	Explores the theoretical performance limits of RIS-enabled communication systems using mathematical methodologies and elaborates on its application in 6G and beyond networks	
[15]	2020	Reviews the latest open research field of smart radio environments empowered by RIS and also provides a theoretic perspective to RIS-empowered smart radio environments	
[16]	2020	Reviews design issues such as channel estimation, and resource allocation of IRS network	
[17]	2020	Reviews the state-of-art of IRS network and debunk the specific three myths about the functionality and performance gain of the IRS-assisted network	
[18]	2019	Highlights the key difference between the DF relay and IRS-assisted networks	
[33]	2020	Overviews different implementations of RIS using meta-surface and reflect-array also discuss the challenges and opportunities of IRS aided networks	
[20]	2020	Reviews the PLS in IRS assisted wireless networks	

propagation environment into a stable, reliable, and efficient communication system entity. Implementation of IRS in communication network has emerged as the most suitable solution to deal with the inherent drawback of wireless environment. This section is dedicated to provide a comprehensive review over the IRS-aided wireless communication networks.

## 2.1 Working principle

To pace up with the fast escalation of the data rate requirements and multimedia traffic, several potential technologies have been introduced in order to fulfil and address the challenges of 5G and beyond networks [38]. Recently, IRS has drawn significant attention of the researchers due to its significant capability in enabling and controllable wireless propagation environment to provide new solutions to achieve high spectral efficiency and energy efficiency for next-generation networks [21, 39]. In particular, IRS consists of an array of passive reflecting units, each one of which can freely bring out some change which might be in the phase,

amplitude, frequency, or even polarization to the incident signal [11]. IRS configures a wireless propagation environment via software-controlled reflection. This consequently gives another way in a general sense to deal with the wireless channel fading impairment issue and possibly accomplishes a breakthrough improvement for the capacity and reliability of the wireless channel. Figure 3 illustrates the IRS-aided wireless communication between the BS and the user terminal where LOS communication is blocked because of the presence of a tree.

The composite channel from the BS to the user via each IRS element is consists of three segments: the BS-IRS link, IRS's reflection, and the IRS-user link. The signal model accounting for all the reflected signals from IRS element is given as [11]  $n \in 1, 2 \dots N$

$$s(t) = \left( \sum_1^N \beta_n e^{j\theta_n} h_{r,n}^* g_n \right) x(t) = \mathbf{h}_r^H \Theta \mathbf{g} x(t), \quad (1)$$



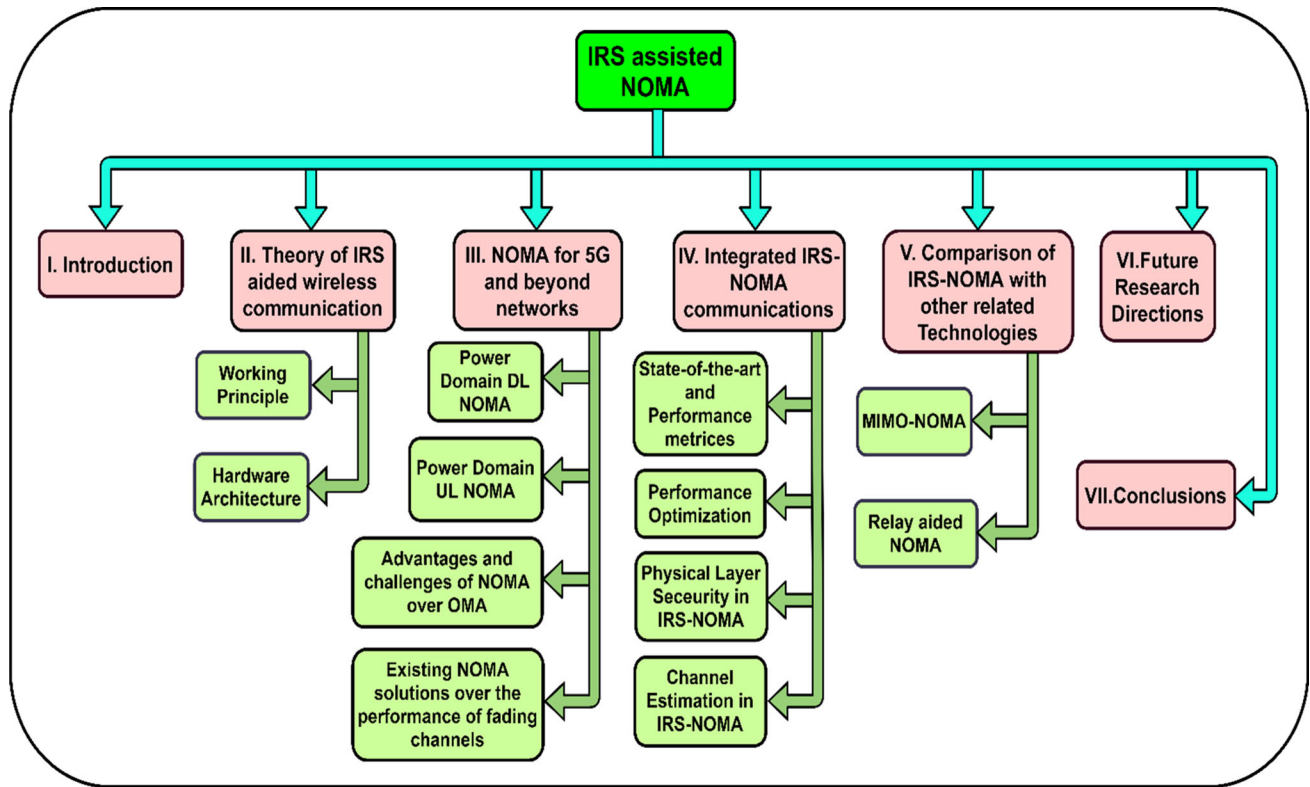


Fig. 2 Framework of the paper

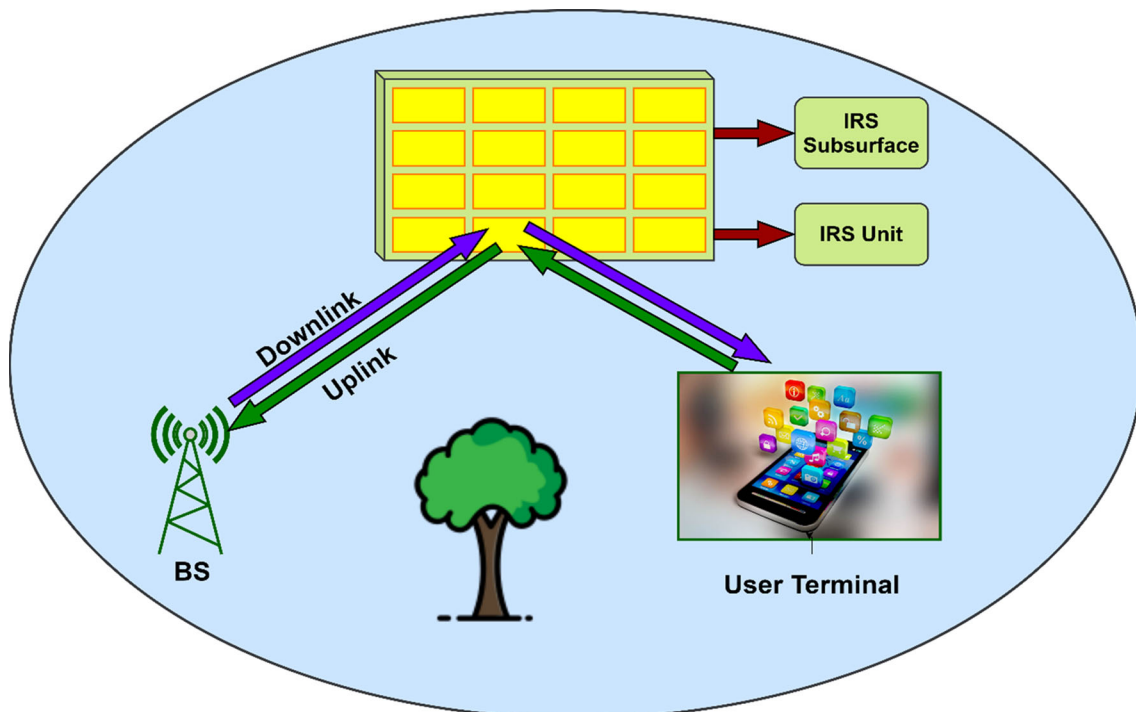


Fig. 3 Architecture of an IRS aided network

where  $\mathbf{h}_r^H = [h_{r,1}^*, h_{r,2}^*, \dots, h_{r,N}^*]$  and  $\mathbf{g} = [g_1, g_2, \dots, g_N]^T \cdot \Theta = \text{diag}(\beta_1 e^{j\theta_1}, \beta_2 e^{j\theta_2}, \dots, \beta_N e^{j\theta_N})$  is a diagonal matrix with  $\beta_n = [0, 1]$  is reflection coefficient and  $\theta = [0, 2\pi]$  is the phase shift.  $x(t)$  is the incident signal and  $N$  denotes the total number of reflecting elements at the IRS. The channel coefficient  $\mathbf{h}_r^H$  and  $\mathbf{g}$  generally depend on the distance-related path loss, large-scale fading, and small-scale fading. By intelligently adjusting these parameters IRS can reconfigure itself to achieve the desired performance.

## 2.2 Hardware architecture of IRS

A typical architecture of the IRS is given in Fig. 4. The first layer comprises of a huge number of tuneable metallic patches imprinted on the dielectric substrate to straightforwardly tune the occurring signal. The subsequent layer is a copper plate utilized to limit the signal energy leakage during the IRS reflections. The third layer is a control circuit board to invigorate the reflecting components just as tune their reflection amplitudes or phase shifts continuously. IRS controller works as a doorway to connect with other network components such as BS or user terminals. In IRS, the reflection coefficient of each element must be tuneable to cater the dynamic wireless perturbations occurring due to user mobility thus requiring reconfigurability. This can be obtained by leveraging electronic devices such as PIN diode, FETs, and MEMS switches. In Fig. 4 example of IRS using PIN diode is shown. By adjusting DC bias voltage ON and OFF mode of the PIN is controlled thus by introducing phase shift of 0 or  $\pi$  [40]. Few significant implementations of IRS in the wireless network is provided in the literature, which includes PLS design [41–44], performance of MIMO networks integrated with IRS [45, 46], performance analysis of fading channels [47, 48], and simultaneous information and power transfer [49, 50]. On comparing the IRS-aided network with other related technologies IRS proves to be more competent. First, compared to active wireless relay-assisted communication [51, 52], during the amplification of the received signal, the noise signal is also amplified which consequently degrades the performance. Also, using relay in HD mode reduces spectral efficiency whereas FD mode needs an effective SIC technique which increases its cost and complexity [53]. A comprehensive comparison of IRS technology with technologies such as backscatter communication and MIMO is studied in [8, 9, 22].

## 3 NOMA- For 5G and beyond networks

Radio spectrum scarcity is one of the major challenges that next-generation and beyond networks have to deal with. MA

has a huge impact on resource utilization, as it allows several users to share the same resource. In the cellular networks, past generations have been implementing several MA methods such as TDMA, FDMA, CDMA, and OFDMA [54]. These techniques are referred as OMA technique because even though sharing the same resource using different approaches, provides orthogonal access to the users. In TDMA all the users are provided a dedicated timeslot in the same frequency band whereas in FDMA, the available bandwidth is divided into frequency sub-slots and is provided to the user. In CDMA, the receiver is assigned a distinct spreading code and data is retrieved by multiplexing with the same pseudo-random sequence at the receiver. In OFDMA, subsets of subcarriers are assigned to the individual users. From the implementation perspective, the evolution of the wireless network generations revolves around these MA schemes. Precisely, FDMA, TDMA, CDMA, and OFDMA are being utilized in the 1G, 2G, 3G, and 4G systems respectively [55]. As these OMA techniques involve orthogonal resource allocation there is no probability of interference among the users so typically low-complexity receivers can be implemented. However, the limitation of employing OMA is that it has severe limitations in terms of supporting a high number of users. This provides a significant challenge to the 5G systems which are supposed to support massive device connectivity with diverse data rates and latency requirements. The creative concept of NOMA is admitted as the newest member in the MA family, in order to meet the stringent demands of 5G and beyond systems [56]. NOMA provides the ability for more users to share the same resource, be it be a time slot in TDMA, a frequency slot in FDMA, or a spreading code in CDMA [57]. In the last few years, NOMA has received considerable attention as a candidate MA technique for LTE, 5G, and beyond 5G systems. With NOMA, multiple UEs are co-scheduled and share the same radio resources in time, frequency, and/or code. Nevertheless, due to heterogeneous latency constraints and channel conditions at receivers, the conventional SIC used in NOMA transmission may be infeasible when implemented in URLLC applications. In [58], to cope with the latency constraints and to explicitly specify the trade-off between latency and reliability the normal approximation of the capacity of finite blocklength codes is adopted to investigate the transmission energy minimization problem for two user links. Due to the heterogeneous latency, the pure NOMA scheme may not achieve the best transmission energy as the traditional homogeneous setting thus a hybrid scheme which can include the TDMA and NOMA as special cases is presented. Similarly in [59], the finite blocklength code capacity formula is adopted to characterize the latency of users where user 1 is a latency-tolerant user which has one packet with a long latency requirement, while user 2 is a latency-sensitive user which has multiple packets with short latency requirements. Then, a hybrid NOMA

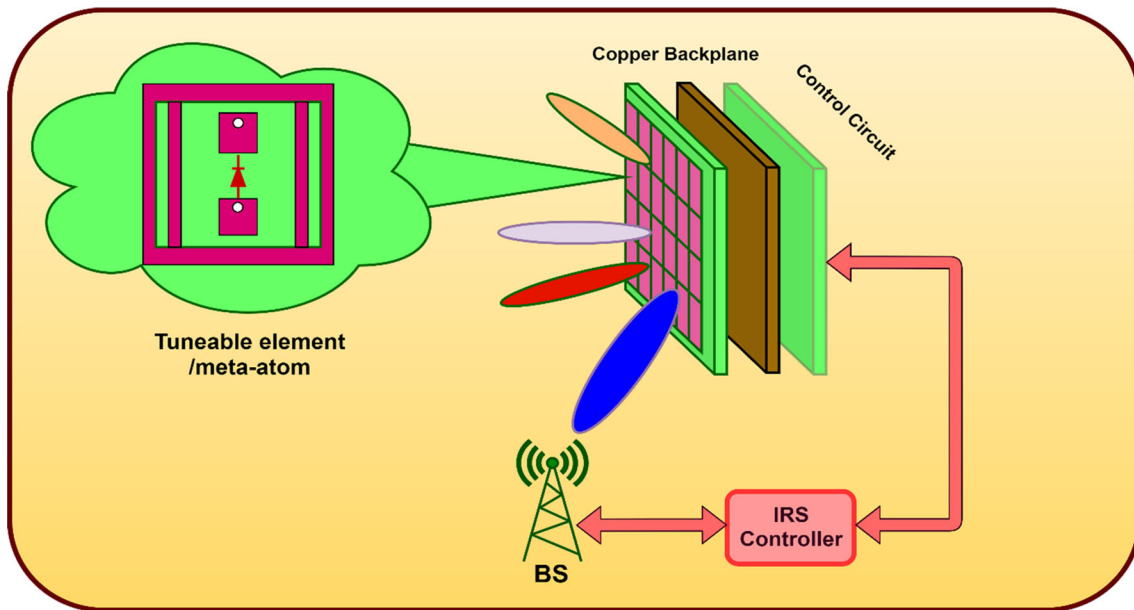


Fig. 4 Hardware architecture representation of IRS

transmission scheme has been proposed to handle the heterogeneous requirements of users. An adaptive power allocation scheme to minimize the average transmit power of a HARQ-CC enabled NOMA system under strict outage constraints of users is proposed in [60]. Considering practical applications, the partial HARQ-CC enabled transmissions in the multi-user scenario are investigated and the power allocation problem for each user pair is solved by the proposed SCA-based algorithm. A novel cooperative SWIPT aided NOMA transmission strategy for MISO and SISO case is proposed in [61]. By jointly optimizing the power splitting ratio and the beamforming vectors, authors have aimed at maximizing the data rate of the “strong user” while satisfying the QoS requirement of the “weak user”. An iterative algorithm based on successive convex approximation is presented to solve the optimization problem. The existing NOMA technique is classified into major two domains: power domain and code domain. In power-domain multiplexing, according to the channel conditions, different users are allocated different power coefficients whereas, in code-domain multiplexing, different users are multiplexed over the same resources and are allocated different codes [62, 63]. In spite of the fact that code-domain multiplexing can possibly improve spectral efficiency, it requires a high data transfer capacity and isn’t effectively appropriate to the current frameworks. Power-domain multiplexing, on the other hand, is straightforward to deploy and no major changes to the existing networks are required. Furthermore, it does not necessitate additional BW to improvise spectral efficiency [64]. In the next section, we focus on the power domain multiplexing to understand the

NOMA and OMA techniques and point out the key differences.

### 3.1 Power domain downlink NOMA

A unique power level is available to the user and multiple users for transmitting their signals while sharing the same time/frequency/code resources, each using their own allocated power. The power levels are decided depending upon the channel conditions. To understand it better let us consider the case of a DL NOMA [65] as shown in Fig. 5a consisting of a BS connecting two users. Suppose,  $s_1$  and  $s_2$  signals are transmitted from BS to user 1 and user 2 with transmitting power  $P_{t1}$  and  $P_{t2}$  respectively. The total transmit power  $P_t$  is predefined for a specific system setting and is given as  $P_t = P_{t1} + P_{t2}$ . As per the adopted power allocation scheme  $P_t$  is divided as  $P_{t1}$  and  $P_{t2}$  for user 1 and user 2 respectively. Based on the CSI feedback received from the users, BS performs SIC ordering on a regular basis and updates this information to users at receiving end. Without loss of consensus, it very well may be expressed that a user with a weaker channel strength is given higher power than a user with a stronger channel strength to increase the throughput. Here in this condition,  $|h_1|^2/N_{f,1} > |h_2|^2/N_{f,2}$ , hence  $P_{t1} < P_{t2}$ , and only user 1 which has higher channel gain performs SIC. User 1 decodes user 2’s signal, and subtracts it from the received signal, before decoding its own signal. User 2 treats user 1’s signal as noise and thus directly decodes its own signal without SIC. For perfect SIC, the obtainable data rate of the NOMA users, transmitted over a bandwidth  $W$  is given

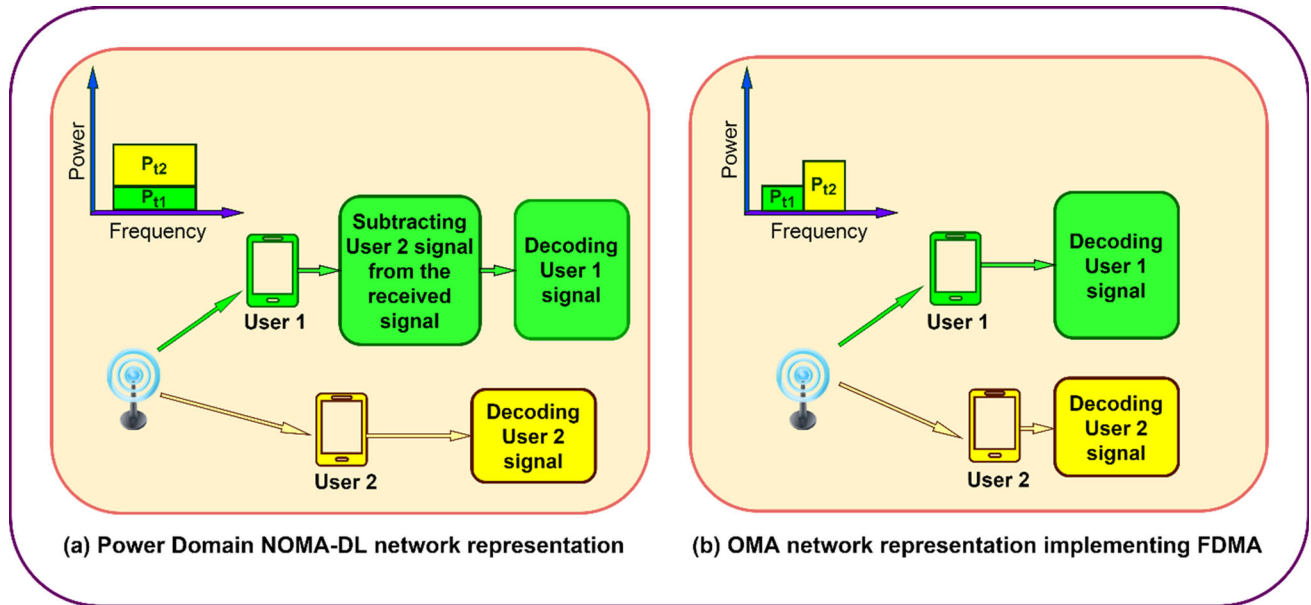


Fig. 5 Representation of NOMA and OMA technique

as [66]

$$R_1 = W \log_2 \left( 1 + \frac{p_{t1}|h_1|^2}{N_{f,1}} \right) \quad (2)$$

$$R_2 = W \log_2 \left( 1 + \frac{p_{t2}|h_2|^2}{p_{t1}|h_1|^2 + N_{f,2}} \right) \quad (3)$$

The achievable channel capacity is given as  $R = R_1 + R_2$ . From the Eqs. (2) and (3) we get that by adjusting the power allocation coefficients  $\alpha_1$  and  $\alpha_2$ , with  $\alpha_1 = P_{t1}/P_t$  and  $\alpha_2 = P_{t2}/P_t$ , BS can control the data rate of each user. Let us consider the same case implementing FDMA to get a comparative understanding of the NOMA transmission. Figure 5b, where the transmission bandwidth  $W$  is equally divided among the user 1 user 2 as  $W/2$  Hz and the power allocation ratio remains the same as of the NOMA technique. Then, the achievable channel capacity of the OMA user  $i$ , can be written as

$$R_1 = \frac{W}{2} \log_2 \left( 1 + \frac{p_{t1}|h_1|^2}{N_{f,1}} \right) \quad (4)$$

$$R_2 = \frac{W}{2} \log_2 \left( 1 + \frac{p_{t2}|h_2|^2}{N_{f,2}} \right) \quad (5)$$

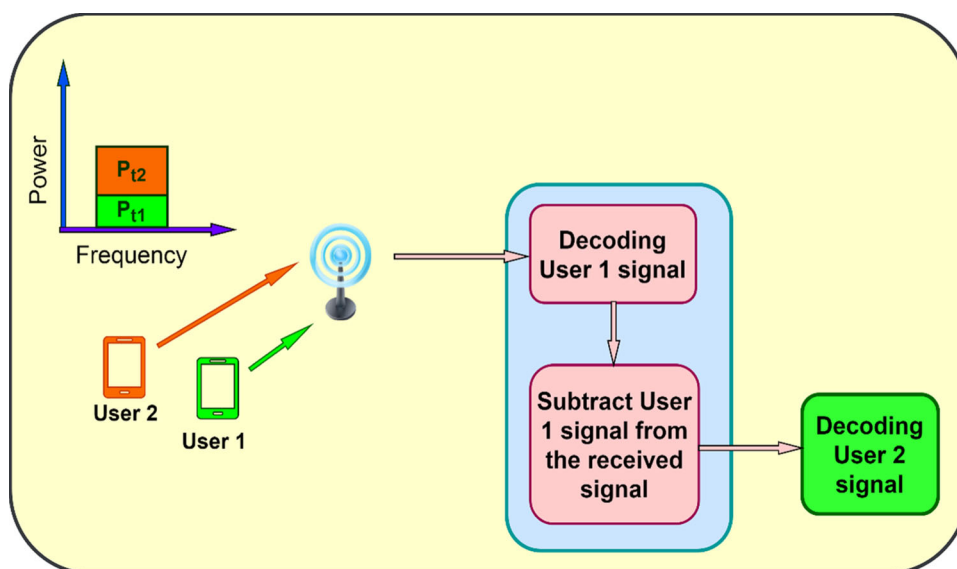
The achievable sum capacity is  $R_{\text{OMA}} = R_1 + R_2$ . From Eqs. (4) and (5) it is clear that, unlike NOMA, OMA does not suffer from interference from another user. However, in terms of the BW, in NOMA both the users take the advantage of full BW. As a result, regardless of the BW allocation

ratio, NOMA greatly surpasses OMA in terms of sum capacity. Therefore, when channel conditions change across users, NOMA in cellular DL delivers high sum capacity and performs a better balance between network efficiency and user fairness than OMA [67].

### 3.2 Power domain uplink NOMA

Multiple users broadcast their own UL signals in the same resource block to the BS in UL-NOMA transmission [68, 69], as shown in Fig. 6. Here users 1 and 2 simultaneously transmit their signals to the BS. User 1 is again marked as the strong user experiencing a higher channel gain compared to user 2, the weak user. Upon receiving the superposed signal, the BS performs SIC to separate the users' signals. The BS initially decodes the user 1's signal by treating user 2's signal as noise, then subtracts user 1's detected signal from the received one. The signal of user 2 is then detected from the remainder. In contrast to DL-NOMA, in UL-NOMA user 1 receives interference from user 2, whereas user 2 is free from the interference from user 1 as user 1's signal is eliminated prior to decoding user 2's signal [57]. This is the key difference between UL and DL NOMA transmission. Also, in terms of the transmit power, in UL-NOMA the transmit power of the users depends upon the conditions of the channel of each user. In contrast to DL-NOMA transmission, if the user's channel conditions are significantly varied, their received SINR can be quite different at the BS independent of broadcast power. SIC [70] is frequently used in wireless networks to significantly increase the network

**Figure. 6** Power domain NOMA-UL network representations



capacity through effective interference management. By initiating the following technique, it empowers the user with the strongest signal to be identified first, having the least interference-contaminated signal. The signal of the strongest user re-encodes and re-modulates and, then subtracted from the composite signal. The 2nd strongest signal, which has now become the strongest signal, follows the same sequence. At last, the most vulnerable user decodes its information without being suffered from any interference in anyway. A deeper understanding of the concept can be achieved by referring to the literature [71] and [72] where SIC is implemented in OFDM and NOMA networks respectively.

### 3.3 Advantages and challenges of NOMA over OMA

The key advantages of the NOMA technique can be summarized as follows:

- (i) *High bandwidth efficiency* As NOMA allows multiple users to exploit the same resource block, makes it extremely spectrum-efficient, resulting in enhanced system throughput [73, 74].
- (ii) *Massive connectivity* By exploiting non-orthogonal characteristics, NOMA fits itself to support the connectivity of billions of smart devices, which 5G and beyond networks are expected to support.
- (iii) *Improved user fairness* NOMA's power allocation allows a system to strike a balance between fairness among the users [75]. As a result, if a proper power allocation is adopted, the cell-edge users can benefit from increased data rates while still preserving the system throughput. Allocation of additional power to the weak users is a critical element of NOMA. In [76] authors

provide a comparison of fairness in resource allocation in both NOMA and OMA and conclude that NOMA outperforms the OMA.

- (iv) *Low latency* The access-grant procedure required in OMA leads to the latency in the networks specifically when it is supporting massive connectivity. UL-NOMA, on the other hand, can achieve grant-free MA by blindly recognizing active users and decoding their data streams [77]. As a result, the grant-free NOMA has lower latency. Furthermore, numerous users are serviced simultaneously in NOMA, a reduction in latency can be achieved.
- (v) *Compatibility* The astounding compatibility with other communication technologies permits NOMA to be promptly integrated with the current and future wireless systems. NOMA, for instance, has been shown to be compatible with the conventional OMA techniques, like TDMA and OFDMA.

Although the NOMA technique proves its brilliance in captivating the wants of 6G and beyond networks, still to get full benefit some limitations and implementation concerns need to be solved. When compared to OMA, NOMA requires additional receiver complexity and energy usages as each user must decode the information of all other users with lower channel gains. When a user makes a SIC error, the following decoding of all other users' information is likely to be incorrect. This implies that to reduce error propagation number of users in each cluster must be less in NOMA which in turn reduces the sum-rate gain. In order to achieve desired functionalities of the power domain concept in NOMA at the receiver, channel gain difference between users should be adequate. This limits the effective number of user pairs



**Table 4** Comparison of NOMA and OMA techniques

Specification	NOMA	OMA
Receiver Complexity	High	Low
Energy Consumption	High	Low
User Connectivity	High	Low
System Throughput	Larger	Smaller

served by clusters. However, NOMA provides the best system throughput ensuring user fairness. On review of the literature such as [78–80] which includes a comprehensive review of the NOMA and OMA techniques, a comparison of these techniques is summarized in Table 4.

### 3.4 Existing NOMA solutions over the performance of fading channels

Ever since the potentials of the NOMA concept was discovered in [56], researchers over the globe have begun to look at the ways to turn the NOMA idea into a next-generation radio access strategy. In wireless communication, it is necessary to consider channel fading in order to verify the performance properly. The performance evaluation of the NOMA systems and comparison with OMA over the several fading channels is widely explored in the literature. Evaluation of NOMA-DL system performance describing power allocation techniques and user pairing to ensure user fairness for a Rayleigh fading channel is provided in [75] and [81] respectively. Outage analysis and power control technique for a NOMA-UL system for Rayleigh fading channel links is reported in [69] and [82] with instantaneous CSI-based user ordering. A multiple user grouping approach is proposed in [83] to enhance the sum throughput in both UL and DL NOMA systems. The implementation of NOMA in cooperative relaying over Rayleigh, Nakagami- $m$ , and  $\alpha$ - $\mu$  fading channels is analyzed in [84, 85] and [86] respectively. To study the influence of composite fading impairments in a DL-NOMA system, in [87] authors have derived the expressions of OP and average BER for shadowed  $\kappa$ - $\mu$  fading links. The effective rate analysis for NOMA-DL with shadowed  $\kappa$ - $\mu$  and  $\alpha$ - $\mu$  distributed fading channels considering instantaneous CSI based user ordering is presented in [88] and [89], respectively. For the downlink NOMA system, the outage performance and sum rate performance are analysed over the  $\alpha$ - $\eta$ - $\mu$  fading channel in [90]. A general expression for the OP in terms of Gaussians hypergeometric function in a NOMA-DL system where users are randomly deployed is obtained in [91].

The performance of a cooperative relaying NOMA network over Fisher-Snedecore  $F$  fading channel and performance comparison with OMA is performed in [92]. In [93], authors analyse OP for the fading links undergoing

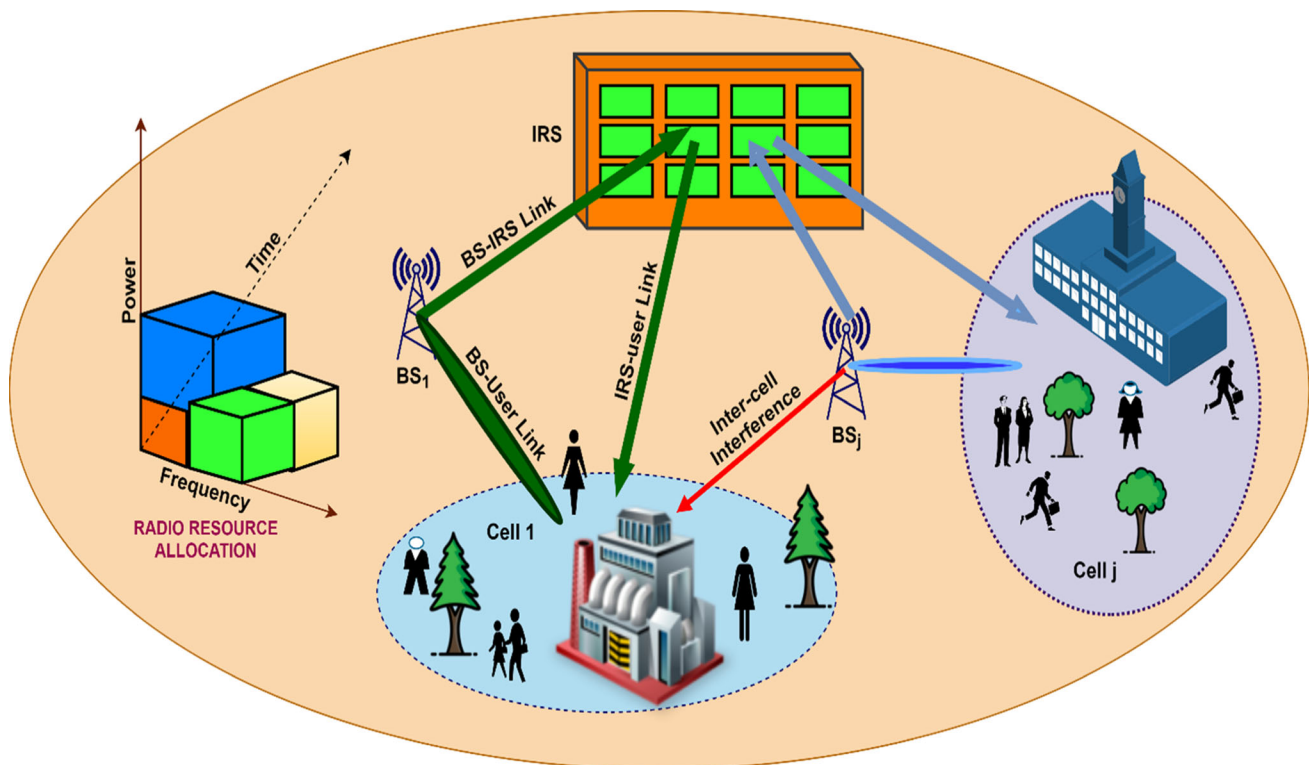
Rayleigh, Rician, Nakagami- $m/q$ /lognormal,  $\kappa$ - $\mu$ , and  $\eta$ - $\mu$ , fading distributions considering multiple user power-domain NOMA-DL and UL communication systems. BER of the Rician and Nakagami- $m$  fading channels implemented in the power domain NOMA network is analysed in [94] and [95] respectively. Table 5 provides the summary of a few significant works related to the performance analysis of NOMA networks.

## 4 Integrated IRS–NOMA communications

From the discussions of the preceding sections, it can be stated that the brilliant capabilities of IRS and NOMA technology make it a promising solution to counter the challenges of next-generation networks. So, this conjecture can be drawn that integration of IRS with the NOMA network can prove a highly captivating technique to suffice the data rate demands of the 6G networks. Integrating NOMA with IRS proposes a new paradigm to provide a potential MA solution to tackle the rigorous requirements on latency, massive connectivity, and data rates for the future 6G networks. For instance, it can be understood as if any networks system implements NOMA without IRS, the performance gain of NOMA will critically depend on the two users' channels  $h_1$  and  $h_2$  respectively. Conventionally, the users' channel conditions are viewed as being fixed to be solely determined by the users' propagation environment. However, the use of IRSs opens up opportunities for intelligently reconfiguring the users' propagation environment in order to facilitate the application of NOMA, which can lead to significant performance gains of NOMA. Specifically, when the channel gain of both the users are same, implementing IRS with NOMA introduces more degree of freedom for system design. As a crucial component of NOMA systems, SIC is key to combating this MA interference. In particular, selecting the SIC decoding order is based on the users' CSI and the users' QoS requirements. IRS increases the design flexibility of NOMA networks, leading to the transition from 'channel condition-based NOMA' to 'QoS-based NOMA'. The use of IRS can reduce the constraints in multiple antenna NOMA networks. In conventional MIMO–NOMA networks, some constraints on the numbers of antennas at the transmitters and receivers may need to be satisfied [96]. With the aid of IRS, such constraints can be relaxed due to the additional passive array gains provided by the IRS. With these potential benefits introduced by the integration of NOMA and IRS, researchers have begun to investigate IRS–NOMA networks. The succeeding section contains a conceptual understanding of the IRS-aided NOMA networks, performance metrics, and their implementation in the existing networks.

**Table 5** Review of the literature over the performance of NOMA networks

Literature	Year	Type of NOMA link	Fading channel	Contribution
[75, 81]	2015, 2016	DL	Rayleigh	Analysis of power allocation strategies and impact of user pairing in the network
[69]	2016	UL	Rayleigh	OP and sum rate analysis
[83]	2016	DL and UL	Rayleigh	Power allocation analysis
[82]	2018	UL	Rayleigh	Outage analysis and power allocation with statistical CSI ordering at BS
[84, 85]	2016, 2020	DL	Nakagami- $m$ , $\alpha$ - $\mu$	Performance of cooperative relaying in NOMA network
[86]	2020	DL	$\eta$ - $\mu$ and $\kappa$ - $\mu$	OP and sum rate analysis
[87]	2019	DL	shadowed $\kappa$ - $\mu$	OP, BER, and average achievable rate
[88, 89]	2020	DL	shadowed $\kappa$ - $\mu$ and $\alpha$ - $\mu$	Effective rate analysis
[90]	2020	DL	$\alpha$ - $\eta$ - $\mu$	OP and sum rate analysis
[92]	2020	DL	Fisher-Snedecore	Ergodic capacity analysis and asymptotic analysis
[93]	2020	DL and UL	Rayleigh, Rician, Nakagami- $m/q$ /lognormal, $\kappa$ - $\mu$ , $\eta$ - $\mu$ , Nakagami-lognormal	OP analysis
[94, 95]	2020, 2019	DL	Rician, Nakagami- $m$	BER analysis

**Fig. 7** Representation of IRS assisted NOMA network

#### 4.1 State-of-the-art and performance metrics

A typical diagram of the IRS–NOMA network is given in Fig. 7, deploying IRS for improving wireless service from single-antenna BS- $j$  to BS1 single antenna cellular users. However, to better understand the IRS–NOMA concept, a scenario of DL-NOMA is considered in which the BS simultaneously transmits the superimposed signal to the  $k$  number of users via IRS assuming that the direct link between users and BS is highly attenuated. Such scenarios particularly exist in the urban environment where the communication between BS and the user is blocked by high-rise infrastructures [18, 6]. Let us consider  $k \in [1, 2, \dots, K]$  set of users. The IRS is having  $m$  reflecting surfaces denoted by  $m \in [1, 2, \dots, M]$ . The received signal at  $k^{th}$  user is given as [6]

$$Y_k = \mathbf{h}_{sr}^H \Theta \mathbf{h}_{rk} \sum_{i=1}^K \sqrt{a_i P_s} x_i + n_k \quad (6)$$

where  $P_s$  denotes the normalized transmission power at the BS.  $\mathbf{h}_{sr} \in \mathbb{C}^{M \times 1}$  and  $\mathbf{h}_{rk} \in \mathbb{C}^{M \times 1}$  are the complex channel coefficient between the BS and IRS, and IRS and user, respectively.  $a_i$  is the power allocation factor for the  $i$ -th user satisfying the condition  $a_1 \geq a_2 \geq \dots \geq a_K$ . Without the loss of generality, we consider the effective cascaded channel gain from the BS to IRS and then to user as  $|\mathbf{h}_{sr}^H \Theta \mathbf{h}_{r1}|^2 \leq |\mathbf{h}_{sr}^H \Theta \mathbf{h}_{r2}|^2 \leq \dots \leq |\mathbf{h}_{sr}^H \Theta \mathbf{h}_{rk}|^2$  [97]. The wireless link is perturbed by the additive white Gaussian noise which is represented by  $n_k$ . As per the NOMA principle, the received SINR at the  $k$ -th user to detect  $n$ -th user's information ( $k \geq n$ ) is given as [6]

$$\text{SINR}_{k \rightarrow n} = \frac{\rho |\mathbf{h}_{sr}^H \Theta \mathbf{h}_{rk}|^2 a_n}{\rho |\mathbf{h}_{sr}^H \Theta \mathbf{h}_{rk}|^2 \sum_{i=n+1}^M a_i + \bar{\omega} \rho |h_I|^2 + 1} \quad (7)$$

where  $h_I$  is complex channel co-efficient,  $\rho$  is the transmit SNR, and  $\bar{\omega} = 1$  for imperfect SIC and 0 for perfect SIC. Using SIC, cancelling all  $k-1$  users signal, received SINR at  $k$ -th user is

$$\text{SINR}_k = \frac{\rho |\mathbf{h}_{sr}^H \Theta \mathbf{h}_{rk}|^2 a_M}{\bar{\omega} \rho |h_I|^2 + 1} \quad (8)$$

The performance metrics primarily include the probabilistic metrics for characterizing the uncertainty in wireless transmissions in individual transceivers, and the ergodic metrics for characterizing the network performance, taking into account the fluctuations in network topology, the distribution of reflectors, etc. Typical performance measures proposed for IRS–NOMA wireless systems include reflection probability [98], coverage or OP [99], bit error probability [100, 101],

and ergodic capacity. A systematic review of these performance metrics are provided in [13]. In [6] using 1-bit coding, exact expressions of OP with perfect and imperfect SIC in IRS–NOMA networks is derived proving that increase in the number of reflecting elements, decreases the OP. Also, it is shown that IRS NOMA outperforms the IRS OMA. The study also shows that deployment of IRS near to BS decreases the OP. For IRS–NOMA and OMA networks [102] analyses the IRS reflection with discrete phase shifts. To reduce the transmitting power and satisfy the data rate requirement of each user, phase shift matrix and beamforming vectors are jointly optimized in [103]. To ensure user fairness and optimize performance rate, a problem is formulated based on user ordering scheme with link experiencing the Rician fading in [8]. In [104], authors have proposed a mathematical framework to find the error rate of an IRS–NOMA system by deriving an accurate expression of pairwise error probability of NOMA users and in [105] BER performance is analysed. For multi-antenna scenarios, the sum rate of IRS–NOMA is improved considering fixed reflecting elements in [106]. Secrecy analysis over the IRS-assisted NOMA network is provided in [107]. An analytical expression for secrecy OP is derived to understand the performance of the system. Also, to get the clarity of the implications of NOMA, asymptotic expressions are derived and the performance is compared with the OMA.

The authors of [108] investigated the influence of coherent and random phase shifting on outage behaviours for IRS–NOMA networks as a potential improvement. Using central limit theorem-based approximation an approximated expression for the OP of the system undergoing Rayleigh fading is derived. As this approximation is not suitable at high SNR values an upper-bounded expression in terms of the incomplete gamma function is also derived. In [109], the OP and ergodic rate of nearby users for IRS–NOMA are examined by developing IRS's the passive beamforming. Authors in [110] investigates new channel statistics of the BS-IRS-user device connection with Nakagami- $m$  fading considering UL and DL scenarios to characterize the system performance. Using moments matching, authors derive a tractable expression of the outage performance under interference cancellation in [111], experiencing Nakagami- $m$  fading. Authors in [112] investigate the secrecy performance of the DL-IRS–NOMA by deriving closed-form expressions for secrecy OP and the ASC in the case of Nakagami- $m$  fading. For UL communication the sum rate maximization problem is solved by jointly optimizing the power at the user and passive beamforming in the IRS–NOMA network in [113]. The efficacy of implementing IRS in DL-NOMA transmission with two users over fading channels is performed in [114]. A summary of the above review is presented in Table 6.



**Table 6** Review of the performance of IRS aided NOMA networks

Year	References	Parameter	Characterization	Fading channel	Type of NOMA Link
2021	[62]	OP and ergodic rate	Exact and asymptotic expressions	Rayleigh	DL
2020	[98]	OP	Closed-form expression	Rayleigh	DL
2021	[104]	Pairwise error probability	Closed-form expression		DL
2021	[105]	Bit error rate	Closed-form expression	Rayleigh	DL
2019	[106]	Secrecy OP	Analytical and asymptotic expressions	Rayleigh	DL
2020	[108]	OP	Approximated and upper bound expression	Rayleigh	DL
2019	[109]	OP and ergodic rate	Closed-form expression	Rayleigh	DL
2020	[110]	OP and ergodic rate	Closed-form expression	Nakagami- <i>m</i>	DL and UL
2020	[111]	OP	Closed-form expression	Nakagami- <i>m</i>	DL and UL
2020	[112]	SOP, ASC	Closed-form expression	Nakagami- <i>m</i>	DL

## 4.2 Performance optimization in IRS–NOMA

Some emerging applications, of the next-generation networks such as the Internet of Things, requires all the connected device to experience similar data rates. To balance the performance of all the devices of the network, the minimum achievable data rate can be maximized however, this brings a deterioration to the devices having good channel conditions. Furthermore, if the weak device's channel conditions are excessively degraded, the data rates obtained under fair power allocation may not be sufficient to meet the QoS requirements of other devices, resulting in poor network performance. By implementing IRS, the channel gains can be boosted to guarantee high network performances. Theoretically, it can be stated that the quality of NOMA to provide improved user fairness and IRS to improve channel gain can be jointly used to provide all the devices reach the same data rate. The increasing demands for larger data rates and faster-speed wireless communications have prompted severe concerns on power consumption, EE, and the coverage range of the future networks. In this section, we deal with the optimization of the performance enhancements provided by the IRS–NOMA networks such as resource allocation, enhanced coverage, high energy efficiency, and high-power efficiency.

### 4.2.1 Resource allocation

The concept of joint power allocation and phase shifts optimization problem for the IRS–NOMA was first presented in [8]. Again in [115], this concept is well explained, where authors propose IRS–NOMA-DL framework and deals with throughput maximization issue considering SIC decoding conditions and IRS reflection coefficients imperatives are planned by together optimizing the channel assignment, reflection coefficients, power allocation, and decoding order.

In [116], an effective SOCP- ADMM based algorithm is proposed assuming perfect SIC for the MISO IRS–NOMA system. In [117], an efficient user ordering scheme having minimal complexity is proposed by considering the phase shifts and transmission rates. In [118], the power-efficient MISO IRS–NOMA system is explored where beamforming vectors and IRS phase shift matrix are jointly optimized using SDR and alternating optimization algorithm. By using the Lagrange duality approach authors in [119] examine the ergodic and delay-restricted capacity for IRS-aided OMA and NOMA networks and optimize resource allocation and phase shift jointly.

### 4.2.2 Energy efficiency

This is a vital performance metric to balance the throughput and power consumption so finding an optimization strategy to achieve the maximal EE is of high importance in the IRS-based networks. Since IRS-based frameworks are equipped for changing the phase shift generated by each reflecting component to constructively concatenate the reflected signals, they can AF the signals without extra power amplifiers. [6] proposes an optimized EE technique by optimizing phase shifts and transmit power while meeting QoS constraints. The EE maximization for the MISO system taking into account, the individual data rate constraints and the total transmit power budget is formulated as [120]

$$\Theta_{w_1 w_2}^{max} \frac{R}{\frac{1}{\eta} \sum_{k=1}^2 ||w_k||^2 + P_c} \quad (9)$$

Such that

$$R_k \geq R_{k,min}, k = 1, 2 \quad (10)$$

$$\sum_{k=1}^2 \|\mathbf{w}_k\|^2 \leq P_{\max} \quad (11)$$

$$0 \leq \theta_n \leq 2\pi, n = 1, \dots, N. \quad (12)$$

where  $R = R1 + R2$ , and  $R_{kmin} = \log_2(1 + \Gamma_{k,min})$  is the minimum data rate requirement for  $k$ -th user  $U_k$ , and where  $\Gamma_{k,min} = 2^{R_{k,min}} - 1$  is the known least SINR for  $U_k$ . Equations (10), (11), and (12) are the specific conditions ensuring the QoS need for each user, limit the transmit power to  $P_{\max}$  and specify the phase shift range respectively with  $\mathbf{w}_k$  being the beamforming vector for the  $k$ -th user. For this formulation, authors have considered MISO network with BS having power amplifier efficiency as  $\eta \in [0, 1]$  with total circuit power at the BS as  $P_c = P_d + P_0$  and where  $P_d$  and  $P_0$  stands for the dynamic and static power consumption respectively. To address this a proficient alternating optimization-based algorithm is presented in [120].

#### 4.2.3 Power efficiency

In addition to EE, power optimization frameworks are prerequisites for the efficient utilization of future wireless networks. By jointly optimizing the transmit beamforming vectors  $\mathbf{w}_{k,i}$  at the BS and reflection coefficient vector  $\varnothing$  at the IRS, subject to the transmission rate requirements at 2  $K$  users in an IRS aided NOMA multi-cluster MISO system, the total minimize power optimization problem is formulated as [116]

$$\varnothing_{\mathbf{w}_{k,i}}^{\min} \sum_{k=1}^K \left( \|\mathbf{w}_{k,c}\|^2 + \|\mathbf{w}_{k,e}\|^2 \right) \quad (13)$$

$$\text{Such that } \log_2(1 + \gamma_{k,c}) \geq r_{k,c} \quad \forall k \in K \quad \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (14)$$

$$\log_2(1 + \min(\gamma_{k,e}, \gamma_{k,c})) \geq r_{k,e} \quad \forall k \in K \quad \varnothing \in \mathcal{F} \quad (15)$$

where  $r_{k,c}$  and  $r_{k,e}$  denote the target data rates of central user  $U_{k,c}$  and cell-edge user  $U_{k,e}$  in the  $k$ -th cluster, respectively.  $\mathcal{F}$  represents the feasible set of the reflection coefficient vector [34]. Authors in [117] and [118] deal with power optimization problems and introduce effective techniques.

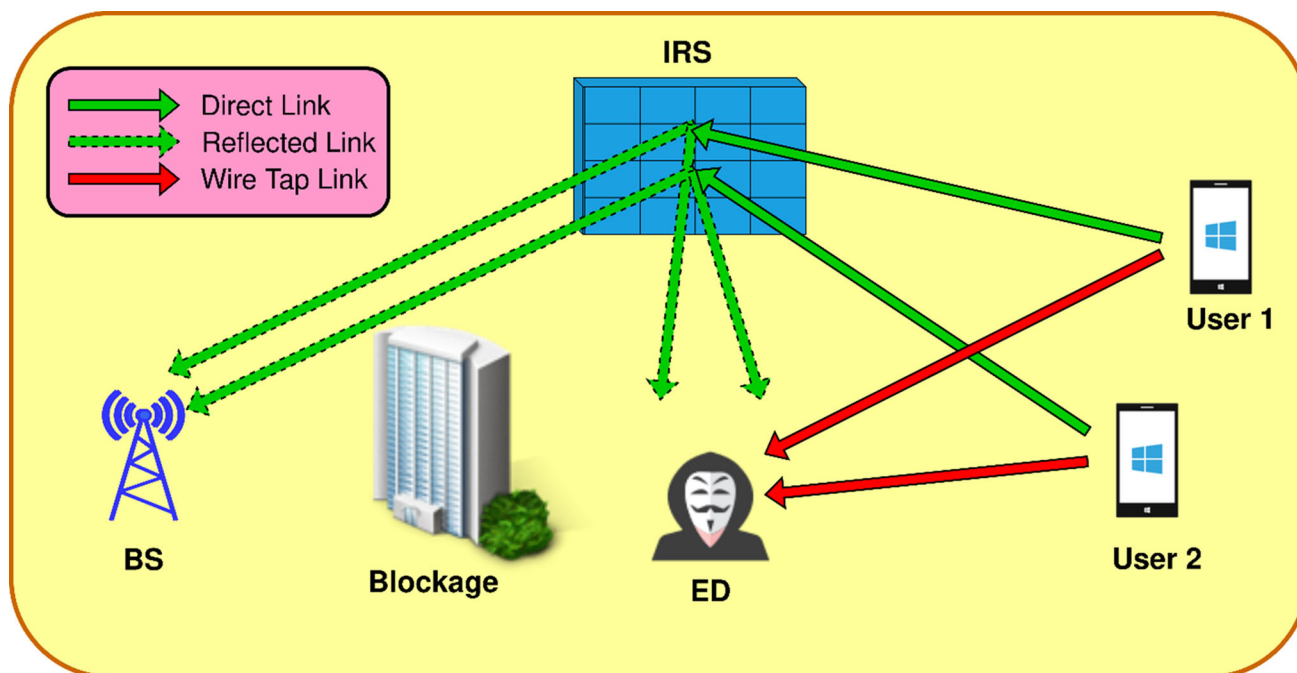
#### 4.2.4 Coverage optimization

It is difficult to ensure uniform signal coverage to all existing devices in conventional networks. This issue becomes more pressing at the higher frequency bands of 5G, 6G, and beyond. The mmWave and terahertz spectrum's short wavelengths can resonate with atmospheric oxygen and

water molecules, causing a considerable portion of the radiated energy to be wasted through kinetic absorption. This harmful characteristic can lead to strong signal attenuation consequently decreasing the communication range. Implementation of IRS at the cell edge can increase the transmission and NOMA can provide multiple user connectivity, proving itself to be an efficient technique to enhance the coverage range. The study in [121] proposes a projected gradient ascent algorithm to tackle the optimization challenges and establish the IRS phases for maximizing the minimal user SINR under optimal linear pre-coder.

#### 4.3 Physical layer security in IRS-NOMA

Because of the potentials of intelligently reconfiguring the wireless propagation environment, IRS has also been applied in NOMA networks to improve wireless communication security and privacy. For the multiuser NOMA scenario, where the quality of the channels from a transmitter to some of the legitimate users is worse than that of the wiretap channel from the transmitter to the ED, it is difficult to guarantee a positive secrecy rate [122]. This thus motivates the application of IRSs to NOMA for physical layer security [123] enhancement. A typical representation of the IRS-NOMA secure communication is provided in Fig. 8. From a theoretic point of view, by implementing IRS in the proximity of the legitimate NOMA users or the EDs and by efficiently designing the passive beamforming, the signals received from reflected and direct paths can be constructively added at the user and destructively at the ED [124, 42]. So by degrading the ED signal reception capability and enhancing the legitimate signal reception capability positive secrecy rate is achieved in IRS-NOMA networks. To clarify this, an UL NOMA scenario is shown in Fig. 8. Two users transmit signals simultaneously to the BS via an IRS, where the IRS adjusts its phase to guarantee that the users' signal constellations fully overlap at the ED but do not non-overlap at the BS [125]. Thus, even if the ED may have a high receive SNR to decode the superimposed signal successfully, the individual signals have to be recovered by random guess, which results in a poor bit error rate performance and degrades the eavesdropping capability significantly. To ensure IRS-NOMA secure communication knowledge of ED's CSI is important for efficient beamforming design, which is practically not feasible to achieve. Authors in [126] investigate secure transmission in an IRS-NOMA network considering a practical eavesdropping scenario with imperfect CSI of the ED. A robust beamforming scheme using artificial noise to guarantee secure NOMA transmission with the IRS is proposed. The findings show that along with IRS use of artificial noise is helpful in degrading the ED's capability and improving the secrecy rate specifically when QoS of the user is a stringent requirement. Considering that only the statistical CSI of the



**Fig. 8** An IRS-aided NOMA secure communication

Eve is available, a joint transmit beamforming and reflection coefficients design problem to maximize the minimum secrecy rate among legitimate users is proposed in [127]. Secrecy outage probability metric is analyzed subject to the total transmit power constraint at the BS, the phase shifts constraints of IRSs and SIC decoding constraints. Authors in [128] investigate the secrecy performance of IRS-aided NOMA networks, where BS communicates with a pair of NOMA users in the presence of an ED and analyse secrecy outage probability and the ASC. Moreover, in [129] also secure transmission problem of IRS-assisted NOMA networks, which demonstrates the great potential of IRS in security enhancement of NOMA communication is studied.

#### 4.4 Channel estimation in IRS–NOMA

As NOMA uses SIC to detect the signals, it is of utmost importance to have the CSI. Most of the NOMA-related applications have been assumed to have perfect CSI so that SIC can be easily implemented to detect the signal at the receiver. Few work focusing on the NOMA application with imperfect CSI can be found in the literature [130–132]. Similarly, channel estimation in IRS–NOMA systems is one of the most critical challenges due to a large number of IRS reflecting unit elements, their distinctive hardware constraints, and the multiuser nature of NOMA. Prior works on IRS applications mainly focus on the design of reflection coefficients under the assumption of perfect CSI, which facilitates in deriving the system performance upper bound but the actual estimation of the CSI is difficult to realize in practice. It

is worth noting that such design is practically challenging due to the lack of transmitting/receiving as well as signal processing capabilities of the passive IRS elements while their numbers can be practically very large, which thus calls for innovative solutions to tackle these new challenges. In the literature, there are two kinds of RIS channel estimation schemes semi-Passive and full passive depending on whether the IRS is integrated with sensing devices or not are proposed in [11]. In semi-passive sensing, devices (e.g., low-cost sensors) are integrated into the reflecting elements of the IRS. Hence, by letting the BS and each user transmit pilot signals, the channels from the BS/users to the IRS can be estimated based on their received pilots. Several literatures are dedicated to develop efficient channel estimation algorithms and protocols. For instance, a novel reflection pattern at the IRS is designed in [133] for OFDM to aid the channel estimation at the access point based on the received pilot signals from the user. With the estimated CSI, the reflection coefficients are then optimized by a low-complexity algorithm based on the resolved strongest signal path in the time domain. Similarly, for an IRS-assisted multiuser communication system a novel three-phase pilot-based channel estimation framework using a small number of pilot symbols is proposed in [134]. Utilizing the correlations among the IRS reflected channels, the minimum pilot sequence length for perfect channel estimation without noise at the BS are characterized, and the linear minimum MSE channel estimators considering noise at the BS is derived. Exploiting the property of BS-RIS to be high dimensional but quasi-static, a dual-link pilot transmission scheme, where the BS transmits downlink pilots and receives

uplink pilots reflected by the RIS is proposed in [135]. Since the quasi-static BS-RIS channel is estimated less frequently than the mobile channel is, the average pilot overhead can be reduced from a long-term perspective. Channel coefficient estimation for IRS-assisted MISO system using deep learning is proposed in [136] which provides less complexity as compared to the linear estimation solutions. Unlike the LS estimates of the IRS-assisted channel vectors [137], authors in [138] exploit the prior knowledge of a large-scale fading statistics at the BS to derive the Bayesian minimum MSE channel estimates under a protocol in which the IRS applies a set of optimal phase shifts vectors over multiple channel estimation sub-phases. The resulting MSE is both analytically and numerically shown to be lower than that achieved by the LS estimates. In [139] authors have proposed PAR-ALLEL FACtor decomposition including an alternating least squares algorithm to iteratively estimate the channel between the base station and RIS, as well as the channels between RIS and users for the DL RIS-assisted MISO communication system. To fully reap the various performance gains brought by IRS-NOMA, efficient channel estimation schemes are required since both active transmit beamforming at the BS and passive reflection beamforming at the IRS rely on CSI. However, this is a tedious task, due to the fact that the IRS does not have any RF chains to enable pilot signal transmission for channel estimation. Also due to a large number of IRS elements and the multi-user nature of NOMA, a large number of IRS-related channel coefficients need to be estimated.

## 5 Performance comparison of IRS-NOMA with other related technologies

NOMA has been integrated with several other communication technologies to extract the maximum benefit from it. There are various pros and cons of each technique which needs to be highlighted. So, the following sub section provides a comparative review of the IRS-NOMA with the contemporary MIMO and the relay technique.

### 5.1 MIMO-NOMA

To ensure the massive connectivity, increased spectrum efficiency and low latency requirements of the 5G and beyond networks MIMO is a promising technology to accomplish these requirements. The key idea is to outfit BSs with multiple antenna arrays that can simultaneously serve numerous terminals in the same time/frequency resource. Integration of MIMO with NOMA will help to exploit the features of both the techniques to reach higher spectral gains which can outperform the conventional systems. The incorporation of MIMO nodes along with the NOMA has expectedly boosted the spectral efficiency [140]. MIMO-NOMA

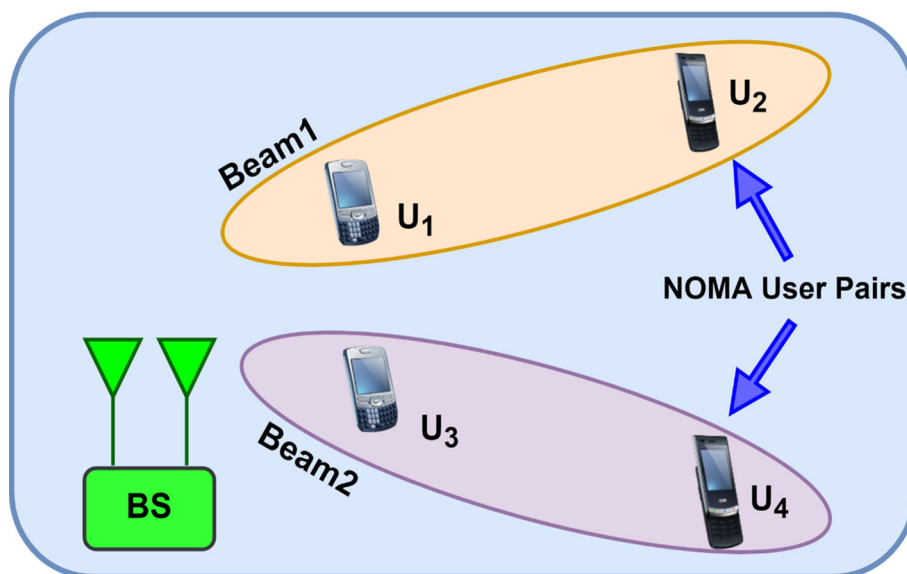
proves itself superior in terms of capacity as compared to MIMO-OMA as well [141]. In [142], the authors have proposed a general framework for MIMO-NOMA system applicable to both DL and UL transmission by applying the concept of signal alignment. This analysis concludes by proving that the performance gain of the system is enhanced using MIMO-NOMA. The design of the MIMO-NOMA is quite challenging, as performance evaluation of the MIMO networks is a tedious process, although it is clear that MIMO-NOMA outperforms MIMO-OMA specifically in the terms of sum channel capacity.

All things considered, in spite of their potential improvements, the MIMO-NOMA system has some limitations. The MIMO nodes have numerous antennas hence the user's channels are in form of vectors or matrices, making it difficult to order users according to their channel conditions so user ordering becomes a complex task in MIMO-NOMA frameworks. To tackle this, random beamforming for NOMA is proposed in [143], which successfully solves this issue by instructing the BS to order the users based on their channel quality inputs. Figure 9 illustrates a typical downlink NOMA link with MIMO with BS having two transmitting antennas. Different transmit beams are formed by MIMO in this technique, and superposed signals are designated to multiple users is applied inside each beam. This strategy can also be thought of as a hybrid of NOMA and MU-MIMO. The performance of MIMO-NOMA network can be scaled up with the increase in antennas. However, this leads to increased energy consumption and poses physical space constraints in actual systems at both the BS and user's devices. Using IRS along with the NOMA helps to overcome these limitations of the MIMO-based transmission.

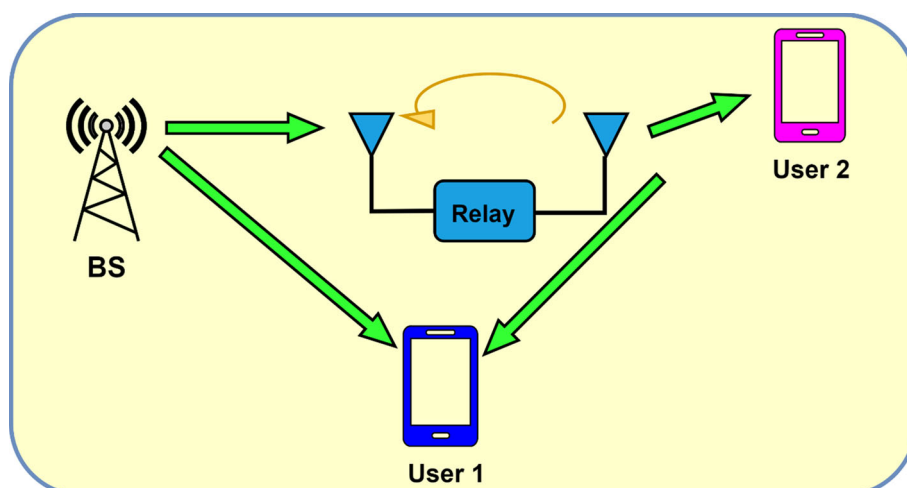
### 5.2 Relay aided NOMA

As of late, the utilization of NOMA in cooperative 5G networks has got a great deal of consideration, because of inferable benefits of cooperative relaying strategies in terms of coverage extension and reliability. Two primary protocols of relay behavior are widely known, namely: DF, in which the relay decodes and re-encodes the information signal before forwarding it, and AF in which the relay simply amplifies the received signal and transmits it to the destination. Major studies have significantly contributed to NOMA through the application of new techniques in 5G networks and beyond networks. It will be promising to incorporate relaying in NOMA systems, explicitly FD, because relaying can offer significantly enhanced throughput and coverage [144, 145]. A typical FD relay assisted DL NOMA link is shown in Fig. 10. Using NOMA, the BS sends the superposed signal of several users to the relay. FD relay provides the received signal to the user in either AF or DF mode at the

**Fig. 9** Representation of NOMA link with MIMO



**Fig. 10** Relay aided FD DL NOMA link



same time. Finally, using SIC each user decodes the superposed signal from the relay. For instance, in [146] authors have investigated relay-aided NOMA network employing partial relay selection among multiple FD-AF relays deriving expressions for the outage and ergodic capacity metric. Similarly, in [147], authors have demonstrated the performance of the FD relay aided integrated CR NOMA network. A relay aided NOMA scheme for the uplink cellular network is proposed in [148]. This new technique does not require CSI at the user so greatly reduces the transmission overheads. In order to improve the system capacity and enhance the reliability of NOMA, cooperative NOMA networks have attracted great attention [150–152]. Adding to this also there are several literature focusing on the performance analysis of relay aided NOMA over the fading channels. Like, a NOMA-based DL-AF relaying network under Nakagami- $m$  fading is studied in [153] considering imperfect CSI.

To analyse the performance closed-form expression for the OP is derived. Again, SC and MRC diversity is applied

over the Nakagami- $m$  fading channel to analyse the performance of a cooperative relay-based NOMA network in [154]. In [155], NOMA-based DF relay network with perfect CSI undergoing Rician fading is analysed. The authors of [156] looked at the outage performance of the NOMA system with DF relay using FD and HD mode, where the close-by user was supposed as a relay to improve transmission reliability of the far-away users. Furthermore, various aspects of the relay based NOMA is also explored in the literature [157–159]. The IRS-assisted NOMA networks provide less hardware complexity than a relay-aided network as low-power and low-complexity electronic circuits ensure their configurability [160]. Relay deployment can be expensive and power-intensive, particularly for realising multiple-antenna designs at millimetre and submillimetre wave frequency bands as they are equipped with several active electronic components, like DACs and ADCs, power and low-noise amplifiers for transmission and reception respectively, and mixers. However, both the technologies



**Table 7** Comparison of IRS, relay and MIMO aided NOMA technology

Technology	Transmission type	Hardware cost	Energy consumption	Operating mechanism
IRS assisted	FD	Low	Low	Passive, reflecting
Relay aided	FD/HD	High	High	Active, receive/transmit
MIMO NOMA	FD	Very High	Very High	Active, receive/transmit

happen to complement each other and have the capacity to hold a place in the 5G and beyond network architectures. Based on the above reviews Table 7 presents the comparison of these technologies.

## 6 Future research directions

Together providing exciting improvements in the performance of the networks by IRS-assisted NOMA, brings new challenges and open research problems. In this section we elaborate on some major research problem that can be considered to unveil some great potential benefits of the IRS-assisted NOMA networks.

### 6.1 Learning based approach for the optimization of IRS co-efficient

In the IRS–NOMA networks, the order of the users could be efficiently controlled by adequately configuring the IRS's meta-atoms. However, to enable this capability, the IRS must be dynamically optimized based on the instantaneous channel realizations. This imposes significant difficulties for the useful execution of IRS–NOMA, one is that such an optimization can become highly complex as the number of reflecting elements increases. Different from the alternating optimizing techniques, machine learning based techniques are proving to be more appealing to instantaneous channel realization based upon the radio environment [161, 162]. In some recent studies, the learning-based techniques have been utilized for the sum rate maximization for all the users of the IRS NOMA networks by optimizing the IRS coefficients. In [163], for an IRS-aided NOMA network to maximize the sum rate of all users by jointly optimizing the performance characteristics, deep learning-based algorithms are proposed. DRL technique is deployed in [164] to optimally tune the IRS phase shifts. Furthermore, some more learning-based techniques can be explored for the development of efficient low-complexity algorithms for optimizing the IRS coefficients and thereby can be considered as a promising research direction.

### 6.2 Physical layer security in IRS NOMA

Due to the broadcast nature of the wireless medium, wireless transmission is confronted with security issues. This draws major attention to the issue of secure transmission in the IRS NOMA networks as well. Several works on PLS for NOMA networks and IRS networks are presented in the literature [165–167]. Ensuring secrecy in the combined IRS NOMA networks is still in its infancy. Hence, it has major scope to explore the possibility of maintaining secrecy in the practical implementation of the IRS–NOMA networks. In [168], the authors investigate the PLS of IRS-aided NOMA-DL network. Similarly, [127] looks into improving the security of an IRS–NOMA network, proposing a distributed IRS–NOMA transmission framework to serve users confidentially in the face of passive eavesdropping.

### 6.3 Comparing IRS NOMA and IRS OMA

The scenarios for using the NOMA or OMA techniques are still not very clear in the literature and needs some dedicated studies to clarify their practical implementation scenarios. In [118] user clustering is used to analyse the performance of IRS–NOMA and IRS-OMA networks. It is demonstrated that if users of certain channel conditions are grouped then, IRS OMA proves to be better than IRS NOMA. For DL and UL IRS-assisted NOMA and OMA network performance is evaluated in [169]. The design of an efficient user pairing algorithm for the IRS NOMA network to extract maximum benefit from this technique also draws a major concern. In [120], authors have proposed a novel energy-efficient algorithm and through simulations demonstrate that this algorithm provides superior performance when compared with the IRS OMA.

### 6.4 Implementation of IRS NOMA in terahertz communication

Terahertz (THz) communications have been broadly imagined as a promising empowering influence to provide sufficient transmission capacity and accomplish ultra-high information rates for 6G wireless networks. Next-generation

wireless communication systems will operate at higher frequencies (mmWave or Terahertz band) to support applications that require sensing of the surrounding environment and accurate localization [170]. Such high frequencies have a limited number of propagation paths (mainly due to large penetration losses, high values of path loss, and low scattering) which may reduce the accuracy of sensing and localization. Creating a smart radio environment using IRS can be considered as a promising solution to this problem, specifically in NLOS communication scenarios [171]. To enhance the spectrum efficiency and coverage ability using NOMA, can add exceptional benefits to the terahertz communication system. Design of efficient sensing and localization model for the IRS-assisted NOMA communication system emerges out to be a subject of further investigation. In [172], using NOMA transmission provides higher average data rates for the scenarios where users are located on variety of distances far from the terahertz access point proving NOMA to be a promising method for terahertz communication networks. IRS-assisted terahertz communication network is investigated in literature [173–175]. Apart from these studies, there is a scope for the study of the IRS-assisted terahertz communication network with the NOMA transmission technique.

### 6.5 IRS NOMA assisted UAV communication

Currently, the aerial radio access networks are emerging as a potential technique to complement the terrestrial networks to develop a seamless next-generation mobile communication. The major component of aerial communication consists of UAVs. Being equipped with communication and signal processing capabilities, UAVs can act as moving BSs or relays to facilitate reliable and efficient communication with multiple users. When the users are roaming continuously, it is challenging for the UAV to periodically reposition itself according to the mobility of the users. In such situations, the feature of IRS to adjust the phase shift can help to establish a reliable communication between the UAV and the multi-users. However, this implementation poses new challenges in designing the system such as it needs to jointly design the three-dimensional UAV trajectory/position, RISs reflection beamforming, and user association to maximize the achievable rate of the users. Also, to fully utilize the NOMA capabilities additional channel condition/QoS-based decoding order design needs to be considered. This makes the joint UAV trajectory/ position, IRSs reflection beamforming, and NOMA decoding order selection problem highly coupled and difficult to resolve. These design considerations have been highlighted in [176, 177]. In [178], ML is implemented to provide real-time control for UAV movement and adjustment of the phase shifts of the IRS and show that using this, the energy consumption of the UAV enabled IRS–NOMA

network can be minimized. However, current researches on this is limited which provides a potential future prospective to the enthusiasts for contributing to exploit the IRS–NOMA technology in enabling efficient UAV communication systems.

## 7 Conclusion

In this paper, we have presented a comprehensive survey of the IRS-assisted NOMA networks as a candidate technology to overcome the challenges of the dense network. We have first explained the basic concept and the advantages of IRS implementation in the communication network. The need of MA techniques and the superiority of the NOMA technique over the OMA transmission is also thoroughly reviewed. Furthermore, the IRS integrated with the NOMA transmission is reviewed proving that an integrated environment provides better energy efficiency and coverage probability. Then, it presents various relevant performance analysis works including capacity analysis, UL/DL data rate analysis, and OP. Afterward, it discusses a review of optimization strategies aimed at achieving certain objectives, namely, maximizing EE, power, and coverage. The paper further provides a comparison of IRS–NOMA technology with the MIMO–NOMA and relay-aided NOMA technologies. As CSI is of paramount importance in both IRS and NOMA techniques, this concept is thoroughly reviewed. Finally, we identify a number of fascinating open challenges for IRS–NOMA wireless networks. This survey is the first of its kind, to the best of the authors' knowledge, that combines the technical aspects and performance analysis of IRS-assisted NOMA networks and sheds light on prospective research directions for the formulations of practical problems in future 6G network systems.

**Authors' contributions** All the authors have equally contributed in this manuscript.

**Funding** No funding was received for this work.

**Availability of data and material** N/A.

**Code availability** N/A.

## Declarations

**Conflicts of interest** Authors declare that there is no conflict of interest.

## References

1. Mahmoud, H. H. H., Amer, A. A., & Ismail, T. (2021). 6G A comprehensive survey on technologies, applications, challenges, and research problems. *Transactions on Emerging Telecommunications Technologies*, 32(4), e4233.

2. Yang, P., Xiao, Y., Xiao, M., & Li, S. (2019). 6G wireless communications: Vision and potential techniques. *IEEE Network*, 33(4), 70–75.
3. Long, W., Chen, R., Moretti, M., Zhang, W., & Li, J. (2021). A promising technology for 6G wireless networks: Intelligent reflecting surface. *Journal of Communications and Information Networks*, 6(1), 1–16.
4. Y. Yuan, Y. Zhao, B. Zong and S. Parolari, "Potential key technologies for 6G mobile communications," [Online] Available: [arXiv:1910.00730](https://arxiv.org/abs/1910.00730) [cs.IT], accessed on April 2021.
5. Alghamdi, R., et al. (2020). Intelligent surfaces for 6G wireless networks: A survey of optimization and performance analysis techniques. *IEEE Access*, 8, 202795–202818.
6. Huang, C., Zappone, A., Alexandropoulos, G. C., Debbah, M., & Yuen, C. (2019). Reconfigurable intelligent surfaces for energy efficiency in wireless communication. *IEEE Transactions on Wireless Communications*, 18(8), 4157–4170.
7. Nguyen, H. V., et al. (2020). A survey on non-orthogonal multiple access: From the perspective of spectral efficiency and energy efficiency. *Energies*, 13(16), 4106.
8. Yang, G., Xu, X. & Liang, Y. (2020). Intelligent reflecting surface assisted non-orthogonal multiple access. In *IEEE proc. of international wireless commun. networking conf. (WCNC)*, Seoul, Kr.
9. Wu, Q., Zhou, X. & Schober, R. (2021). IRS-assisted wireless powered NOMA: Do we really need different phase shifts in DL and UL?, [Online] Available: [arXiv:2102.08739v4](https://arxiv.org/abs/2102.08739v4) [cs.IT]. Accessed on May 2021.
10. Zhao, J. (2021). A survey of intelligent reflecting surfaces (IRSs): Towards 6G wireless communication networks. [Online] Available: <https://arxiv.org/pdf/1907.04789.pdf>. accessed on May 2021.
11. Wu, Q., Zhang, S., Zheng, B., You C. & Zhang, R. (2021). Intelligent reflecting surface aided wireless communications: A tutorial. [Online] Available: <https://arxiv.org/pdf/2007.02759v2.pdf> [cs.IT]. Accessed on May 2021.
12. Wu, Q., & Zhang, R. (2020). Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network. *IEEE Communications Magazine*, 58(1), 106–112.
13. Gong, S. et al. (2021). Towards smart wireless communications via intelligent reflecting surfaces: A contemporary survey. [Online] Available: [arXiv:2102.07794v2](https://arxiv.org/abs/2102.07794v2) [cs.IT]. Accessed on May 2021.
14. Basar, E., Di Renzo, M., de Rosny, J., Debbah, M., Alouini, M.-S., & Zhang, R. (2019). Wireless communications through reconfigurable intelligent surfaces. *IEEE Access*, 7, 116753–116773.
15. Di Renzo, M. et al. (2021). Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and road ahead. [Online] Available: [arXiv preprint arXiv:2004.09352](https://arxiv.org/abs/2004.09352). Accessed on May 2021.
16. Yuan, X., Zhang, Y.-J., Shi, Y., Yan, W. & Liu, H. (2021). Reconfigurable-intelligent-surface empowered 6G wireless communications: Challenges and opportunities. [Online] available: [arXiv:2001.00364](https://arxiv.org/abs/2001.00364). Accessed on May 2021.
17. Bjornson, E., Zdogan, O. & Larsson, E. G. (2021). Reconfigurable intelligent surfaces: Three myths and two critical questions. [Online] available: [arXiv:2006.03377](https://arxiv.org/abs/2006.03377). Accessed on May 2021.
18. Di Renzo, K. M., et al. (2020). Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison. *IEEE Open Journal of the Communications Society*, 1, 798–807.
19. ElMossallamy, M. A., Zhang, H., Song, L., Seddik, K. G., Han, Z., & Li, G. Y. (2020). Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities. *IEEE Transactions on Cognitive Communications and Networking*, 6(3), 990–1002.
20. Almohamad, A., et al. (2020). Smart and secure wireless communications via reflecting intelligent surfaces: A short survey. *IEEE Open Journal of the Communications Society*, 1, 1442–1456.
21. Wu, Q. & Zhang, R. (2018). Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design. In 2018 IEEE global communications conference (GLOBECOM) (pp. 1–6), Abu Dhabi, UAE.
22. Hu, S., Rusek, F., & Edfors, O. (2018). Beyond massive MIMO: The potential of data transmission with large intelligent surfaces. *IEEE Transactions on Signal Processing*, 66(10), 2746–2758.
23. Hu, S., Rusek, F. & Edfors, O. (2017). The potential of using large antenna arrays on intelligent surfaces. In 2017 IEEE 85th vehicular technology conference (VTC Spring), 2017 (pp. 1–6), Sydney, Australia.
24. Jung, M., Saad, W., & Kong, G. (2021). Performance analysis of active large intelligent surfaces (LISs): Uplink spectral efficiency and pilot training. *IEEE Transactions on Communications*, 69(5), 3379–3394.
25. Yuan, X., & He, Z. (2020). Cascaded channel estimation for large intelligent metasurface assisted massive MIMO. *IEEE Wireless Communications Letters*, 9(2), 210–214.
26. Di Renzo, M. et al. (2019). Smart radio environments empowered by AI reconfigurable meta-surfaces: An idea whose time has come. *EURASIP J. Wireless Commun. Netw.*, Article no. 129.
27. Tan, X., Sun, Z., Koutsonikolas, D. & Jornet, J. M. (2018). Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays. In *IEEE Conference on Computer Communications* (pp. 270–278), Honolulu, HI, USA.
28. Tan, X., Sun, Z., Jornet, J. M. & Pados, D. (2016). Increasing indoor spectrum sharing capacity using smart reflect-array. In *IEEE international conference on communications (ICC)* (pp. 1–6), Kuala Lumpur, Malaysia.
29. Johansson, H. & Mishra, D. (2019). Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer. In *IEEE international conference on acoustics, speech and signal processing (ICASSP)*, Brighton, UK.
30. Huang, C., Zappone, A., Debbah, M. & Yuen, C. (2018). Achievable rate maximization by passive intelligent mirrors. In *Proc. IEEE ICASSP* (pp. 3714–3718), Calgary, AB, Canada.
31. Basar, E. (2020). Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G. *IEEE Transactions on Communications*, 68(5), 3187–3196.
32. Liaskos, C., Tsioliaridou, A., Nie, S., Pitsillides, A., Ioannidis, S. & Akyildiz, I. (2019). An interpretable neural network for configuring programmable wireless environments. In *IEEE 20th international workshop on signal processing advances in wireless communications (SPAWC)*, (pp. 1–5) Cannes, France.
33. Liu, F. et al. (2018). Programmable metasurfaces: State of the art and prospects. In 2018 IEEE International Symposium on Circuits and Systems (ISCAS) (pp. 1–5), Florence, Italy.
34. Liang, Y.-C., Long, R., Zhang, Q., Chen, J., Cheng, H. V., & Guo, H. (2019). Large intelligent surface/antennas (LISA): Making reflective radios smart. *Journal of Communications and Information Networks*, 4(2), 40–50.
35. Yadav, P., Kumar, R., & Kumar, S. (2020). Effective capacity analysis over generalized lognormal shadowed composite fading channels. *Internet Technology Letters*. <https://doi.org/10.1002/itl2.171>
36. Yadav, P., Kumar, S. & Kumar, R. (2020). Effective capacity analysis over  $\alpha$ - $\kappa$ - $\mu$ /gamma composite fading channel. In *2nd international conference on advances in computing, communication control and networking (ICACCCN)*, Greater Noida, India, 2020, pp. 587–592.



37. Yadav, P., Kumar, S., & Kumar, R. (2021). A review of transmission rate over wireless fading channels: Classifications, applications, and challenges. *Wireless Personal Communications*. <https://doi.org/10.1007/s11277-021-08968-1>
38. Boccardi, F., Heath, R. W., Lozano, A., Marzetta, T. L., & Popovski, P. (2014). Five disruptive technology directions for 5G. *IEEE Communications Magazine*, 52(2), 74–80.
39. Wu, Q., & Zhang, R. (2019). Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming. *IEEE Transactions on Wireless Communications*, 18(11), 5394–5409.
40. Cui, T. J., et al. (2014). Coding metamaterials, digital metamaterials and programmable metamaterials. *Light Science & Applications*, 3, e218.
41. Yu, X., Xu, D. & Schober, R. (2019). Enabling secure wireless communications via intelligent reflecting surfaces. *IEEE Global Communications Conference (GLOBECOM)* (pp. 1–6), Waikoloa, HI, USA.
42. Cui, M., Zhang, G., & Zhang, R. (2019). Secure wireless communication via intelligent reflecting surface. *IEEE Wireless Communications Letters*, 8(5), 1410–1414. <https://doi.org/10.1109/LWC.2019.2919685>
43. Chen, J., Liang, Y.-C., Pei, Y., & Guo, H. (2019). Intelligent reflecting surface: A programmable wireless environment for physical layer security. *IEEE Access*, 7, 82599–82612. <https://doi.org/10.1109/ACCESS.2019.2924034>
44. Xu, D., Yu, X., Sun, Y., Ng, D. W. K. & Schober, R. (2019). Resource allocation for secure IRS-assisted multiuser MISO systems. In *2019 IEEE Globecom Workshops (GC Wkshps)* (pp. 1–6), Waikoloa, HI, USA.
45. Ye, J., Guo, S., & Alouini, M.-S. (2020). Joint reflecting and precoding designs for SER minimization in reconfigurable intelligent surfaces assisted MIMO systems. *IEEE Transactions on Wireless Communications*, 19(8), 5561–5574.
46. Pan, C., et al. (2020). Multicell MIMO communications relying on intelligent reflecting surfaces. *IEEE Transactions on Wireless Communications*, 19(8), 5218–5233.
47. Waqar, O. (2021). Performance analysis for IRS-aided communication systems with composite fading/shadowing direct link and discrete phase shifts. *Transactions on Emerging Telecommunications Technologies*, 32(10), e4320.
48. Nagarajan, D., & Balakrishnan, R. (2021). Error probability and throughput analysis of IRS-assisted wireless system over generalized  $\kappa$ - $\mu$  fading channels. *Wireless Personal Communications*, 120, 1929–1944.
49. Wu, Q., & Zhang, R. (2020). Weighted sum power maximization for intelligent reflecting surface aided SWIPT. *IEEE Wireless Communications Letters*, 9(5), 586–590. <https://doi.org/10.1109/LWC.2019.2961656>
50. Pan, C., et al. (2020). Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer. *IEEE Journal on Selected Areas in Communications*, 38(8), 1719–1734. <https://doi.org/10.1109/JSAC.2020.3000802>
51. Rasethunsa, T. R., Kumar, S., & Kaur, M. (2021). On the performance of DF-based multi-hop system over  $\alpha$ - $\kappa$ - $\mu$  and  $\alpha$ - $\kappa$ - $\mu$ -extreme fading channels. *Digital Signal Processing*. <https://doi.org/10.1016/j.dsp.2020.102909>
52. Agarwal, D., Bansal, A., & Kumar, A. (2018). Analyzing selective relaying for multiple relay based differential DF-FSO network with pointing errors. *Transaction on Emerging Telecommunications Technologies*, 29(9), 1–17.
53. Björnson, E., Özdogan, Ö., & Larsson, E. G. (2020). Intelligent reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying? *IEEE Wireless Communications Letters*, 9(2), 244–248.
54. Vaezi, M., Ding, Z., & Poor, H. V. (2019). *Multiple access techniques for 5G wireless networks and beyond*. Springer.
55. Aldababsa, M., Toka, M., Gökçeli, S., Kurt, G. K., & Kucur, O. (2018). A tutorial on nonorthogonal multiple access for 5G and beyond. *Wireless Communications and Mobile Computing*. <https://doi.org/10.1155/2018/9713450>
56. Saito, Y. et al. (2013). Non-orthogonal multiple access (NOMA) for cellular future radio access. In *IEEE vehicular technology conference (VTC Spring)*, Dresden, Germany.
57. Islam, S. R., Zeng, M., Dobre, O. A. & Kwak, K. (2019). Nonorthogonal multiple access (NOMA): How it meets 5G and beyond. *Wiley 5G Ref*, pp. 1–28.
58. Xu, Y., Shen, C., Chang, T., Lin, S., Zhao, Y., & Zhu, G. (2020). Transmission energy minimization for heterogeneous low-latency NOMA downlink. *IEEE Transactions on Wireless Communications*, 19(2), 1054–1069.
59. Xu, Y., Shen, C., Cai, D., & Zhu, G. (2020). Latency constrained non-orthogonal packets scheduling with finite block-length codes. *IEEE Transactions on Vehicular Technology*, 69(10), 12312–12316.
60. Xu, Y., Cai, D., Fang, F., Ding, Z., Shen, C., & Zhu, G. (2020). Outage constrained power efficient design for downlink NOMA systems with partial HARQ. *IEEE Transactions on Communications*, 68(8), 5188–5201.
61. Xu, Y., et al. (2017). Joint beamforming and power-splitting control in downlink cooperative SWIPT NOMA systems. *IEEE Transactions on Signal Processing*, 65(18), 4874–4886.
62. Yuan, Z., Yu, G., & Li, W. (2015). Multi-user shared access for 5G. *Telecommunications Network Technology*, 5(5), 28–30.
63. Nikopour, H. & Baligh, H. (2013) Sparse code multiple access. In *IEEE 24th Annual international symposium on personal, indoor, and mobile radio communications (PIMRC'13)*, London, UK.
64. Dai, L., Wang, B., Yuan, Y., Han, S., & C. I. and Z. Wang. (2015). Nonorthogonal multiple access for 5G: solutions, challenges, opportunities and future research trends. *IEEE Communications*, 53(9), 74–81.
65. Benjebbovu, A., Li, A., Saito, Y., Kishiyama, Y. Harada, A. & Nakamura, T. (2013). System-level performance of downlink NOMA for future LTE enhancements. In *IEEE global communications conference (GLOBECOM)*, Atlanta, USA, 2013, pp. 66–70.
66. Islam, S. M. R., Avazov, N., Dobre, O. A. & Kwak, K. (2017). Power-domain non-orthogonal multiple access (NOMA) in 5G systems: Potentials and challenges. *IEEE Communications Surveys and Tutorials*, 19(2), pp. 721–742, Second quarter 2017.
67. Wang, P., Xiao, J., & Ping, L. (2006). Comparison of orthogonal and non-orthogonal approaches to future wireless cellular systems. *IEEE Vehicular Technology Magazine*, 1(3), 4–11.
68. Al-Imari, M., Xiao, P., Imran, M. A. & Tafazolli, R. (2014). Uplink nonorthogonal multiple access for 5G wireless networks. In *Proc. IEEE intern. sympos. on wireless commun. systems*, pp. 781–785.
69. Zhang, N., Wang, J., Kang, G., & Liu, Y. (2016). Uplink nonorthogonal multiple access in 5G systems. *IEEE Communications Letters*, 20(3), 458–461.
70. Al-Imari, M., Xiao, P., Imran, M. A. & Tafazolli, R. (2014). Uplink nonorthogonal multiple access for 5G wireless networks. In: *Proc. IEEE intern. sympos. on wireless commun. systems* (pp. 781–785) Barcelona, Spain.
71. Vergados, D. D., & Miridakis, N. I. (2013). A survey on the successive interference cancellation performance for single-antenna and multiple-antenna OFDM systems. *IEEE Commun. Surveys Tutorials*, 15(1), 312–335.
72. Higuchi, K., & Benjebbour, A. (2015). Non-orthogonal multiple access (NOMA) with successive interference cancellation. *IEEE Transactions on Communications*, E98-B(3), 403–414.

73. RiazulIslam, S. M., Zeng, M., & Dobre, O. A. (2017). NOMA in 5G Systems: Exciting Possibilities for Enhancing. *IEEE Tech Focus*, 1(2), 1–6.
74. Liu, Y., Qin, Z., ElKashlan, M., Ding, Z., Nallanathan, A., & Hanzo, L. (2017). Nonorthogonal multiple access for 5G and beyond. *Proceedings of the IEEE*, 105(12), 2347–2381.
75. Timotheou, S., & Krikidis, I. (2015). Fairness for non-orthogonal multiple access in 5G systems. *IEEE Signal Processing Letters*, 22(10), 1647–1651.
76. Wei, Z., Guo, J., Kwan Ng, D. W., & Yuan, J. (2017). Fairness comparison of uplink NOMA and OMA. In: *IEEE 85th vehicular technology conference: VTC2017-Spring*.
77. Tao, Y., Liu, L., Liu, S., & Zhang, Z. (2015). A survey: Several technologies of non-orthogonal transmission for 5G. *China Communications*, 12, 1–15.
78. Shin, W., Vaezi, M., Lee, B., Love, D. J., Lee, J., & Poor, H. V. (2017). Non-orthogonal multiple access in multi-cell networks: Theory, performance, and practical challenges. *IEEE Communications Magazine*, 55(10), 176–183.
79. Ding, Z., Liu, Y., Choi, J., Sun, Q., ElKashlan, M., Chih-Lin, I., & Poor, H. V. (2017). Application of non-orthogonal multiple access in LTE and 5G networks. *IEEE Communications Magazine*, 55(2), 185–191.
80. Zhang, Z., Sun, H. & Lei, X. (2018). Non-orthogonal multiple access. [https://doi.org/10.1007/978-3-319-32903-1\\_57-1](https://doi.org/10.1007/978-3-319-32903-1_57-1).
81. Ding, Z., Fan, P., & Poor, H. V. (2016). Impact of user pairing on 5G non-orthogonal multiple-access downlink transmissions. *IEEE Transactions on Vehicular Technology*, 65(8), 6010–6023.
82. Ali, M. S., Tabassum, H., & Hossain, E. (2016). Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems. *IEEE Access*, 4, 6325–6343.
83. Liu, Y., Derakhshani, M., & Lambbotharan, S. (2018). Outage analysis and power allocation in uplink non-orthogonal multiple access systems. *IEEE Communications Letters*, 22(2), 336–339.
84. Ding, Z., Peng, M., & Poor, H. V. (2015). Cooperative non-orthogonal multiple access in 5G systems. *IEEE Communications Letters*, 19(8), 1462–1465.
85. Men, J., Ge, J., & Zhang, C. (2016). Performance analysis for downlink relaying aided non-orthogonal multiple access networks with imperfect CSI over Nakagami-m fading. *IEEE Access*, 5, 998–1004.
86. Sharma, P., Kumar, A., & Bansal, M. (2020). Performance analysis of downlink NOMA over  $\eta - \mu$  and  $\kappa - \mu$  fading channels. *IET Communications*, 14(3), 522–531.
87. ElHalawany, B. M., Jameel, F., da Costa, D. B., Dias, U. S., & Wu, K. (2019). Performance analysis of downlink NOMA systems over  $\kappa - \mu$  shadowed fading channels. *IEEE Transactions on Vehicular Technology*, 69(1), 1046–1050.
88. Kumar, V., Cardiff, B., Prakriya, S. & Flanagan, M. F. (2020). Effective rate of downlink NOMA over  $\kappa - \mu$  shadowed fading with integer fading parameters. In *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)* (pp. 1–7).
89. Kumar, V., Cardiff, B., Prakriya, S., & Flanagan, M. F. (2020). Delay violation probability and effective rate of downlink NOMA over  $\alpha - \mu$  fading channels. *IEEE Transactions on Vehicular Technology*, 69(10), 11241–11252.
90. Alqahtani, A. S. & Alsusa, E. (2020). Performance analysis of downlink NOMA system over  $\alpha - \eta - \mu$  generalized fading channel. In *IEEE 91st Veh. Technol. Conf. (VTC-Spring)* (pp. 1–5), Antwerp, Belgium.
91. Papanikolaou, V. K., Karagiannidis, G. K., Mitsiou, N. A., & Diamantoulakis, P. D. (2020). Closed-form analysis for NOMA with randomly deployed users in generalized fading. *IEEE Wireless Commun. Lett.*, 9(8), 1253–1257.
92. Rabie, K et al. (2021). On the performance of non-orthogonal multiple access over composite fading channels. [Online] Available: <https://arxiv.org/pdf/2004.07860.pdf>. Accessed on May 2021.
93. Agarwal, A., Chaurasiya, R., Rai, S., & Jagannatham, A. K. (2020). OP analysis for NOMA downlink and uplink communication systems with generalized fading channels. *IEEE Access*, 8, 220461–220481.
94. Khansa, A. A., Chen, X., Gui, G. & Sari, H. (2020). A BER analysis of NOMA on Rician fading channels. In *2020 IEEE Latin-American conference on communications (LATINCOM)*, (pp. 1–6), Santo Domingo, Dominican Republic.
95. Assaf, T., Al-Dweik, A., Moursi, M. E., & Zeineldin, H. (2019). Exact BER performance analysis for downlink NOMA systems over nakagami-m fading channels. *IEEE Access*, 7, 134539–134555.
96. Liu, Y., et al. (2018). Multiple-antenna-assisted non-orthogonal multiple access. *IEEE Wireless Commun.*, 25(2), 17–23.
97. Yue, X. & Liu, Y. (2021). Performance analysis of intelligent reflecting surface assisted NOMA networks. [Online] Available: [arXiv:2002.09907v4](https://arxiv.org/abs/2002.09907v4) [cs.IT]. Accessed on May 2021.
98. Renzo, M. D. & Song, J. (2019). Reflection probability in wireless networks with metasurface-coated environmental objects: An approach based on random spatial processes. *EURASIP Journal on Wireless Communications and Networking* 99.
99. Jung, M., Saad, W., Jang, Y., Kong, G., & Choi, S. (2019). Reliability analysis of large intelligent surfaces (LISs): Rate distribution and OP. *IEEE Wireless Commun. Lett.*, 8(6), 1662–1666.
100. Al-Jarrah, M., Al-Dweik, A., Alsusa, E., Iraqi, Y. & Alouini, M.-S. (2021). IRS-assisted UAV communications with imperfect phase compensation. [Online] Available: TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.13153211.v1>, Accessed on May 2021.
101. Ferreira, R., Facina, M., De Figueiredo, F., Fraidenraich, G., & De Lima, E. (2020). Bit error probability for large intelligent surfaces under double-Nakagami fading channels. *IEEE Open Journal of the Communications Society*, 1, 750–759.
102. Zheng, B., Wu, Q., & Zhang, R. (2020). Intelligent reflecting surface-assisted multiple access with user pairing: NOMA or OMA? *IEEE Communications Letters*, 24(4), 753–757.
103. Fu, M., Zhou, Y. & Shi, Y. (2019). Intelligent reflecting surface for downlink non-orthogonal multiple access networks. In *IEEE proc. of global commun. conf. (GLOBECOM)*, Waikoloa, USA.
104. Bariah, L., et al. (2021). Large intelligent surface assisted non-orthogonal multiple access for 6G networks: Performance analysis. *IEEE Internet of Things Journal*, 8(7), 5129–5140.
105. Thirumavalavan, V. C. & Jayaraman, T. S. (2020). BER analysis of reconfigurable intelligent surface assisted downlink power domain NOMA system. In *2020 international conference on COMMunication systems & NETWORKS (COMSNETS)* (pp. 519–522) Bengaluru, India.
106. X. Mu, Y. Liu, L. Guo, J. Lin, and N. Al-Dhahir, “Exploiting intelligent reflecting surface in multi-antenna aided NOMA systems,” [Online] Available: <https://arxiv.org/abs/1910.13636v1>, accessed on May 2021.
107. Yang, L., & Yuan, Y. (2020). Secrecy OP analysis for RIS-assisted NOMA systems. *IET*, 56(23), 1254–1256.
108. Ding, Z., Schober, R., & Poor, H. V. (2020). On the impact of phase shifting designs on IRS-NOMA. *IEEE Wireless Communications Letters*, 9(10), 1596–1600. <https://doi.org/10.1109/LWC.2020.2991116>
109. Hou, T., Liu, Y., Song, Z., Sun, X. Chen, Y. & Hanzo, L. (2021). Reconfigurable intelligent reflecting surface aided NOMA networks. [Online] Available: <https://arxiv.org/abs/1912.10044v1>. Accessed on May 2021.
110. Cheng, Y., Li, K. H., Liu, Y., Teh, K. C., & Poor, H. V. (2021). Downlink and uplink intelligent reflecting surface aided networks:

- NOMA and OMA. *IEEE Transactions on Wireless Communications*, 20(6), 3988–4000.
111. Tahir, B., Schwarz, S., & Rupp, M. (2021). Analysis of uplink IRS-assisted NOMA under nakagami-m fading via moments matching. *IEEE Wireless Communications Letters*, 10(3), 624–628.
  112. Tang, Z., Hou, T., Liu, Y., Zhang, J. & Hanzo, L. (2021). Physical layer security of intelligent reflective surface aided NOMA networks. [Online] Available: <https://arxiv.org/pdf/2011.03417.v1>. Accessed on May 2021.
  113. Zeng, M., Li, X., Li, G., Hao, W., & Dobre, O. A. (2021). Sum rate maximization for IRS-assisted uplink NOMA. *IEEE Communications Letters*, 25(1), 234–238.
  114. Guo, Y., Qin, Z., Liu, Y., & Al-Dhahir, N. (2021). Intelligent reflecting surface aided multiple access over fading channels. *IEEE Transactions on Communications*, 69(3), 2015–2027.
  115. Zuo, J., Liu, Y., Qin, Z., & Al-Dhahir, N. (2020). Resource allocation in intelligent reflecting surface assisted NOMA systems. *IEEE Transactions on Communications*, 68(11), 7170–7183.
  116. Mu, X., Liu, Y., Guo, L., Lin, J. & Al-Dhahir, N. (2021). Exploiting intelligent reflecting surface in multi-antenna aided NOMA systems. [Online] Available: <https://arxiv.org/abs/1910.13636v1>. Accessed on May 2021.
  117. Fu, M., Zhou, Y., Shi, Y., & Letaief, K. B. (2021). Reconfigurable intelligent surface empowered downlink non-orthogonal multiple access. *IEEE Transactions on Communications*, 69(6), 3802–3817.
  118. Zhu, J., Huang, Y., Wang, J., Navaie, K., & Ding, Z. (2021). Power efficient IRS-assisted NOMA. *IEEE Transactions on Communications*, 69(2), 900–913.
  119. Mu, X., Liu, Y., Guo, L., Lin, J., & Al-Dhahir, N. (2021). Capacity and optimal resource allocation for IRS-assisted multi-user communication systems. *IEEE Transactions on Communications*, 69(6), 3771–3786.
  120. Fang, F., Xu, Y., Pham, Q.-V., & Ding, Z. (2020). Energy-efficient design of IRS-NOMA networks. *IEEE Transactions on Vehicular Technology*, 69(11), 14088–14092.
  121. Nadeem, Q.-U.-A., Kammoun, A., Chaaban, A., Debbah, M., & Alouini, M.-S. (2020). Asymptotic max-min SINR analysis of reconfigurable intelligent surface assisted MISO systems. *IEEE Transactions on Wireless Communications*, 19(12), 7748–7764.
  122. Lv, L., Jiang, H., Ding, Z., Yang, L., & Chen, J. (2020). Secrecy-enhancing design for cooperative downlink and uplink NOMA with an untrusted relay. *IEEE Transactions on Communications*, 68(3), 1698–1715.
  123. Yadav, P., Kumar, S., & Kumar, R. (2021). A comprehensive survey of physical layer security over fading channels: Classifications, applications, and challenges. *Transactions on Emerging Telecommunication Technologies*. <https://doi.org/10.1002/ett.4270>
  124. Yan, S. et al. (2021). Intelligent reflecting surface for wireless communication security and privacy. [Online] Available: [arXiv:2103.16696](https://arxiv.org/abs/2103.16696). Accessed on Oct 2021.
  125. Lv, L. et al. (2021) Secure non-orthogonal multiple access: An interference engineering perspective. [Online] Available: [arXiv:2003.13488](https://arxiv.org/abs/2003.13488) [cs.IT]. Accessed on Oct 2021.
  126. Zhang, Z., Lv, L., Wu, Q., Deng, H., & Chen, J. (2021). Robust and secure communications in intelligent reflecting surface assisted NOMA networks. *IEEE Communications Letters*, 25(3), 739–743.
  127. Zhang, Z., Chen, J., Wu, Q., Liu, Y., Lv, L., & Su, X. (2021). Securing NOMA networks by exploiting intelligent reflecting surface. [Online] Available: [http://arxiv.org/abs/2104.03460v3](https://arxiv.org/abs/2104.03460v3), Accessed on Oct 2021.
  128. Tang, Z. et al. (2021). Physical layer security of intelligent reflective surface aided NOMA networks. [Online] Available: [arXiv:2011.03417v1](https://arxiv.org/abs/2011.03417v1) [eess.SP], Accessed on 2021.
  129. Li, N., Li, M., Liu, Y., Yuan, C., & Tao, X. (2021). Intelligent reflecting surface assisted NOMA with heterogeneous internal secrecy requirements. *IEEE Wireless Communications Letters*, 10(5), 1103–1107.
  130. Fang, F., Zhang, H., Cheng, J. & Leung, V. C. M. (2017). Energy-efficient resource scheduling for NOMA systems with imperfect channel state information. In *2017 IEEE international conference on communications (ICC), 2017* (pp. 1–5P, Paris, France).
  131. Thakre, A., & Sumathi, S. (2019). Impact of imperfect channel state information on downlink sum-rate of two user mmwave non orthogonal multiple access. In *International conference on communication and electronics systems (ICCES)* (pp. 1–6), Coimbatore, India, 2019.
  132. Murti, F. W., Siregar, R. F., Royyan, M. & Shin, S. Y. (2021) Exploiting non-orthogonal multiple access in downlink coordinated multipoint transmission with the presence of imperfect channel state information. [Online] available: [arXiv:1812.10266](https://arxiv.org/abs/1812.10266) [cs.IT]. Accessed on Oct 2021.
  133. Zhang, R., & Zheng, B. (2020). Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization. *IEEE Wireless Communications Letters*, 9(4), 518–522.
  134. Wang, Z., Liu, L., & Cui, S. (2020). Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis. *IEEE Transactions on Wireless Communications*, 19(10), 6607–6620.
  135. Hu, C., Dai, L., Han, S., & Wang, X. (2021). Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications. *IEEE Transactions on Communications*. <https://doi.org/10.1109/TCOMM.2021.3072729>
  136. Kundu, N. K., & McKay, M. R. (2021). Channel estimation for reconfigurable intelligent surface aided MISO communications: From LMMSE to deep learning solutions. *IEEE Open Journal of the Communications Society*, 2, 471–487.
  137. De Carvalho, E., & Jensen, T. L. (2020). An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator. In *IEEE international conference on acoustics, speech and signal processing (ICASSP)*, Barcelona, Spain.
  138. Nadeem, Q.-U.-A., Alwazani, H., Kammoun, A., Chaaban, A., Debbah, M., & Alouini, M.-S. (2020). Intelligent reflecting surface-assisted multi-user MISO communication: channel estimation and beamforming design. *IEEE Open Journal of the Communications Society*, 1, 661–680.
  139. Wei, L., Huang, C., Alexandropoulos, G. C. & Yuen, C. (2020) Parallel factor decomposition channel estimation in RIS-assisted multi-user MISO communication. In *2020 IEEE 11th sensor array and multichannel signal processing workshop (SAM)* (pp. 1–5), Hangzhou, China, 2020.
  140. Sun, Q., Han, S., Chin-Lin, I., & Pan, Z. (2015). On the ergodic capacity of MIMO NOMA systems. *IEEE Wireless Communications Letters*, 4(9), 405–408.
  141. Liu, Y., Pan, G., Zhang, H., & Song, M. (2016). On the capacity comparison between MIMO-NOMA and MIMO-OMA. *IEEE Access*, 4, 2123–2129.
  142. Ding, Z., Schober, R., & Poor, H. V. (2016). A general MIMO framework for NOMA downlink and uplink transmission based on signal alignment. *IEEE Transactions on Wireless Communications*, 15(6), 4438–4454.
  143. Kishiyama, Y., & Higuchi, K. (2013). Non-orthogonal access with random beamforming and intra-beam SIC for cellular MIMO downlink. In *IEEE Veh. Tech. Conf.*, Las Vegas, NV, US.
  144. Kader, M. F., Shin, S. Y., & Leung, V. C. (2018). Full-duplex non-orthogonal multiple access in cooperative relay sharing for 5G systems. *IEEE Transactions on Vehicular Technology*, 67(7), 5831–5840.



145. Zhong, C., & Zhang, Z. (2016). Non-orthogonal multiple access with cooperative full-duplex relaying. *IEEE Communications Letters*, 20(12), 2478–2481.
146. Tregancini, A., Olivo, E. E. B., Osorio, D. P. M., de Lima, C. H. M., & Alves, H. (2019). Performance analysis of full-duplex relay-aided NOMA systems using partial relay selection. *IEEE Transactions on Vehicular Technology*, 69(1), 622–635.
147. Do, D.-T., Nguyen, M.-S.V., Jameel, F., Jäntti, R., & Ansari, I. S. (2020). Performance evaluation of relay-aided CR-NOMA for beyond 5G communications. *IEEE Access*, 8, 134838–134855.
148. Shin, W., Yang, H., Vaezi, M., Lee, J., & Poor, H. V. (2017). Relay-Aided NOMA in Uplink Cellular Networks. *IEEE Signal Processing Letters*, 24(12), 1842–1846.
149. Lee, I., & Keem, J. (2015). Capacity analysis of cooperative relaying systems using non-orthogonal multiple access. *IEEE Communications Letters*, 19(11), 1949–1952.
150. Xu, M., Ji, F., Wen, M. W., & Duan, W. (2016). Novel receiver design for the cooperative relaying system with non-orthogonal multiple access. *IEEE Communications Letters*, 20(8), 1679–1682.
151. Shokair, M., Saad, W., & Ibraheem, S. M. (2018). On the performance of downlink multiuser cognitive radio inspired cooperative NOMA. *Wireless Personal Communications*, 101, 875–895.
152. Wang, Z., Peng, Z., Pei, Y., Wang, H. (2020). Performance analysis of cooperative NOMA systems with incremental relaying. *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 4915638.
153. Men, J., Ge, J., & Zhang, C. (2017). Performance analysis for downlink relaying aided non-orthogonal multiple access networks with imperfect CSI over Nakagami-m fading. *IEEE Access*, 5, 998–1004.
154. Kumar, V., Cardiff, B., & Flanagan, M. F. NOMA-based cooperative relaying with receive diversity in Nakagami-m Fading Channels. [Online] Available: [arXiv:2011.11692v1](https://arxiv.org/abs/2011.11692v1). Accessed on May 2021.
155. Jha, P. K., Kumar, S. S. S. & Kumar, D. S. (2017). Achievable rate analysis of cooperative relay assisted opportunistic-NOMA under Rician fading channels with channel state information,” *Wireless Personal Communication*.
156. Yue, X., Liu, Y., Kang, S., Nallanathan, A., & Ding, Z. (2018). Exploiting full/half-duplex user relaying in NOMA systems. *IEEE Transactions on Communications*, 66(2), 560–575.
157. Yue, X., Liu, Y., Kang, S., Nallanathan, A., & Ding, Z. (2018). Spatially random relay selection for full/half-duplex cooperative NOMA networks. *IEEE Transactions on Communications*, 66(8), 3294–3308.
158. Tang, X., An, K., Guo, K., et al. (2019). On the performance of two-way multiple relay non-orthogonal multiple access-based networks with hardware impairments. *IEEE Access*, 7, 128896–128909.
159. Peng, Z., & Wang, Z. (2019). Secrecy performance analysis of relay selection in cooperative NOMA systems. *IEEE Access*, 7, 86274–86287.
160. Tang, W., et al. (2020). Wireless communications with programmable metasurface: New paradigms, opportunities, and challenges on transceiver design. *IEEE Wireless Communications*, 27(2), 180–187.
161. Zappone, A., Renzo, M. D., Debbah, M., Thanh, T. L., & Qian, X. (2019). Model-aided wireless artificial intelligence: Embedding expert knowledge in deep neural networks towards wireless systems optimization. *IEEE Vehicular Technology Magazine*, 14(3), 60–69.
162. Taha, A., Zhang, Y., Mismar, F. B., & Alkhateeb, A. (2020). Deep reinforcement learning for intelligent reflecting surfaces: Towards standalone operation. In *IEEE 21st int. workshop signal process. Adv. Wirel. Commun. (SPAWC)*, Atlanta, GA, USA, 2020.
163. Gao, X., Liu, Y., Liu, X. & Song, L. (2021). Machine learning empowered resource allocation in IRS aided MISO-NOMA networks. [Online] available: <https://arxiv.org/abs/2103.11791>. Accessed on May 2021
164. Shehab, M., Ciftler, B. S., Khattab, T., Abdallah, M. M. & Trinchero, D. (2021). Deep reinforcement learning powered IRS-assisted downlink NOMA,” [Online] Available:[arXiv:2104.01414v1](https://arxiv.org/abs/2104.01414v1) [cs.IT]. Accessed on May 2021
165. Qin, Z., Liu, Y., Ding, Z., Gao, Y. & El Kashlan, M. (2016) Physical layer security for 5G non-orthogonal multiple access in large-scale networks. In *2016 IEEE international conference on communications (ICC)* (pp. 1–6), Kuala Lumpur , Malasia, 2016.
166. Hu, L., Zheng, X., & Chen, C. (2019). Physical layer security in nonorthogonal multiple access wireless network with Jammer selection. *Security and Communication Networks*. <https://doi.org/10.1155/2019/78693>
167. Wijewardena, M., Samarasinghe, T., Hemachandra, K. T., Atapattu, S., & Evans, J. S. (2021). Physical layer security for intelligent reflecting surface assisted two-way communications. *IEEE Communications Letters*. <https://doi.org/10.1109/LCOMM.2021.3068102>
168. Tang, Z., Hou, T., Liu, Y., Zhang, J. & Hanzo, L. (2021). Physical layer security of intelligent reflective surface aided NOMA networks. Available:[arXiv:2011.03417](https://arxiv.org/abs/2011.03417). Accessed on May 2021.
169. Chen, Y et.al. (2021). Downlink and uplink intelligent reflecting surface aided networks: NOMA and OMA,” [Online] Available: [arXiv:2005.00996v2](https://arxiv.org/abs/2005.00996v2) [eess.SP]. Accessed on May 2021.
170. Lima, C. D., et al. (2021). Convergent communication, sensing and localization in 6G systems: An overview of technologies, opportunities and challenges. *IEEE Access*, 9, 26902–26925.
171. Sarieddeen, H., Saeed, N., Al-Naffouri, T. Y. & Alouini, M.-S. (2019) Next generation terahertz communications: A rendezvous of sensing, imaging and localization. [arXiv:1909.10462v1](https://arxiv.org/abs/1909.10462v1) [eess.SP]
172. Ülgen, O., Erküçük, S., & Baykaş, T. (2020). Non-orthogonal multiple access for terahertz communication networks. In: 11th IEEE annual ubiquitous computing, electronics & mobile communication conference (UEMCON), New York
173. Ma, X., et al. (2020). Intelligent reflecting surface enhanced indoor terahertz communication systems. *Nano Communication Networks*, 24(2), 100284.
174. Chen, W., Chen, Z., & Ma, X. (2020). Channel estimation for intelligent reflecting surface aided multi-user MISO terahertz system. *Terahertz Science and Technology (TST)*, 30(2), 51–60.
175. Chen, Z. et al., Intelligent reflecting surfaces assisted terahertz communications toward 6G,”[Online] available: <https://arxiv.org/abs/2104.02897v2>. Accessed on June 2021.
176. Jiao, S., Fang, F., Zhou, X., & Zhang, H. (2020). Joint beamforming and phase shift design in downlink UAV networks with IRS-assisted NOMA. *Journal of Communications and Information Networks*, 5(2), 138–149.
177. Mu, X., Liu, Y., Guo, L., Lin, J., & Poor, H. V. (2021). Intelligent reflecting surface enhanced multi-UAV NOMA networks. *IEEE Journal on Selected Areas in Communications*, 39(10), 3051–3066.
178. Liu, X., Liu, Y., & Chen, Y. (2021). Machine learning empowered trajectory and passive beamforming design in UAV-RIS wireless networks. *IEEE Journal on Selected Areas in Communications*, 39(7), 2042–2055.



**Sandeep Kumar** received his B. Tech. in electronics and communication from Kurukshetra University, India in 2004 and Master of Engineering in Electronics and Communication from Thapar University, Patiala, India in 2007. He received his Ph.D. from Delhi Technological University, Delhi, India in 2018. He is currently working as Member (Senior Research Staff) at Central Research Laboratory, Bharat Electronics Limited Ghaziabad, India. He has received various awards and certificates of appreciation for his research activities. His research interests include the study of wireless channels, performance modeling of fading channels and cognitive radio networks. He is a senior member of IEEE and also serving as a reviewer for IEEE, Elsevier and Springer journals etc.



**Poonam Yadav** was born in India. She received her Bachelor of Engineering in Electronics from Dr. Babasaheb Ambedkar Marathwada University, India in 2005 and M. Tech in Electronic Design and Technology from Uttar Pradesh Technical University, India in 2013. She is currently pursuing PhD from North Eastern Regional Institute of Science and Technology, Itanagar, India. She is presently working as Assistant Professor in Mahatma Gandhi Mission's College of Engineering and Technology, Noida, India. Her research interest includes wireless communication and fading channel modelling.



**Manpreet Kaur** was born in Punjab, India in 1983. She received her B. Tech. in Electronics and Communication Engineering from Punjab Technical University, Punjab, India in 2005 and Master of Engineering in Electronics and Communication from Thapar University, Patiala, India in 2007. She is currently working as Member (Senior Research Staff) at Central Research Laboratory, Ghaziabad, Bharat Electronics Limited. Her research interest includes the wireless communication and cognitive radios networks.



**Rajesh Kumar** was born in India. He received his Bachelor of Engineering in ECE from Bangalore University in 2001, Master of Engineering Science in Control System from NIT, Patna in 2007 and PhD in the field of sensors from Jamia Millia Islamia, New Delhi (A Central University) in 2012. He is having 17+ years of teaching experience in various capacities at RPSIT, Patna (Bihar); AKGEC, Ghaziabad (UP) & ABESEC, Ghaziabad (UP) at undergraduate and postgraduate level of Engineering Education. He is presently working as Associate Professor in the Department of Electronics and Communication Engineering at North Eastern Regional Institute of Science and Technology (Under MHRD, Gov. of India), Nirjuli, Itanagar, India. He has published numerous research articles in Journals and Conferences at National and International level. His area of research is in the field of Sensing technology using MEMS, RF MEMS, Biosensors, WSN and Mobile communication.

# An analytical insight of discussions and sentiments of Indians on Omicron-driven third wave of COVID-19 using twitter data

Deepika Vatsa Deepika Vatsa (✉ [vatsa.deepika.email@gmail.com](mailto:vatsa.deepika.email@gmail.com))

Bennett University

Ashima Yadav Ashima Yadav

Bennett University

---

## Research Article

**Keywords:** Coronavirus, COVID-19, India Pandemic, Sentiment analysis, Topic modeling, Twitter

**Posted Date:** April 6th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1508291/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# An analytical insight of discussions and sentiments of Indians on Omicron-driven third wave of COVID-19 using twitter data

Deepika Vatsa<sup>a,\*</sup>, Ashima Yadav<sup>b</sup>

<sup>a</sup>*School of Computer Science Engineering and Technology, Bennett University, Greater Noida, 201310, Uttar Pradesh, India*

<sup>b</sup>*School of Computer Science Engineering and Technology, Bennett University, Greater Noida, 201310, Uttar Pradesh, India*

## ARTICLE INFO

### Keywords:

Coronavirus

COVID-19

India

Pandemic

Sentiment analysis

Topic modeling

Twitter

## ABSTRACT


Microblogging has become one of the most crucial tool for expressing and sharing the opinions and views of everyday life events. Digital channels are being used to monitor public health issues on the Internet. Twitter is a very popular source that provides tweets related to the sentiment of the public during the COVID-19 pandemic. Many researchers have used tweets to monitor the opinion of the people towards the coronavirus vaccine, mental health problems, treatment received by the doctors, impact of lockdown, etc. However, these works were mostly limited to the first and second waves of the pandemic. In this work, we aim to study the impact of the third wave of the pandemic, which started in December 2021 in India. We accomplished this by collecting tweet data set of two months, i.e., December 2021 and January 2022, discussing COVID-19 and having country code as "IN". We employed the Latent Dirichlet Allocation (LDA) technique for topic modeling and labeled each tweet message with the topic words that best describe it. We also utilized sentiment labels for each tweet and analyzed the distribution of different topics across different sentiment labels. This helped us to analyze the perspectives and sentiments of the people with respect to different topic discussions. Our analysis discovered that the two most discussed topics were "precautionary measures" like get well soon, stay safe, wear mask, etc., and "vaccine" where people have discussed about its effectiveness and vaccination drive in India. We found that people mostly had neutral sentiments for the former topic while for the latter, overall sentiment polarity was negative, reflecting peoples' mistrust in the COVID-19 vaccine.

## 1. Introduction

Twitter has become one of the popular sources for gathering public opinion on health research. According to the 2019 survey results [1], there were 290.5 million monthly users actively using Twitter, and this count will increase to 340 million by 2024. Hence, Twitter can get real-time opinions and attitudes about people located in different parts of the world. The outbreak of COVID-19 has become a cause of concern for policymakers and scientists. In March 2020, the World Health Organization (WHO) declared COVID-19 as a pandemic [2]. Since then, the dreaded disease has caused a devastating effect on the entire world, resulting in more than 5,878,328 deaths worldwide [3].

Previous studies have applied sentiment analysis during different outbreaks and epidemics. Baker et al. [4] utilized machine learning-based techniques to study the spread of influenza based on Arabic tweets. The authors studied and conducted experiments using several machine learning-based methods like Support Vector Machine, Decision Trees, Naïve Bayes, and K-Nearest Neighbor to analyze around 54,065 influenza-related tweets in Arabic. Culotta et al. [5] detected N1H1 influenza-related tweets and compared the results with the Centers for Disease Control and Prevention by applying different classification methods. Experimental results show that the multiple linear regression model achieved the highest accuracy of 84.3%. Some studies have developed models which collected data for multiple infectious diseases like measles, Ebola, swine flu, listeria, etc., from Twitter [6, 7]. The authors have developed a hybrid model consisting of clue-based lexicons that separated the opinionated text from the factual reports and utilized machine learning classifiers to classify the multiple infections.

\* Corresponding author Bennett University, India

 [vatsa.deepika.email@gmail.com](mailto:vatsa.deepika.email@gmail.com) (D. Vatsa); [ashimayadavdtu@gmail.com](mailto:ashimayadavdtu@gmail.com) (A. Yadav)

ORCID(s): 0000-0002-2032-8076 (D. Vatsa); 0000-0002-1467-1601 (A. Yadav)

The first case of the COVID-19 virus was reported in December 2019 in Wuhan [8]. Since then, many of us are left with the question of “when will this be over?”. With each passing year, we see different variants of the deadly virus, resulting in multiple waves of the pandemic. This creates a huge psychological impact in people’s minds as it has been more than two years, and people cannot socialize due to worldwide curfews that have confined them into their respective homes. Currently, the entire world is going through the third wave of the coronavirus, with its new variant, known as Omicron. The first case of Omicron was reported on 02 December 2021 in India [9]. Past studies have focused on the first and second waves of the virus, where researchers have analyzed the impact of the pandemic on people [10]. However, many people have either gotten tired of taking precautions, staying home, and relying on the digital world or are getting used to it.

Hence, in this paper, we aim to study the state of the mind of the people by analyzing the varying pattern of public sentiments over time during the third wave of the virus (Omicron) among citizens of India. We also identify whether the sentiments of people change after the third wave of the pandemic or not. Since now the third wave is getting over in India, we aim to study the after-effects of this wave on the psychology of Indian citizens. In order to gain insights into the experience of the people and uncover public concerns during the third wave of the virus, we apply the topic modeling technique, which extracts the popular topics that are getting significant attention from the public and study the sentiments associated with each of them. We also show the temporal trend in the sentiments of the people. This study will be helpful to the policymakers and the healthcare professionals as they can take timely actions for the well-being of their citizens during any pandemic.

The main contributions of our work are summarized as follows:

1. To the best of our knowledge, this is the first work which focuses on sentiment analysis and retrieving topic discussions on the Omicron-led third COVID-19 wave in India.
2. We performed LDA-based topic modeling on Twitter data geo-located as India to extract the important topics which were prevalent during the third wave of the pandemic.
3. We also analyzed the sentiment trend across different topics on complete two-month data and week-wise data.
4. Finally, we summarized the prominent topics that gained major public attention during the third wave of COVID-19 pandemic in India. We found that the discussions were majorly negative towards “vaccine” topic and neutral towards “precautionary measures” topic.

The rest of the paper is organized as follows: Section 2 reviews the crucial work on sentiment analysis and topic modeling related to the COVID-19 pandemic. Section 3 discusses the proposed methodology. Section 4 focuses on the experimental results and analysis. Section 5 presents the discussion about the analysis. Section 7 concludes the paper with future remarks.

## 2. Related Work

This section discusses the previous work related to sentiment analysis and topic modeling. Past studies have focused on applying sentiment analysis to study different diseases and during disease outbreaks in public. We also explore the popular work that utilizes topic modeling techniques to discover abstract topics from large textual documents.

### 2.1. Sentiment analysis

Basiri et al. [11] studied the sentiment intensities of Twitter users for the coronavirus by fusing deep learning techniques like CNN, BiGRU, FastText, DistilBERT, and one machine learning classifier NBSVM. The study aimed to detect the correlation of the Tweets generated at the pandemic with the news and events that gained significant attention from the public. Although the authors targeted eight different countries for this study, selecting the right keyword for searching information and filtering the tweets was independent of the country. Hence, the study could not provide an accurate estimation of the sentiment trend of the people in each country. Priyadarshini et al. [12] analyzed the psychology of the people during lockdown by performing sentiment analysis on the tweets after two and four weeks of lockdown. The study helped to analyze the mental well-being of the citizens during the lockdown. The results show that the people were optimistic and supported the lockdown strategies imposed by the government.

Yousefinaghani et al. [13] extracted the sentiments of the people towards the COVID-19 vaccine by retrieving the tweets and comparing their progression based on time, themes, geographical distribution, and other characteristics. The results show that the people were more interested in discussing about vaccine rejections. People were vaccine-hesitant rather than favoring them or being optimistic towards them. The limitation of their work was that the approach used



by the authors to categorize the sentiments of tweets might have missed some important posts as the entire corpus was not reviewed. Similarly, Liu et al. [14] identified the themes and studied the temporal trends in the COVID-19 vaccine-related tweets in different countries and amongst different states of the US. The study majorly focused on analyzing the sentiments of the citizens before and after the announcement of the Pfizer vaccine. Based on the geographical analysis, the fluctuation patterns of sentiment were influenced by the number of positive cases or reported deaths in that area. Nezhad et al. [15] presented a study that aimed to assess the Persian tweets to analyze the sentiments of Iranian citizens towards the COVID-19 vaccines. The authors compared the sentiments of homegrown vaccine (named Barekat) and imported vaccines like Pfizer, Moderna, AstraZeneca, and Sinopharm. The authors observed that Iranian citizens reflected more positive sentiments towards the imported vaccines as compared to the homegrown vaccine.

Huerta et al. [16] explored the sentiment polarity trend in Massachusetts during the pandemic. The tweets were majorly focusing on increasing the risk in the health of the citizens and anxiety expressions. Das et al. [17] applied CNNs for training the classifier on the COVID-19 tweets. The authors simulated Bayesian regression based model for predicting the future cases of the virus and the recovery rate with respect to the latest scenario.

## 2.2. Topic modeling

Ridhwan et al. [18] applied emotion analysis using a pre-trained RNN classifier and sentiment analysis using the VADER tool to classify the sentiments into different categories. The authors also applied topic modeling to find the prevalent discussion topics during the COVID-19 pandemic in Singapore. The results show that during the lockdown, the positive sentiment was dominant. However, emotions like fear and joy varied over time due to the developments involved during the pandemic. Chekijian et al. [19] examined the emergency care given to the patients during the pandemic. The authors applied a topic modeling approach to analyze the comments of the patients and uncover the concerns of patient experiences in the hospital. The results show that patients were having many issues regarding their safety, treatment protocols, family or visitors' restrictions, and limitation of testing.

Melton et al. [20] investigated the sentiments of the people towards the COVID-19 vaccine by applying topic modeling on text-based data collected from 13 Reddit communities. The authors applied a topic modeling technique that identified popular topics from the combined dataset and the polarity-wise topics. The polarity analysis was conducted using lexical-based methods, which suggested that most people were showing positive sentiments towards the vaccine-related news. This sentiment remained static over an initial period. However, later on, negative sentiments emerged that were majorly focused on the side-effects of these vaccines as citizens were not confident about them.

Garcia et al. [21] analyzed the tweets of Brazil and the USA for four months by applying different machine learning classifiers: Naïve Bayes, Logistic Regression, Random Forest, Linear SVM, MLP, and AdaBoost. Topic modeling was applied to extract the ten main topics related to both countries, out of which seven topics were common in both. The negative sentiments were prevalent in topics like case statistics, economic impact, and proliferation care. In Portuguese text, the positive sentiment was mainly directed towards politics. Similarly, for English tweets, the highest number of positive messages were written for treatment received by the doctors. The major limitation of their work was that the machine learning-based classifiers failed to generalize and handle the enormous data found on social media platforms.

## 3. Proposed methodology

### 3.1. Data Collection and pre-processing

Lopez et al. [22] provided COVID-19 related tweet data set on Github. The authors have continuously collected data set using standard Tweeter API since 22 January 2020. The authors used certain keywords like coronavirus, COVID, mask, vaccine, etc. to collect tweets for the data set [22]. The data set is organised by each hour of the day. It is pre-processed and contains summary details like mentions, hashtags, sentiment scores, Named Entity Recognition (NER) data of tweets. The sentiment scores and NER data are generated using state-of-the-art Twitter Sentiment and Named Entity Recognition (NER) algorithms. For sentiment scores, Cliche's Twitter Sentiment algorithm [23] is used, which is an ensemble model of multiple Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) Networks. The method performed the best in the 11th international workshop on Semantic Evaluation SemEval-2017, where the task was sentiment analysis of Twitter data. For each tweet, the algorithm generates a vector of predicted scores for three sentiment classes: positive, neutral, and negative. Then the tweet is assigned to the sentiment class having highest predicted probability. The algorithm is found to have an accuracy of 75-80% on a sample of COVID-19 tweets, as observed by Rustam et al. [24]. For NER of English language tweets, the authors used the state-of-the-art English NER model provided by Akbik et al. [25].

In India, the first case of the Omicron variant was reported on 2 December 2021 in Bangalore. The Omicron-driven third wave reached its peak on 21 January 2022 with nearly 3,47,000 recorded cases, after which a decline in the number of cases was observed [26]. Thus, in this study for analysis we have used the data of two months, from 1st December 2021 till 31st January 2022. Specifically, we used CSV files from the folder *Summary\_Details*, where each CSV file is named as *yyyy\_mm\_dd\_hr\_Summary\_Details.csv*. For instance, *2022\_01\_01\_00\_Summary\_Details.csv* file contains data of first hour of 1 January 2022. Each CSV file consists of eight columns having headers as: "Tweet\_ID", "Language", "Geolocation\_coordinate", "RT", "Likes", "Retweets", "Country", "Date Created". Since we require tweet data of Indians, we pulled out tweet IDs with country code - "IN" and language set to "EN". For the extracted tweet IDs, we then retrieved the corresponding sentiment labels from the CSV files in *Summary\_Sentiments* folder in the data set. For topic modeling, we first obtained the tweet message by using the Hydrator application on extracted tweet IDs. The data is preprocessed before applying LDA over it. That is, the data is changed to lower case; then punctuation marks, stop words, and special characters are removed. We also removed hashtags, mentions, and URLs from it. Also, some keywords like Omicron, COVID, COVID19, COVID-19, corona are likely to be present in most of the tweet messages and are thus removed to get crisp and concrete topics. Finally, Lemmatization is performed to get the base word in each tweet message.

### 3.2. Topic Modeling

Topic modeling or topic discovery is a sub-problem in natural language processing (NLP). The aim is to discover abstract topics discussed in a set of documents, and then classifies any individual document in the set depending upon its relevance to each of the discovered topics. A topic is a set of words taken together to suggest a theme. It is crucial and comes handy when one wants to analyse a huge amount of textual data. It is useful for summarization, similarity estimation, novelty detection, and categorizing a massive collection of documents. In this work, we applied Latent Dirichlet Allocation (LDA) [27] technique on tweet messages to retrieve the topic words.

Assuming our dataset is a collection of multiple documents denoted by  $D = d_1, d_2, \dots, d_n$ . Each document  $d_i$  is a mixture of different topics, where each topic is a probabilistic mixture over different words that are combined to form a document. Topic modeling is used to explain the hidden information in any document. This can be achieved by grouping the words in such a way that each group represents a topic in a document. Hence, we apply the Latent Dirichlet allocation (LDA), which is a Bayesian hierarchical probabilistic generative model for finding the hidden thematic structure in the unstructured text.

The LDA model utilizes two matrices, namely:  $\theta(t_d)$  and  $\phi(w_t)$ , which are defined as follows:

$$\begin{aligned}\theta(t_d) &= P(\text{topic } t \mid \text{document } d) = \text{Probability distribution of topics in the documents} \\ \phi(w_t) &= P(\text{word } w \mid \text{topic } t) = \text{Probability distribution of words in the topics}\end{aligned}$$

Hence,

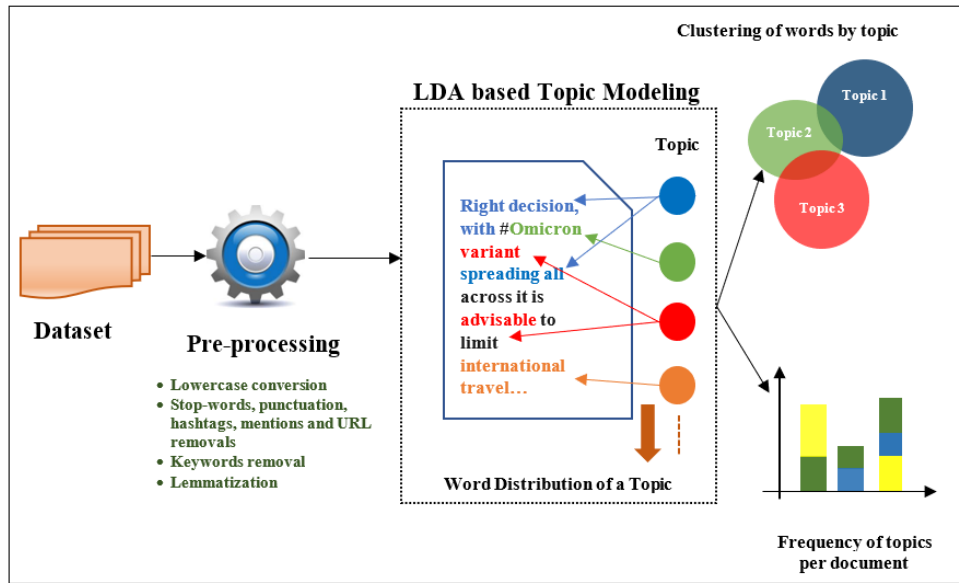
$$P(\text{word } w \mid \text{document } d) = \phi(w_t) * \theta(t_d)$$

Assuming we have a total number of topics as  $T$ , then the probability distribution of words in the documents  $P(w|d)$  is explained as below:

$$P(w|d) = \sum_T P(t|d) * P(w|t) \quad (1)$$

Where,  $*$  represents the dot product and the weights of  $\theta(t_d)$  and  $\phi(w_t)$  are assigned randomly. The entire process is summarized as below:

1. For every document  $d$ , initialize each word randomly to a topic from the distribution of topics based on their assigned weights.
2. For every document  $d$ : For every word  $w$  in  $d$ : Calculate  $P(t|d)$  and  $P(w|t)$
3. Considering all words and their topics, reassign the topic to the word  $w$  based on the dot product of  $P(t|d)$  and  $P(w|t)$  as shown in Eq (1).
4. Repeat the above step for the entire document until the assigned topics are not changed.



**Figure 1:** Architecture of the proposed methodology

In topic modeling, we eliminate the list of stop words as they carry no inherent meaning and some repeated keywords. This is done to ensure that the topics generated by the topic modeling approach are not dominated by the stop words or some repeated keywords, and meaningful topics are generated by the algorithm. Figure 1 illustrates the proposed methodology of the work.

## 4. Experimental Analysis

We have analysed the data set in two ways: firstly, on complete two-month data, and secondly, on weekly data of two months. Analysis of the complete data set gives us an overview of peoples' sentiments and their opinions for the third wave of COVID disease in India. While week-wise analysis of eight weeks data lets us understand the trend and shifts in sentiments and topics discussed by people during the emergence and peak of the third wave in India.

### 4.1. Complete data analysis

On the complete data set of two months, we have analyzed the retrieved topics, sentiment labels, and topic distribution among sentiments. Table 1 shows the topic words retrieved from the overall tweet data of two months from 1<sup>st</sup> December 2021 - 31<sup>st</sup> January 2022. Topic modeling on complete data set yielded six topics. Second table in Table 1 presents the topic themes interpreted by the set of topic words discovered in each topic.

Topic words like "get", "well", "soon", "stay", "safe", "wear", "mask", "up", etc. are categorized under the theme "Precautionary measures". Topic words related to exams like "student", "online", "exam", "sir", "lockdown", etc. are categorized under the theme "Online exam". Statistics related words like "Case", "positive", "report", "update", "test", etc. are categorized as "COVID statistics report". Words like "medicine", "pandemic" when used in conjunction with "artificial", "intelligence", "benefit" are categorized as theme "AI in medicine". Vaccine related words like "vaccine", "dose", "vaccination", etc. are categorized as "Vaccine".

Figure 2 shows topic distribution, sentiment distribution, and topic distribution among different sentiment labels on complete twitter data set of two months. As can be seen from the topic bar plot, topic 0 having theme "precautionary measures" was the most discussed while topic 3 having theme "AI in medicine" is the least discussed topic. From the sentiment labels bar plot, out of a total of 42,157 tweets, more than 18,000 tweets carry negative sentiment while 7886 were positive tweets. Thus, the number of negative tweets are more than twice the number of positive tweets. From the sentiments bar plot, we can say that the people of India mostly shared negative or neutral opinion with respect to the third COVID wave in India. Looking at the distribution of topics among different sentiment labels allows us to understand how the sentiments vary for each topic. For instance, out of all topics, topic 1 having theme: "Online exam" carries the most number of negative tweets while topic 3 having theme: "AI in medicine" carries the least number of

**Table 1**

Topics retrieved from the complete tweet data of two month: December 2021 - January 2022. Second table presents topic themes interpreted by the set of retrieved words in each topic.

Topic Words	
Topic 0	['mask' 'virus' 'get' 'stay' 'well' 'soon' 'people' 'safe' 'take' 'wear']
Topic 1	['exam' 'student' 'online' 'case' 'sir' 'lockdown' 'please' 'pandemic' 'due' 'situation']
Topic 2	['case' 'india' 'variant' 'new' 'virus' 'death' 'positive' 'update' 'report' 'test']
Topic 3	['medicine' 'intelligence' 'artificial' 'great' 'relief' 'pandemic' 'benefit' 'here' 'how' 'world']
Topic 4	['vaccine' 'dose' 'vaccination' 'india' 'year' 'dos' 'crore' 'pm' 'vaccinated' 'via']
Topic 5	['hai' 'nhm' 'employee' 'hp' 'ke' 'pandemic' 'ho' 'india' 'protesting' 'last']

Topic Themes	
Topic 0	Precautionary measures
Topic 1	Online exam
Topic 2	COVID statistics report
Topic 3	AI in medicine
Topic 4	Vaccine
Topic 5	Protest

negative tweets. It shows that Indian students were much affected by the shifting of exams from offline to online mode due to the pandemic. Among all positive tweets, topic 0 theme: “precautionary measure” is the most discussed, while topic 5 (theme: “protest”) is the least discussed topic. It shows that people of India were vigilant and were taking proper precautions. The proportion of tweets assigned topic 3 (having theme: “AI in medicine”) is almost the same in all three sentiment label categories. Thus people had overall neutral sentiments over topic “AI in medicine”.

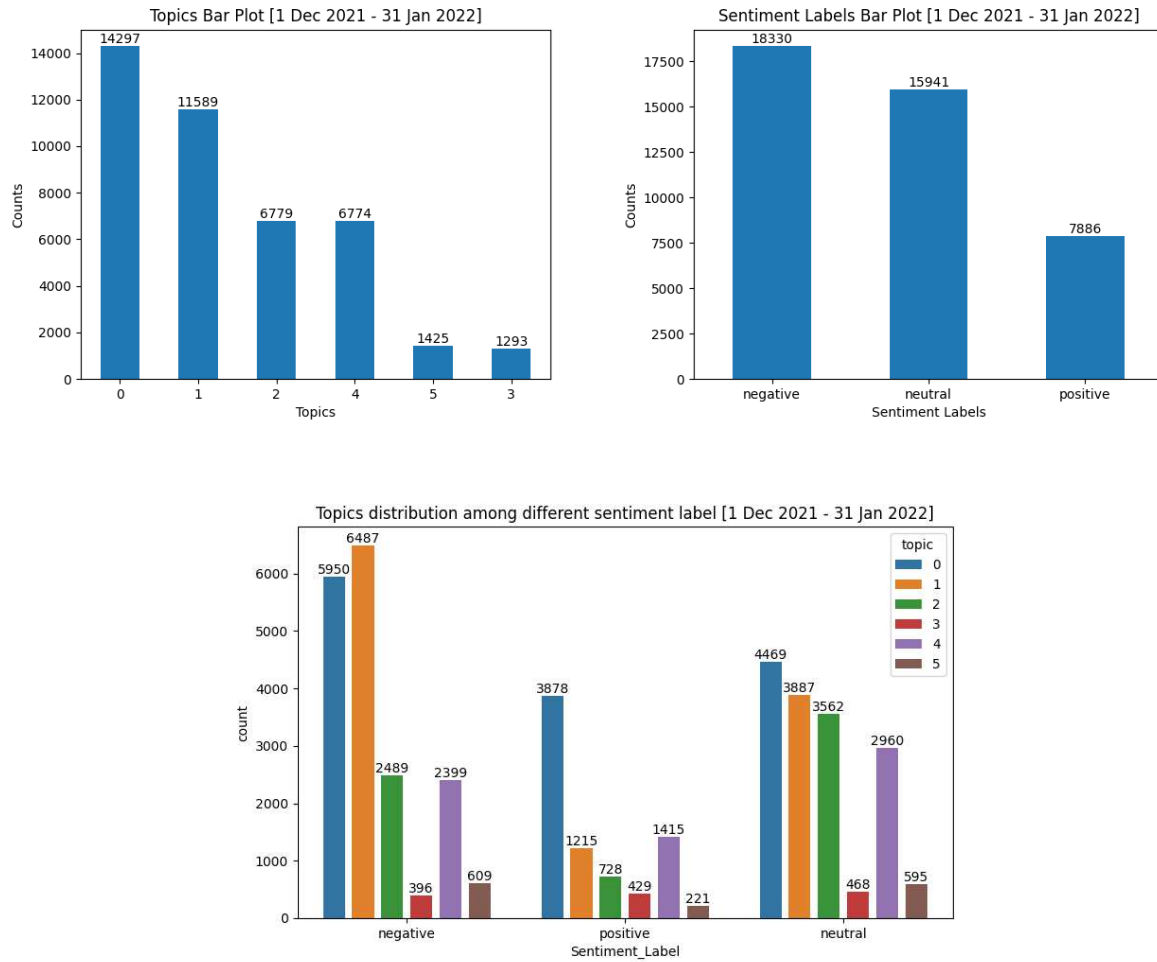
#### 4.2. Week-wise data analysis

We have sectioned the complete two month data into eight weeks data and analyzed the retrieved topics, sentiment labels, and topic distribution among sentiments on week-wise data set. Figure 3 shows sentiment label bar plots for each week. The trend of distribution of negative, neutral, and positive sentiment tweets is seen to be almost the same in each week with the most number of negative labels, followed by neutral labels, followed by the least number of positive labels. Exceptionally, for week 6, the difference in the number of negative and neutral labels is 17 which can be considered negligible.

Figure 4 shows retrieved topic words, topics bar plot and topic distribution among sentiments for week 1 and week 2. Week 1 corresponds to date range 1<sup>st</sup> Dec 2021 - 7<sup>th</sup> Dec 2021 and week 2 corresponds to date range 8<sup>th</sup> Dec 2021 - 14<sup>th</sup> Dec 2021. For week 1, topic 0 is the most discussed topic having topic theme interpreted as “new COVID variant”. Sentiment distribution over this topic shows approximately twice the number of negative and neutral tweets on this topic than the number of positively labeled tweets. For week 2, the most discussed topic is topic 1, where topic theme is “Vaccine”. Sentiment distribution of topic 1 shows that there are more negatively sentiment labeled tweets than positive ones, which reflect that people were not optimistic about the effectiveness of the vaccine developed for COVID virus on the new variant. Topic words in topic 4 and 5 suggest an odd topic, not associated with COVID disease. Topic words suggest the topic theme to be “bail granted to bapuji”. It is to note that apart from tweets related to COVID disease, a large number of tweets come from topic 4 and 5 in week 2.

Figure 5 shows the analysis results for week 3 and week 4. Week 3 corresponds to date range 15<sup>th</sup> Dec 2021 - 21<sup>st</sup> Dec 2021 and week 4 corresponds to date range 22<sup>nd</sup> Dec 2021 - 31<sup>st</sup> Dec 2021. For week 3, the most discussed topic is topic 5 having discussion words as “government”, “health”, “pandemic”, “nhm” (stands for national health mission) suggesting the work done by government employees for public health during pandemic. Again the number of negatively labeled tweets outnumber positively labelled ones for this topic. Other topics discussed are all related to vaccination. The trend of proportion of negative and positive tweets for each topic is the same, i.e., the number of negative tweets are almost twice the number of positive tweets. For week 4, the most discussed topic is topic 2 having topic words as “booster”, “vaccine”, “dose”, “India” suggesting discussion on booster vaccine dose in India for fighting the virus. The

## Short Title of the Article

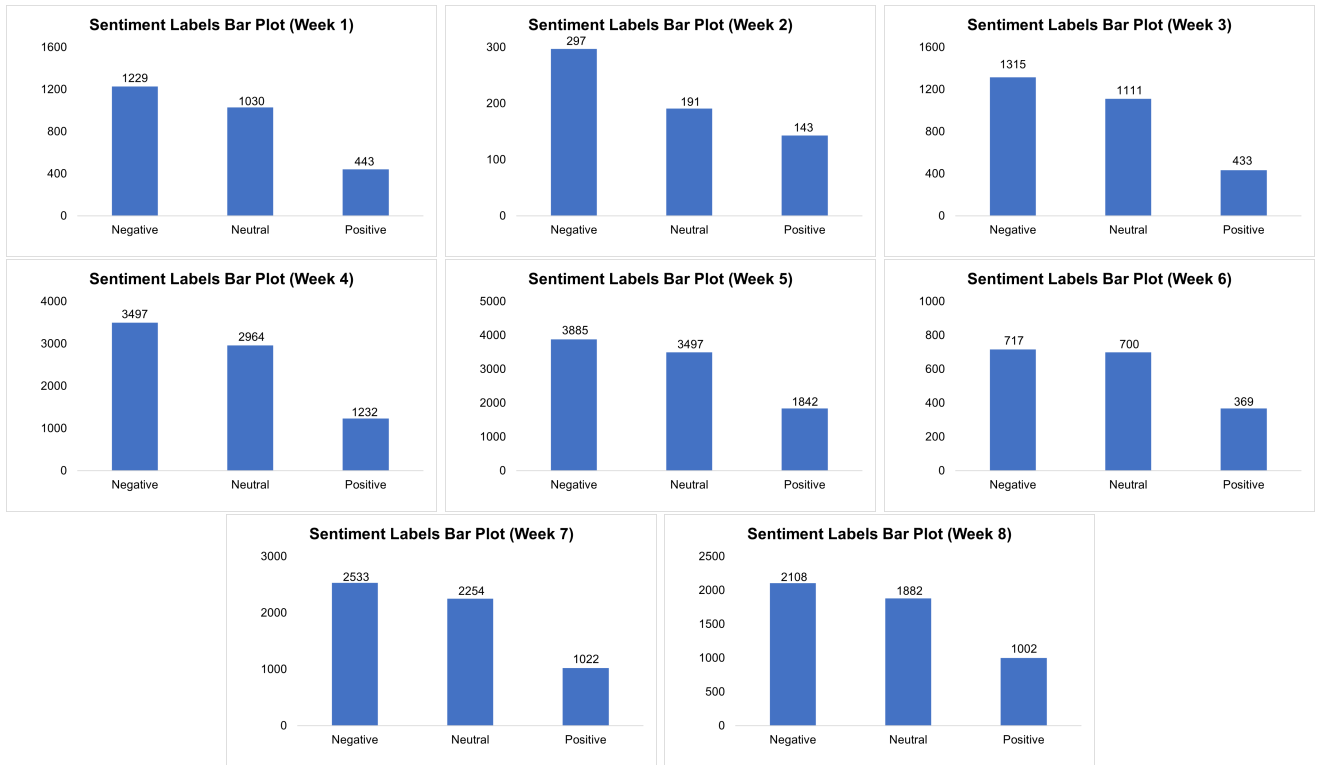


**Figure 2:** Sentiment and topic distribution of tweet data of two month from 1<sup>st</sup> Dec 2021 to 31<sup>st</sup> Jan 2022

number of negative tweets for this topic is 953 which is much higher than the number of positive tweets that is 329. Topic 3 having topic words as “lockdown”, “get”, “well”, “soon”, “speedy” “recovery”, etc. suggesting topic theme as “precautionary measures and recovery” have comparatively more proportion of positive tweets than other topics.

Figure 6 shows the analysis results for week 5 and week 6. Week 5 corresponds to date range 1<sup>st</sup> Jan 2022 - 7<sup>th</sup> Jan 2022 and week 6 corresponds to date range 8<sup>th</sup> Jan 2022 - 14<sup>th</sup> Jan 2022. Week 5 marks the beginning of new year 2022 and that was the time when the government of India imposed restrictions in various parts of the country. The most discussed topic in week 5 is topic 2 having topic words as “get”, “well”, “soon”, “lockdown”, “new”, “year”, “positive”, “case” suggesting tweets having discussion around lockdown, recovery, new year, and positive cases. While for almost all topics, the tweet sentiment distribution is like the number of negative tweets is around twice or more than twice the number of positive tweets, for topic 2, the number of positive tweets is 636 which is slightly less than the number of negative tweets which is 779. For week 6 also, topic distribution over sentiments is mostly negative than positive, except for topic 4 suggesting topic theme as “recovery” where the number of positive tweets, 86 is slightly less than the number of negative tweets, i.e., 93.

Figure 7 represents analysis results for week 7 and week 8. Week 7 corresponds to date range 15<sup>th</sup> Jan 2022 - 21<sup>st</sup> Jan 2022 and week 8 corresponds to date range 22<sup>nd</sup> Jan 2022 - 31<sup>st</sup> Jan 2022. For week 7, the most discussed topic is topic 2 having theme “vaccine in India”. The trend for all topics is the same that is negative tweets outnumber the positive tweets. For week 8, the most discussed topic is topic 3 having words “pandemic”, “vaccine”, “India”. Looking at the topic distribution among sentiments, topic 4 having “recovery” theme is seen to have around the same proportion for positive and negative tweets. Least number of positive tweets is for topic 5 having topic words related to “elections during pandemic”.



**Figure 3:** Distribution of sentiment labels across eight weeks starting from 1<sup>st</sup> December 2021 to 31<sup>st</sup> January 2022

## 5. Discussion

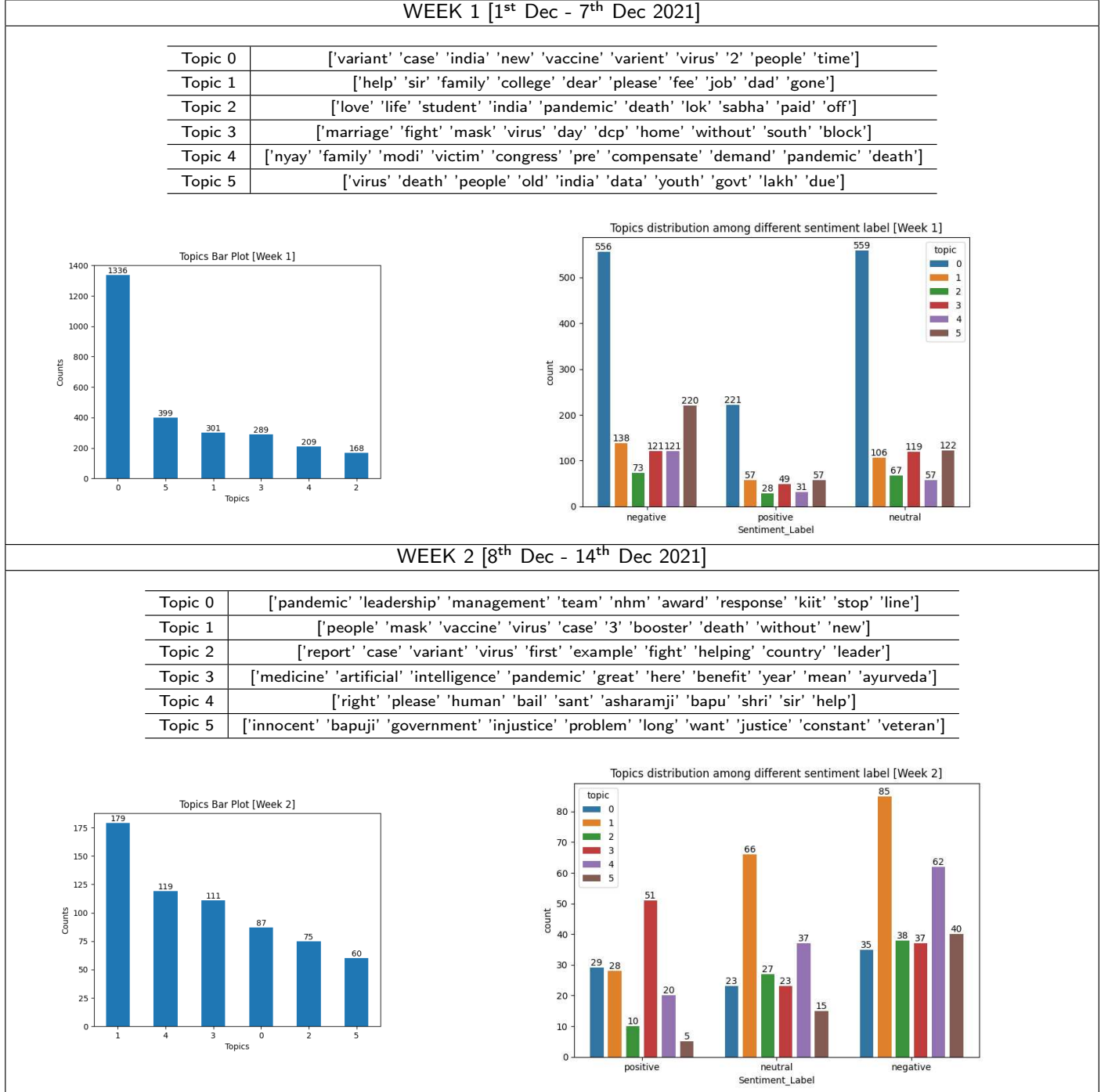
In this work, we have analyzed the Twitter data set related to COVID disease for two months December 2021 - January 2022. The most discussed topics and their associated number of negative and positive sentiment tweets of week-wise data are summarised in Table 2. Most discussed topics across weeks fall under the themes - “precautionary measures” and “Vaccine”. In the table, for all topics, the number of negative tweets is more than triple the number of positive tweets except for weeks 5 and 6, where the difference in the number of negative and positive tweets is not significant. It should be noted that this is due to the fact that in weeks 5 and 6, mostly precautionary measures tweets were posted.

A major challenge that we faced during the analysis of week-wise data was that the extracted words of different topic themes were overlapping each other, thus it was difficult to find concrete topic themes.

The analysis results show that people of India were having more negative sentiments over topics like new COVID variant, vaccines, booster dose, etc. during the third wave of COVID. It is observed that the most of the discussions on twitter across weeks were vaccine-related (See Table 2). Weekly sentiment labels for topic “vaccine” show that the tweets were mainly negative in general. Both COVID vaccine, Covaxin and Covishield had been proven to be effective as in the third wave very few people required hospitalizations those who were vaccinated. Also, vaccination drive of both vaccines, Covaxin and Covishield was started on 16<sup>th</sup> January 2021 and was done at a quite fast pace. By 1<sup>st</sup> December 2021, 33% of Indian population (1.39 Billion in 2021) was fully vaccinated and 24% was partially vaccinated with one dose [28]. The negative sentiments of people during the third COVID wave reflect low level of trust of people in the vaccine of COVID disease and the vaccination drive in India.



## Short Title of the Article

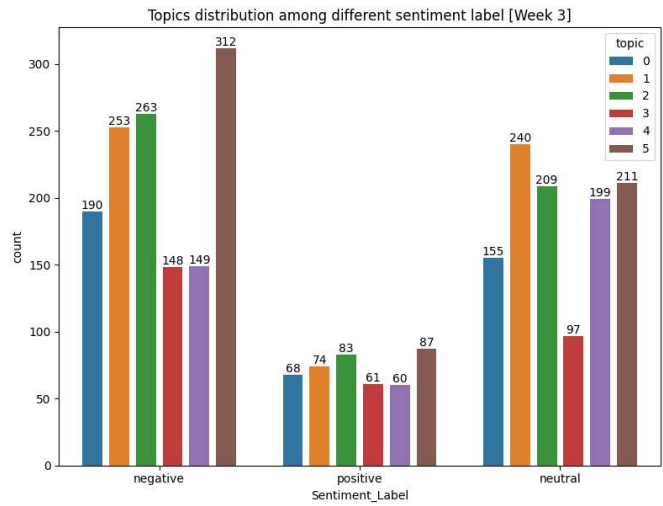
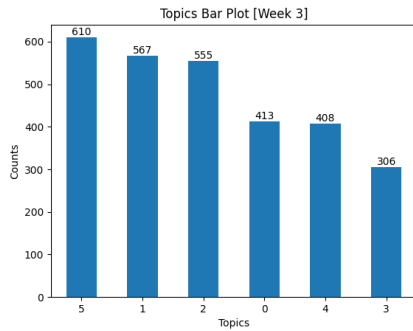


**Figure 4:** Topic distribution and topic-wise sentiment analysis of tweet data of week 1 and 2 of two month data set [1<sup>st</sup> Dec 2021 - 31<sup>st</sup> Jan 2022]

## Short Title of the Article

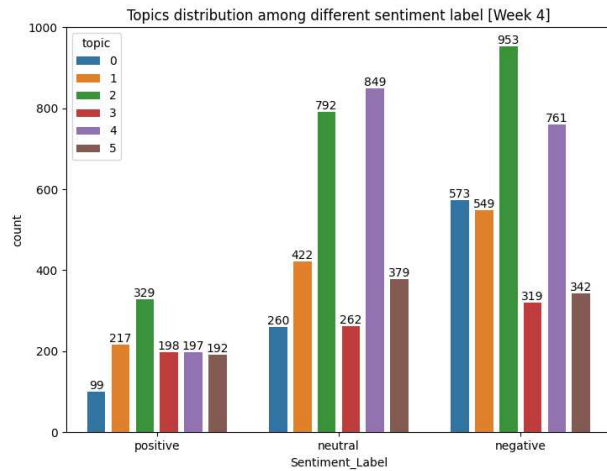
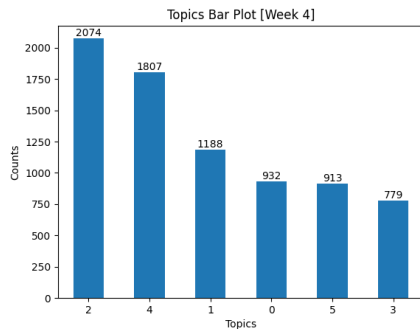
### WEEK 3 [15<sup>th</sup> Dec - 21<sup>st</sup> Dec 2021]

Topic 0	['vaccine' 'people' 'mask' 'need' 'india' 'u' 'careful' 'pandemic' 'jamat' 'time']
Topic 1	['case' 'variant' 'new' 'india' 'virus' 'vaccinated' 'fully' 'mask' 'state' 'report']
Topic 2	['please' 'india' 'also' 'note' 'mask' 'day' 'country' 'exam' 'time' 'vaccination']
Topic 3	['mask' 'time' 'vaccine' 'take' 'people' 'word' 'world' 'get' 'must']
Topic 4	['india' 'marriage' 'vaccine' 'case' 'new' 'time' 'virus' 'variant' 'u' 'tablighi']
Topic 5	['nhm' 'virus' 'worker' 'government' 'stop' 'mask' 'service' 'employee' 'pandemic' 'health']



### WEEK 4 [22<sup>nd</sup> Dec - 31<sup>st</sup> Dec 2021]

Topic 0	['sir' 'exam' 'student' 'online' 'aktu' 'offline' 'forcing' 'examination' 'want' 'still']
Topic 1	['mask' 'stay' 'safe' 'distancing' 'social' 'wear' 'please' 'home' 'protocol' 'virus']
Topic 2	['vaccine' 'year' 'india' 'hp' 'day' 'booster' 'world' 'dose' 'virus' 'people']
Topic 3	['lockdown' 'soon' 'get' 'well' 'hai' 'recovery' 'mask' 'speedy' 'govt' '7']
Topic 4	['case' 'variant' 'india' 'new' 'election' 'amid' 'virus' 'delhi' 'state' 'night']
Topic 5	['people' 'positive' 'pm' 'must' 'year' 'new' 'relief' 'said' 'modi' 'mask']

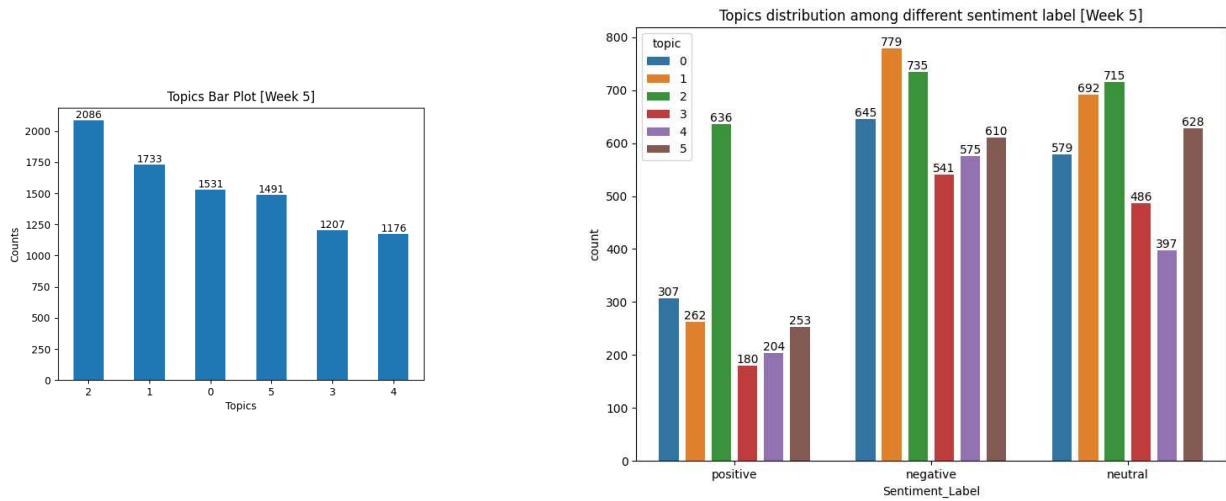


**Figure 5:** Topic distribution and topic-wise sentiment analysis of tweet data of week 3 and 4 of two month data set [1<sup>st</sup> Dec 2021 - 31<sup>st</sup> Jan 2022]

## Short Title of the Article

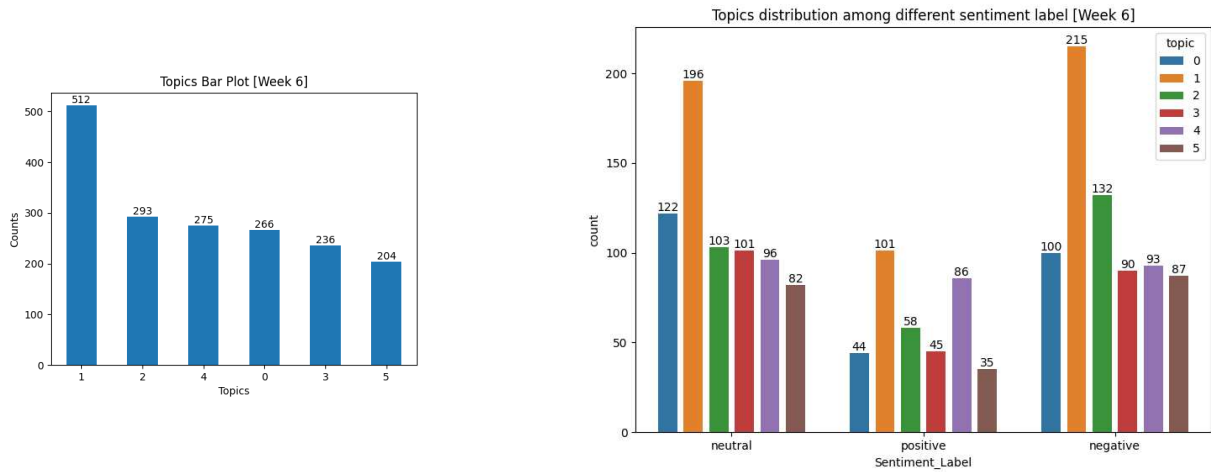
### WEEK 5 [1<sup>st</sup> Jan - 7<sup>th</sup> Jan 2022]

Topic 0	['mask' 'stay' 'safe' 'india' 'please' 'wear' 'home' 'time' 'day' 'variant']
Topic 1	['india' 'vaccine' 'people' 'case' 'positive' 'virus' 'test' 'day' 'go' 'u']
Topic 2	['soon' 'get' 'well' 'lockdown' 'take' 'sir' 'year' 'new' 'case' 'positive']
Topic 3	['case' 'new' '24' 'death' 'last' 'hour' 'india' 'report' 'mask' 'pandemic']
Topic 4	['year' 'college' 'time' 'new' 'people' 'virus' 'case' 'student' 'protocol' 'get']
Topic 5	['vaccine' 'vaccination' 'virus' 'dose' 'people' 'year' 'vaccinated' 'age' 'india' 'first']



### WEEK 6 [8<sup>th</sup> Jan - 14<sup>th</sup> Jan 2022]

Topic 0	['exam' 'positive' 'online' 'test' 'case' 'please' 'pandemic' 'time' 'day' 'health']
Topic 1	['get' 'soon' 'well' 'case' 'vaccine' 'student' 'sir' 'take' 'dose' 'mask']
Topic 2	['mask' 'stay' 'lockdown' 'safe' 'wear' 'india' 'please' 'people' 'public' 'without']
Topic 3	['india' 'case' 'vaccine' 'pandemic' 'pm' 'app' 'variant' 'day' '2022' 'mask']
Topic 4	['recovery' 'speedy' 'virus' 'wish' 'get' 'soon' 'wishing' 'well' 'ji' 'positive']
Topic 5	['virus' 'case' 'new' 'day' 'india' 'test' 'today' 'positive' 'take' 'like']

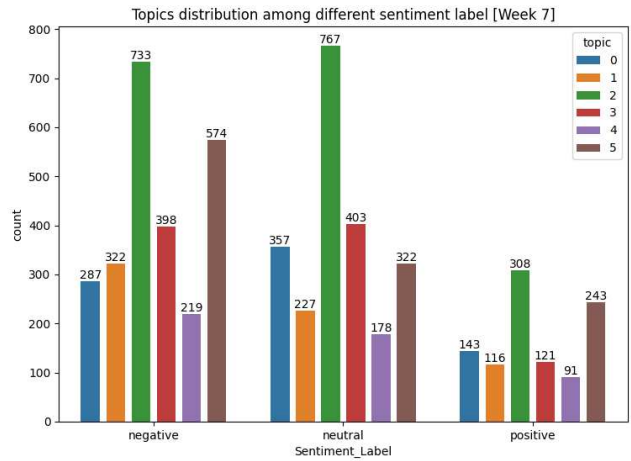
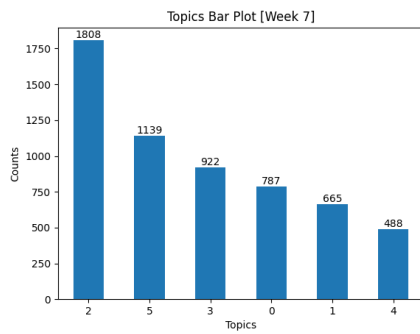


**Figure 6:** Topic distribution and topic-wise sentiment analysis of tweet data of week 5 and 6 of two month data set [1<sup>st</sup> Dec 2021 - 31<sup>st</sup> Jan 2022]

## Short Title of the Article

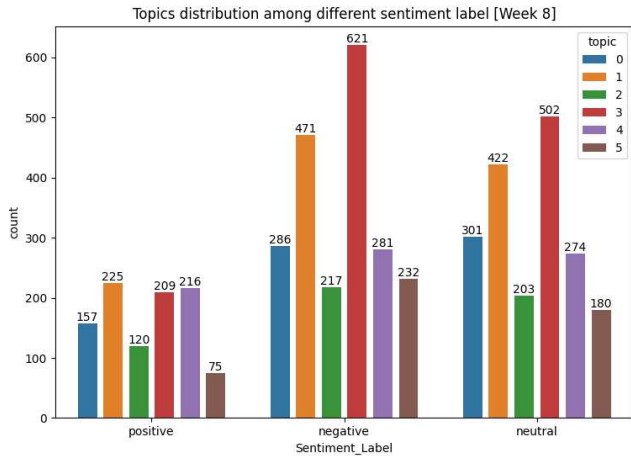
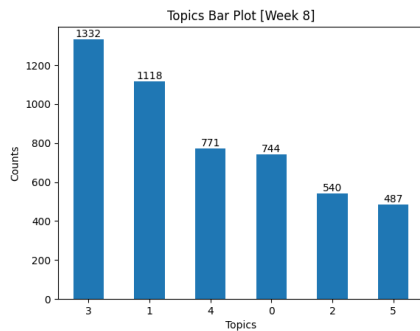
### WEEK 7 [15<sup>th</sup> Jan - 21<sup>st</sup> Jan 2022]

Topic 0	['case' 'exam' 'stay' 'want' 'respected' 'chief' 'minister' 'board' 'increase' 'online']
Topic 1	['sir' 'fast' 'get' 'request' 'pandemic' 'examination' 'wave' 'growing' 'people' 'u']
Topic 2	['vaccine' 'india' 'case' 'vaccination' 'year' 'dos' 'update' 'dose' 'day' 'today']
Topic 3	['lockdown' 'mask' 'vaccine' 'people' 'wear' 'virus' 'vaccinated' 'get' 'pandemic' 'even']
Topic 4	['virus' 'food' 'pandemic' 'agenda' '80' 'world' 'strength' 'free' 'delivered' 'citizen']
Topic 5	['exam' 'student' 'online' 'please' 'soon' 'take' 'well' 'get' 'sir' 'offline']



### WEEK 8 [22<sup>nd</sup> Jan - 31<sup>st</sup> Jan 2022]

Topic 0	['update' 'school' 'virus' 'new' 'death' 'time' 'must' 'case' 'year' '2022']
Topic 1	['mask' 'exam' 'stay' 'virus' 'please' 'u' 'take' 'safe' 'wear' 'home']
Topic 2	['mask' 'case' 'look' 'state' 'india' 'leadership' 'without' 'forward' 'new' 'united']
Topic 3	['pandemic' 'vaccine' 'india's' 'via' 'fight' 'people' 'doesn't' 'vaccinated' 'exist' 'year']
Topic 4	['soon' 'get' 'well' 'day' 'test' 'please' 'sir' 'positive' 'republic' 'india']
Topic 5	['virus' 'people' 'pandemic' 'new' 'case' 'death' 'election' 'student' 'help' 'protocol']



**Figure 7:** Topic distribution and topic-wise sentiment analysis of tweet data of week 7 and 8 of two month data set [1<sup>st</sup> Dec 2021 - 31<sup>st</sup> Jan 2022]

**Table 2**

Summary of analysis on most discussed topic in each week using week wise data set

Week	Most discussed topic	Negative tweets	Positive tweets
Week 1	New Variant in India	556	221
Week 2	Precautionary measures, Vaccine	85	28
Week 3	Work done by Govt for public health	312	87
Week 4	Vaccine	953	329
Week 5	Precautionary measures, New year, Positive cases	735	636
Week 6	Precautionary measures, Vaccine	215	196
Week 7	Vaccination in India	733	308
Week 8	Vaccine in India	621	209

## 6. Conclusion

This study identifies the topics and sentiments of Indian citizens about the Omicron-driven third wave of COVID-19 using the Twitter data. Analysis over the two-month Twitter data examines various topics discussed and the changes in these topics over time to understand better the trend of public opinion and perception about the third wave. It also examines the public sentiments about the different topics on the complete and weekly sectioned data set. Among the six distinct topics, the most discussed topics remained “precautionary measures” and “vaccine”. While the proportion of negative and positive sentiments over the topic “precautionary measures” is almost similar, negative sentiments outnumber the positive sentiments by a large extent over the latter topic. This suggests that though the third wave was considered to be the “mild” wave of COVID-19 in India, the people of India were still in fear of catching the disease again and thus were taking precautions. It should be noted that the developed COVID-19 vaccines, i.e., Covaxin and Covishield had proven to be quite effective, as from the people who were vaccinated, very few of them required hospitalizations. Also, the vaccination drive taken by the government of India was fast in pace. Still, the negative sentiments toward the topic discussions on “vaccine” reflect a low level of trust of Indians in either the efficacy of the developed vaccine to fight the new variant or the vaccination drive carried out in India or both. A more close attention to the tweets related to vaccine discussion is needed to understand this.

Such kind of study is extremely helpful for public health agencies to understand the major concerns of people and their varied reactions to different issues. In this study, public health agencies can refer to the distinct topic themes and their associated number of negative and positive tweets to understand the concerns of the general public of India and provide them with more information needed in case of topics where most people have reacted negatively.

In the future, we could expand this work to explore the emotion pattern and behavioral changes of people surrounding the third wave of pandemic across different countries.

## 7. Declarations

**Ethical Approval and Consent to participate:** Not applicable

**Human and Animal Ethics:** Not applicable

**Consent for publication:** Not applicable

**Availability of supporting data:**

The datasets analysed during the current study are available in the “COVID19\_Tweets\_Dataset” repository at [https://github.com/lopezbec/COVID19\\_Tweets\\_Dataset](https://github.com/lopezbec/COVID19_Tweets_Dataset)

**Competing interests:** The authors declare that they have no competing interests

**Funding:** Not applicable

**Authors’ contributions:** DV conceptualized the idea, extracted and interpreted the data set, performed methodological experiments, and analysed results. AY wrote the first draft. All authors interpreted the results, edited and approved the final manuscript.

**Acknowledgments:** Not applicable



**Authors' information:** DV is currently an Assistant Professor in the School of Computer Science Engineering and Technology, Bennett University, Greater Noida, Uttar Pradesh, India. She received her Ph.D. degree in Computational Biology from the Indian Institute of Technology Delhi, India. Her current research interest includes Systems biology, natural language processing, machine learning, and sentiment analysis.

AY is currently an Assistant Professor in the School of Computer Science Engineering and Technology, Bennett University, Greater Noida, Uttar Pradesh, India. She received her Ph.D. degree in Deep learning from Delhi Technological University, New Delhi, India. She has been awarded the “Research Excellence Award” by Delhi Technological University, Delhi, India, in the years 2020 and 2021. Her current research interest includes deep learning, natural language processing, machine learning, fake news detection, Deepfakes identification, Emotion Recognition, and sentiment analysis.

## References

- [1] Statista: Number of Twitter users worldwide from 2019 to 2024, <https://www.statista.com/statistics/303681/twitter-users-worldwide/>. Accessed: 22-02-2022.
- [2] WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020, <https://tinyurl.com/2p8fytjc>. Accessed: 22-02-2022.
- [3] WHO Coronavirus (COVID-19) Dashboard, <https://covid19.who.int/>. Accessed: 22-02-2022.
- [4] Q. B. Baker, F. Shatnawi, S. Rawashdeh, M. Al-Smadi, Y. Jararweh, Detecting Epidemic Diseases Using Sentiment Analysis of Arabic Tweets, JUCS - Journal of Universal Computer Science 26 (2020) 50–70.
- [5] A. Culotta, Towards Detecting Influenza Epidemics by Analyzing Twitter Messages, in: Proceedings of the First Workshop on Social Media Analytics, SOMA '10, Association for Computing Machinery, New York, NY, USA, 2010, p. 115–122. URL: <https://doi.org/10.1145/1964858.1964874>. doi:10.1145/1964858.1964874.
- [6] K.-W. Fu, H. Liang, N. Saroha, Z. T. H. Tse, P. Ip, I. C.-H. Fung, How people react to Zika virus outbreaks on Twitter? A computational content analysis, Am. J. Infect. Control 44 (2016) 1700–1702.
- [7] H. Liang, I. C.-H. Fung, Z. T. H. Tse, J. Yin, C.-H. Chan, L. E. Pechta, B. J. Smith, R. D. Marquez-Lameda, M. I. Meltzer, K. M. Lubell, K.-W. Fu, How did Ebola information spread on twitter: broadcasting or viral spreading?, BMC Public Health 19 (2019) 438.
- [8] Coronavirus disease (COVID-19) pandemic, <https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/novel-coronavirus-2019-ncov>. Accessed: 22-02-2022.
- [9] India's first Omicron cases detected in Karnataka, <https://www.hindustantimes.com/india-news/indias-first-omicron-cases-detected-in-karnataka-101638445884205.html>. Accessed: 23-02-2022.
- [10] A. Yadav, D. K. Vishwakarma, A Language-Independent Network to Analyze the Impact of COVID-19 on the World via Sentiment Analysis, ACM Trans. Internet Technol. 22 (2021).
- [11] M. E. Basiri, S. Nemati, M. Abdar, S. Asadi, U. R. Acharya, A novel fusion-based deep learning model for sentiment analysis of COVID-19 tweets, Knowledge-Based Systems 228 (2021) 107242.
- [12] I. Priyadarshini, P. Mohanty, R. Kumar, R. Sharma, V. Puri, P. K. Singh, A study on the sentiments and psychology of twitter users during COVID-19 lockdown period, Multimed. Tools Appl. (2021) 1–23.
- [13] S. Yousefinaghani, R. Dara, S. Mubareka, A. Papadopoulos, S. Sharif, An analysis of COVID-19 vaccine sentiments and opinions on twitter, International Journal of Infectious Diseases 108 (2021) 256–262.
- [14] S. Liu, J. Liu, Public attitudes toward COVID-19 vaccines on English-language Twitter: A sentiment analysis, Vaccine 39 (2021) 5499–5505.
- [15] Z. Bokaei Nezhad, M. A. Deihimi, Twitter sentiment analysis from Iran about COVID 19 vaccine, Diabetes & Metabolic Syndrome: Clinical Research & Reviews 16 (2022) 102367.
- [16] D. Thorpe Huerta, J. B. Hawkins, J. S. Brownstein, Y. Hswen, Exploring discussions of health and risk and public sentiment in Massachusetts during COVID-19 pandemic mandate implementation: A Twitter analysis, SSM Popul. Health 15 (2021) 100851.
- [17] S. Das, A. K. Kolya, Predicting the pandemic: sentiment evaluation and predictive analysis from large-scale tweets on Covid-19 by deep convolutional neural network, Evol. Intell. (2021) 1–22.
- [18] K. Mohamed Ridhwan, C. A. Hargreaves, Leveraging Twitter data to understand public sentiment for the COVID-19 outbreak in Singapore, International Journal of Information Management Data Insights 1 (2021) 100021.
- [19] S. Chekijian, H. Li, S. Fodeh, Emergency care and the patient experience: Using sentiment analysis and topic modeling to understand the impact of the COVID-19 pandemic, Health Technol. (Berl.) 11 (2021) 1073–1082.
- [20] C. A. Melton, O. A. Olusanya, N. Ammar, A. Shaban-Nejad, Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence, Journal of Infection and Public Health 14 (2021) 1505–1512. Special Issue on COVID-19 – Vaccine, Variants and New Waves.
- [21] K. Garcia, L. Berton, Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA, Applied Soft Computing 101 (2021) 107057.
- [22] C. E. Lopez, C. Gallemlere, An augmented multilingual Twitter dataset for studying the COVID-19 infodemic, Soc. Netw. Anal. Min. 11 (2021) 102.
- [23] M. Cliche, BB\_twtr at SemEval-2017 task 4: Twitter sentiment analysis with CNNs and LSTMs, in: Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017), Association for Computational Linguistics, Vancouver, Canada, 2017, pp. 573–580. URL: <https://aclanthology.org/S17-2094>. doi:10.18653/v1/S17-2094.

- [24] F. Rustam, M. Khalid, W. Aslam, V. Rupapara, A. Mehmood, G. S. Choi, A performance comparison of supervised machine learning models for COVID-19 tweets sentiment analysis, PLOS ONE 16 (2021) 1–23.
- [25] A. Akbik, T. Bergmann, D. Blythe, K. Rasul, S. Schweter, R. Vollgraf, FLAIR: An easy-to-use framework for state-of-the-art NLP, in: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations), Association for Computational Linguistics, Minneapolis, Minnesota, 2019, pp. 54–59. URL: <https://aclanthology.org/N19-4010>. doi:10.18653/v1/N19-4010.
- [26] Omicron peaked on January 21 with 3,47,000 daily cases, url<https://economictimes.indiatimes.com/news/india/omicron-peaked-on-january-21-with-347000-daily-cases/articleshow/89335010.cms>.
- [27] D. M. Blei, A. Y. Ng, M. I. Jordan, Latent dirichlet allocation, Journal of machine Learning research 3 (2003) 993–1022.
- [28] E. Mathieu, H. Ritchie, E. Ortiz-Ospina, M. Roser, J. Hasell, C. Appel, C. Giattino, L. Rodés-Guirao, A global database of COVID-19 vaccinations, Nature Human Behaviour 5 (2021) 947–953.

# An attention-based deep-learning system with fMRI functional connectivity optimized frequency EEG microstates classifies distinct temporal cortical communications of different cognitive tasks

Swati Agrawal

NMR Research Centre Institute of Nuclear Medicine and Allied Sciences (INMAS)

Vijayakumar Chinnadurai (✉ [vijayakumar.inmas@gov.in](mailto:vijayakumar.inmas@gov.in))

NMR Research Centre Institute of Nuclear Medicine and Allied Sciences (INMAS)

Rinku Sharma

Delhi Technological University

---

## Research Article

**Keywords:** Attention-based deep learning, Frequency-microstates, Hemodynamic functional connectivity, EEG informed fMRI. Multifrequency coupling, Target detection

**Posted Date:** April 15th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1550374/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Temporal analysis of global cortical communication of cognitive tasks in coarse EEG information is still challenging due to the underlying complex neural mechanisms. This study proposes an attention-based time-series deep learning framework that processes fMRI functional connectivity optimized quasi-stable frequency microstates for classifying distinct temporal cortical communications of the cognitive task. Seventy volunteers were subjected to visual target detection tasks, and their electroencephalogram (EEG) and functional MRI (fMRI) were acquired simultaneously. At first, the acquired EEG information was preprocessed, and the band passed to delta, theta, alpha, beta, gamma bands and then subjected to quasi-stable frequency-microstate estimation. Subsequently, time-series elicitation of each frequency microstates is optimized with graph theory measures of simultaneously eliciting fMRI functional connectivity between frontal, parietal, and temporal cortices. The distinct neural mechanisms associated with each optimized frequency-microstate were analyzed using microstate-informed fMRI. Finally, these optimized, quasi-stable frequency microstates were employed to train and validate the attention-based Long Short-Term Memory (LSTM) time-series architecture for classifying distinct temporal cortical communications of the target from other cognitive tasks. Based on the stability of transition probabilities of the optimized microstates, three different temporal sampling windows (200, 300, and 500 ms/segment) have been employed to collect the input microstates information for these time-series deep-learning systems. Our results revealed twelve distinct frequency microstates capable of deciphering target detections' temporal cortical communications from other task engagements. Particularly, fMRI functional connectivity measures of target engagement were observed significantly correlated with the right-diagonal delta ( $r=0.31$ ), anterior-posterior theta ( $r=0.35$ ), left-right theta ( $r=-0.32$ ), alpha ( $r=-0.31$ ) microstates. Further, neuro-vascular information of microstate-informed fMRI analysis revealed the association of delta/theta and alpha/beta microstates with cortical communications and local neural processing, respectively. The classification accuracies of the attention-based LSTM were higher than the traditional LSTM architectures. Particularly, the attention-based LSTM sampled microstates for 300 ms revealed a higher classification accuracy of 96% compared with other attention-based LSTM models with 200 ms (95%) and 500 ms (93%) temporal sampling windows. In conclusion, the study demonstrates reliable temporal classifications of global cortical communication of distinct tasks using an attention-based deep learning framework utilizing fMRI functional connectivity optimized quasi-stable frequency microstates.

## 1. Introduction

### 1.1 Decoding distant cortical communications from cortical EEG

The human brain dynamically engages distinct neural populations between distant brain regions, and their spatiotemporal oscillations often modulate systematically with behavioural and cognitive tasks. Many distantly located local brain circuitry performs specific localized jobs during such a neural

engagement. The synchrony or lack thereof between these remote brain regions brings effective global brain communication for information processing and is measured as functional connectivity dynamics in acquired task functional MRI and cortical EEG information. Many researchers (Kiehl and Liddle 2003; Laufs et al. 2003; Mantini 2007) employed simultaneous EEG information acquired with fMRI imaging to understand these functional connectivities, neural origins, and correlated brain states. These researches revealed that the elicitation of local neural systems is observed as high-frequency dynamics in EEG cortical oscillations. Cortical high-frequency gamma oscillation mainly explains local high-level neural information and positively correlates with the fMRI BOLD strength. The distant cortical and long-range coordination emerges as the lower EEG cortical frequency oscillation (Buzsáki 2009). The alpha and beta power modulate the BOLD response's latency and strength to gamma power changes (Magri et al. 2012). However, many researchers (Murta et al. 2017) subsequently observed that multi-frequency cortical EEG interaction explains hemodynamic task elicitation better than single EEG power information. JC Pang et al (Pang and Robinson 2018) revealed that the inverse correlation of alpha and BOLD originates from high- and low-frequency components of the same underlying neural engagement caused by modulation in corticothalamic and intra-thalamic feedback. Despite these clear insights of distinct global and local neural processing associated with every task engagement, decoding their temporal dynamics computationally from spatially coarse-grained cortical times-series EEG information is still a challenge.

## **1.2 Microstates and global cortical communication**

Microstates are the cluster centers unique in every cognitive task engagement in healthy and disease populations. Every microstate topography is associated with a "quasi-stable" functional state (Gschwind et al. 2016; Michel and Koenig 2018), explaining the brain's specific neural interaction. These quasi-stable patterns span around 100 ms are the most robust approach to bringing distant functional communication in cortical EEG information. Many researchers (Yuan et al. 2012; Khanna et al. 2015) observed that the time course of microstate metrics when correlated with the fMRI BOLD signal, reveals functional networks similar to the resting-state networks. Further, the microstate dynamics are observed to measure transitions between global cortical communications characterized by specific local neural alpha inhibitions (Milz et al. 2017; Croce et al. 2020; Kaur et al. 2020; von Wegner et al. 2021). The microstate estimation inherently employs the EEG frequency range (Koenig et al. 2002; Pascual-Marqui et al. 2014) of 2-20Hz, making their time-series dynamics explain the cortical alpha inhibitory/excitatory modulations (Milz et al. 2016). However, the cognitive task engagement's distant and local cortical communication manifests through other EEG rhythms (Hipp et al. 2011; Ribary et al. 2017). Thus, the task-induced modulation of cortical communication and associated local neural engagement are manifested as a combination of different frequencies (Akam and Kullmann 2014). More recently, the beta-band and the coverage feature of the EEG microstate analysis have been revealed as the essential features for the classification of epilepsy and PNES patients with reasonably high accuracy and precision (Ahmadi et al. 2020).

## **1.3 Deep learning approaches for time-series EEG analysis**



The EEG time-series information is higher-dimensional data, and the cognitive information is spread across its timelines. Hence, the feature information derived from a single time point of EEG time series data is inadequate to explain any cognitive process. Thus, Recurrent Neural networks (RNN) perform better in extracting sequential information embedded in higher dimensional EEG time-series information. However, the traditional RNN system suffers in learning long-term dynamics due to vanishing/exploding gradient problems. Long Short-Term Memory (LSTM) architecture addresses this exploding gradient obstacle by learning both long- and short-term dependencies. Recently, the attention mechanism has been introduced (Vaswani, A. 2017) to improve the performance of deep learning models; it highlights the more informative feature and subsequently gives higher weights to the corresponding original feature sequence. It has been embedded with the LSTM architecture in several EEG studies (Zhang et al. 2018, 2020b; Karim et al. 2019; Xie et al. 2019; Yan et al. 2019; Jin et al. 2020; Kim and Choi 2020; Rashid et al. 2020; Yao et al. 2020; Jiang et al. 2021; Zheng and Chen 2021) by effectively selecting the feature information and observed with significantly improved efficiency and performance accuracy of deep learning systems.

## **1.4 Present Study**

The present study proposes an attention-based LSTM computational model that employs optimized frequency microstates to decipher the distant cortical communication of visual target detection tasks. The temporal dynamics of the frequency microstate metrics are correlated with fMRI hemodynamic functional connectivity measures to optimize the cortical EEG quasi-stable frequency patterns with the local/global brain communication elicited by the task. The hemodynamic functional connectivity is assessed by employing the graph-theoretical analysis on simultaneously acquired fMRI information. The significantly correlated frequency microstates are further subjected to the robust correlation analysis to understand their multi-frequency coupling elicited during intercortical interaction during the task engagement. The local and global neural mechanisms underlying these frequency quasi-stable microstates were further estimated through EEG-informed-fMRI analysis. Finally, a hybrid deep learning framework consisting of LSTM with attention is employed to classify the target detection task engagement from the temporal dynamics of these optimized frequency microstate quasi-stable patterns. Six deep learning architectures (three LSTM and three attention-based LSTM networks) are trained with all optimized, quasi-stable frequency microstates collected from their associated time segments (200, 300, and 500 ms/segment, respectively) and passed to their respective input cells (15, 10, and 6) of the first LSTM layer. Six deeply layered architectures employed in this study possessed three stacked LSTM layers, and the final LSTM layer was integrated with an attention layer in three attention-based LSTM networks. The performance metrics such as precision, accuracy, and recall are estimated for all six deep learning architectures and validated using a 10-fold cross-validation approach.

## **2. Methods And Materials**

### **2.1 Participants**

Seventy healthy right-handed volunteers (30 males and 40 females; mean age: 23 years; age range, 20–32 years) were selected from the academic environment. All participants gave written informed consent and did not have any psychiatric, or neurological disorders or medication. The experiment was conducted following the World Medical Association (Declaration of Helsinki), and the local ethical committee approved all measurements. The vision of participants was corrected using MR-compatible lenses whenever required.

## 2.2 Task design

The visual target detection task comprised of a sequence of simple geometrical shapes such as squares, circles, stars, and triangles filled with primary colors were presented, as shown in Fig. 1. The paradigm consisted of 105 sub-trials over five trials, where 32 target stimuli are distributed pseudo-randomly over the sub-trials. Each trial started with the presentation of a 'Target' object for 3 seconds. In subsequent slides, a collection of geometric shapes, one item at a time, is presented for 3 seconds to the volunteer. The volunteer's task was to detect the previously shown target and respond quickly by pressing a button with the right thumb. A single cross in the slide's center is presented for 3 seconds during the fixation period following each task stimulus. All the volunteers are instructed to avoid any motor response during fixation and distractor stimuli. These task stimuli are projected onto MR compatible lenses fitted on the head coil inside MR.

## 2.3 Simultaneous EEG-fMRI acquisition

In the present study, simultaneous EEG-fMRI data acquisition is carried out using a 3T Siemens Magnetom Skyra scanner (Siemens, Erlangen) and MR compatible 32-channel Brain Amp system with an EEG cap. EEG signals are recorded at a sampling rate of 5 kHz, and the impedance of all scalp electrodes is maintained below five kOhms throughout the recording.

The fMRI were acquired using Echo Planar Imaging (EPI) sequence with TR = 3000 ms; TE = 36 ms; voxel size = 3.6\*3.6\*3.0 mm; matrix = 64\*64; FoV read = 230 mm; flip angle = 90°, 36 axial slices. The axial slices are acquired parallel to the Anterior-Posterior (AC-PC) line in an interleaving manner with a slice thickness of 3.0 mm. The high-resolution structural images of the brain are acquired using T1 MPRAGE sequence with parameters, voxel size = 1.0\*1.0\*1.0 mm; TR = 2000 ms; flip angle = 90°; FoV = 240 mm; matrix = 512\*512; slice thickness = 1.0 mm, 160 axial slices. The axial slices are acquired parallel to the Anterior-Posterior (AC-PC) line in an interleaving manner with a slice thickness of 5.0 mm.

## 2.4 Data preprocessing

### 2.4.1 EEG data

The detailed pipeline employed in this study for cleaning both MRI artifacts and other artifacts is the same as our earlier study (Kaur et al. 2020). It employs the FMRIB plugin of EEGLAB that uses combined adaptive thresholding (Christov 2004) and the Teager energy operator (Kim et al. 2004), and the Harvard Automated Processing Pipeline for EEG (HAPPE) (Gabard-Durnam et al. 2018) for noise-free time-

frequency analyses. Then, the data is downsampled to 250 Hz and then re-referenced to the common average reference. Finally, the artifact corrected data of all the participants are then segregated specific to the target, distractor, and fixation blocks.

## **2.4.2 fMRI data**

The fMRI data preprocessing is performed using Statistical Parametric Mapping version 12. The data is corrected for slice-timing differences and spatially realigned and excluded if movement exceeds 3 mm. It follows by registering the functional scans to standard MNI template space. The preprocessed images are subjected to spatially smoothing with 5\*5\*5 mm full width half maximum Gaussian kernel.

## **2.5 Data Analysis**

The present study develops a computational framework that classifies distinct task engagement's temporal global cortical communication through unique temporal EEG quasi-stable information that decodes the neural basis of the distant cortical communications. Figure 2 explains these processes in detail. The following section will elaborate on each one of these steps in detail.

### **2.5.1 Estimation of frequency-microstates and their association with task's cortical communications**

At first, artifact-corrected EEG data is band-pass filtered to segregate it into frequency-band limited data comprising of (1–4) Hz for delta, (4–8) Hz for theta, (9–14) Hz for an alpha, (15–35) Hz for beta and (35–48) Hz for gamma. Then, the Global Field Power (GFP) of each frequency information is computed and subjected to the modified K-means clustering algorithm (Pascual-Marqui et al. 1995) to identify every frequency microstate topographic prototype. A detailed description of this estimation is given in the supplementary file (section S.1). Then, each frequency microstate prototype is back fitted in every individual's data and estimated re-expressed sequences of microstate classes. Finally, statistics about the sequence of microstate classes, such as their frequency of occurrence or average duration, are calculated. This quasi-stable frequency-microstate patterns elicitation information is subsequently mapped with global functional connectivity of each task engagement assessed from simultaneously acquired fMRI information.

#### **Global functional connectivity estimation from simultaneously acquired fMRI information**

At first, General Linear Model-based analysis was employed using simultaneously acquired fMRI information to identify the task's neural correlates. At the subject level analysis, BOLD responses of each task engagement (target, distractor, and fixation) are modeled by a canonical hemodynamic response function with temporal and dispersion derivatives with six realignment parameters for each run. In second-level GLM modeling, group average maps were computed using one-sample t-tests, cluster corrected ( $p < 0.05$ ) across subjects. Subsequently, the second-level GLM model results for every task engagement are passed as Regions of Interest (ROI) to graph theory analysis to estimate the global/local functional connectivity (Whitfield-gabrieli and Nieto-castanon 2012). A detailed description of this

estimation is given in the supplementary file (section S.2). Finally, the graph theory metrics such as global and local efficiency of the functionally connected regions are estimated for each subject, and the ROI-to-ROI connectivity matrix is thresholded at  $p\text{-FDR} < 0.05$  in a two-sided analysis.

### **Optimization of frequency-microstate elicitations with fMRI functional connectivity measures and their validation through Microstate informed fMRI**

This study examines whether cortical quasi-stable frequency EEG elicitation could be employed to explain temporal global cortical communications of different task engagements. For this purpose, the number of occurrences of each delta, theta, alpha, beta, and gamma EEG-microstates of every individual during task engagement (target, distractor, and fixation) is subjected to the robust correlation with global functional connectivities metrics measured from the simultaneously measured fMRI information.

This study further validated these significantly correlating frequency microstates by subjecting them to the EEG-informed fMRI analysis (Huster et al. 2012; Abreu et al. 2018) and studied their neural mechanisms. Since the frequency microstates that correlate with fMRI functional connectivity metrics are different for target, distractor, and fixation, three separate EEG-informed fMRI models were constructed for every task engagement. A detailed description of this estimation is given in the supplementary file (section S.3).

## **2.5.2 Classification using attention-based LSTM deep learning model**

This study primarily aims to develop a computational framework, attention-based Long Short-Term Memory (LSTM), to classify the global cortical communication using frequency quasi-stable oscillations associated with the global brain communications observed in fMRI functional connectivity. LSTM has been a popular recurrent neural network for learning sequential features in time series data and classifying EEG information (Wang et al. 2018; Nagabushanam et al. 2020). The attention mechanisms have been recently conceptualized (Vaswani, A. 2017) and integrated with the LSTM framework and found helpful in classifying the data that involves remembering and aggregating feature embeddings in time-series information (Zhang et al. 2020a). Due to the discrepant and task-dependent nature of the task's EEG information, attention-based LSTM architecture is more suited and employed in our study for classifying the task's EEG signals. Hence, at first, each frequency-microstate strongly associated with fMRI functional connectivity measures was correlated with every individual preprocessed EEG information. Then, the feature vector consisting of the correlation value for each quasi-stable frequency microstates belonging to every task engagement is formed and used as a training, testing, and validation dataset for the attention-based LSTM deep learning model. A detailed description of the deep attentional LSTM Model and the fine-tuning process is given in the supplementary file (section S.4). Figure 3 clearly explains these processes in detail. Table 1 provides the information on the fine-tuning parameter adopted for the deep learning model.

As the sampling rate of the EEG data is 250Hz, 250 feature vectors were compiled for every second, with a single vector covering every 4 ms neural activity. Each task engagement was carried out for 3 seconds; there are 750 feature vectors staggered for every task block. To avoid feeding too many features to the architecture, feature vectors are combined across task blocks, and three distinct segregation of input feature vectors (200ms/segment, 300ms/segment, and 500ms/segment) belong to every task block were explored. The study independently employed three distinct attention-based LSTM architectures in input layers with 15, 10, and 6-LSTM cells. The final LSTM layer of all the three architecture was ensued by an attention layer, succeeded by a fully connected layer having a sigmoid activation function to predict the probability of each task engagement. Finally, to estimate the efficacy of attention mechanisms in the original LSTM system, all the three attention-based LSTM architectures were compared with their attention counterpart.

Table 1  
Final Tuning parameter of LSTM model.

Hyperparameters	Tuned parameters
Hidden layer size	256
Batch size	64
Training epoch numbers	1000
Rate dropout	Input Layer: 0, 1st LSTM Layer: 0.2, 2nd LSTM Layer: 0.1, 3rd LSTM Layer: 0.2
Recurrent depth	3
Learning rate	0.001

**Validation approach:**

Finally, the proposed deep learning architecture is validated by employing a 10-fold cross-validation approach with no overlap of training and testing segments. True Positive (TP), True Negative (TN), False Negative (FN), and False Positive (FP) were used to calculate the performance metrics. They are formulated as Precision = TP/(TP + FP), Accuracy = (TP + TN)/(TP + TN + FP + FN) and Recall = TP/(TP + FN).

### 3. Results

The present study presents the times series computational frameworks that classify different task engagement based on temporal modulation of distant brain communications through optimized frequency EEG microstates. The frequency EEG microstates were optimized by associating them with simultaneously eliciting distant hemodynamic functional connectivity measures. Further, the study explored the EEG-informed fMRI approach that has been employed to understand the insights into the



neuronal mechanisms associated with frequency microstates that correlate with global task communications. The following sections present the results of each of the above steps in detail.

### **3.1 Frontal, parietal, and temporal cortical interaction elicited during different task engagements:**

The neural correlates associated with each task engagement are assessed through GLM models in fMRI information with the double-sided t-test of  $p < 0.5$ , FDR corrected. The results suggest that target engagement enhanced the hemodynamic response in the frontal (frontal orbital cortex (FOC), frontal pole (FP), superior frontal gyrus (SFG), parietal (angular gyrus (AG), Precuneus cortex (PC)), and temporal (Inferior and middle temporal Gyrus (ITG, MTG)) cortices. Further, it involves significant engagement of the cingulate gyrus (CG), lateral occipital cortex, occipital pole (OP), paracingulate gyrus, and insular cortex (IC) regions. The neural correlates of distractor and fixation have also revealed distinct intercortical engagement. The detailed list of neural correlates of each task engagement is tabulated in Supplementary Table (ST.1).

The presence of this distinct inter cortical communication during each task engagements are further supported by the graph theoretical functional connectivity metrics such as global (GE) and local efficiency (LE). The target engagement significantly engaged inter and intra-cortical communication at frontal cortex (frontal pole [GE:0.893, LE:0.89], frontal orbital cortex [GE: 0.886, LE:0.89], and superior frontal gyrus [GE:0.886, LE:0.894]), parietal cortex (left angular gyrus [GE:0.872, LE:0.893], right angular gyrus [GE:0.9, LE:0.887], precuneus cortex [GE:0.87, LE:0.894]) and temporal cortical regions (inferior temporal gyrus [GE: 0.88, LE: 0.889], middle temporal gyrus [GE:0.889, LE: 0.891], temporal occipital fusiform cortex [GE:0.88, LE:0.89]) at  $p < 0.05$  with FDR correction.

Functional connectivity elicitation remained distinct during this target detection task and had minimal overlap in neural mechanisms, regions, and intracortical interaction with each task engagement. Notably, frontoparietal and frontotemporal interaction engagement was distinct at the neuronal level for target, distractor and fixation engagement. The detailed information of graph-theoretical measures estimated for each task engagement is given in the supplementary file (section S.2).

### **3.2 Quasi-stable frequency-microstates and their association with task's fMRI functional connectivity measures**

As mentioned in the data analysis section, the preprocessed band passed EEG information of every task engagement (target, distractor, fixation) is subjected to the frequency-microstate estimation. Four dominant frequency microstates are estimated for every EEG frequency band of each task engagement. Figure 4 illustrates spatial topographical patterns of each frequency-microstate topography associated with every task engagement.

Further, the number of occurrences of every task's frequency-microstate prototype is estimated by back-fitting these frequency-microstates to every volunteer's respective frequency information. Figure 5 compares the mean number of occurrences of all volunteers of each frequency-microstate across every task. Most of the researchers in the literature use the alphabetical approach to label each microstate. This nomenclature brings difficulty in following similar observations across the literature. Hence, in this study, the unique directional pattern of each microstate's moderate activation band (green colour band) is used for labelling these microstates. In this study, based on the microstate's green colour band direction, they are labelled as anterior-posterior (AP), left-right (LR), left diagonal (LD), and right diagonal (RD) microstate prototypes. For example, the delta microstate with a green band travelling between anterior-posterior is called the "anterior-posterior delta microstate". As the study primarily focuses on identifying the frequency-microstate that manifests the cortical communication elicits during the task engagement, every frequency microstate's number of occurrences is robustly correlated with the fMRI functional connectivity measures. The significantly ( $p < 0.01$ ) correlating frequency microstates were colour-coded (+ve correlation: green, -ve correlation: red) in Fig. 5.

The significantly associated, optimized frequency microstates with graph-theoretical measures of every task engagement are illustrated in Fig. 6. A total of twelve frequency microstates (target: 4, distractor: 4 and fixation: 4) significantly correlated with the functional connectivity measures of simultaneously acquired fMRI information. During target engagement, the number of occurrences of right diagonal delta-microstate positively correlates with the global efficiency of fMRI functional connectivity measures. On the other hand, the local efficiency of fMRI connectivity measures correlates negatively with both left-right theta and alpha microstate occurrences and positively with anterior-posterior theta microstate.

### **3.3 Neurovascular analysis of optimized frequency microstates: EEG Informed fMRI analysis**

Engagement of task generally elicits local neural clustering (high-frequency quasi-stable oscillations) at distinct brain regions, responsible for efficient local information processing, together with distant cortical intercommunication to facilitate global communication (low-frequency quasi-stable oscillations). The neuro-vascular analysis through EEG-informed fMRI explains these insights through synchronizing neural information (optimized quasi-stable EEG oscillations) with hemodynamic information (vascular) that is elicited from a specific task engagement. For this purpose, each one of twelve optimized frequency microstates is processed in independent EEG-informed fMRI models, which modelled each microstate as independent regressors ( $p < 0.01$ , FDR corrected) to estimate their neuro-vascular information. Supplementary figures (SF.1(a-c)) show neuro-vascular information of each twelve optimized microstates.

Figure 7 shows the neuro-vascular coupling of each significant frequency-microstate at the neural correlates of the respective task engagement. The figure also shows the functional connectivity between each neural correlate of task engagement. It is further evident from Fig. 7 that fMRI functional connectivity optimized frequency microstate associates with almost most of the brain regions involved in

each task engagement. Target engagement elucidated the right-diagonal delta-microstate synchronously with the BOLD response in the frontal cortex. Further, anterior-posterior and left-right theta-microstates are found to synchronize with the BOLD response of frontal, temporal, parietal, and occipital regions except for right-lateralized SFG and AG. The role of theta-microstate in BOLD-synchronization of these regions characterizes its significant association with target engagement. The alpha-microstate is observed de-synchronizing with the parietal, PCG's BOLD response, and synchronizing with the frontal, occipital region during target engagement. Our findings also reveal the effects of multiple frequencies on specific brain regions, such that the BOLD response of FP, PCG, SFG, IC, PC, and AG are modulated with delta, theta, and alpha-microstates. Specifically, the study observes the de-synchronization of alpha-microstate with the delta and theta microstates such as PCG, SFG, PC, and synchronization between delta and theta microstates as IC and FP. Hence, the relationships mentioned above reveal that the multi-frequency interactions modulate the BOLD response of task-engaged brain regions.

### **Performance of Deep Attentional LSTM Model:**

The performance of the proposed deep attentional model revealed that hybrid deep learning architecture allows it to apply the attention layer that finds meaningful patterns using LSTM by overcoming the fixed-length input sequences. Figure 8 compares the performance metrics, precision, accuracy, and recall of all the six deep learning models (LSTM and attention-LSTM) in classifying the different task engagement based on the stacked feature vectors that consist of correlation information of each optimized frequency microstate. Three distinct segregation of input feature vectors (200ms/segment, 300ms/segment, and 500ms/segment) with 15, 10, and 6-LSTM cells in input layers were analyzed within the context of each type of network (LSTM and attention-LSTM). As can be observed, LSTM combined with attention appears to perform significantly better in three of the performance metrics. In this case, the choice of nodes (15, 10, and 6) that depends on the EEG time series window (200ms, 300ms, and 500 ms) seems reasonable since it reflects a significant variation in the precision, accuracy, and recall. The results reflected improved performance for using ten nodes from six nodes; a significant decline in performance metrics is observed for using 15 nodes from 10 nodes. In this regard, the proposed attention-based LSTM architecture has proven to enhance the model performance. However, the choice of nodes for the first layer of LSTM architecture leads to variation in deep learning architecture performance.

## **4. Discussion**

The present study brings more insights into understanding and optimizing frequency microstate information estimated from cortical, coarse EEG information with distant fMRI functional connectivity measures associated with different task engagements and then utilizes them to train a times series of deep learning frameworks. The study employs an attention-based stacked LSTM for effectively remembering and aggregating feature embeddings in the time-series classification of stacked temporal dynamics of the frequency microstate quasi-stable patterns. The results reveal that combining quasi-stable frequency microstates (optimized by the functional connectivity) with an attention-based LSTM algorithm better classifies target engagement, distractors and fixation. Further, the study has also

employed EEG informed fMRI to understand the optimized frequency microstates' neuro-vascular insights and mechanisms. The following section discusses those observations in detail.

## **4.1 Association of distinct frequency microstates with each task's cortical functional connectivities**

The study observed four distinct cortical frequency-microstates elicitations for each of three different task engagements (thus, twelve distinct frequency microstates) correlated strongly with their hemodynamic functional connectivity measures independently. Nonoverlap of association of these frequency microstates clearly states the underlying difference in distant cortical communications associated with each task engagement. In addition, the study also has observed a distinct neuro-vascular functional association of slow and faster oscillations in the brain regions involved in each task's cortical communications. Some of the essential observations related to these multi-frequency quasi-stable EEG frequency associations with each task's distant neural interactions are summarized below.

### **Elicitations of slower and faster quasi-stable microstates oscillations and their sync/desynchronization with task's global/local neural engagements:**

Complex cognitive engagement requires global interactions of different brain regions enabling the large-scale integration of local neural information. The integration of neural engagement of spatially distant regions constituting the large-scale networks is primarily moulded with low-frequency synchronized oscillations due to their long-range communications and integrative roles in various brain functions.

Further, the strength of large-scale networks has also been found highest for lower frequencies and seen gradually decreasing with increases in the frequency range (Wu et al. 2008; Gohel and Biswal 2015; Li et al. 2015). Therefore, the dominance of large-scale networks for the lower range of frequencies confirms their functional significance. The present study's results (Fig. 6) distinctly revealed synchronization (positive correlation) of the slow frequency oscillations with global, distant brain communication and desynchronized (negative correlation) with local neural elicitation across all the task engagement. Similarly, the high-frequency quasi-stable oscillation distinctly synchronized with the elicitation of local neural engagement and desynchronized with global communications across all the task engagements. A similar observation (Fig. 7) is revealed in the neurovascular coupling from EEG-informed fMRI analysis at the brain regions engaged in all tasks. They demonstrated significant desynchronization between slow and faster quasi-stable oscillations in most brain regions involved during every task engagement. Further, the region-wise neuro-vascular insights brought slower quasi-stable oscillations associations with large-scale frontoparietal and frontotemporal functional networks. These observations are supported by the proposal of Polich et al (Polich 2007) and Harper et al, (Harper et al. 2017) explaining the role of delta and theta band activity underlying the frontoparietal and frontotemporal functional networks.

The accumulating literature suggests the association of local neuronal processing with the global cortical communication between neural assemblies by coupling multiple oscillatory frequencies (Canolty and

Knight 2010) and referred to as cross-frequency coupling. The most well-studied example of cross-frequency coupling is the theta-gamma coupling, which explains the engagement of gamma frequency in certain phases of theta cycles. Further, (Schroeder and Lakatos 2009) suggested that low-frequency oscillations may provide an essential role in engaging gamma rhythms during attention. (Lisman and Jensen 2013) discussed the relationship of alpha and theta frequency oscillations in the cortex and revealed the possibility of the theta-gamma code's contribution to memory and sensory processes. Thus, the functional interaction of more extensive networks oscillating at lower frequencies and local neuronal ensembles oscillating at higher frequencies has been revealed for cortical communication and integration (Fell and Axmacher 2011; Lisman and Jensen 2013; Knyazev et al. 2019). Hence, the studies mentioned above better understand the modulation of faster oscillations from the slow EEG oscillations during cognitive engagements.

The present study's quasi-stable cortical oscillation's neuro-vascular insights and their association with local and global neural information are consistent with these studies and explain the role of multi-frequency interactions in explaining the complex cognitive engagements' neuronal mechanisms.

## **4.2 Performance of attention-based deep learning system in classifying time series quasi-stable cortical frequency EEG information**

All six proposed deep learning architectures demonstrated in Fig. 3 revealed comparable accuracy, precision, and recall rate for classifying target detection task engagements. However, two primary aspects distinguished the performance of each proposed deep learning architecture. They are the temporal sampling window and incorporation of the attention mechanism.

### **4.2.1 Temporal sampling window of quasi-stable frequency information optimizes the performance of deep learning architectures**

Temporal sampling window size plays a significant role in achieving the best classification accuracy of deep learning architectures. This crucial aspect of time-locked microstates events facilitated a better temporal data handling of multi-frequency interaction in the cognitive task's neural engagements. The present study employed three different windows of temporal sampling (six nodes: 500ms, ten nodes: 300 ms, and fifteen nodes: 200 ms) to explore optimized, quasi-stable microstate's ability to classify task engagement. The highest classification accuracy, up to a 96% accuracy for ten node attention-based architecture, confirms the temporal dynamics of quasi-stable frequency oscillations optimally with 300ms. Several target identification-related EEG studies support this observation (Bledowski et al. 2004; Bansal et al. 2014; Sclocco et al. 2014; Pandey et al. 2016; Arvaneh et al. 2019) revealed that the peak of task event-related potentials following the stimulus onset in between 100 ms to 300 ms at multiple.

Specifically, a recent EEG study (Harper et al. 2017) showed event-related synchronization of theta and delta bands occurring around 300 ms after the onset of target stimuli.

## **4.2.2 Effect of attention mechanism with traditional LSTM architectures**

The overall better performance of all the deep learning architectures justified the optimization of quasi-stable frequency-microstate information. However, the attention mechanisms further improved this in the deep learning system. The combined effect of attention phenomena and the LSTM architecture allowed the deep learning architecture to dynamically emphasize the task-relevant neural information in the time series sequence of EEG data and give less attention to other irrelevant information. The improvement in the performance of the attention-based LSTM system can be seen precisely in Fig. 8. The plots suggest that the attention mechanism is optimum for extracting the most relevant task neural features and improves the LSTM's performance compared to independent LSTM.

## **5. Significance Of The Study**

There is extensive literature (Supplementary Table (ST.2)) on employing a deep learning approach to decode task engagement using EEG elicitations. However, most of those works are restricted to the sensory-motor tasks (hand, leg movements, imagery tasks) whose engagement can be localized in a few cortical regions. However, minimal research is engaged to decode the cognitive task engagement's functional connectivity of distant and distinct cortical engagement using the deep learning framework. Further, despite many research studies that microstates are a promising neural signature, their association with the neural mechanisms of task engagement is still not clearly understood (Stam 2005; Tognoli and Kelso 2009; Kim et al. 2021). In addition, not many works in the literature explain the quasi-stable nature of the different EEG frequency oscillations either. The present work employs attention-based LSTM architecture to decode the temporal dynamics of cognitive task engagement through fMRI functional connectivity optimized frequency microstates. Recently, (Sikka et al. 2020) investigated the temporal dynamics of traditional microstates using recurrent neural networks. However, their work did not address the quasi-stable frequency microstate's neural mechanism and modulation during task engagement. Our present work further brings more insights into the attention mechanism's ability to improvise the classification of cognitive task engagement based on the optimized neural signatures. To our current knowledge, the present study is one of the few works that employs simultaneous EEG-fMRI information to optimize the neural signatures for improvising the performance of deep learning architectures.

## **6. Conclusions**

The present study proposes an attention-based deep learning framework that processes temporal dynamics of the twelve distinct, fMRI functional connectivity optimized, quasi-stable frequency microstates to classify different cognitive task engagement. It further utilizes neurovascular insights of



these optimized frequency microstates through EEG-informed fMRI analysis to understand the local and distant cortical interaction revealed by the optimized frequency microstate. This optimized neural information was passed as input information at three distinct temporal samplings windows (200, 300, and 500 ms/segment) to train and validate the attention-based LSTM architecture. The results suggest that the classification accuracies of the attention-based LSTM architectures were better than the traditional LSTM architectures due to the ability of the attention mechanisms in deep learning systems in localizing temporal feature information. Notably, the attention-based LSTM model with 300 ms temporal sampling revealed a higher classification accuracy than other architectures. Hence, the study demonstrates an attention-based deep learning framework to perform a robust classification of complex, distant cortical engagement and communication caused by cognitive task engagements based on the novel, quasi-stable frequency microstates.

## Declarations

**Availability of data and materials:** The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

**Competing interests:** The authors declare that they have no competing interests.

**Funding:** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Authors' contributions:** The study was designed by SA and VC. Data was acquired and processed by SA. Data was interpreted and analyzed by VC, SA, and RS. The manuscript was written by SA with the help of VC. All authors reviewed the manuscript.

**Acknowledgments:** Not applicable.

## References

1. Abreu R, Leal A, Figueiredo P (2018) EEG-informed fMRI: A review of data analysis methods. *Front Hum Neurosci* 12:1–23. <https://doi.org/10.3389/fnhum.2018.00029>
2. Ahmadi N, Pei Y, Carrette E, et al (2020) EEG-based classification of epilepsy and PNES: EEG microstate and functional brain network features. *Brain Informatics* 7:. <https://doi.org/10.1186/s40708-020-00107-z>
3. Akam T, Kullmann DM (2014) Oscillatory multiplexing of population codes for selective communication in the mammalian brain. *Nat Rev Neurosci* 15:111–122. <https://doi.org/10.1038/nrn3668>
4. Arvaneh M, Robertson IH, Ward TE (2019) A P300-Based Brain-Computer Interface for Improving Attention. *Front Hum Neurosci* 12:. <https://doi.org/10.3389/fnhum.2018.00524>

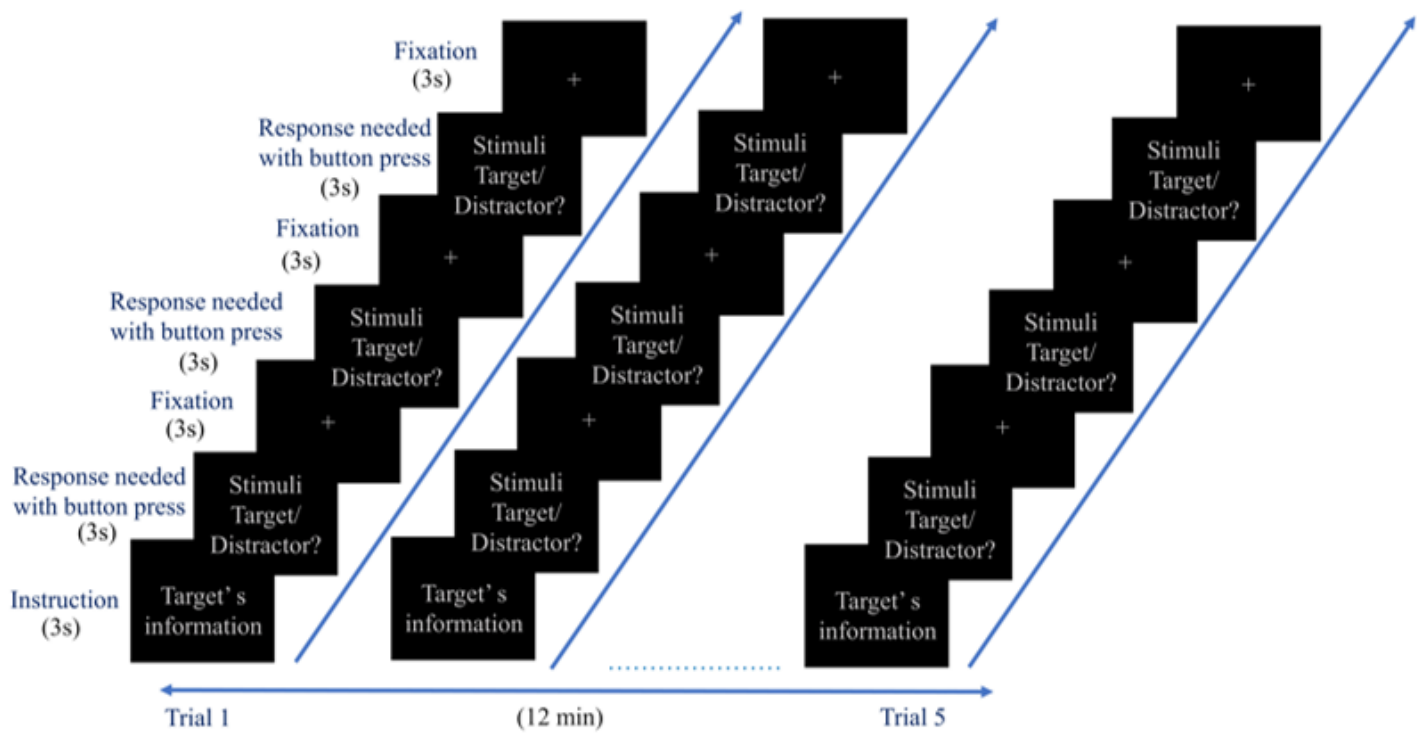
5. Bansal AK, Madhavan R, Agam Y, et al (2014) Neural dynamics underlying target detection in the human brain. *J Neurosci* 34:3042–3055. <https://doi.org/10.1523/JNEUROSCI.3781-13.2014>
6. Bledowski C, Prvulovic D, Hoechstetter K, et al (2004) Localizing P300 generators in visual target and distractor processing: A combined event-related potential and functional magnetic resonance imaging study. In: *Journal of Neuroscience*
7. Buzsáki G (2009) *Rhythms of the Brain*
8. Canolty RT, Knight RT (2010) The functional role of cross-frequency coupling. *Trends Cogn Sci* 14:506–515. <https://doi.org/10.1016/j.tics.2010.09.001>
9. Christov II (2004) Real time electrocardiogram QRS detection using combined adaptive threshold. *Biomed Eng Online* 3:1–9. <https://doi.org/10.1186/1475-925X-3-28>
10. Croce P, Quercia A, Costa S, Zappasodi F (2020) EEG microstates associated with intra- and inter-subject alpha variability. *Sci Rep* 10:1–11. <https://doi.org/10.1038/s41598-020-58787-w>
11. Fell J, Axmacher N (2011) The role of phase synchronization in memory processes. *Nat Rev Neurosci* 12:105–118. <https://doi.org/10.1038/nrn2979>
12. Gabard-Durnam LJ, Leal ASM, Wilkinson CL, Levin AR (2018) The harvard automated processing pipeline for electroencephalography (HAPPE): Standardized processing software for developmental and high-artifact data. *Front Neurosci* 12:1–24. <https://doi.org/10.3389/fnins.2018.00097>
13. Gohel SR, Biswal BB (2015) Functional integration between brain regions at rest occurs in multiple-frequency bands. *Brain Connect* 5:23–34. <https://doi.org/10.1089/brain.2013.0210>
14. Gschwind M, Hardmeier M, Van De Ville D, et al (2016) Fluctuations of spontaneous EEG topographies predict disease state in relapsing-remitting multiple sclerosis. *NeuroImage Clin* 12:466–477. <https://doi.org/10.1016/j.nicl.2016.08.008>
15. Harper J, Malone SM, Iacono WG (2017) Theta- and delta-band EEG network dynamics during a novelty oddball task. *Psychophysiology* 54:. <https://doi.org/10.1111/psyp.12906>
16. Hipp JF, Engel AK, Siegel M (2011) Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron* 69:387–396. <https://doi.org/10.1016/j.neuron.2010.12.027>
17. Huster RJ, Debener S, Eichele T, Herrmann CS (2012) Methods for simultaneous EEG-fMRI: An introductory review. *J Neurosci* 32:6053–6060. <https://doi.org/10.1523/JNEUROSCI.0447-12.2012>
18. Jiang M, Gu J, Li Y, et al (2021) HADLN: Hybrid Attention-Based Deep Learning Network for Automated Arrhythmia Classification. *Front Physiol* 12:. <https://doi.org/10.3389/fphys.2021.683025>
19. Jin Y, Wu D, Guo W (2020) Attention-based LSTM with filter mechanism for entity relation classification. *Symmetry (Basel)* 12:1–16. <https://doi.org/10.3390/sym12101729>
20. Karim F, Majumdar S, Darabi H, Harford S (2019) Multivariate LSTM-FCNs for time series classification. *Neural Networks* 116:237–245. <https://doi.org/10.1016/j.neunet.2019.04.014>
21. Kaur A, Chinnadurai V, Chaujar R (2020) Microstates-based resting frontal alpha asymmetry approach for understanding affect and approach/withdrawal behavior. *Sci Rep* 10:1–25. <https://doi.org/10.1038/s41598-020-61119-7>

22. Khanna A, Pascual-Leone A, Michel CM, Farzan F (2015) Microstates in resting-state EEG: Current status and future directions. *Neurosci Biobehav Rev* 49:105–113.  
<https://doi.org/10.1016/j.neubiorev.2014.12.010>
23. Kiehl KA, Liddle PF (2003) Reproducibility of the hemodynamic response to auditory oddball stimuli: A six-week test-retest study. *Hum Brain Mapp* 18:42–52. <https://doi.org/10.1002/hbm.10074>
24. Kim K, Duc NT, Choi M, Lee B (2021) EEG microstate features according to performance on a mental arithmetic task. *Sci Rep* 11:1–14. <https://doi.org/10.1038/s41598-020-79423-7>
25. Kim KH, Yoon HW, Park HW (2004) Improved ballistocardiac artifact removal from the electroencephalogram recorded in fMRI. *J Neurosci Methods* 135:193–203.  
<https://doi.org/10.1016/j.jneumeth.2003.12.016>
26. Kim Y, Choi A (2020) Eeg-based emotion classification using long short-term memory network with attention mechanism. *Sensors (Switzerland)* 20:1–22. <https://doi.org/10.3390/s20236727>
27. Knyazev GG, Savostyanov AN, Bocharov A V., et al (2019) Cross-frequency coupling in developmental perspective. *Front Hum Neurosci* 13:1–10. <https://doi.org/10.3389/fnhum.2019.00158>
28. Koenig T, Prichep L, Lehmann D, et al (2002) Millisecond by millisecond, year by year: Normative EEG microstates and developmental stages. *Neuroimage* 16:41–48.  
<https://doi.org/10.1006/nimg.2002.1070>
29. Laufs H, Kleinschmidt A, Beyerle A, et al (2003) EEG-correlated fMRI of human alpha activity. *Neuroimage* 19:1463–1476. [https://doi.org/10.1016/S1053-8119\(03\)00286-6](https://doi.org/10.1016/S1053-8119(03)00286-6)
30. Li X, Kehoe EG, McGinnity TM, et al (2015) Modulation of Effective Connectivity in the Default Mode Network at Rest and During a Memory Task. *Brain Connect* 5:60–67.  
<https://doi.org/10.1089/brain.2014.0249>
31. Lisman JE, Jensen O (2013) The Theta-Gamma Neural Code. *Neuron* 77:1002–1016.  
<https://doi.org/10.1016/j.neuron.2013.03.007>
32. Magri C, Schridde U, Murayama Y, et al (2012) The amplitude and timing of the BOLD signal reflects the relationship between local field potential power at different frequencies. *J Neurosci* 32:1396–1407. <https://doi.org/10.1523/JNEUROSCI.3985-11.2012>
33. Mantini (2007) Electrophysiological signature of brain resting state networks. *Proc Natl Acad Sci U S A* 104:13170–13175
34. Michel CM, Koenig T (2018) EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review. *Neuroimage* 180:577–593.  
<https://doi.org/10.1016/j.neuroimage.2017.11.062>
35. Milz P, Faber PL, Lehmann D, et al (2016) The functional significance of EEG microstates-Associations with modalities of thinking. *Neuroimage* 125:643–656.  
<https://doi.org/10.1016/j.neuroimage.2015.08.023>
36. Milz P, Pascual-Marqui RD, Achermann P, et al (2017) The EEG microstate topography is predominantly determined by intracortical sources in the alpha band. *Neuroimage* 162:353–361.  
<https://doi.org/10.1016/j.neuroimage.2017.08.058>

37. Murta T, Chaudhary UJ, Tierney TM, et al (2017) Phase–amplitude coupling and the BOLD signal: A simultaneous intracranial EEG (icEEG) - fMRI study in humans performing a finger-tapping task. *Neuroimage* 146:438–451. <https://doi.org/10.1016/j.neuroimage.2016.08.036>
38. Nagabushanam P, Thomas George S, Radha S (2020) EEG signal classification using LSTM and improved neural network algorithms. *Soft Comput* 24:9981–10003. <https://doi.org/10.1007/s00500-019-04515-0>
39. Pandey AK, Kamarajan C, Manz N, et al (2016) Delta, theta, and alpha event-related oscillations in alcoholics during Go/NoGo task: Neurocognitive deficits in execution, inhibition, and attention processing. Elsevier B.V.
40. Pang JC, Robinson PA (2018) Neural mechanisms of the EEG alpha-BOLD anticorrelation. *Neuroimage* 181:461–470. <https://doi.org/10.1016/j.neuroimage.2018.07.031>
41. Pascual-Marqui RD, Lehmann D, Faber P, et al (2014) The resting microstate networks (RMN): cortical distributions, dynamics, and frequency specific information flow. 1–14
42. Pascual-Marqui RD, Michel CM, Lehmann D (1995) Segmentation of Brain Electrical Activity into Microstates; Model Estimation and Validation. *IEEE Trans Biomed Eng* 42:658–665. <https://doi.org/10.1109/10.391164>
43. Polich J (2007) Updating P300: An integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118
44. Rashid M, Sulaiman N, P. P. Abdul Majeed A, et al (2020) Current Status, Challenges, and Possible Solutions of EEG-Based Brain-Computer Interface: A Comprehensive Review. *Front Neurobot* 14:1–35. <https://doi.org/10.3389/fnbot.2020.00025>
45. Ribary U, Doesburg SM, Ward LM (2017) Unified principles of thalamo-cortical processing: the neural switch. *Biomed Eng Lett* 7:229–235. <https://doi.org/10.1007/s13534-017-0033-4>
46. Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci* 32:9–18. <https://doi.org/10.1016/j.tins.2008.09.012>
47. Sclocco R, Tana MG, Visani E, et al (2014) EEG-informed fMRI analysis during a hand grip task: Estimating the relationship between EEG rhythms and the BOLD signal. *Front Hum Neurosci* 8:1–13. <https://doi.org/10.3389/fnhum.2014.00186>
48. Sikka A, Jamalabadi H, Krylova M, et al (2020) Investigating the temporal dynamics of electroencephalogram (EEG) microstates using recurrent neural networks. *Hum Brain Mapp* 41:2334–2346. <https://doi.org/10.1002/hbm.24949>
49. Stam CJ (2005) Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. *Clin Neurophysiol* 116:2266–2301. <https://doi.org/10.1016/j.clinph.2005.06.011>
50. Tognoli E, Kelso JAS (2009) Brain coordination dynamics: True and false faces of phase synchrony and metastability. *Prog Neurobiol* 87:31–40. <https://doi.org/10.1016/j.pneurobio.2008.09.014>
51. Vaswani, A. et al (2017) Attention Is All You Need. *IEEE Ind Appl Mag* 8:8–15. <https://doi.org/10.1109/2943.974352>

52. von Wegner F, Bauer S, Rosenow F, et al (2021) EEG microstate periodicity explained by rotating phase patterns of resting-state alpha oscillations. *Neuroimage* 224:117372. <https://doi.org/10.1016/j.neuroimage.2020.117372>
53. Wang P, Jiang A, Liu X, et al (2018) LSTM-based EEG classification in motor imagery tasks. *IEEE Trans Neural Syst Rehabil Eng* 26:2086–2095. <https://doi.org/10.1109/TNSRE.2018.2876129>
54. Whitfield-gabrieli S, Nieto-castanon A (2012) Conn: A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. 2:. <https://doi.org/10.1089/brain.2012.0073>
55. Wu CW, Gu H, Lu H, et al (2008) Frequency specificity of functional connectivity in brain networks. *Neuroimage* 42:1047–1055. <https://doi.org/10.1016/j.neuroimage.2008.05.035>
56. Xie Y, Liang R, Liang Z, et al (2019) Speech Emotion Classification Using Attention-Based LSTM. *IEEE/ACM Trans Audio Speech Lang Process* 27:1675–1685. <https://doi.org/10.1109/TASLP.2019.2925934>
57. Yan J, Chen S, Deng S (2019) A EEG-based emotion recognition model with rhythm and time characteristics. *Brain Informatics* 6:. <https://doi.org/10.1186/s40708-019-0100-y>
58. Yao Q, Wang R, Fan X, et al (2020) Multi-class Arrhythmia detection from 12-lead varied-length ECG using Attention-based Time-Incremental Convolutional Neural Network. *Inf Fusion* 53:174–182. <https://doi.org/10.1016/j.inffus.2019.06.024>
59. Yuan H, Zotev V, Phillips R, et al (2012) Spatiotemporal dynamics of the brain at rest - Exploring EEG microstates as electrophysiological signatures of BOLD resting state networks. *Neuroimage* 60:2062–2072. <https://doi.org/10.1016/j.neuroimage.2012.02.031>
60. Zhang G, Davoodnia V, Sepas-Moghaddam A, et al (2020a) Classification of Hand Movements from EEG Using a Deep Attention-Based LSTM Network. *IEEE Sens J* 20:3113–3122. <https://doi.org/10.1109/JSEN.2019.2956998>
61. Zhang L, Zhu G, Mei L, et al (2018) Attention in convolutional LSTM for gesture recognition. *Adv Neural Inf Process Syst* 2018-Decem:1953–1962
62. Zhang Z, Ye S, Liao P, et al (2020b) Enhanced Capsule Network for Medical image classification. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS* 2020-July:1544–1547. <https://doi.org/10.1109/EMBC44109.2020.9175815>
63. Zheng X, Chen W (2021) An Attention-based Bi-LSTM Method for Visual Object Classification via EEG. *Biomed Signal Process Control* 63:102174. <https://doi.org/10.1016/j.bspc.2020.102174>

## Figures



**Figure 1**

Schematic of visual target detection task design.



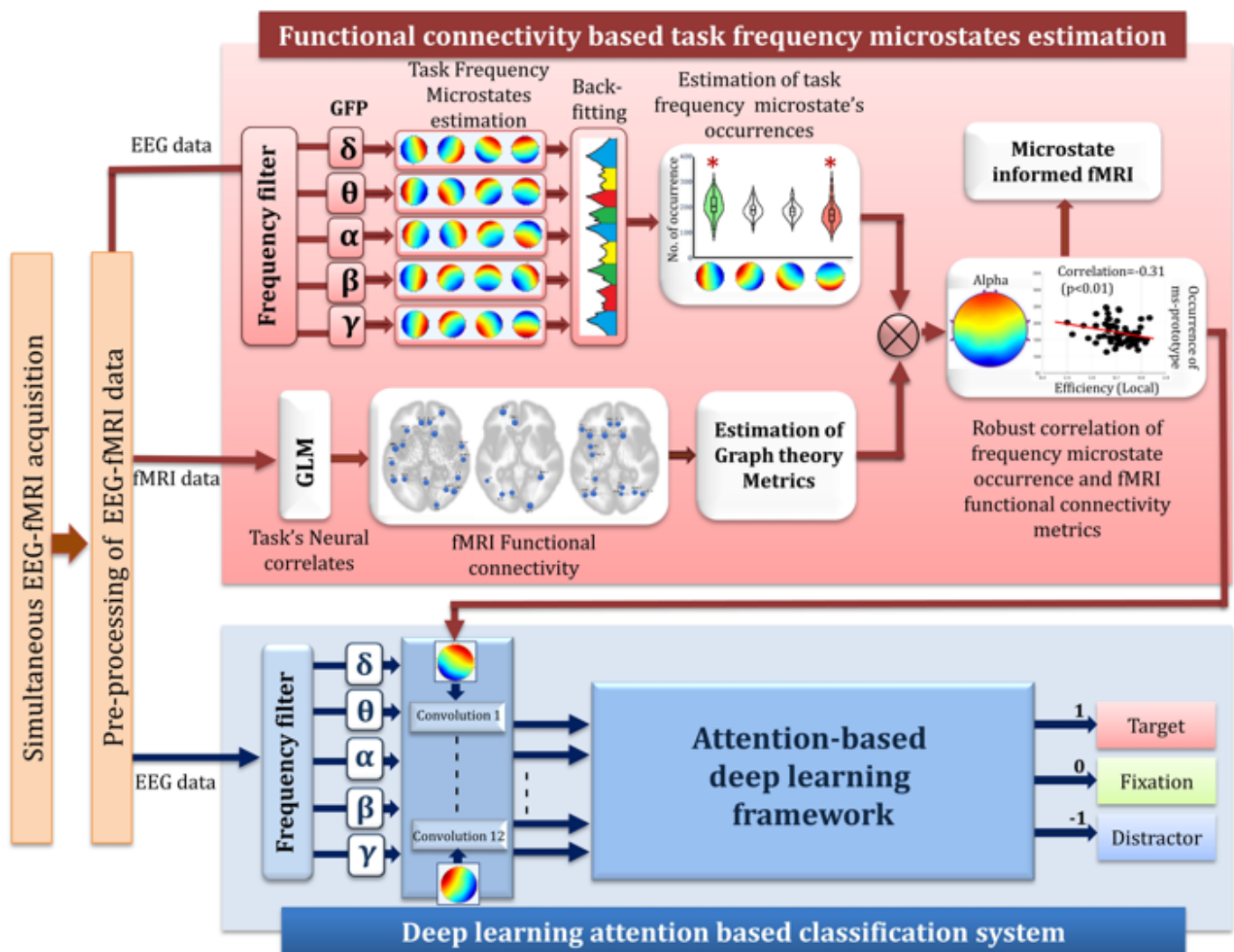


Figure 2

Methodological framework of the study.

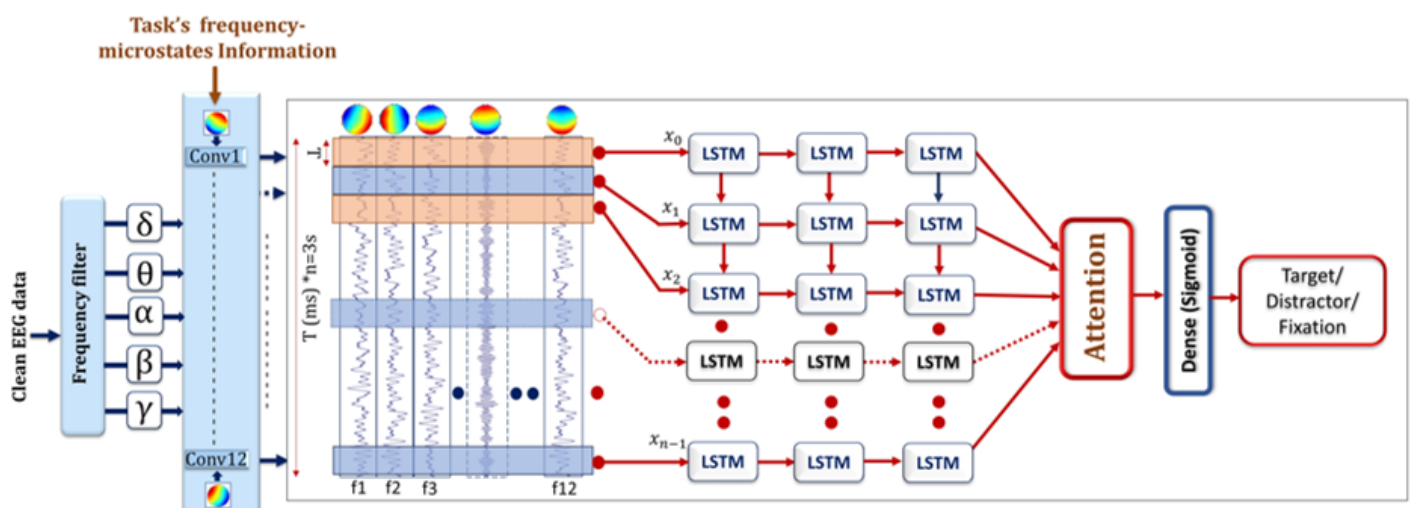


Figure 3

Attention-based LSTM deep learning framework for classification of visual target detection task.

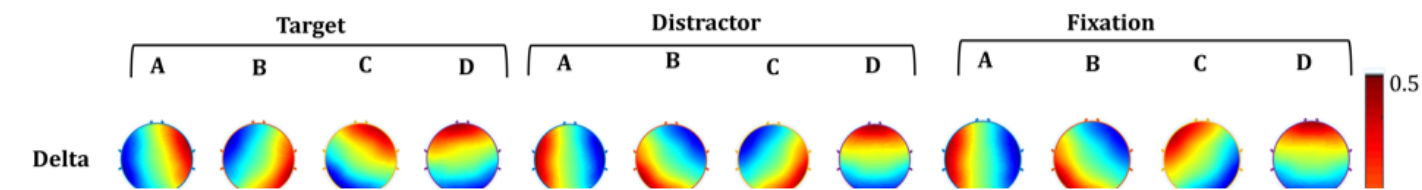
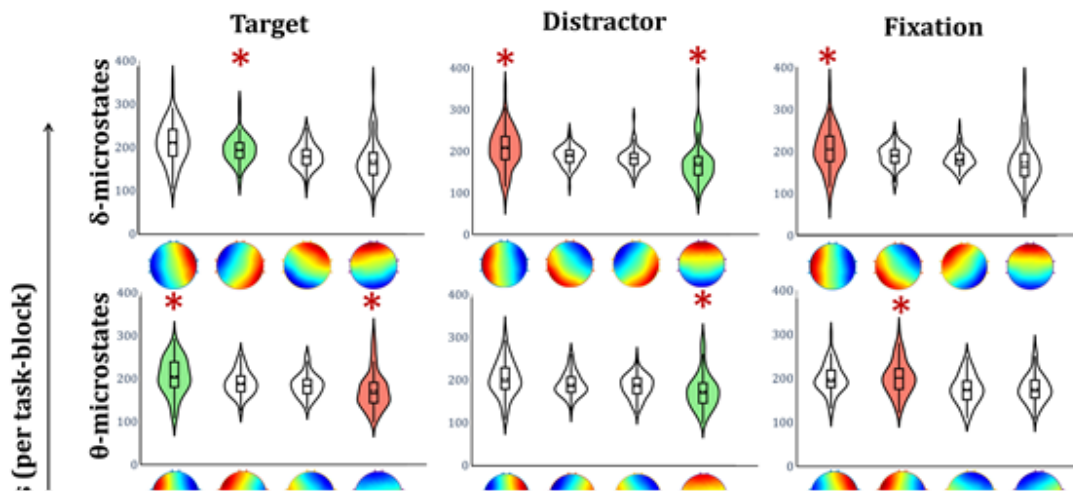


Figure 4

A topographical representation of frequency-microstates for the target, distractor, and fixation.

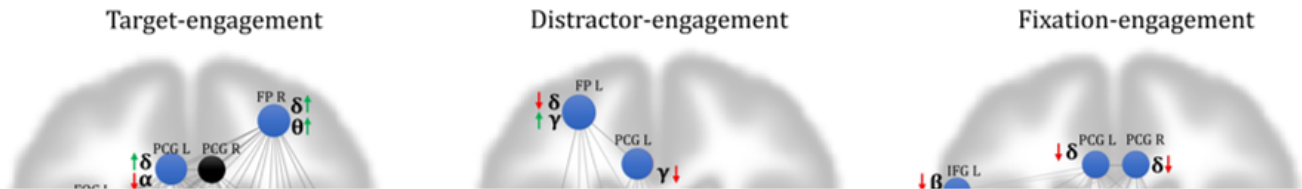


**Figure 5**

Estimated metrics of frequency-microstates. The violin graph plots the mean occurrence of each frequency-microstates estimated for each frequency band. The violin plot's green and red colour shades reveal the correlation (positive and negative) of the number of occurrences of quasi-stable elicitation with the task's fMRI functional connectivity measures. The red star on violin plots specifies the significance ( $p < 0.01$ ).

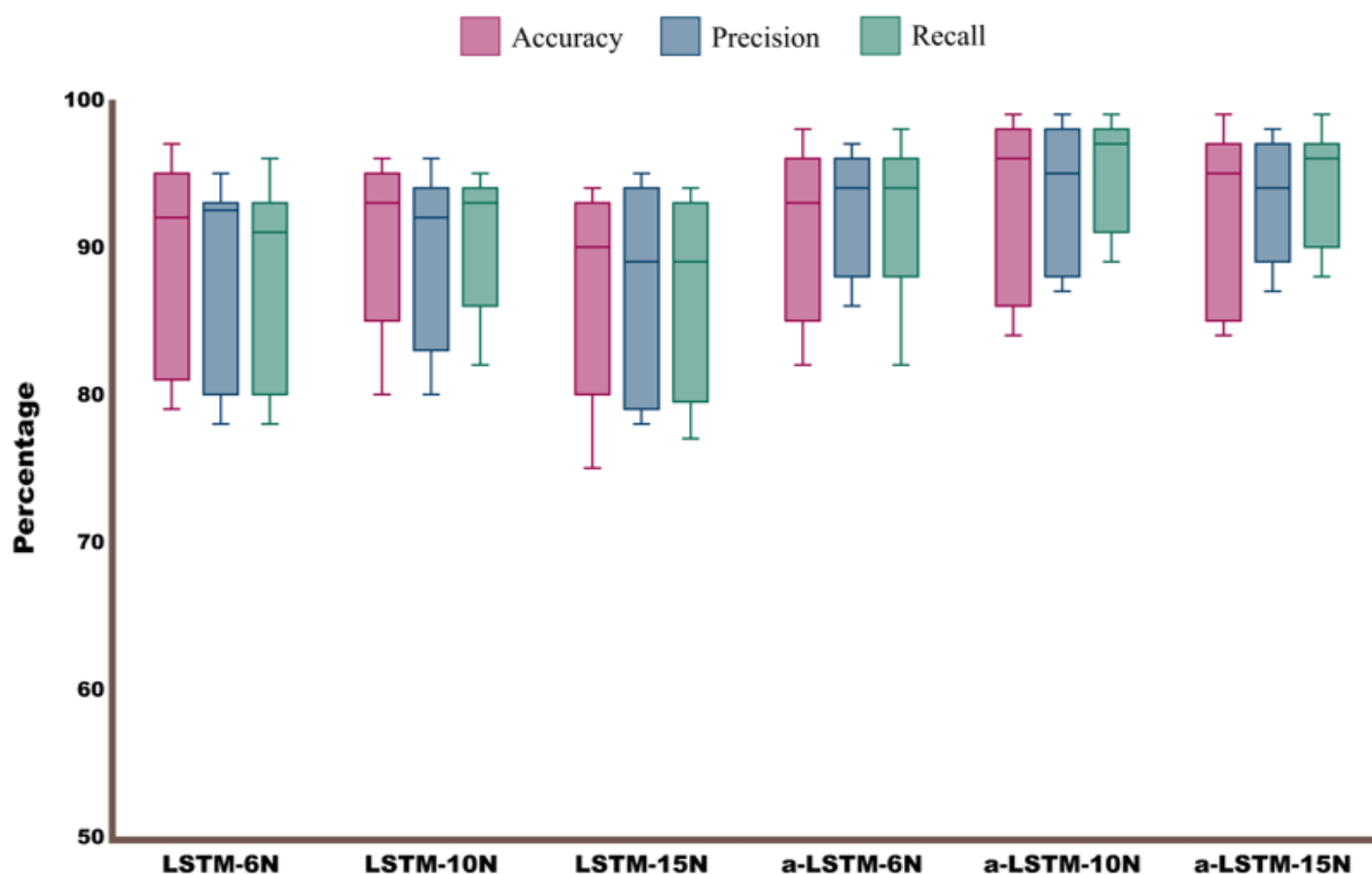
**Figure 6**

Significantly associated, optimized, frequency-microstates with graph-theoretical measures of a) Target, b) Distractor, and c) Fixation task engagement. Frequency microstates are labelled as anterior-posterior (AP), left-right (LR), left diagonal (LD), and right diagonal (RD).



**Figure 7**

Neuro-vascular coupling of each significant, fMRI functional connectivity optimized frequency-microstates of every task. The Regions shown are neural correlates observed for every task. The neuro-vascular association of optimized frequency microstate with each region is drawn through the arrow next to it (Synchronization: Up green arrow, De-synchronization: Down Red arrow). The region coloured black has no neurovascular association with any one of the optimized frequency microstates.



**Figure 8**

Comparison of performance of the deep learning frameworks between attention-based LSTM and traditional LSTM. The plot reveals the classification performance metrics (accuracy, precision, and recall) for frameworks corresponding to three distinct architectures with 15, 10, and 6-LSTM cells (15N, 10N, 6N) of 200ms/segment, 300ms/segment, and 500ms/segment, respectively.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryfile.docx](#)

# Applicability of ML-IoT in Smart Healthcare Systems: Challenges, Solutions & Future Direction

Nikhil Sharma, Prashant Giridhar Shambharkar  
Department of Computer Science & Engineering  
Delhi Technological University  
Delhi, India  
nikhilsharma1694@gmail.com  
prashant.shambharkar@dtu.ac.in

**Abstract**—In today's world, as population is at high peak and due to changing life style of people, individuals are suffering from various chronic disease. With shift towards modern methodology, involvement of human efforts has decreased, as know a day's people need to finish particular amount of task within few hours and with less effort. No doubt technology has made less intervention of human but it has certain limitations too. Due to less physical involvement, humans are more prone to diseases. Internet of things (IoT) plays a very crucial role in health care sector. Using various sensors, it become possible to trace the medical health condition of the human, and a message can be forwarded to nearby hospitals which helps the patients with ease. In this paper, three different diseases like heart disease, diabetes and novel COVID-19 are discussed where different machine learning algorithms are reviewed with involvement of IoT sensors.

**Keywords**— *Smart Healthcare Systems, Internet of Things, Machine Learning, Privacy, Security.*

## I. INTRODUCTION

The continuous productivity and innovation in the field of healthcare industry will fulfil needs of the people. We have already witnessed the implication of real time data. In regards to Covid 19, the information related to patient's health like symptoms or body temperature is exceedingly helpful [1]. By 2050, 22% of the overall population will reach to the age 60 [2]. Along with health-related emergency, the people affected by the chronic diseases are increasing day by day, which will directly create the pressure on the shoulder of the healthcare industry. In the future, health industry may need to tackle with large number of chronic problems, if we are not taking it seriously in the present time. Recently, Covid 19 has already accentuated the value of intelligent medical-care and accurate e-medical care involving diver's kind of physiological and medical information to diagnose the virus. According to prediction in 2019, the world smart health market has reaches to 143.6 billion USD, and between 2020 to 2027, an average growth rate may further be expanded by 16.2% [3]. Smart healthcare system uses devices such as mobile internet, Internet of Things (IoT), wearable appliances, to collect health information, preparing the health records of the patients. Hospital, physicians, research bodies, staff belongs to the diverse actor's category of intelligent medical treatments. It includes a dynamic framework with multiple features like assessment and evaluation, diseases identification and prevention, as well as management of medical research, healthcare and patient decision making. The features of intelligent healthcare include the automated networks such as cloud networking, artificial intelligence, 5G, Big Data,

mobile internet, biotechnology, and IoT. Sensors are also become part of our lives as they are embedded into different devices using automation, computer technology, and automated signal processing. The data produced by the sensors help doctors to quickly recognize the critical situations of the patient. Medical signals like electromyograms (EMGs), Blood pressure (BP), Electroglottographs (EGGs), body temperature, heart rate (HR), and electroencephalograms (ECGs) receives in the form of 1D & 2D signals [4]. To monitor the patient, these medical signals will be used by the healthcare monitoring system. IoT is the major source through which both consumer and doctor get connected. The body check-ups like EEGs, BP readings, Glucose receptors, ultrasounds, and ECGs, help to monitor patients' wellness. Many hospitals are using smart beds which automatically identify the movement of the patient and act accordingly by adjusting the location and angle of the bed. Internet of medical things (IoMT) play vital role in establishing smart healthcare industry. Sometimes to diagnose a disease, it is not possible to depend upon only one type of medical signals. At different levels like classification level, feature level, and data level, these signals can be fused. Due to this, many challenges may be experienced by the system like classification fusion, data buffering, synchronization while receiving signals from divers' sensors, and feature normalization. The technology like Artificial Intelligence (AI), wireless local area network (WLAN), deep learning (DL), and Machine learning (ML) bring revolution in the field of intelligent healthcare to guarantee satisfaction to both stakeholders and patients. For many medical situations, IoMT (Internet of Medical Things) systems deliver the excellent assistance like for health conditions, implantable device like pacemakers is used to control the heartbeat. For improving medical-care experience, the assisting wearables devices like smartwatches are used to monitor the heart rate [5]. Security is the only need for the success of IoMT system [6]. The 11 set of security required by this system to deliver the integrity, non-repudiation, confidentiality, authentication, and data availability. As these security demands are fulfilled by the traditional system as well, even though this system failed to provide security. Because of other system, specification requirement and power consumption lead to the system failure [7]. Based on cryptography, several techniques have been introduced by the researchers which are designed for IoT and IoMT systems. The techniques include keyless noncryptographic techniques, asymmetric cryptography, and symmetric cryptography.

The remainder of the paper is arranged as follows: Section 2 discussed about Literature review; Section 3 introduced Applicability of ML-IoT (Machine Learning &



Internet of Things) in Smart Healthcare System; Section 4 explained about different Challenges, Solutions and Future Directions in adopting ML-IoT in Smart Healthcare System followed by Conclusion in Section 5.

## II. LITERATURE REVIEW

In both Wireless body area networks (WBANs) and IoMT, different security techniques have been assorted by Yaacoub et al [8] to categorize the countermeasures into intrusion detection system (IDS), availability, and authorization. The open challenges like handling emergencies, flexibility, and single point failure have been discussed by the Vyas et al. [9]. Bhushan et al. [10] discussed about the problems allied to security and design for securing patient's record in the cloud. The technology like blockchain, AI, and ML are considered to improve the system performance as well as secure IoMT systems [11]. It also provides tolerance to some issues and attacks like single point of failure, and Denial of service (DoS) attack. As compared to traditional authentication techniques, ML techniques can minimize the error rate by 64% of physical layer authentication [12].

To calculate BP, Xiao et al [3] employed a multi sensor platform with one channel ECG and two channel pressure pulse wave (PPW) signals. Using these signals, 35 informative and physiological features were extracted. In mine area, disaster is one of the major risks to the labourer's life. Therefore, to guarantee an accuracy, and minimize the threats to the labourer's life, a real time monitoring system have been introduced by Gu et al. [13]. For this system, covering theories, IoT, situation awareness, and multi sensor data fusion discussed by the author. For identifying the situation level, a model based on Random Forest Support Vector Machine (SVM) have been employed in it. For heart disease prediction, an ensemble approach based on data fusion was introduced by Muzammal et al. [14]. The information collected using Body sensor networks (BSNs) were inserted into ensemble classifier for prediction purposes.

To enhance the reliability & performance of predicting sleep apnea, Steenkiste et al. [15] proposed a model, which is used to gather and integrate multi-sensor data using backward shortcut connection approach. The multi-sensor data includes abdominal respiratory belt, heart rate, oxygen saturation, and thoracic respiratory belt. The DL based models like long short-term memory (LSTM) and Convolutional neural network (CNN) were used to analyse the performance of the model. To enable the smart health services (includes divers' sensors, robot, and data processing technologies), Lin et al [16] introduced a hybrid BSN architecture based on multi sensor fusion. A multi-sensor fusion also uses AI based interpretable neural network (MFIN) to ensure that all the decision made by the robot is correct. To improve the detection of mental stress, Al-Shargie et al. [17] integrate seven EEG electrodes with the seven channels of Functional near infrared spectroscopy (fNIRS). This fusion results the generation of 12 subjects because of the measurements of EEG and fNIRS signals. Under two diver's situations i.e., stress and control, these subjects solved the arithmetic problems. To improve video resolution, Mergin et al. [18] suggested the fusion of ECG and EEG videos by making use of three transforms like Hybrid transform, Discrete wavelet transform (DWT), and

Discrete cosine transform (DCT). To calculate the effect of fusion both, mean squared error (MSE) rate and peak signal-to-noise ratio (PSNR) were used. To measure BP from ECG sensor data, a predictive model based on multi-level data fusion has been introduced by Simjanoska et al. [19]. In this model, the information was integrated in five different level. In level one and two, different methods have been used to take out the features from the input information. In level 3, the seven divers classifiers output was inserted into meta classifier. For each BP type, the knowledge from multi target regression methods was fused into the level 4. In level 5, a single predictor for mean arterial pressure (MAP), diastolic BP (DBP), and systolic BP (SBP) were acquired.

For monitoring patient medical status, a smart healthcare system has been introduced by Islam et al. [20]. This system contains five sensors. First three sensors (i.e., CO sensor (MQ-9), CO sensor (MQ-9), and Room temperature sensor (DHT11)) are deployed to detect the condition of living environment; and the other two sensors are used for monitoring the condition of the patient, using body temperature sensor (MQ-135) and heart rate sensor. ESP32 is the processing device and the communication medium used for sharing the information of patient to a web server is Wi-Fi.

For monitoring the detect falls, body temperature, and HR of the patients, a wearable remote healthcare monitoring system (RHM) has been suggested by Elango et al. [21]. This proposed framework uses three sensing elements i.e., accelerometer (MPU 6050), body temperature sensor (LM35), and heartbeat sensor along with NodeMCU as processing device. A patient monitoring system using IoT tool has been proposed by Chigozirim et al. [22]. With the help of this prototype, doctors can easily monitor body temperature and HR through body temperature sensors and pulse. NodeMCU and ATmega328P microcontroller have been used as edge devices. In this framework, different sensors are used for collecting the information and transferred the same to the internet using Wi-Fi connection.

## III. APPLICABILITY OF ML-IOT (MACHINE LEARNING & INTERNET OF THINGS) IN SMART HEALTHCARE SYSTEM

IoMT systems consist of multiple interconnected devices that are used for handling and transferring the information to improve health of the patients. Nowadays, healthcare industry has become the fastest growing field with numerous inventions. With the use of IoT technology, this sector has gained popularity in minimum period of time. According to the McKinsey study, by 2025, this sector will have a financial impact of \$ 11.1 trillion [23]. IoT gadgets provide good capabilities for data sharing, intelligence, and data analytics. For IoMT systems, AI and ML has become an effective and powerful solution [24]. In recent time, we all witnesses the impact of COVID 19 and other disease. The following sub-sections discuss about the major solutions in healthcare industry using ML.

### A. Novel Coronavirus (COVID-19)

In recent period, we all witnesses the impact of this public health crisis. This virus harms almost every sector in the whole world. Still many fields are struggling to recuperate from this ailment. According to the worldwide statistic, 98 million cases have found, and 2 million people lost their lives because of this virus [25]. Using modern

technologies, innumerable works have been carried out to minimize impact of the virus. This section discusses the role of ML algorithm to diagnose COVID-19 by using IoT environment. To diagnose COVID-19, an IoT based framework which uses the concept of SVM and CNN has been proposed by Le et al. [26]. With the help of IoT sensors, this system fetches the information from patients and by using 5G networks, transmits that data to the cloud storage. CIoTVID framework has been suggested by Gonzalez et al. [27]. This is a IoT based system that is used for the detection of Coronavirus. This system consists of diver's level of sensorization, through which the data can be evaluated and help in decision making process. The different symptoms collected by data collection layer from patients includes respiration rate, oxygen saturation, and voice signals. Afterward, to diagnose COVID-19, a deep learning-based system has been proposed by Ahmed et al. [28] to diagnose coronavirus in IoT environment. From chest X-ray samples, the collaborated approach like Faster-RCNN and ResNet-101 is used by this framework to diagnose COVID-19 cases. Otoom et al. [29] suggested ML based scheme to find out or monitor the infected patients by COVID-19 in an IoT environment. Using IoT devices, the patient's real time symptom data is collected and transmitted to the cloud storage. Rashidy et al. [30] proposed an end-to-end environment-based DL architecture, which is used to monitor and diagnose the patients who are affected by the Coronavirus. Humanoid software has been developed by Karmore et al. [31]. This software helps to detect the coronavirus in an IoT environment and recognize whether the person is infected with the virus or not. For navigation process, a camera module and IR sensors are used by the robotic systems. For the diagnose purpose, chest X-ray scan and E-Health sensor kit were deployed. The hardware components like ECG sensors, NodeMCU, temperature sensors, and Raspberry Pi are used by the advanced humanoid robot. An IoT based system has been proposed by Cacovean et al. [32] for detecting COVID-19. This system uses ML approach in diagnosis process. The participant's data is being collected with the use of wearable gadgets such as HR sensors, temperature, and GPS. After that, all the retrieved information of the participants is transmitted to the cloud server of oracle through Bluetooth for the processing purpose. Kumar et al. [33] proposed a framework uses IoT technology and sensors to observe the patients who are infected by COVID-19. IoT sensor are used for retrieving the real-time information of patient, and passes for pre-processing purposes to the Bayesian network.

### B. Heart Disease

For every aged person, heart disease has become a critical disease. According to the statistic, among all the death cases per year, 30% of the cases belongs to this category only [34]. To minimize the impact of heart disease in smart healthcare environment, the researchers are mainly put their attention on the advancement of the Decision support system (DSS). Machine learning plays vital role in the advancement of heart disease diagnosis in IoT environment. The patient's monitoring framework has been proposed by Sarmah et al. [35] for heart patients. The sensor like body-worn were used to gather all the information about the patient and securely transfer that information to the cloud storage for further operations. The data examined with the help of improved Deep learning modified Neural Networks (DLMNN). For the patients infected by heart disease, Ali et

al. [36] suggested a smart healthcare monitoring system which uses the notion of feature fusion and ensemble learning. In this framework, the extracted feature from both patient history as well as sensor information should be combined using feature fusion approach. It selects the most germane feature which is responsible for heart disease and eliminate superfluous and gratuitous features by using information gain approach. Deperlioglu et al. [37] proposed a system which uses autoencoder network for diagnose heart disease in IoHT (Internet of Healthcare Things) environment. In this framework, a central system has been deployed which synchronize devices and cloud communication. Beacons are used for transferring the information in the cloud environment. A ML-based system has been proposed by Khan et al. [38] for diagnosis heart disease in IoMT environment. This ML-based framework developed by using an adaptive neuro-fuzzy inference system and modified salp swarm optimization. To predict heart disease, a IoT based framework has been introduced by Khan et al [39], which uses modified deep convolutional neural network (MDCNN). To diagnose this ailment, a smart healthcare system called HealthFog which uses ensemble learning in Fog computing and IoT environment has been proposed by Tuli et al. [40]. Using FogBus (Fog based cloud environment), the performance of advanced system can be easily evaluated on the basis of execution time, latency, accuracy, and energy consumption. The FogBus contains broker node, cloud data center, and worker node. Nguyen et al. [41] proposed a ML based ailment diagnoses scheme in IoT environment. The patient's information has been collected using ECG gadgets and by using Wi-Fi media all that information transmitted to the cloud storage. From the raw information, most suitable classification feature gets extracted after the pre-processing of data using Wavelet-based Kernel Principal Component Analysis method (wkPCA). Based on input data to diagnose heart disease, the extricated feature fed into Backpropagation Neural Network (BNN).

### C. Diabetes

In the present time, many individuals are getting affected by diabetes. It is very severe disease which causes human's life per year. According to the statistic, in 2019 about 463 million peoples had diabetes. The number of cases are expected to increase to 578 million by the year 2030 [42]. So, to tackle with this situation some necessary step should be taken to diagnosis this ailment using modern technologies like ML and IoT. This section discussed about the utilization of these techniques in diabetes. To monitor and predict the patients affected by this ailment, Rghioui et al. [43] proposed a system which uses ML approaches in IoT environment. The glucometer is attached to NodeMCU to collect the information from the patients. Using IoT platform, a large amount of information gets transmitted to the cloud storage where data processed using ML techniques. After this process, decision report shared with the doctor for the treatment process. A diabetes prediction system has been introduced by Allugunti et al. [44], which uses the concept decision tree and IoT to detect the patients who are infected by this ailment in real time. A health monitoring framework have been suggested by Efât et al [45]. This system helps in monitoring the pulse rate, sleep time, sugar level, and food intake by the diabetic patients. Using wearable sensors, the patient health information is transmitted to neural network through Bluetooth. Using this advanced system, the data gets

categorized according to the seriousness of diabetes cases. This system also sent an alert call/message to the parents or guardian of patient in an emergency. By using ML and IoT technology, diabetes prediction and monitoring system has been introduced by Rghioui et al. [46]. The hardware components used in this system are GSM modem, Arduino, and blood glucose meter. Using these components, data for experiment purposes has been extracted and processes in microcontroller and through GSM modem, the processed information delivered to mediators. An intelligent system using ML architecture has been proposed by Rghioui et al. [47] for monitoring the patients infected with diabetes disease. This system used to monitor body temperature, glucose level, and physical activity in IoT network. Different sensors like motion sensor, temperature sensor, and glucometer are used for gathering the information from patients. Smartphones are used to transfer data collected using sensors to the database server through 5G networks. Furthermore, a healthcare monitoring system using ML modalities has been proposed by Godi et al. [48] to monitor and diagnose this disease through IoT network. The wearables devices are used to collect the information from patients from anywhere like homes, and hospitals. For cloud based IoT environment, Kaur et al. [49] developed CI-DPF systems which was designed to predict diabetes disease. Smart sensors were used to collect the information related to the blood glucose level of the patients and transmitted that information to cloud storage. After that the stored data processed using IoT devices.

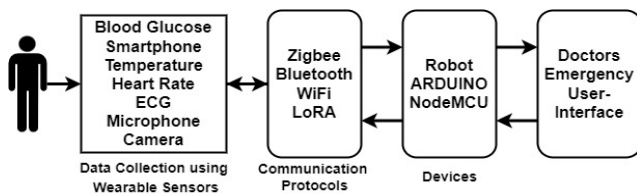


Fig. 1. Health Monitoring data collected using Wearable Devices

#### IV. CHALLENGES, SOLUTIONS AND FUTURE DIRECTIONS IN ADOPTING ML-IoT IN SMART HEALTHCARE SYSTEM

In the above section, different assistive systems have been discussed which were used to monitor and predict disease using demanding technologies like IoT, AI and ML. These modern techniques are useful in delivering a smart healthcare system. This system must face various challenges in implementation process like integrating data received using different sensors or wearable devices. The different data types have been generated from divers' sensors. Therefore, it is necessary to convert signals into a meaningful information that is received from heterogeneous sensors attached to the patients for monitoring their health. Different data fusion techniques have been used to integrate the information received from multi-sensory gadgets. It provides the streamed signals for ameliorating dependability and reducing the bandwidth needed for communication purposes with the cloud layer [50-51]. The major challenges in designing AI and IoT based smart healthcare system includes device management, privacy and security, efficient utilization of AI technology, device communication, and sensor's interoperability.

In IoMT devices, multiple devices are used to diagnose and identify an illness. The information has been gathered from heterogeneous sensors which consists of different

issues like connectivity problems, drained batteries, and hardware glitches [52]. Sometimes, an unexplained error has been occurred while using demanding medical sensors like smartwatches and smartphones. The regular complexities also create challenges related to variations in platforms, battery power, and difference between physical attributes. Many above problems can be overcome using IoMT devices and multimodal signal.

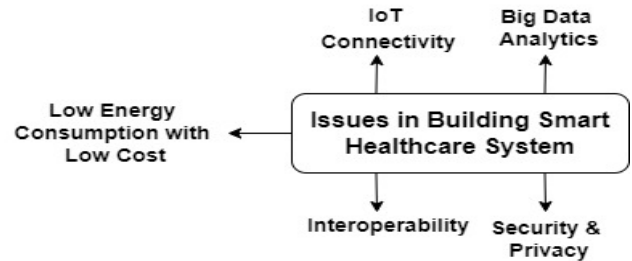


Fig. 2. Issues in Building Smart Healthcare Systems

Following are the open issues and challenges in building smart healthcare.

##### A. IoT connectivity

A smart healthcare system is a collection of billions of gadgets whose aim is to provide connection to each and every device available in the architecture. In this system, IoT devices can use any network for the communication. IoT devices include cellular network, Wi-Fi, and Bluetooth. In smart healthcare system, ensuring connectivity between IoT devices expresses multiple challenges like:

- Providing connectivity to each and every device used in wide range of the network.
- Guaranteeing connection to the high mobility gadgets (such as carrying patients, and high-speed ambulance) in the network.

##### B. Interoperability

The ability to interconnect various devices and networks to exchange information with each other is called Interoperability. Smart healthcare is built with the combination of IoT devices. From different domains like remote healthcare monitoring, remote surgery, and ECG monitoring. The role of interoperability is to establish connection between these IoT devices by employing divers' communication technologies. Because of lack of universal standards, it becomes a challenge to connect various devices belong from divers' domains [53]. To provide solution to this challenge, it is necessary to identify it at each level whether belong from network, application, communications, or devices. An AI based integrated techniques is required to setup communication between these multiple IoT devices. For example, ETSI (European Telecommunication Standards Institute), 3GPP (Third Generation Partnership Project), and OMA (Open Mobile Alliance) are standardization bodies which collaborate with standards like oneM2M (Machine to Machine) and FIWARE to work on interoperability issues.

##### C. Low Energy Consumption with Low Cost

Smart healthcare system is the combination of sensors which are small in size and, are used for data communication and storage purposes. All the sensors needed continuous

power to perform their operation, which brings a severe issue related to battery life and cost. The sensors employed in healthcare system must consume low energy and low cost to overcome this problem. One of the solutions to this issue is intelligent algorithms which helps the sensors to operate and share all patient related information with each other by consuming low power.

#### D. Big Data Analytics

The multiple sensors and devices are used in healthcare system which generates enormous amount of information for analysis purpose. The information is allied to the patient private data which are collected using smart devices. Therefore, to analyse all these information, intelligent techniques and algorithm must be needed. For example, DL algorithms can be adopted to effectively analysed data collected using different gadgets and sensors.

#### E. Security and Privacy

Security and privacy are the key factor needed in smart healthcare system. In this system, multiple devices are connected with each other using internet facilities, which makes security a major concern in the system. It is difficult to develop complicated security algorithms and protocols because IoT devices has low battery life and limited processing. Due to this reason, in future nearly 70% of smart devices are at high risk of attack. Therefore, 5G network must be needed in the future to develop smart healthcare system. This can be achieved by following steps discussed below:

- It is necessary to provide privacy aware communication for user information.
- Between cloud-based centres and IoT based smart healthcare devices, a secure communication system must be needed for information authenticity and integrity.
- By providing strong privacy, user trust and approval can be guarantee.
- The details related to the new and present attacks must be carried more effectively in the system.

#### V. CONCLUSION

In this paper, different types of chronic diseases have been discussed. Involvement of different types of sensors used in health care unit, helps in assisting the health status of patient. Effective utilization of resources results in better time management between hospital assistant and patient. Second half of paper gives an insight of literature review discussed by different researchers. Combined technique of IoT and machine learning helps in monitoring patients' condition. As the confidential information is to be kept secret, security mechanisms must be implemented in the system. Later section of the paper, discusses the challenges, solutions and future research direction. Further, open scope still exists by implementing security mechanisms in the system as key elements of security must be preserved while implementing a system.

#### REFERENCES

- [1] Hossain, M. S., Muhammad, G., & Guizani, N. (2020). Explainable AI and mass surveillance system-based healthcare framework to combat COVID-19 like pandemics. *IEEE Network*, 34(4), 126–132. <https://doi.org/10.1109/mnet.011.2000458>
- [2] World Health Organization. (n.d.). World Report on Ageing and Health. World Health Organization. Retrieved March 1, 2022, from <https://www.who.int/publications-detail-redirect/9789241565042>
- [3] Miao, F., Liu, Z.-D., Liu, J.-K., Wen, B., He, Q.-Y., & Li, Y. (2020). Multi-sensor fusion approach for cuff-less blood pressure measurement. *IEEE Journal of Biomedical and Health Informatics*, 24(1), 79–91. <https://doi.org/10.1109/jbhi.2019.2901724>
- [4] Yang, F., Zhao, X., Jiang, W., Gao, P., & Liu, G. (2019). Multi-method fusion of cross-subject emotion recognition based on high-dimensional EEG features. *Frontiers in Computational Neuroscience*, 13. <https://doi.org/10.3389/fncom.2019.00053>
- [5] Jordan, N. (2019, August 23). 7 staggering stats on healthcare IOT Innovation. Medium. Retrieved March 1, 2022, from <https://medium.datadriveninvestor.com/7-staggering-stats-on-healthcare-iot-innovation-fe6b92774a5c>
- [6] Sengupta, J., Ruj, S., & Das Bit, S. (2020). A comprehensive survey on attacks, security issues and blockchain solutions for IOT and liot. *Journal of Network and Computer Applications*, 149, 102481. <https://doi.org/10.1016/j.jnca.2019.102481>
- [7] A Steger. How the Internet of Medical Things Is Impacting Healthcare. Accessed: Nov. 18, 2020. [Online]. Available: <https://healthtechmagazine.net/article/2020/01/how-internet-medicalthings-impacting-healthcare-perfcon>
- [8] Yaacoub, J.-P. A., Noura, M., Noura, H. N., Salman, O., Yaacoub, E., Couturier, R., & Chehab, A. (2020). Securing internet of medical things systems: Limitations, issues and recommendations. *Future Generation Computer Systems*, 105, 581–606. <https://doi.org/10.1016/j.future.2019.12.028>
- [9] Vyas, A., & Pal, S. (2020). Preventing security and privacy attacks in WBANS. *Handbook of Computer Networks and Cyber Security*, 201–225. [https://doi.org/10.1007/978-3-030-22277-2\\_8](https://doi.org/10.1007/978-3-030-22277-2_8)
- [10] Bhushan, D., & Agrawal, R. (2019). Security challenges for designing wearable and IOT Solutions. *A Handbook of Internet of Things in Biomedical and Cyber Physical System*, 109–138. [https://doi.org/10.1007/978-3-030-23983-1\\_5](https://doi.org/10.1007/978-3-030-23983-1_5)
- [11] Pesapane, F., Suter, M. B., Codari, M., Patella, F., Volonté, C., & Sardanelli, F. (2020). Regulatory issues for artificial intelligence in Radiology. *Precision Medicine for Investigators, Practitioners and Providers*, 533–543. <https://doi.org/10.1016/b978-0-12-819178-1.00052-6>
- [12] Xiao, L., Li, Y., Han, G., Liu, G., & Zhuang, W. (2016). Phy-layer spoofing detection with reinforcement learning in wireless networks. *IEEE Transactions on Vehicular Technology*, 65(12), 10037–10047. <https://doi.org/10.1109/tvt.2016.2524258>
- [13] Gu, Q., Jiang, S., Lian, M., & Lu, C. (2019). Health and Safety Situation Awareness Model and emergency management based on multi-sensor signal fusion. *IEEE Access*, 7, 958–968. <https://doi.org/10.1109/access.2018.2886061>
- [14] Muzammal, M., Talat, R., Sodhro, A. H., & Pirbhulal, S. (2020). A multi-sensor data fusion enabled ensemble approach for medical data from Body Sensor Networks. *Information Fusion*, 53, 155–164. <https://doi.org/10.1016/j.inffus.2019.06.021>
- [15] T. Van Steenkiste, D. Deschrijver, and T. Dhaene, "Sensor fusion using backward shortcut connections for sleep apnea detection in multi-modal data," 2019, arXiv:1912.06879. [Online]. Available: <http://arxiv.org/abs/1912.06879>
- [16] Lin, K., Li, Y., Sun, J., Zhou, D., & Zhang, Q. (2020). Multi-sensor fusion for Body Sensor Network in medical human–robot interaction scenario. *Information Fusion*, 57, 15–26. <https://doi.org/10.1016/j.inffus.2019.11.001>
- [17] Al-Shargie, F. (2019). Fusion of fmirs and EEG signals: Mental Stress Study. <https://doi.org/10.31224/osf.io/kaqew>
- [18] Mergin, A. A., & Premi, M. S. (2017). Pixel level fusion of medical signals using DCT, DWT and hybrid(DWT-DCT) transform based on maximum selection rule-A comparison. *2017 International Conference on Computing Methodologies and Communication (ICCMC)*. <https://doi.org/10.1109/iccmc.2017.8282596>
- [19] Simjanoska, M., Kochev, S., Tanevski, J., Bogdanova, A. M., Papa, G., & Efimov, T. (2020). Multi-level information fusion for learning

- a blood pressure predictive model using sensor data. *Information Fusion*, 58, 24–39. <https://doi.org/10.1016/j.inffus.2019.12.008>
- [20] Islam, M. M., Rahaman, A., & Islam, M. R. (2020). Development of Smart Healthcare Monitoring System in IOT environment. *SN Computer Science*, 1(3). <https://doi.org/10.1007/s42979-020-00195-y>
  - [21] Elango, K., & Muniandi, K. (2020). A low-cost wearable remote healthcare monitoring system. *Role of Edge Analytics in Sustainable Smart City Development*, 219–242. <https://doi.org/10.1002/9781119681328.ch11>
  - [22] Chigozirim, A., Vivian, N., Uchenna, N., & Oreoluwa, A. (2020). A patient monitoring system using internet of things technology. *Ingénierie Des Systèmes d'Information*, 25(3), 351–357. <https://doi.org/10.18280/isi.250309>
  - [23] J. Manyika, M. Chui, P. Bisson, J. Woetzel, R. Dobbs, J. Bughin, and D. Aharon. Unlocking the Potential of the Internet of Things. Accessed: Jan. 16, 2021. [Online]. Available: <https://www.mckinsey.com/businessfunctions/mckinseydigital/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world>
  - [24] Li, W., Chai, Y., Khan, F., Jan, S. R., Verma, S., Menon, V. G., Kavita, & Li, X. (2021). A comprehensive survey on machine learning-based big data analytics for IOT-enabled Smart Healthcare System. *Mobile Networks and Applications*, 26(1), 234–252. <https://doi.org/10.1007/s11036-020-01700-6>
  - [25] Worldometer. COVID-19 Coronavirus Pandemic. Accessed: Jan. 22, 2021. [Online]. Available: <https://www.worldometers.info/coronavirus/>
  - [26] Le, D.-N., Velmurugan, ., Parvathy, S., Gupta, . D., Khanna, A., Rodrigues, J. J., Shankar, . K., & Shankar, K. (2021). IOT enabled depthwise separable convolution neural network with deep support vector machine for covid-19 diagnosis and classification. <https://doi.org/10.22541/au.162912530.05037483/v1>
  - [27] Ramallo-González, A. P., González-Vidal, A., & Skarmeta, A. F. (2021). Ciotvid: Towards an open IOT-platform for infective pandemic diseases such as covid-19. *Sensors*, 21(2), 484. <https://doi.org/10.3390/s21020484>
  - [28] Ahmed, I., Ahmad, A., & Jeon, G. (2021). An IOT-based deep learning framework for early assessment of covid-19. *IEEE Internet of Things Journal*, 8(21), 15855–15862. <https://doi.org/10.1109/jiot.2020.3034074>
  - [29] Otoom, M., Otoum, N., Alzubaidi, M. A., Etoom, Y., & Banihani, R. (2020). An IOT-based framework for early identification and monitoring of COVID-19 cases. *Biomedical Signal Processing and Control*, 62, 102149. <https://doi.org/10.1016/j.bspc.2020.102149>
  - [30] El-Rashidy, N., El-Sappagh, S., Islam, S. M., El-Bakry, H. M., & Abdelrazek, S. (2020). End-to-end Deep Learning Framework for coronavirus (COVID-19) detection and monitoring. *Electronics*, 9(9), 1439. <https://doi.org/10.3390/electronics9091439>
  - [31] Karmore, S., Bodhe, R., Al-Turjman, F., Kumar, R. L., & Pillai, S. (2021). IOT based humanoid software for identification and diagnosis of covid-19 suspects. *IEEE Sensors Journal*, 1–1. <https://doi.org/10.1109/jsen.2020.3030905>
  - [32] CACOVEAN, D., IOANA, I., & NITULESCU, G. (2020). IOT system in diagnosis of covid-19 patients. *Informatica Economica*, 24(2/2020), 75–89. <https://doi.org/10.24818/issn14531305/24.2.2020.07>
  - [33] K.R., K., M., I., V.R., N., Magesh, S., Magesh, G., & Marappan, S. (2020). Monitoring and analysis of the recovery rate of covid-19 positive cases to prevent dangerous stage using IOT and sensors. *International Journal of Pervasive Computing and Communications*. <https://doi.org/10.1108/ijpcc-07-2020-0088>
  - [34] S. Kaptoge, L. Pennells, D. De Bacquer, M. T. Cooney, M. Kavousi, G. Stevens, L. M. Riley, S. Savin, T. Khan, S. Altay, and P. Amouyel, "World health organization cardiovascular disease risk charts: Revised models to estimate risk in 21 global regions," *Lancet Global Health*, vol. 7, no. 10, 2019, Art. no. e1332.
  - [35] Sarmah, S. S. (2020). An efficient IOT-based patient monitoring and heart disease prediction system using Deep Learning Modified Neural Network. *IEEE Access*, 8, 135784–135797. <https://doi.org/10.1109/access.2020.3007561>
  - [36] Ali, F., El-Sappagh, S., Islam, S. M. R., Kwak, D., Ali, A., Imran, M., & Kwak, K.-S. (2020). A smart healthcare monitoring system for heart disease prediction based on Ensemble Deep Learning and feature fusion. *Information Fusion*, 63, 208–222. <https://doi.org/10.1016/j.inffus.2020.06.008>
  - [37] Deperioglu, O., Kose, U., Gupta, D., Khanna, A., & Sangaiah, A. K. (2020). Diagnosis of heart diseases by a secure internet of health things system based on Autoencoder Deep Neural Network. *Computer Communications*, 162, 31–50. <https://doi.org/10.1016/j.comcom.2020.08.011>
  - [38] Khan, M. A., & Algarni, F. (2020). A healthcare monitoring system for the diagnosis of heart disease in the IOMT cloud environment using MSSO-ANFIS. *IEEE Access*, 8, 122259–122269. <https://doi.org/10.1109/access.2020.3006424>
  - [39] Khan, M. A. (2020). An IOT framework for heart disease prediction based on MDCNN classifier. *IEEE Access*, 8, 34717–34727. <https://doi.org/10.1109/access.2020.2974687>
  - [40] Tuli, S., Basumatary, N., Gill, S. S., Kahani, M., Arya, R. C., Wander, G. S., & Buyya, R. (2020). HealthFog: An ensemble deep learning based smart healthcare system for automatic diagnosis of heart diseases in integrated IOT and fog computing environments. *Future Generation Computer Systems*, 104, 187–200. <https://doi.org/10.1016/j.future.2019.10.043>
  - [41] Nguyen, T.-H., Nguyen, T.-N., & Nguyen, T.-T. (2019). A deep learning framework for heart disease classification in an IoTs-based system. *A Handbook of Internet of Things in Biomedical and Cyber Physical System*, 217–244. [https://doi.org/10.1007/978-3-030-23983-1\\_9](https://doi.org/10.1007/978-3-030-23983-1_9)
  - [42] Saeedi, P., Petersohn, I., Salpea, P., Malanda, B., Karuranga, S., Unwin, N., Colagiuri, S., Guariguata, L., Motala, A. A., Ogurtsova, K., Shaw, J. E., Bright, D., & Williams, R. (2019). Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Research and Clinical Practice*, 157, 107843. <https://doi.org/10.1016/j.diabres.2019.107843>
  - [43] Rghioui, A., Naja, A., Mauri, J. L., & Oumnad, A. (2021). An IOT based diabetic patient monitoring system using machine learning and node MCU. *Journal of Physics: Conference Series*, 1743(1), 012035. <https://doi.org/10.1088/1742-6596/1743/1/012035>
  - [44] Allugunti, V. R., Kishor Kumar Reddy, C., Elango, N. M., & Anisha, P. R. (2020). Prediction of diabetes using internet of things (IOT) and decision trees: SLDPS. *Advances in Intelligent Systems and Computing*, 453–461. [https://doi.org/10.1007/978-981-15-5679-1\\_43](https://doi.org/10.1007/978-981-15-5679-1_43)
  - [45] Efat, M. I., Rahman, S., & Rahman, T. (2020). IOT based Smart Health Monitoring System for diabetes patients using neural network. *Cyber Security and Computer Science*, 593–606. [https://doi.org/10.1007/978-3-030-52856-0\\_47](https://doi.org/10.1007/978-3-030-52856-0_47)
  - [46] Rghioui, A., Naja, A., & Oumnad, A. (2020). Diabetic patients monitoring and data classification using IOT Application. *2020 International Conference on Electrical and Information Technologies (ICEIT)*. <https://doi.org/10.1109/iceit48248.2020.9113171>
  - [47] Rghioui, A., Lloret, J., Sendra, S., & Oumnad, A. (2020). A smart architecture for diabetic patient monitoring using machine learning algorithms. *Healthcare*, 8(3), 348. <https://doi.org/10.3390/healthcare8030348>
  - [48] Godi, B., Viswanadham, S., Muttipati, A. S., Prakash Samantray, O., & Gadiraju student, S. R. (2020). E-healthcare monitoring system using IOT with machine learning approaches. *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)*. <https://doi.org/10.1109/iccsea49143.2020.9132937>
  - [49] Kaur, P., Sharma, N., Singh, A., & Gill, B. (2018). CI-DPF: A cloud IOT based framework for diabetes prediction. *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. <https://doi.org/10.1109/iemcon.2018.8614775>
  - [50] Wu, R.-T., & Jahanshahi, M. R. (2018). Data Fusion Approaches for Structural Health Monitoring and system identification: Past, present, and future. *Structural Health Monitoring*, 19(2), 552–586. <https://doi.org/10.1177/1475921718798769>
  - [51] Kene, A. P., & Choudhury, S. K. (2019). Analytical Modeling of Tool Health Monitoring System using multiple sensor data fusion approach in hard machining. *Measurement*, 145, 118–129. <https://doi.org/10.1016/j.measurement.2019.05.062>
  - [52] Lin, K., Song, J., Luo, J., Ji, W., Shamim Hossain, M., & Ghoneim, A. (2017). Green video transmission in the Mobile Cloud Networks. *IEEE Transactions on Circuits and Systems for Video Technology*, 27(1), 159–169. <https://doi.org/10.1109/tcsvt.2016.2539618>

- [53] Kharel, J., Reda, H. T., & Shin, S. Y. (2018). Fog computing-based Smart Health Monitoring System deploying Lora Wireless Communication. *IETE Technical Review*, 36(1), 69–82. <https://doi.org/10.1080/02564602.2017.1406828>



# CAPACITANCE CHARACTERISTICS BEHAVIOR OF 0.5 ORDER FC USING CFOA BASED FC MULTIPLIER

Rakesh VERMA , Neeta PANDEY , Rajeshwari PANDEY 

Department of Electronics and Communication Engineering, Delhi Technological University, Bawana Road, Shahbad Daulatpur Village, Rohini, 110042 Delhi, India

dce.rkverma@gmail.com, n66pandey@rediffmail.com, rajeshwaripandey@gmail.com

DOI: 10.15598/aece.v20i1.3621

Article history: Received Jul 07, 2019; Revised Mar 14, 2021; Accepted Mar 26, 2021; Published Mar 31, 2022.  
This is an open access article under the BY-CC license.

**Abstract.** This paper presents a method for capacitance scaling of Fractional Capacitor (FC) which is implemented using Current Feedback Operational Amplifiers (CFOA) based Capacitance Multipliers (C Multipliers). The circuit facilitates the change in FC value without changing component values in R-C network used for FC modelling or fabricating a new FC. The performance of the proposed circuit is examined for non ideal effects of CFOA. Effective impedance of scaled FC is examined through MATLAB simulations. The functionality of the realized scalers is verified using SPICE simulations where the FC is modelled using domino RC ladder network. Simulation results for impedance magnitude and phase responses are presented for various scaling factors and are compared with theoretical counterparts. The circuit application of proposed FC scaler is demonstrated through implementation of fractional order lossy and lossless integrators; and may be extended to fractional order filters, oscillators, controllers etc.

therefore, look for alternate schemes for placing a small capacitor on-chip and use a multiplier circuit. Gyrator, Generalized Impedance Converter (GIC), and Negative Impedance Converter (NIC) are also used for tuning of Capacitance multiplier (C multiplier) circuit. Such C multiplier circuits have been deployed for appropriate tuning of filters [1], [2], [3],[4], [5], [6], [7], [8], [9], [10], [11], [12] and [13] oscillators [14], PLLs [15] and series resonators [16]. Commercially available active elements such as Operational Transconductance Amplifier (OTA) [17], [18] and [19] and Op-amp [1] and [15] and AD 844 (CFOA) [20], and [21], [22] and [23] based C multipliers are reported in the literature.

Besides this, researchers are also focusing on fractional domain signal processing and generation applications especially in biomedical instrumentation and control systems. Analogous to capacitor in integer order circuits, the fractional order circuits use FC. The impedance of FC with Capacitance value ( $C_\alpha$ ) and non-integer order ( $\alpha$ ,  $0 < \alpha < 1$ ) is given by  $(s^\alpha C_\alpha)^{-1}$ .

## Keywords

**Capacitance Scaling, Current Feedback Operational Amplifier (CFOA), Capacitance Multipliers, Fractional Capacitor.**

## 1. Introduction

Monolithic integration of circuits and systems has witnessed a tremendous boost due to continuous downsizing of device dimensions. Low frequency applications such as sensing and subsequent processing of biomedical signals and integration of loop filter used in Phase Locked Loop (PLL) could not be benefited from this as these require large value capacitors. Researchers,

In the available literature, the FCs is modelled using various structures [24] such as passive Resistor-Capacitor (RC) elements arranged in the form of semi-infinite tree, domino ladder, nested ladder, and symmetric network. Researchers have also explored FC emulators based on Metal Oxide Semiconductor/Complementary Metal Oxide Semiconductor (MOS/CMOS) [25] and OTA [26]. Few physical realizations of FC are also reported in [27], however, being in primitive stage these are not commercially available. The desired value of FC for given order  $\alpha$  may be obtained by computing component (R and C) values [24] or bias currents and capacitor values [25] and [26] or by physical implementation of specific FC [27].

**Tab. 1:** Comparison on scaling methods of FC.

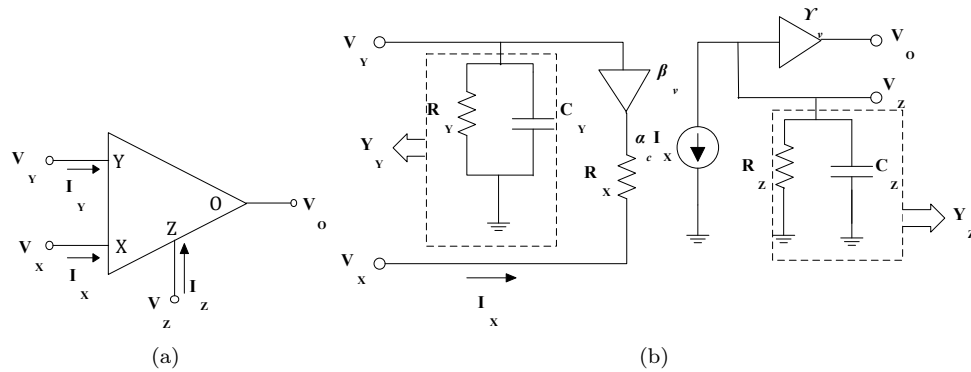
Ref.	Active element	No. of active elements	No. of passive element	FC	Solution	Additional circuit required	Experimental verification
[28]	Op-Amp	2	4	Electrolytic chemical process	GIC	No	Yes
[29]	Op-Amp	2	4	Domino ladder network	GIC	No	Yes
[30]	OTA	4	1	RC ladder network	IIMC	Impedance inverter	Yes
[31]	ECCII	3-4	3	RC ladder network	Synthetic inductor	Impedance inverter	No
[32]	CCII/ ECCII/ VGA	1/ 2/ 1	6	RC ladder network	FDNR	Impedance inverter	No
[33]	VCA	1	3	RC network	FDNR	Impedance inverter	No
[34]	OTA	4	0	RC network	Gyrator	Impedance inverter	No
[35]	OTA	17	5	Active simulation of RC network	SFG	No	No
Proposed	CFOA	1-2	2	RC ladder network	FC scaler	No	Yes

ECCII: Electronically Controllable CCII

VGA: Variable Gain Amplifier

VCA: Voltage Controlled Gain Amplifier

SFG: Signal Flow Graph

**Fig. 1:** (a) CFOA symbol and (b) its non-ideal model.

Any change in the FC value is a tedious task as it requires either recomputation of the component or bias current values or calls for a new physical realization. The capacitance scaling may be achieved using GIC [28] and [29] and Inverted Impedance Multiplier Circuit (IIMC) [30] topologies. These, however, use large number of passive elements and/or active blocks. Therefore, FC scaling through Capacitance multiplier (C multiplier) is examined in this work as an alternate solution and compared with other scaling methods having different forms of FC component Tab. 1.

This paper is organized as follows: Sec. 2. presents proposed CFOA based capacitance multiplier circuit having different form of scaling factors. It also includes investigation of impact of FC order and scaling factor of multiplier circuits on impedance values. The resulted impedances are also compared with unscaled normal capacitor. Section 3. illustrates non-ideal ef-

fect on Fractional Order (FO) capacitance scalars. The performances of realized scalars with applications are demonstrated by simulation and experimental results in Sec. 4. and Sec. 5. , respectively. Section 1. is the conclusion part.

## 2. CFOA Based C Multiplier Circuit

In this section, capacitance scaling of FC using CFOA based C multiplier circuits is presented. The symbol and equivalent non-ideal model of CFOA are shown in Fig. 1. Its terminal characteristics can be described by Eq. (1):

$$I_Y = 0, \quad V_X = V_Y, \quad I_Z = I_X, \quad V_0 = V_Z, \quad (1)$$

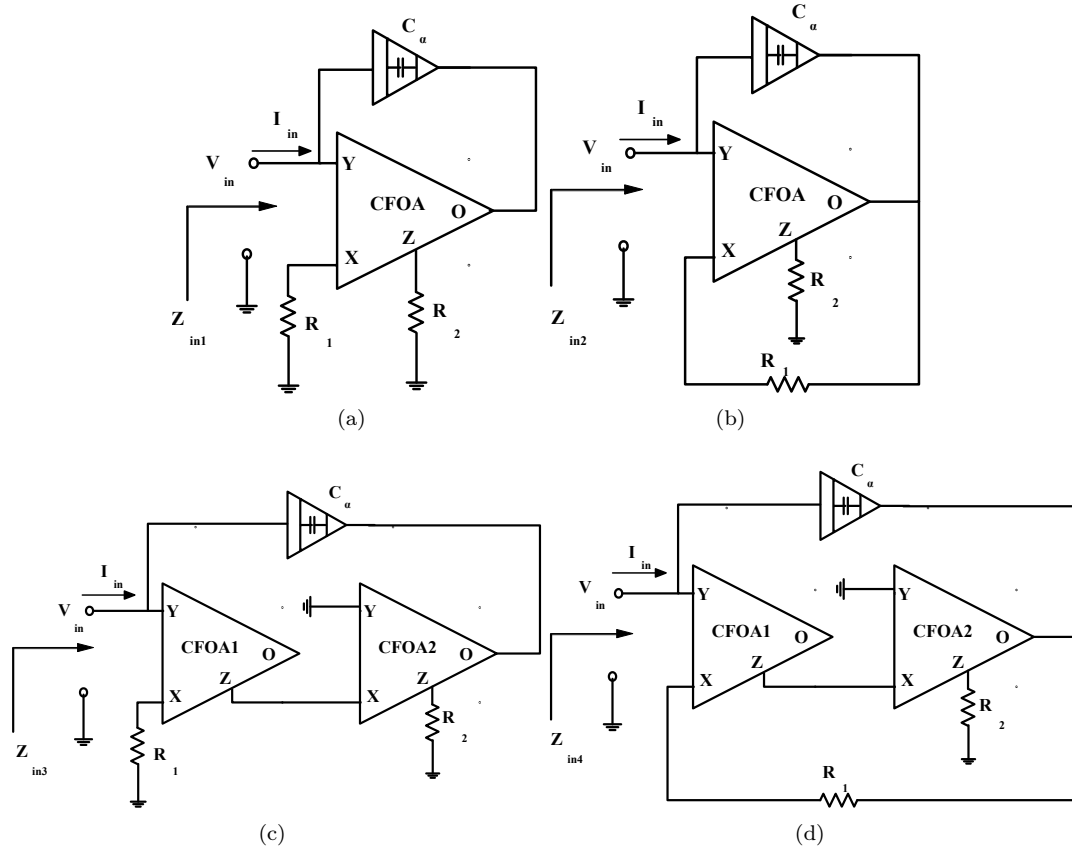


Fig. 2: CFOA based FC scaling circuits.

where  $V_X, V_Y, V_Z, V_O$  and  $I_X, I_Y, I_Z, I_O$  correspond to voltages and currents at  $X, Y, Z, O$ -terminals, respectively. In practice, the terminal characteristics may deviate from Eq. (1) due to non-idealities which appear in form of tracking errors and parasitic elements. The modified terminal characteristics are given as:

$$\begin{aligned} I_Y &= \left( sC_Y + \frac{1}{R_Y} \right) V_Y = Y_Y V_Y, \\ I_Z &= \alpha_c I_X + \left( sC_Z + \frac{1}{R_Z} \right) V_Z = I_X + Y_Z V_Z, \\ V_X &= \beta_v V_Y + I_X R_X, \quad V_O = \gamma_v V_Z, \end{aligned} \quad (2)$$

where  $\alpha_c, \beta_v, \gamma_v$  correspond to current and voltage transfer gains due to tracking errors. Parasitic elements appear in form of resistance  $R_X$ , parallel resistance capacitance combination  $R_Y // C_Y$  and  $R_Z // C_Z$  at terminals  $X, Y$ , and  $Z$ , respectively.

Figure 2 shows proposed CFOA based FC scaling circuits. The topologies in Fig. 2(a), Fig. 2(b), and Fig. 2(c) are realized by generalizing  $C$  multipliers reported in [23] while topology in Fig. 2(d) is obtained using topology reported in [36]. Routine analysis of the circuits in Fig. 2(a), Fig. 2(b), Fig. 2(c), and Fig. 2(d)

yields in the following impedance functions:

$$Z_{in1}(s) = \frac{1}{s^\alpha \left( 1 - \frac{R_2}{R_1} \right) C_\alpha}, \quad (3)$$

$$Z_{in2}(s) = \frac{\left( 1 + \frac{R_2}{R_1} \right)}{s^\alpha C_\alpha}, \quad (4)$$

$$Z_{in3}(s) = \frac{1}{s^\alpha \left( 1 + \frac{R_2}{R_1} \right) C_\alpha}, \quad (5)$$

$$Z_{in4}(s) = \frac{\left( 1 - \frac{R_2}{R_1} \right)}{s^\alpha C_\alpha}, \quad (6)$$

where order  $\alpha$  of FC provides  $-\alpha\pi \cdot 2^{-1}$  phase shift, therefore  $\alpha = 0.5$  has  $-45^\circ$  phase shift.

Equations (3), Eq. (4), Eq. (5) and Eq. (6) show that FC is scaled by factors  $K_i$  ( $i = 1, \dots, 4$ ) where  $K_1 = (1 - R_2 \cdot R_1^{-1})$ ,  $K_2 = 1/(1 + R_2 \cdot R_1^{-1})$ ,  $K_3 = (1 + R_2 \cdot R_1^{-1})$ , and  $K_4 = 1/(1 - R_2 \cdot R_1^{-1})$ .

It may be observed that the impedance functions of realized scalars are majorly influenced by two factors

$$Z_{in1}(s) |_{n=1} = \frac{1}{Y_Y + s^\alpha C_\alpha \left[ 1 - \frac{\alpha_c \beta_v}{(R_1 + R_X)(G_2 + Y_Z)} \right]}, \quad (7)$$

$$Z_{in2}(s) |_{n=2} = \frac{1}{Y_Y + s^\alpha C_\alpha \left[ 1 - \frac{\alpha_c \beta_v \gamma_v}{(R_1 + R_X)(G_2 + Y_Z) + \alpha_c \gamma_v} \right]}, \quad (8)$$

$$Z_{in3}(s) |_{n=3} = \frac{1}{Y_{Y1} + s^\alpha C_\alpha \left[ 1 + \frac{\alpha_{c1} \alpha_{c2} \beta_{v1}}{(R_1 + R_{X1})(1 + R_{X2} Y_{Z1})(G_2 + Y_{Z2})} \right]}, \quad (9)$$

$$Z_{in4}(s) |_{n=4} = \frac{1}{Y_{Y1} + s^\alpha C_\alpha \left[ 1 + \frac{\alpha_{c1} \alpha_{c2} \beta_{v1} \gamma_{v2}}{(R_1 + R_{X1})(1 + R_{X2} Y_{Z1})(G_2 + Y_{Z2} - \alpha_{c1} \alpha_{c2} \gamma_{v2})} \right]}. \quad (10)$$

(i)  $\alpha$ , and (ii)  $R_2 \cdot R_1^{-1}$ . To achieve higher scaling factor, larger resistor ratio ( $R_2 \cdot R_1^{-1}$ ) is needed for topologies in Fig. 2(a) and Fig. 2(c) whereas similar results may be obtained by selecting  $R_2 \cdot R_1^{-1}$  closer to unity for topology in Fig. 2(d). Further, all the topologies show an increasing trend in impedance for decreasing  $\alpha$  ( $\alpha < 1$ ) for fixed resistor ratio.

MATLAB simulations for change in impedance magnitude with respect to (w.r.t.)  $\alpha$  and ( $R_2 \cdot R_1^{-1}$ ) for circuits in Fig. 2(a), Fig. 2(b), Fig. 2(c) and Fig. 2(d) are demonstrated in Fig. 3(a), Fig. 3(b), Fig. 3(c) and Fig. 3(d). It is useful to examine the effect of combined variation of  $\alpha$  and ( $R_2 \cdot R_1^{-1}$ ). The change in impedance magnitude value is calculated from its original impedance value and plotted against  $\alpha$  and resistor ratio as shown in Fig. 3.

To examine the cumulative effect of variation of  $\alpha$  and  $R_2 \cdot R_1^{-1}$  on change in impedance magnitude, MATLAB simulations are carried out for circuits of Fig. 2(a), Fig. 2(b), Fig. 2(c) and Fig. 2(d) and corresponding results are plotted in Fig. 3(a), Fig. 3(b), Fig. 3(c) and Fig. 3(d). Here % change of  $|Z|$  is calculated from original value w.r.t.  $\alpha$  and resistor ratio. The simulation results corroborate with the theoretical results. It may be noted that smaller  $\alpha$  values have larger impact on impedance magnitude for a resistor ratio  $R_2 \cdot R_1^{-1}$ , i.e. close to unity for Fig. 2(a)/ much larger than unity for Fig. 2(b)/ negligible for Fig. 2(a), Fig. 2(b), Fig. 2(c) and Fig. 2(d).

### 3. Non-Ideal Analysis of FC Scalers

To analyse the behaviour of proposed circuits in presence of CFOA non-idealities (Fig. 1(b)), the input

impedance functions of topologies of Fig. 2 are recomputed as Eq. (7), Eq. (8), Eq. (9) and Eq. (10), where subscript  $n$  corresponds to nonideal; and subscripts 1 and 2 with current transfer gain ( $\alpha_{c1,2}$ ), voltage transfer gains ( $\beta_v, \gamma_v$ ); and parasitics,  $Y_Y$ ,  $R_X$ ,  $Y_Z$  correspond to CFOA1 and CFOA2.

It may be noted from the graphs in Fig. 2(a), Fig. 2(b), Fig. 2(c) and Fig. 2(d) that Y-terminal of CFOA/CFOAs is either connected to input or to ground, therefore the performance remains unaltered due to parasitic associated with this terminal. The parasitic at X terminal of CFOA may be accommodated by adjusting the value of external resistor connected to it. The overall impact of CFOA parasitic elements on FC scaler behaviour may be ignored by considering the frequency of operation much below than parasitic pole associated with Z terminal. Therefore, ( $K_i$ ,  $i = 1, 2, 3, 4$ ) for topologies of Fig. 2(a), Fig. 2(b), Fig. 2(c) and Fig. 2(d) are recomputed as:

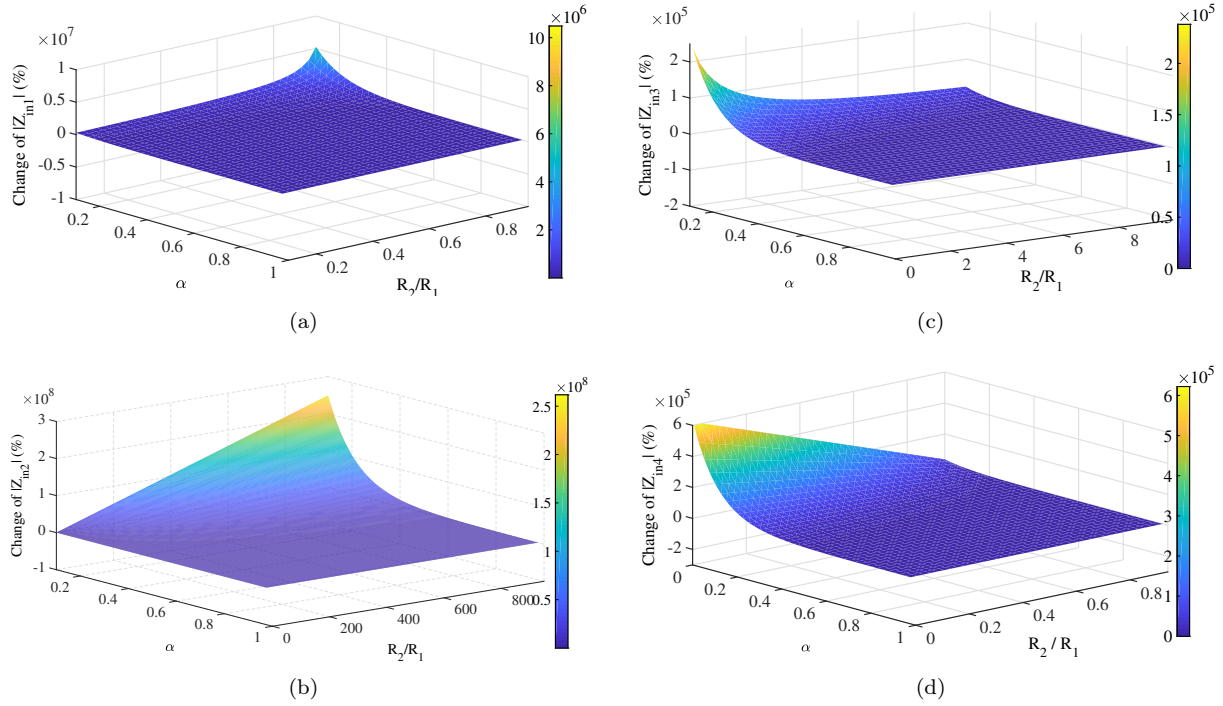
$$K_1 |_{n=1} = 1 - \alpha_c \beta_v \frac{R_2}{R_1}, \quad (11)$$

$$K_2 |_{n=2} = \frac{1 + \left( \alpha_c \gamma_v - \alpha_c \beta_v \gamma_v \frac{R_2}{R_1} \right)}{1 + \alpha_c \gamma_v \frac{R_2}{R_1}}, \quad (12)$$

$$K_3 |_{n=3} = 1 + \alpha_{c1} \alpha_{c2} \beta_{v1} \frac{R_2}{R_1}, \quad (13)$$

$$K_4 |_{n=4} = \frac{1 - \left( \alpha_{c1} \alpha_{c2} \gamma_{v2} - \alpha_{c1} \alpha_{c2} \beta_{v1} \gamma_{v2} \frac{R_2}{R_1} \right)}{1 - \alpha_{c1} \alpha_{c2} \gamma_{v2} \frac{R_2}{R_1}}. \quad (14)$$

The parasitic effects of CFOA based FC multipliers (Fig. 2) are demonstrated using SPICE simulation results as shown in Fig. 5. The typical values of parasitic



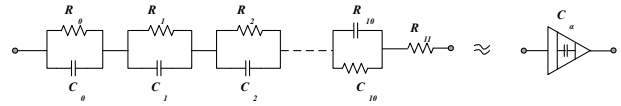
**Fig. 3:** Percent change in impedance magnitude with respect to  $\alpha$  and  $R_2 \cdot R_1^{-1}$ .

impedances of CFOA are  $R_x = 50 \Omega$ ,  $R_Y = 2 \text{ M}\Omega$ ,  $C_Y = 2 \text{ pF}$ ,  $R_Z = 3 \text{ M}\Omega$ ,  $C_Z = 4.5 \text{ pF}$ . It is clear that the parasitic element on X-terminal affects magnitude responses more with maximum errors of 3 dB, -1.18 dB and -4.14 dB for the circuits of Fig. 2(a), Fig. 2(b), Fig. 2(c) and Fig. 2(d), respectively in the operational frequency range of  $\alpha$  order FC while it produces -0.6 dB magnitude error with Y and Z-terminal parasitic elements of Fig. 2(b). The phase responses of FC multipliers as shown in Fig. 2(a) and Fig. 2(b) are more influenced by the parasitic of Z-terminal at lower frequency and Y-and X-terminals at higher frequency. Whereas the (Y-, X-) and (Y-, Z-) terminals parasitic elements have a few more effect on the phase responses of Fig. 2(a), Fig. 2(b), Fig. 2(c) and Fig. 2(d), respectively.

## 4. Application and Simulation Results

The functionality of the realized scalers is verified using SPICE simulations using CFOA model [37]. The FC is implemented using infinite order domino RC ladder network truncated Fig. 4 to 12 blocks [38]. The component values of FC model having  $\alpha = 0.5$  and  $C_{\alpha=3.75} \mu\text{F} \cdot \text{s}^{-0.5}$  are  $R_0 = 330 \text{ k}\Omega$ ,  $R_1 = 82 \text{ k}\Omega$ ,  $R_2 = 33 \text{ k}\Omega$ ,  $R_3 = 12 \text{ k}\Omega$ ,  $R_4 = 4.7 \text{ k}\Omega$ ,  $R_5 = 2 \text{ k}\Omega$ ,  $R_6 = 736 \Omega$ ,  $R_7 = 270 \Omega$ ,  $R_8 = 120 \Omega$ ,  $R_9 = 47 \Omega$ ,  $R_{10} = 8.2 \Omega$ ,  $R_{11} = 18.2 \Omega$ ,  $C_0 = 4.7 \mu\text{F}$ ,  $C_1 = 3.1 \mu\text{F}$ ,

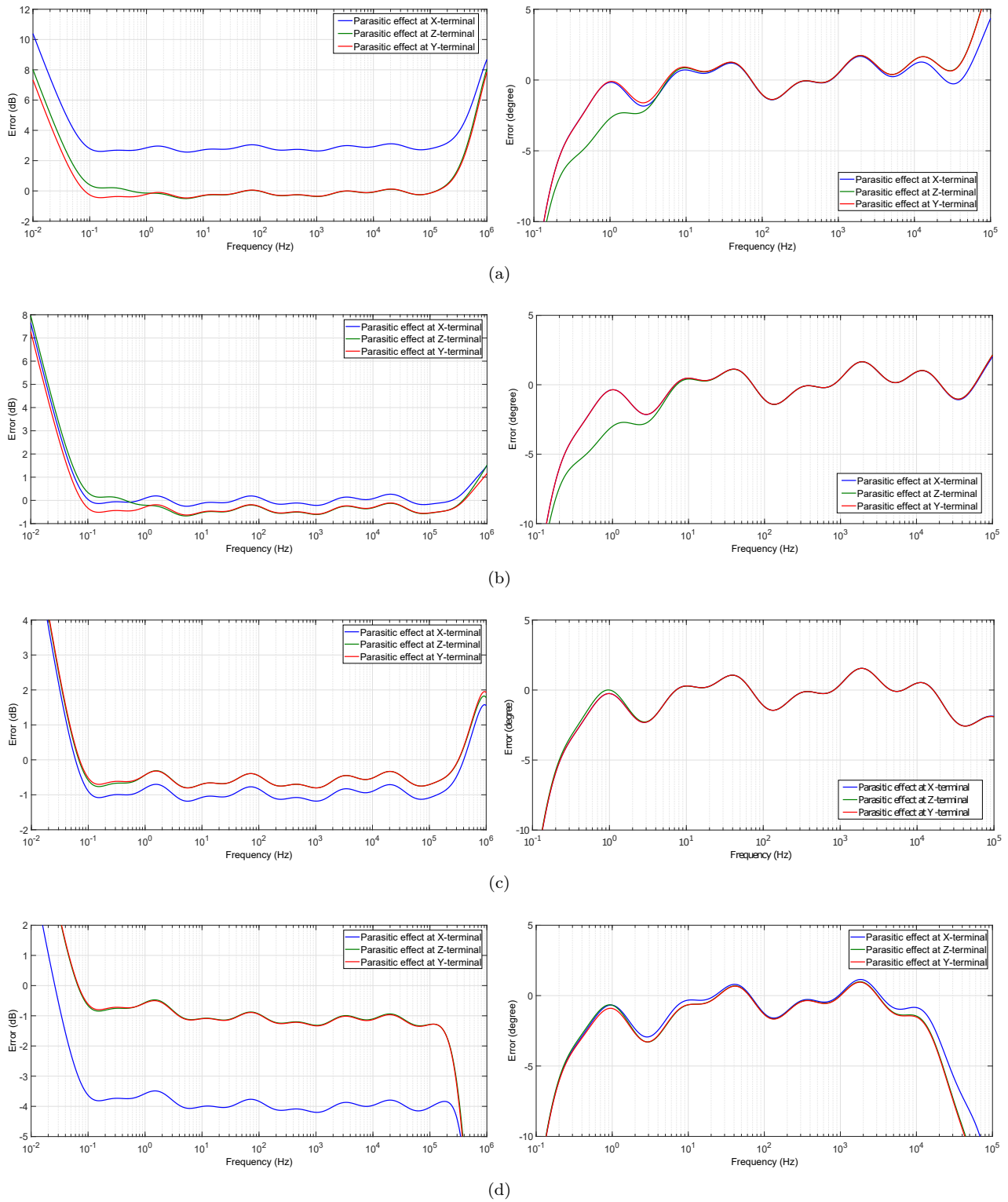
$C_2 = 1 \mu\text{F}$ ,  $C_3 = 470 \text{ nF}$ ,  $C_4 = 168 \text{ nF}$ ,  $C_5 = 68 \text{ nF}$ ,  $C_6 = 27 \text{ nF}$ ,  $C_7 = 10 \text{ nF}$ ,  $C_8 = 4.7 \text{ nF}$ ,  $C_9 = 1 \text{ nF}$ ,  $C_{10} = 2.2 \text{ nF}$ .



**Fig. 4:** Truncated RC domino ladder network realizing FC. [36]

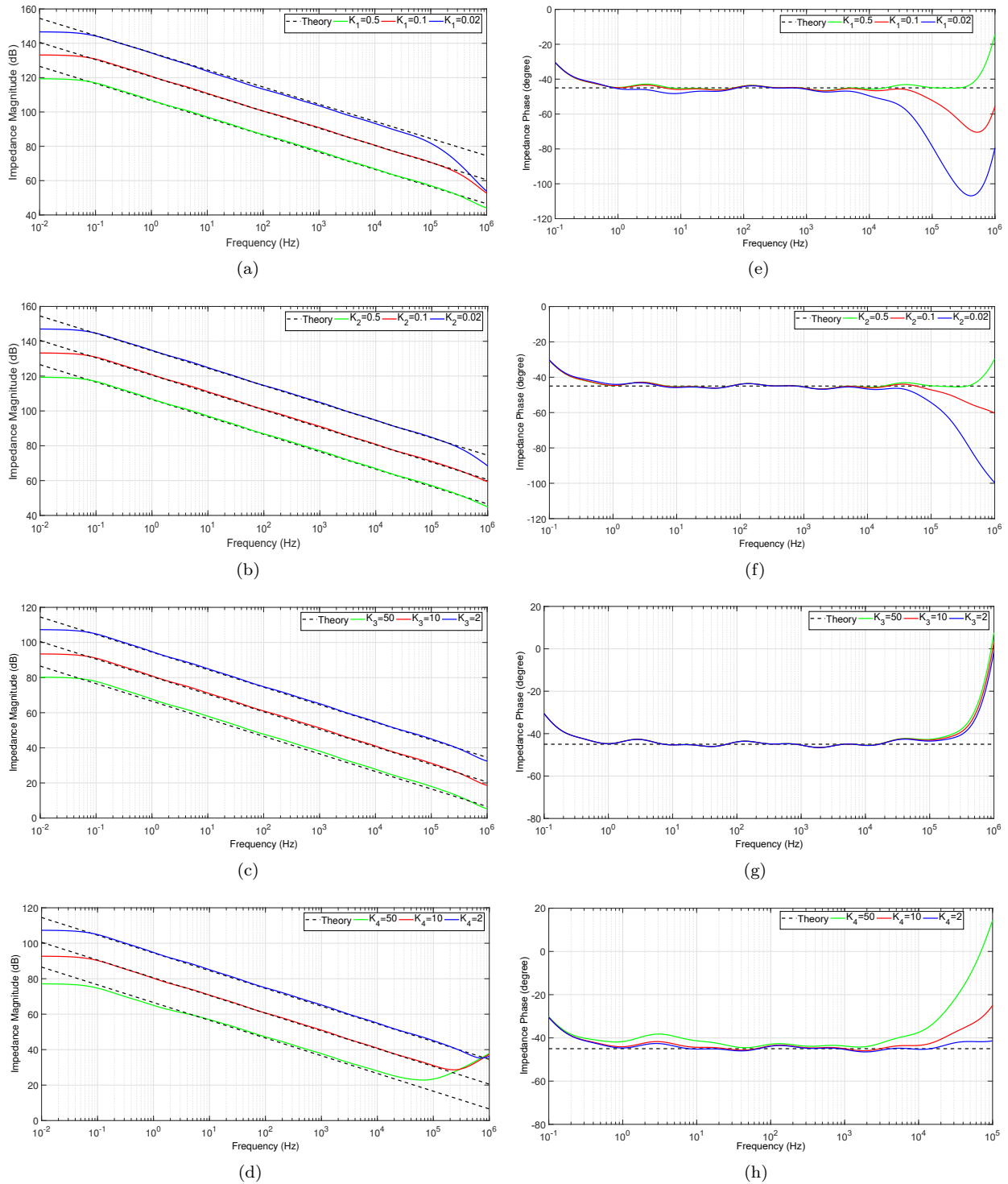
Simulations are performed for different scaling factors for topologies in Fig. 2 for examining impedance magnitude and phase response and corresponding results are depicted in Fig. 6. The excitation voltage is set at 100 mV for the simulation results. The excitation voltage is set 100 mV for the simulation results. Table 2 enlists simulation setting for capacitance scaling factors and component settings used therein and performance of circuits. In the view of non-ideal effects of CFOA on the realized circuits, it may be observed that Fig. 2(b) and Fig. 2(c), have more linearity than Fig. 2(a) and Fig. 2(d). In Fig. 2(a) and Fig. 2(d), it increases the range of operation for lesser value of  $R_2 \cdot R_1^{-1}$ .

To illustrate the use of proposed scaler, CFOA based lossy/lossless integrator circuit is constructed as shown in Fig. 9. The notation  $C_{\text{eff}}$  indicates the effective capacitance value of FC (that is  $C_{\text{eff}} = K_i \cdot C_{\alpha}$ ). The transfer functions of lossy and lossless fractional



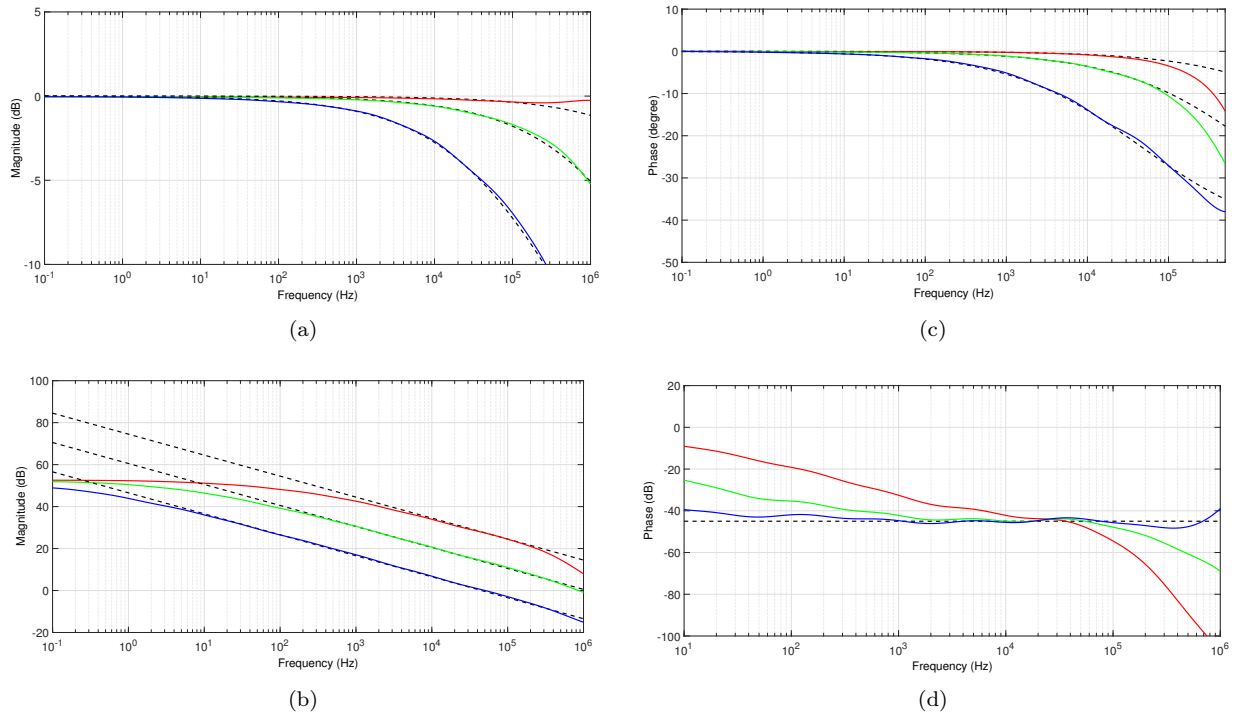
**Fig. 5:** The magnitude and phase errors for (a) Fig. 2(a), (b) Fig. 2(b), (c) Fig. 2(c) and (d) Fig. 2(d).



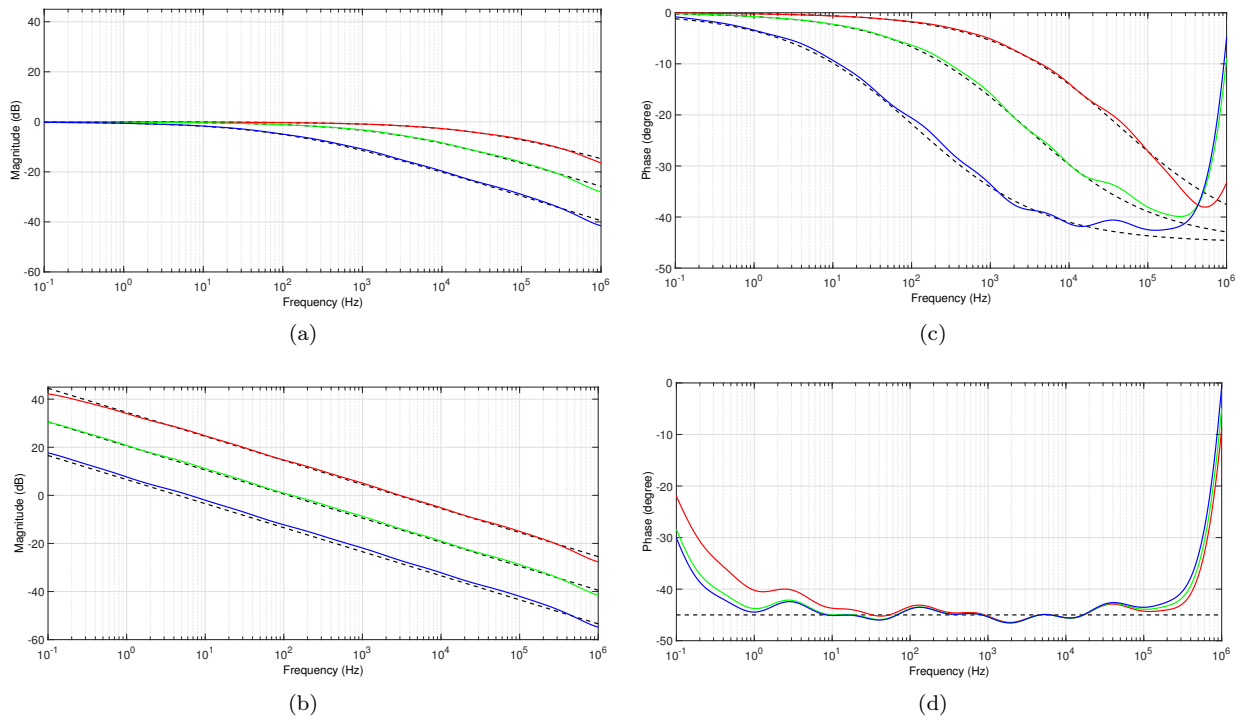


**Fig. 6:** Simulated impedance magnitude (a) – (d) and phase (e) – (h) responses for circuits of Fig. 2(a), Fig. 2(b), Fig. 2(c) and Fig. 2(d).





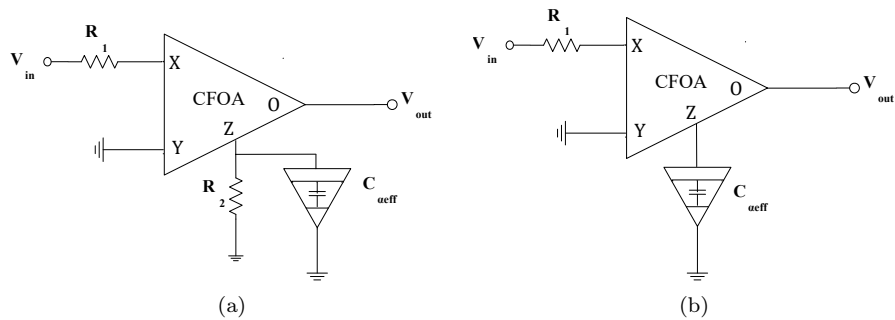
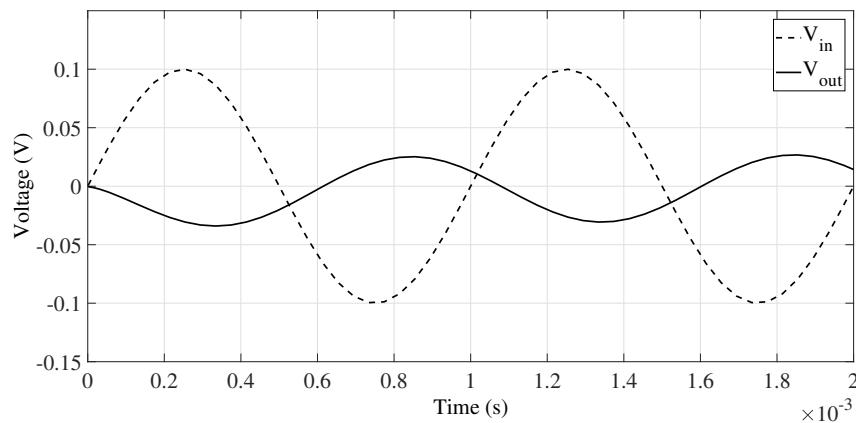
**Fig. 7:** Magnitude (a, b) and phase (c, d) responses of fractional lossy and lossless integrators for multiplication factor  $< 1$ .



**Fig. 8:** Magnitude (a, b) and phase (c, d) responses of fractional lossy and lossless integrators for multiplication factor  $> 1$ .

**Tab. 2:** The components values and performance of FO capacitance scalars.

Components setting and performance evaluation	Fig. 2(a)			Fig. 2(b)			Fig. 2(c)			Fig. 2(d)		
Multiplication factor	0.02	0.1	0.5	0.02	0.1	0.5	2	10	50	2	10	50
$R_1$ (k $\Omega$ )	1	1	1	1	1	1	1	1	1	1	1	1
$R_2$ (k $\Omega$ )	0.98	0.9	0.5	49	9	1	1	9	49	0.5	0.9	0.98
$(C\alpha)_{eff}F(v/s^\alpha)$	75 n	0.375 $\mu$	1.875 $\mu$	75 n	0.375 $\mu$	1.875 $\mu$	7.5 $\mu$	37.5 $\mu$	187.5 $\mu$	7.5 $\mu$	37.5 $\mu$	187.5 $\mu$
Frequency range of magnitude response (Hz) (within 1.5 dB deviation)	0.05–50.1 k	0.04–330 k	0.042–588 k	0.046–392 k	0.042–1 Meg	0.042–935 k	0.04–625 k	0.04–676 k	0.04–741 k	0.042–970 k	0.052–218 k	1.4–14.8 k
Frequency range of phase response (Hz) (within 2.5° deviation)	13.3–6 k	0.44–53 k	0.43–426 k	0.57–44 k	0.45–107 k	0.42–525 k	0.4–202 k	0.4–154 k	0.4–120 k	0.44–28 k	4.6–14.5 k	19.5–3 k

**Fig. 9:** CFOA based fractional (a) lossy and (b) lossless integrators.**Fig. 10:** Transient responses of fractional lossy integrator.

integrators can be expressed as follows:

$$T_{\text{lossy}}^\alpha = \frac{V_o(s)}{V_{in}(s)} = -\frac{R_2}{R_1} \cdot \frac{1}{1 + s^\alpha R_2 C_{\text{eff}}}, \quad (15)$$

$$T_{\text{lossless}}^\alpha = \frac{V_o(s)}{V_{in}(s)} = -\frac{1}{s^\alpha R_1 C_{\text{eff}}}. \quad (16)$$

The simulated magnitude and phase responses of fractional order lossy/lossless integrators using 0.5 order FC scaler of Fig. 2(b) and Fig. 2(c) with scaling factors of (0.02, 0.1, 0.5) and (2, 10, 50) are depicted in Fig. 7 and Fig. 8, respectively. It may be observed that the simulated magnitude responses for lossy and lossless integrators follow theoretical values with deviations of (0.45 dB, 0.7 dB, 1.5 dB) and (0.6 dB,

0.8 dB, 1.5 dB) up to frequencies (478 kHz, 1.58 MHz, 2.7 MHz) / (380 kHz, 457 kHz, 1.7 MHz) and (1.5 kHz–346 kHz, 165 Hz–660 kHz, 2.8 Hz–891 kHz) / (410 kHz, 483 kHz, 1.7 MHz) for scaling factors (0.02, 0.1, 0.5) / (2, 10, 50), respectively. Further, phase deviations are well within 2.5° for frequencies up to (190 kHz, 200 kHz, 295 kHz) / (370 kHz, 280 kHz, 215 kHz) for lossy integrator and that for lossless integrator in the frequency range of (7.2 kHz–64 kHz, 1 kHz–94 kHz, 186 Hz–234 kHz) / (5.6 Hz–343 kHz, 3.3 Hz–278 kHz, 0.43 Hz–214 kHz). The transient response of fractional lossy integrator circuit is shown in Fig. 10 where input 100 mV amplitude and 1 kHz frequency sinusoidal signal is applied. The output is plotted for  $K_3 = 10$ , which verifies phase difference of  $-5.5^\circ$  closer to  $-34^\circ$  theoretical value.

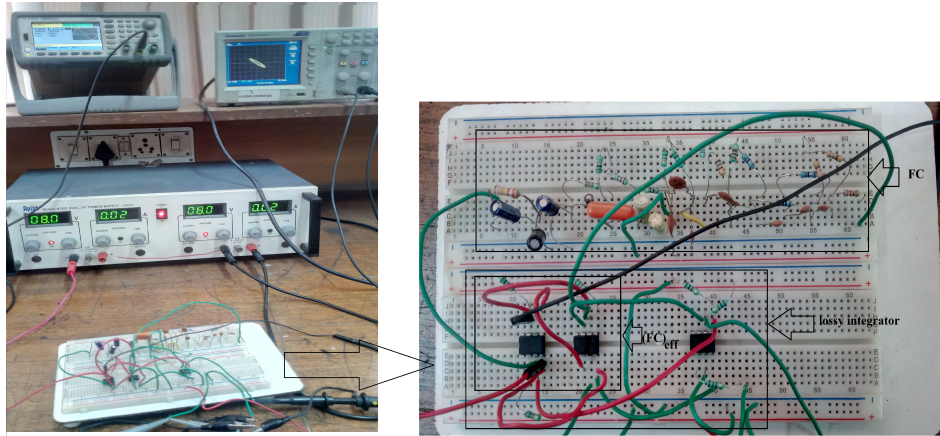


Fig. 11: Experimental setup for FO lossy integrator circuit.

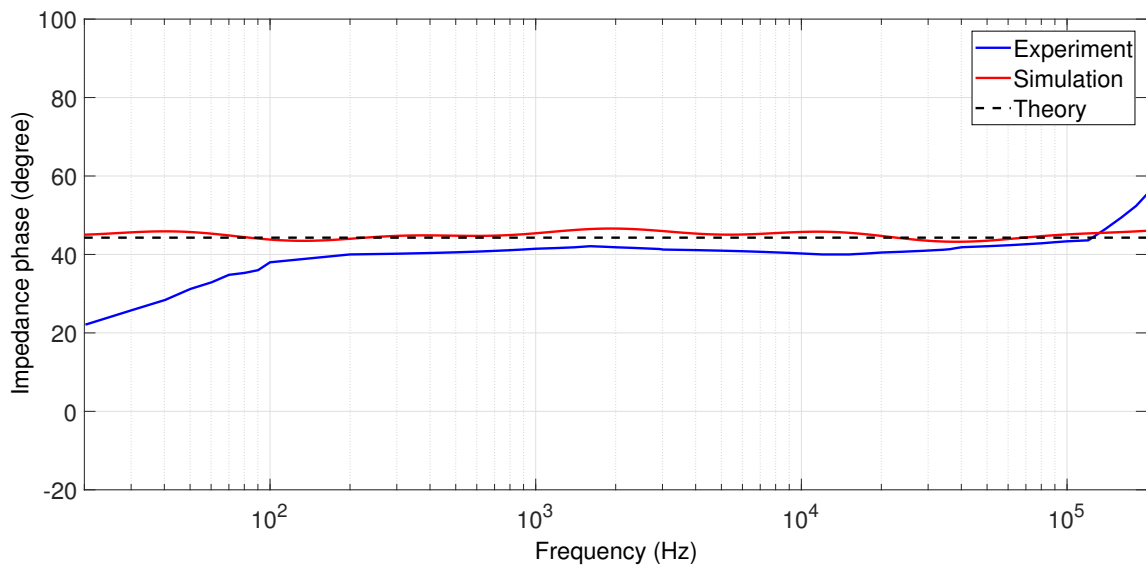


Fig. 12: Impedance phase response of 0.5 order FC.

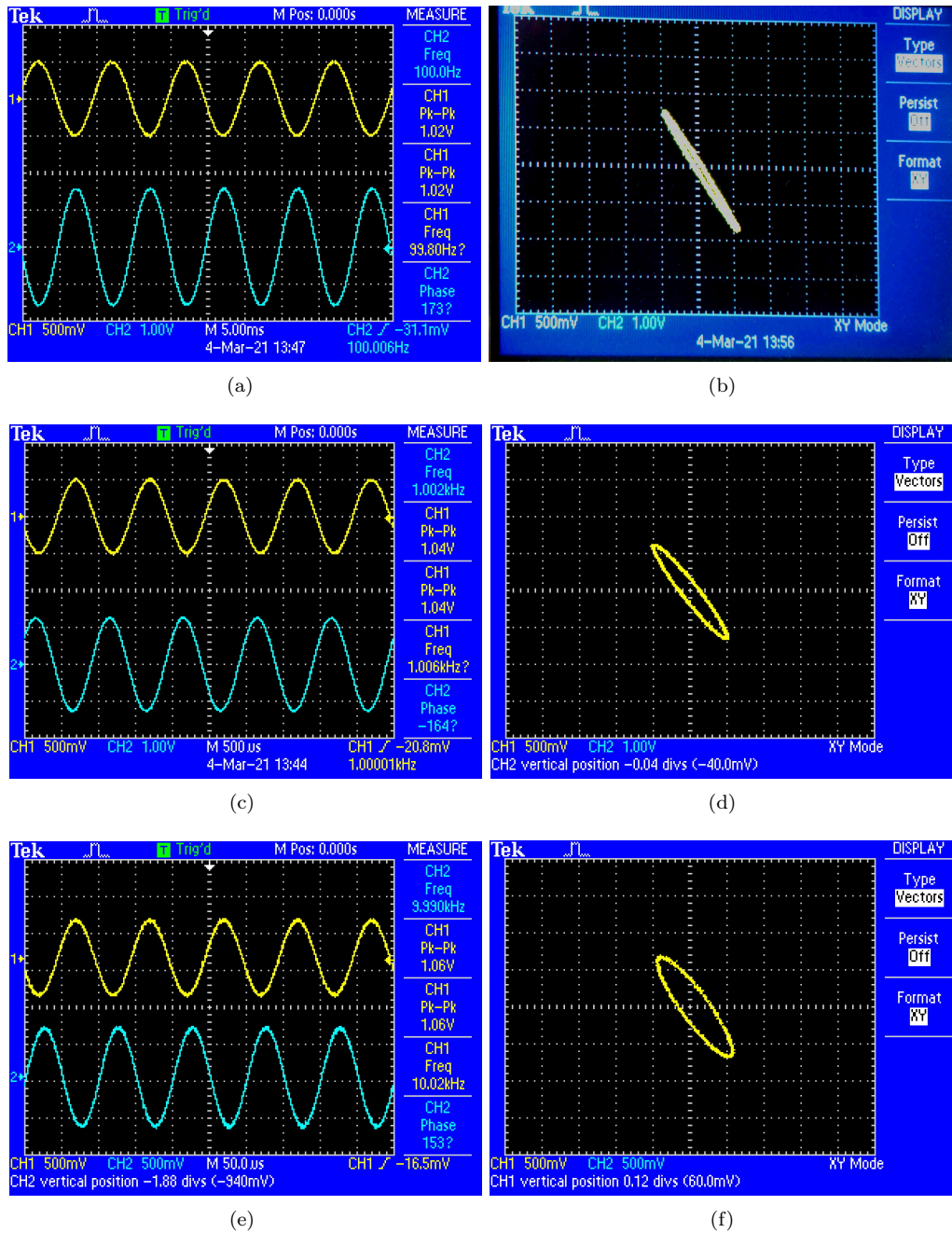
## 5. Experimental Verification

The realization of FC having  $\alpha$  closed to 0.5 and  $C_\alpha = 3.75 \mu\text{F}/s^\alpha$  using infinite RC ladder network is considered for experimental verification. The components values of ladder network (Fig. 4) are  $R_0 = 330 \text{ k}\Omega$ ,  $R_1 = 82 \text{ k}\Omega$ ,  $R_2 = 33 \text{ k}\Omega$ ,  $R_3 = 12 \text{ k}\Omega$ ,  $R_4 = 4.7 \text{ k}\Omega$ ,  $R_5 = 2 \text{ k}\Omega$ ,  $R_6 = 736 \Omega$ ,  $R_7 = 270 \Omega$ ,  $R_8 = 120 \Omega$ ,  $R_9 = 47 \Omega$ ,  $R_{10} = 8.2 \Omega$ ,  $R_{11} = 18.2 \Omega$ ,  $C_0 = 4.7 \mu\text{F}$ ,  $C_1 = 3.1 \mu\text{F}$ ,  $C_2 = 1 \mu\text{F}$ ,  $C_3 = 470 \text{ nF}$ ,  $C_4 = 168 \text{ nF}$ ,  $C_5 = 68 \text{ nF}$ ,  $C_6 = 27 \text{ nF}$ ,  $C_7 = 10 \text{ nF}$ ,  $C_8 = 4.7 \text{ nF}$ ,  $C_9 = 1 \text{ nF}$ ,  $C_{10} = 2.2 \text{ nF}$ . The impedance phase response of FC is plotted as shown in Fig. 12. The impedance phase value of FC (Fig. 4) is calculated with phase differences between FC and  $R_{11}$  using two voltage probes of oscilloscopes.

The experiment is performed using commercially available AD844AN (CFOA). The hardware setup for FO capacitance scalar, FC, and lossy integrator circuits

(Fig. 2(c), Fig. 5, and Fig. 6) is shown in Fig. 11. The approach to test the characteristics of Fig. 2(c) employing Fig. 9 is implemented for multiplication factor  $K_3 = 3.2$  ( $R_1 = 1 \text{ k}\Omega$  and  $R_2 = 2.2 \text{ k}\Omega$ ), order  $\alpha = 0.5$  and capacitance  $C_\alpha = 3.75 \mu\text{F}/s^\alpha$  of FC (the used component values as given in Sec. 4. ).

The component values of lossy integrator ( $R_1 = 0.22 \text{ k}\Omega$  and  $R_2 = 1 \text{ k}\Omega$ ) provides dc gain equals to 13.15 dB and its response is shown in Fig. 13. The performance is observed for sinusoidal input supply  $V_{\text{peak-peak}} = 1 \text{ V}$  at 100 Hz, 1 kHz, and 10 kHz frequency points. Figure 13 also shows the responses of sinusoidal input and their Lissajous figures. The dc supply voltage in CFOA is fixed at  $\pm 8 \text{ V}$ . Moreover, Fig. 13 displays gain of 10 dB, 8 dB, 2.4 dB and phase of 1730, 1640, 1530 where calculated gain = 11.34 dB, 8.03 dB, 1.6 dB and phase = 1700, 159.10, 145.80 at corresponding 100 Hz, 1 kHz, 10 kHz frequency points. Thus, it can be stated that the experimental work verifies the theoretical study.



**Fig. 13:** Responses (channel 2) of Fig. 9 choosing dc gain = 13.15 dB for sinusoidal input (channel 1) at (a) 100 Hz, (c) 1 kHz, and (e) 10 kHz frequency points and (b, d and f) their Lissajous pattern.

## 6. Conclusion

Four fractional capacitance scaler topologies obtained through generalization of CFOA based capacitance multipliers are presented in this paper. The effect of non idealities of CFOA on proposed scalers is investigated. Functionality of these scalers is tested through MATLAB and SPICE simulations for various scaling factors. The application of proposed scaler is illustrated through fractional order lossy and lossless integrators.

## Author Contributions

R.V., N.P. and R.P. conceived of the presented idea. R.V. developed the theory and performed the computations. N.P. and R.P. verified the analytical methods. N.P. and R.P. encouraged R.V., to investigate practical application and supervised the findings of this work. All authors discussed the results and contributed to the final manuscript.



## References

- [1] MARCELLIS, A. D., G. FERRI and V. STORNELLI. NIC-based capacitance multipliers for low-frequency integrated active filter applications. In: *2007 Ph.D Research in Microelectronics and Electronics Conference*. Bordeaux: IEEE, 2007, pp. 225–228. ISBN 978-1-4244-1000-2. DOI: 10.1109/RME.2007.4401853.
- [2] LI, Y.-A. A series of new circuits based on CFTAs. *AEUE - International Journal of Electronics and Communications*. 2012, vol. 66, iss. 7, pp. 587–592. ISSN 1434-8411. DOI: 10.1016/j.aeue.2011.11.011.
- [3] AYTEN, U. E., M. SAGBAS, N. HERENC SAR and J. KOTON. Novel Floating General Element Simulators Using CBTA. *Radioengineering*. 2012, vol. 21, iss. 1, pp. 11–19. ISSN 1805-9600.
- [4] MYDERRIZI, I. and A. ZEKI. Electronically tunable DXCCII-based grounded capacitance multiplier. *AEU - International Journal of Electronics and Communications*. 2014, vol. 68, iss. 9, pp. 899–906. ISSN 1434-8411. DOI: 10.1016/j.aeue.2014.04.013.
- [5] ALPASLAN, H. DVCC-based floating capacitance multiplier design. *Turkish Journal of Electrical Engineering & Computer Sciences*. 2017, vol. 25, iss. 1, pp. 1334–1345. ISSN 1303-6203. DOI: 10.3906/elk-1509-112.
- [6] ALPASLAN, H. and E. YUCE. Bandwidth expansion methods of inductance simulator circuits and voltage-mode biquads. *Journal of Circuits, Systems and Computers*. 2011, vol. 20, iss. 3, pp. 557–572. ISSN 1793-6454. DOI: 10.1142/S0218126611007451.
- [7] UNHAVANICH, S., O. ONJAN and W. TANGSRIRAT. Tunable Capacitance Multiplier with a Single Voltage Differencing Buffered Amplifier. In: *Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS)*. Hong Kong: International Association of Engineers, 2016, pp. 1–4. ISBN 978-988-14047-6-3.
- [8] SILAPAN, P., C. TANAPHATSIRI and M. SIRIPRUCHYANUN. Current Controlled CCTA Based- Novel Grounded Capacitance Multiplier with Temperature Compensation. In: *IEEE Asia-Pacific Conference on Circuits and Systems (APCCAS)*. Macao: IEEE, 2008, pp. 1490–1493. ISBN 978-1-4244-2342-2. DOI: 10.1109/APCCAS.2008.4746314.
- [9] YUCE, E. On the realization of the floating simulators using only grounded passive components. *Analog Integrated Circuits and Signal Processing*. 2006, vol. 49, iss. 2, pp. 161–166. ISSN 0925-1030. DOI: 10.1007/s10470-006-9351-7.
- [10] JANTAKUN, A., N. PISUTTHIPONG and M. SIRIPRUCHYANUN. Single element based-novel temperature insensitive/electronically controllable floating capacitance multiplier and its application. In: *International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. Chiang Mai: IEEE, 2010, pp. 37–41. ISBN 978-1-4244-5607-9.
- [11] YUCE, E. A novel floating simulation topology composed of only grounded passive components. *International Journal of Electronics*. 2010, vol. 97, iss. 3, pp. 249–262. ISSN 1362-3060. DOI: 10.1080/00207210903061907.
- [12] PROMMEE, P. and M. SOMDUNYAKANOK. CMOS-based current-controlled DDCC and its applications to capacitance multiplier and universal filter. *AEU - International Journal of Electronics and Communications*. 2011, vol. 65, iss. 1, pp. 1–8. ISSN 1434-8411. DOI: 10.1016/j.aeue.2009.12.002.
- [13] KARTCI, A., U. E. AYTEN, N. HERENC SAR, R. SOTNER, J. JERABEK and K. VRBA. Application possibilities of VDCC in general floating element simulator circuit. In: *European Conference on Circuit Theory and Design (ECCTD)*. Trondheim: IEEE, 2015, pp. 1–4. ISBN 978-1-4799-9877-7. DOI: 10.1109/ECCTD.2015.7300064.
- [14] AL-ABSI, M. A., E. S. AL-SUHAIBANI and M. T. ABUELMA'ATTI. A new compact CMOS C-multiplier. *Analog Integrated Circuits and Signal Processing*. 2017, vol. 90, iss. 3, pp. 653–658. ISSN 1573-1979. DOI: 10.1007/s10470-016-0822-1.
- [15] TANG, Y., M. ISMAIL and S. BIBYK. Adaptive Miller capacitor multiplier for compact on-chip PLL filter. *Electronics Letters*. 2003, vol. 39, iss. 1, pp. 43–45. ISSN 1350-911X. DOI: 10.1049/el:20030086.
- [16] YUCE, E. Floating inductance, FDNR and capacitance simulation circuit employing only grounded passive elements. *International Journal of Electronics*. 2006, vol. 93, iss. 10, pp. 679–688. ISSN 1362-3060. DOI: 10.1080/00207210600750208.
- [17] JAIKLA, W. and M. SIRIPRUCHYANAN. An Electronically Controllable Capacitance Multiplier with Temperature Compensation. In: *International Symposium on Communications and Information Technologies*. Bangkok:

- IEEE, 2006, pp. 356–359. ISBN 0-7803-9740-1. DOI: 10.1109/ISCIT.2006.340064.
- [18] KHAN, I. A. and M. T. AHMED. OTA-based integrable voltage/current-controlled ideal C-multiplier. *Electronics Letters*. 1986, vol. 22, iss. 7, pp. 365–366. ISSN 1350-911X. DOI: 10.1049/el:19860248.
- [19] AHMED, M. T., I. A. KHAN and N. MINHAJ. Novel electronically tunable C-multipliers. *Electronics Letters*. 1995, vol. 31, iss. 1, pp. 9–11. ISSN 1350-911X. DOI: 10.1049/el:19950018.
- [20] YUCE, E. and S. MINAEI. A Modified CFOA and Its Applications to Simulated Inductors, Capacitance Multipliers, and Analog Filters. *IEEE Transactions on Circuits and Systems I: Regular Papers*. 2008, vol. 55, iss. 1, pp. 266–275. ISSN 1558-0806. DOI: 10.1109/TCSI.2007.913689.
- [21] ABUELMA'ATTI, M. T., S. K. DHAR and Z. J. KHALIFA. New two-CFOA-based floating immittance simulators. *Analog Integrated Circuits and Signal Processing*. 2008, vol. 55, iss. 1, pp. 266–275. ISSN 1558-0806. DOI: 10.1109/TCSI.2007.913689.
- [22] ABUELMA'ATTI, M. T. and S. K. DHAR. New CFOA-based floating lossless negative immittance function emulators. In: *IEEE Region 10 International Conference (TENCON)*. Macao: IEEE, 2015, pp. 1–4. ISBN 978-1-4799-8641-5. DOI: 10.1109/TENCON.2015.7372799.
- [23] KHAN, A. A., S. BIMAL, K. K. DEY and S. S. ROY. Current conveyor based R- and C- multiplier circuits. *AEU-International Journal of Electronics and Communications*. 2002, vol. 56, iss. 5, pp. 312–316. ISSN 1434-8411. DOI: 10.1078/1434-8411-54100121.
- [24] PODLUBNY, I., I. PETRAS, B. M. VINAGRE, P. O. LEARY and L. DORCAK. Analogue Realizations of Fractional-Order Controllers. *Nonlinear Dynamics*. 2002, vol. 29, iss. 1, pp. 281–296. ISSN 1573-269X. DOI: 10.1023/A:1016556604320.
- [25] BERTSIAS, P., C. PSYCHALINOS, A. S. ELWAKIL and A. G. RADWAN. Low-voltage and low-power fractional-order parallel tunable resonator. *Microelectronics Journal*. 2019, vol. 88, iss. 1, pp. 108–116. ISSN 1879-2391. DOI: 10.1016/j.mejo.2019.05.002.
- [26] TSIRIMOKOU, G., C. PSYCHALINOS, A. S. ELWAKIL and K. N. SALAMA. Electronically Tunable Fully Integrated Fractional-Order Resonator. *IEEE Transactions on Circuits and Systems II: Express Briefs*. 2018, vol. 65, iss. 2, pp. 166–170. ISSN 1558-3791. DOI: 10.1109/TC-SII.2017.2684710.
- [27] AGAMBAYEV, A., S. P. PATOLE, M. FARHAT, A. ELWAKIL, H. BAGCI and K. N. SALAMA. Ferroelectric Fractional-Order Capacitors. *Chem-ElectroChem*. 2017, vol. 4, iss. 11, pp. 2807–2813. ISSN 2196-0216. DOI: 10.1002/celec.201700663.
- [28] TRIPATHY, M. C., D. MONDAL, K. BISWAS and S. SEN. Experimental studies on realization of fractional inductors and fractional-order bandpass filters. *International Journal of Circuit Theory and Applications*. 2015, vol. 43, iss. 9, pp. 1183–1196. ISSN 1097-007X. DOI: 10.1002/cta.2004.
- [29] ADHIKARY, A., S. SEN and K. BISWAS. Practical Realization of Tunable Fractional Order Parallel Resonator and Fractional Order Filters. *IEEE Transactions on Circuits and Systems I: Regular Papers*. 2016, vol. 63, iss. 8, pp. 1142–1151. ISSN 1558-0806. DOI: 10.1109/TCSI.2016.2568262.
- [30] VERMA, R., N. PANDEY and R. PANDEY. Realization of a higher fractional order element based on novel OTA based IIMC and its application in filter. *Analog Integrated Circuits and Signal Processing*. 2018, vol. 97, iss. 1, pp. 177–191. ISSN 1573-1979. DOI: 10.1007/s10470-018-1315-1.
- [31] DOMANSKY, O., SOTNER, R. and L. LANGHAMMER. Reconfigurable Impedance Converter for Synthesis of Integer and Fractional-Order Synthetic Elements. In: *International Conference on Telecommunications and Signal Processing (TSP)*. Athens: IEEE, 2018, pp. 1–5. ISBN 978-1-5386-4695-3. DOI: 10.1109/TSP.2018.8441376.
- [32] DOMANSKY, O., R. SOTNER, J. PETRZELA, L. LANGHAMMER and T. DOSTAL. Higher order differentiator block for synthesis of controllable frequency dependent elements. In: *27th International Conference Radioelektronika (RADIOELEKTRONIKA)*. Brno: IEEE, 2017, pp. 1–5. ISBN 978-1-5090-4591-4. DOI: 10.1109/RADIOELEK.2017.7937587.
- [33] SOTNER, R., J. JERABEK, L. LANGHAMMER, J. KOTON, D. ANDRIUKAITIS and A. MERFELDAS. Higher order differentiator block for synthesis of controllable frequency dependent elements. In: *43rd International Conference on Telecommunications and Signal Processing (TSP)*. Milan: IEEE,

2020, pp. 578–582. ISBN 978-1-7281-6376-5. DOI: 10.1109/TSP49548.2020.9163553.

- [34] KUBANEK, D., J. KOTON, J. DVORAK, N. HERENC SAR and R. SOTNER. Optimized Design of OTA-Based Gyrator Realizing Fractional-Order Inductance Simulator: A Comprehensive Analysis. *Applied Sciences*. 2021, vol. 11, iss. 1, pp. 1–19. ISSN 2076-3417. DOI: 10.3390/app11010291.
- [35] PROMMEE, P., P. PIENPICHAYAPONG, N. MANOSITTHICHAI and N. WONGPROMMOON. Realization of Tunable Fractional-order Device based on Ladder Network Approximation. In: *17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. Phuket: IEEE, 2020, pp. 547–550. ISBN 978-1-7281-6486-1. DOI: 10.1109/ECTI-CON49241.2020.9158320.
- [36] VERMA, R., N. PANDEY and R. PANDEY. Novel CFOA based capacitance multiplier and its application. *AEU - International Journal of Electronics and Communications*. 2019, vol. 107, iss. 1, pp. 192–198. ISSN 1434-8411. DOI: 10.1016/j.aeue.2019.05.010.
- [37] MADIAN, A. H., S. A. MAHMOUD and A. M. SOLIMAN. Realization of Tunable Fractional-order Device based on Ladder Network Approximation. In: *13th IEEE International Conference on Electronics, Circuits and Systems*. Nice: IEEE, 2006, pp. 600–603. ISBN 1-4244-0394-4. DOI: 10.1109/ICECS.2006.379860.
- [38] ADHIKARY, A., P. SEN, S. SEN and K. BISWAS. Design and Performance Study of Dynamic Fractors in Any of the Four Quadrants. *Circuits, Systems and Signal Processing*. 2016, vol. 35, iss. 6, pp. 1909–1932. ISSN 1531-5878. DOI: 10.1007/s00034-015-0213-3.

## About Authors

**Rakesh VERMA** (corresponding author) received B.Tech. degree in Electronics and Communication Engineering (ECE) from Birla Institute of Applied Sciences, Bhimtal and M.Tech. degree in Control and Instrumentation Engineering from Delhi Technological University, Delhi. He served as Assistant Professor in ECE department, National Institute of Technology, Delhi. He received his Ph.D. from Delhi Technological University, Delhi in 2019. His research interest is focused on fractional order current mode circuits.

**Neeta PANDEY** is Professor in ECE department, Delhi Technological University. She received her M.E. in Microelectronics from Birla Institute of Technology and Science (BITS) Pilani and Ph.D. from Guru Gobind Singh Indraprastha University, Delhi. She has served in Central Electronics Engineering Research Institute, Pilani; Indian Institute of Technology, Delhi; Priyadarshini College of Computer Science, Noida; and Bharati Vidyapeeth's College of Engineering, Delhi in various capacities. She is having approximately 27 years of teaching and research experience in Electronics and Communication Engineering. Her research interests are in analog and digital Very-Large-Scale Integration (VLSI) design. She is life member of ISTE and senior member of IEEE, and IEEE WIE for 13 years.

**Rajeshwari PANDEY** received her B.Tech. (Electronics and Telecommunication) from J. K. Institute of Applied Physics, University of Allahabad and her M.E. (Electronics and Control) from BITS, Pilani, Rajasthan, India and Ph.D. from Faculty of Technology, Delhi University, India. She has served BITS Pilani, Applied Environmental Research Foundation (AERF), Noida and Priyadarshini College of Computer Science, Noida in various capacities. Currently, she is Professor in Department of Electronics and Communication Engineering, Delhi Technological University. Her research interests include analog integrated circuits and microelectronics. She is life member of ISTE, IETE and IEEE, and IEEE WIE.





## Characterization of finished surface through thermal additive centrifugal abrasive flow machining for better surface integrity

Parvesh Ali<sup>a</sup>, R S Walia<sup>b\*</sup>, Qasim Murtaza<sup>c</sup>, & Ranganath Muttanna Singari<sup>d</sup>

<sup>a, c, d</sup>Mechanical Department, Delhi Technological University, Delhi 110 042, India

<sup>b</sup>Department of Production and Industrial Engineering, Punjab Engineering College, Chandigarh 160 012, India

Received: 07 February 2019 ; Accepted: 8 October 2021

Abrasive Flow Machining (AFM) process has been a useful technique for deburring and polishing of the surface and edges through the abrasive laden media. The surface material has been removed in form of micro chips due to abrasion action of sharp cutting edges abrasive particles. A large amount of force and energy has been lost due to frictional forces between the surface and abrasive particles in AFM process. A new hybrid form of AFM process named as thermal additive centrifugal abrasive flow machining (TACAFM) has been discussed in the present investigation, which utilized the spark energy to melt the surface material. A lesser amount of force has been required by the abrasive particles to remove the molten material from the surface and also minimized the energy loss. In the present investigation central composite design response surface methodology has been used to plan and conduct the experiments using Design Expert<sup>®</sup> 11 software. Experiments have been performed to analyze the effect of input process variables such as current intensity, duty cycle, abrasive concentration, rotational speed of the electrode and extrusion pressure on scatter of surface roughness, micro hardness and % improvement in  $R_a$  of the workpiece. Also the finished surface of the brass work piece has been characterized for the microstructure study using SEM and XRD analysis. From the experimental results it has been found that duty cycle has the most significant effect towards Scatter of surface roughness with a contribution of 17.5 % while current has been contributed largest as 85.17 % towards micro hardness. Also it has been observed that current has contributed largest as 21.88% against the % improvement in  $R_a$ . The optimum scatter of surface roughness, micro hardness and % improvement in  $R_a$  has been observed as 0.15  $\mu\text{m}$ , 345.95 HV and 39.52 % respectively.

**Keywords:** Abrasive Flow Machining (AFM), Abrasives, Finishing, Hardness, Scanning Electron Microscopy (SEM), X-Ray Diffraction (XRD), Machining, Spark

### 1 Introduction

In the present situation mechanical components having better functional performance and dimensional accuracy are highly demanded. It has developed competitiveness among various industries to produce better products with economic viability and good surface integrity. In recent past, industries used conventional techniques such as lapping, honing and broaching. However, these processes were very slow and had difficulty in machining the complex shapes. This leads to the evolution of nonconventional finishing processes, such as Abrasive Flow Machining (AFM). AFM is a micro/ nano finishing process used for deburring and polishing of the complex surfaces which cannot be achieved by conventional techniques<sup>1, 2</sup>. It used an abrasive laden media which is a combination of polymer, gel and abrasive particles, and passed through the restrictive passage with a high extrusion pressure. The abrasive laden

media is viscous enough to hold the abrasive particles together during the finishing. The sharp cutting edge particles get in contact with the finishing surface and remove material in the form of micro/nano chips. AFM process has an advantage of less processing time and takes approximately only 10 percent of finishing time compared to the conventional technique<sup>3</sup>. In present scenario, AFM has numerous applications and is widely used to finish intricate shapes. Subramanian and Balashanmugam have reported the use of AFM technique for a wide variety of products including aircraft shuttle valves, propellers, dies etc. and stated that it is a flexible process and can be used for variety of materials<sup>4</sup>. Han *et al.*<sup>5</sup> used AFM process for the finishing of 15-5 PH stainless steel internal channels and observed the residual stress profile over the surface. It was found that AFM process produces compressive stress over the surface due to the generation of lesser temperature as compared to grinding and turning. Tzeng *et al.*<sup>6</sup> finished stainless steel (SUS 304) having

\*Corresponding author (E-mail:waliaravinder@yahoo.com)

micro slit of width in the range of  $0.23 \pm 0.02$  mm and concluded that the optimum machining quality was obtained at abrasive grain size of  $150 \mu\text{m}$ , 50 percent of abrasive concentration, and 6.7 MPa of extrusion pressure. Sushil *et al.*<sup>7</sup> finished Al/SiC MMCs using AFM technique and optimized the parameters for material removal. The author concluded that workpiece material and extrusion pressure has a significant effect on material removal.

Furthermore, AFM process parameters such as pressure, abrasive concentration in media and viscosity of media were studied to optimize material removal and surface finish by many researchers<sup>8-20</sup>. Wanet *et al.*<sup>8</sup> analyzed slip line velocity and wall shear stress for different values of extrusion pressure with two elliptical cross sections. It was stated that for low variation in cross section, zero order methodology was used and for larger variation order chronology was taken into consideration. Chen *et al.*<sup>9</sup> proposed a model on CFD-ACE software for different passageways of media flow and concluded that the helical passageways were better in comparison to polygonal passageways. Fu *et al.*<sup>10</sup> proposed a simulation on AFM process and concluded that an irregular stream line occurs at the leading as well as trailing edges in the media flow direction. Mali *et al.*<sup>11</sup> claimed through experimental results that abrasive mesh size has a significant effect on the material removal. Lv *et al.*<sup>12</sup> studied about the erosion mechanism of hard brittle materials during the finishing process and concluded that radial and lateral cracks were the main modes of fracture during the finishing. Wang *et al.*<sup>13</sup> sharpened the cutting edges of different coated and uncoated milling cutter using AFM process. The researchers observed that micro sized diamond film increases the roughness of rake and flank faces which correspond to reduction in radius of cutting edges. Shao and Cheng<sup>14</sup> studied about the surface roughness and topography profile during the AFM finishing process and observed that volume of abrasive media develops a huge difference in the machining system.

A number of development and modification was done in AFM process to enhance machining process productivity such as how to increase abrasive motion effectiveness, material removal, and surface integrity. Sankar *et al.*<sup>15-16</sup> developed Rotational AFM, which involved workpiece rotation through the gears to produce a centrifugal effect in the flowing media for the better contact of the abrasive particles with the

finishing surface. The researcher also used a drill bit in the media path, which made the media to flow through the flutes of drill and increased number of active abrasive particles. This resulted in improvement of material removal from the surface. Brar *et al.*<sup>17</sup> developed Helical AFM process and optimized process parameters using Taguchi method. It was found from the experimental results that using stationary drill bit in the media flow path significantly improved material removal. Walia *et al.*<sup>18</sup> provided centrifugal motion to the abrasive laden media by rotating the flute rod in the centre of media and concluded that Centrifugal assisted AFM gave better surface in comparison to conventional AFM after specific number of cycles. Marzban *et al.*<sup>19</sup> used spin motion of the media along with the workpiece rotation during the finishing process in AFM and observed increase in material removal. Tzeng *et al.*<sup>20</sup> finished a micro channel by using self-modulating abrasive media. The author claimed that the highly concentrated coarse abrasive particles gave good level of finish.

It was concluded that AFM process cannot be used for larger surface irregularities because it removes material uniformly from the surface<sup>21</sup>. Although this process provided good surface finish but it had a constraint of low material removal. To remove these limitations the researchers are now trying to hybridize the process with other conventional and non-conventional techniques. Through various ongoing advanced finishing processes it seems that there is a scope of developing new techniques by hybridizing AFM process with other non-conventional processes to get better surface integrity. Therefore, the main objective of this paper was to develop such a hybrid process which can increase the material removal and produce better finished products. The developed technique is termed as Thermal additive Centrifugal Abrasive Flow Machining (TACAFM) process in which Centrifugal force assisted AFM (CFAAFM) is clubbed with EDM process to enhance the material removal and surface finish. The developed technique is operated at low pressure in comparison to the Conventional AFM process because the flow is restricted more in TACAFM process due to electrode rotation. Response surface Methodology was used to conduct the experiments on the developed Thermal additive Centrifugal AFM (TACAFM) setup for variable parameters to optimize process parameter for scatter of surface roughness, micro-hardness and percentage improvement in  $R_a$ .

## 2 Materials and Methods

### 2.1 Mechanism of Material Removal in TACAFM Process

AFM has a constraint of low material removal. It can be reduced by hybridizing AFM process with other nonconventional processes. In AFM process a huge amount of force was exerted by the abrasive particles for removing the material from the surface. Also a part of energy was lost due to friction between the sharp cutting edges and surface. This leads to the development of such a hybrid process which can reduce the energy loss and improves the material removal.

Thermal additive Centrifugal AFM (TACAFM) was a hybridization of conventional AFM with a combination of Centrifugal assisted AFM and Electrical discharge machining process. The process used EDM principle for producing spark between the rotating electrode and the workpiece surface. In EDM process, spark was generated when two current carrying poles were short circuited. During EDM machining process, metal was eroded from both the poles and formed crater on the workpiece surface. It involved controlled erosion through electrically conductive materials by the starting of rapid and repetitive discharge between both the poles having a gap of less than 0.25 mm between them. In TACAFM process rotating electrode was taken as negative and workpiece was taken as positive terminal. Both the poles were separated by a small gap with a non conductive environment between them, which was developed by the flow of non conductive media (mixture of media and kerosene based oil).

As the pulsed power supply was given to both the poles (rotating electrode and workpiece), free electrons on the rotating electrode were subjected to the electrostatic force. As the free electrons moved towards the work surface, they collide with the media molecules, which break it into positive and negative ion and characterized as plasma. Thus all the positive and negative ions moved towards their opposite polarity. This type of movement of ions could be seen as spark. Figure 1 shows the principle of TACAFM process and ions formation in the plasma channel. If during rotation, electrode make contact with the surface a large intensity current was induced and melted material from the both the poles. As the spark was generated inside the hollow cavity due to potential difference between the rotating electrode and workpiece, high temperature was developed and melted/softened the surface material which could be easily carried out by the abrasive particles.

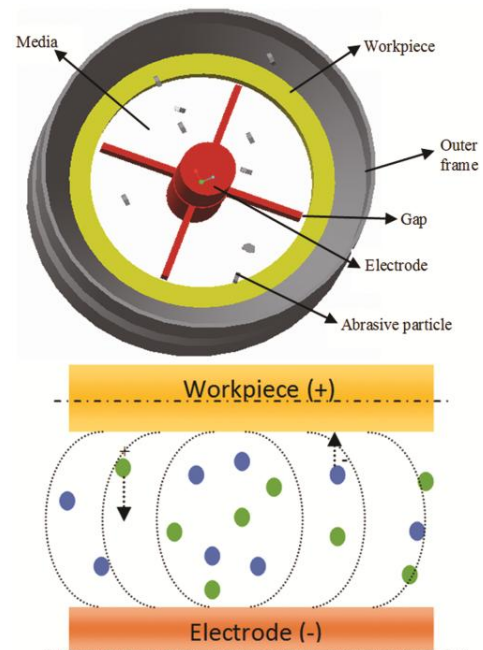


Fig. 1 — Mechanism of material removal in the developed TACAFM process.

### 2.2 Experimental Setup

The arrangement of the TACAFM process is as shown in Fig. 2. The experimental setup included fixture, 3 phase induction motor, power drive, bearings, gears, manual analog controller (25 Ampere), EDM power supply. The experimental setup of TACAFM process and fixture arrangement is shown in Fig. 2 (a and b) respectively. The fixture was made of nylon and its function was to hold the workpiece and guiding the media between the two opposite media cylinders. The fixture was divided into three parts as shown in Fig. 2(c) and included electrode, bearings and power supply terminals. The rotating electrode was attached through gear which was further rotated by the 3-phase induction motor using intermediate gears as shown in Fig. 2(d). The driving gear was made of mild steel and was fitted to the shaft connected with the 3 phase induction motor. The driving gear was connected with two intermediate gears made of metalon and aluminium and further rotates the gear connected with the electrode.

Basic purpose for making one of the intermediate gears non conductive was to prevent the reverse of the current to the motor. Power drive was used to control the rotation speed of the 3 phase induction motor by changing its frequency of rotation. The function of EDM power supply was to convert the main AC supply to the pulsed DC supply required for the

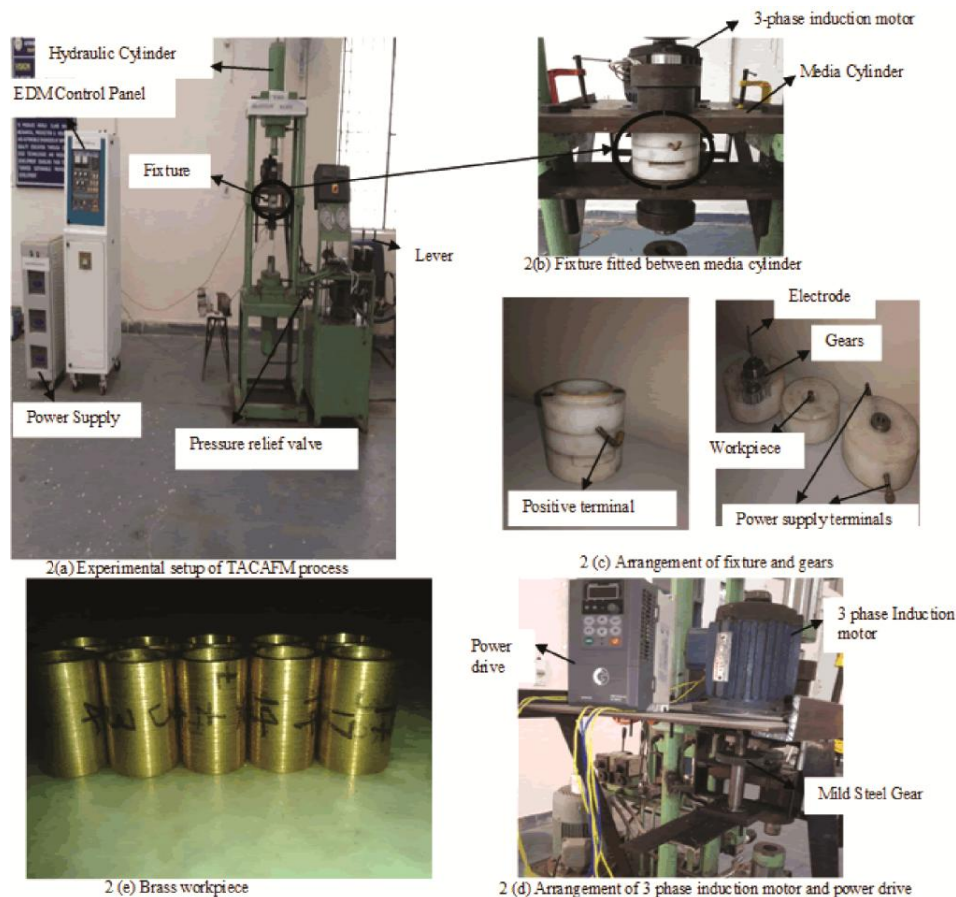


Fig. 2 — Experimental set up of developed TACAFM process.

generation of spark. It also sensed the potential difference between the rotating electrode and the workpiece and helped manual analog controller to sense the gap between both the poles. Whenever the gap was maintained between both the poles spark was generated inside the hollow cavity.

The specification of the pulse generator is as given in Table 1.

### 2.3 Workpiece

In the present invention brass was used as the workpiece material because it has better malleability and consist low melting point with better flow characteristics. Also it is widely used alloy in the industry such as gears, bearings, valves, marine construction; die making etc. It has a good corrosion resistant property. The brass workpiece and its geometry are as shown in Fig. 2(e) and Fig. 3 respectively.

### 2.4 Media

The media used for the experimentation is a Non Newtonian media and consist polymer, gel, abrasive particles and kerosene based oil in it. It is visco-

Table 1— Specification of pulse generator

Current	25 Ampere
Power Supply	415V, 3 Ph, 50 Hz
Control System	Microprocessor based
Power	3 KVA
MRR Gr-St	180 mm <sup>3</sup> /min
MRR Cu-St	140 mm <sup>3</sup> /min
Best Surface finish	0.5 $\mu$ Ra
Min Electrode Wear	0.3 %
Pulse on Time	1-2000 / 10 steps micro sec
Max Electrode Wt.	50 Kgs

elastic in nature and retains its viscosity for a longer duration on the increase of temperature. This increases the abrasive holding capacity of the media and more number of abrasive particles participates in abrasion process. The media also has very less conductivity which permits the lesser spark energy to interact with the workpiece surface to prevent the deterioration in the surface integrity. The properties of the media are as follows:

Viscosity = 0.32 MPa-s, Thermal conductivity = 0.22 W/meter-K, Density = 1219 kg/m<sup>3</sup>

Based on literature survey and pilot experimentations major process parameters as per CCD response surface methodology were selected and trial experiments were performed for analyzing the effect of variable parameters over the responses, i.e. scatter of surface roughness (SSR), micro-hardness and % improvement in  $R_a$ . The variable parameters with different levels were presented in Table 2.

The current, duty cycle, rotational speed, extrusion pressure and abrasive concentration were taken as

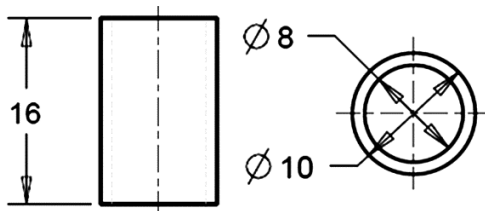


Fig. 3 — Drawing of workpiece geometry.

input process variables. The process performance was measured as scatter of surface roughness, micro hardness and % improvement in  $R_a$ . A set of 32 experiments according to the CCD of Response surface methodology technique was performed and measured responses were recorded in Table 3. Surface

Table 2 — Control parameters with their limits

Factors	Levels				
	-2	-1	0	+1	+2
A – Current (Ampere)	0	4	8	12	16
B – Duty Cycle (Fraction)	0.63	0.68	0.73	0.78	0.83
C – Rotation (rpm)	100	150	200	250	300
D – Pressure (MPa)	5	10	15	20	25
E – Abrasive Concentration (Fraction)	0.2	0.3	0.4	0.5	0.6

Constant Parameters: Polymer to Gel ratio :1:1, Abrasive particles-  $Al_2O_3$ , Mesh Size:180, Temperature -  $32 \pm 2$  °C, Workpiece Material- Brass, Media Flow Volume – 290  $cm^3$

Table 3— Central composite design for the measured experimental results and actual factors

Std	Run	A (Amp)	B (Fraction)	C (rpm)	D (MPa)	E (Fraction)	Initial $R_a$ ( $\mu m$ )	Final $R_a$ ( $\mu m$ )	% Improved $R_a$	Scatter of Surface roughness ( $\mu m$ )	Micro hardness (HV)
1	3	4	0.68	150	10	0.5	1.21	0.82	31.85	0.82	198.36
2	13	12	0.68	150	10	0.3	1.46	1.06	27.34	0.45	312.86
3	4	4	0.78	150	10	0.3	1.56	1.19	23.22	0.38	221.12
4	12	12	0.78	150	10	0.5	1.33	0.85	36.02	0.42	355.48
5	2	4	0.68	250	10	0.3	1.38	0.95	30.92	0.32	252.12
6	32	12	0.68	250	10	0.5	1.22	0.75	38.45	0.64	332.58
7	15	4	0.78	250	10	0.5	1.16	0.67	41.84	0.43	245.74
8	22	12	0.78	250	10	0.3	1.19	0.69	42.35	0.42	322.14
9	10	4	0.68	150	20	0.3	1.62	1.16	28.67	0.85	212.12
10	5	12	0.68	150	20	0.5	1.41	0.85	40.06	0.39	328.45
11	6	4	0.78	150	20	0.5	1.12	0.79	29.25	0.47	212.97
12	1	12	0.78	150	20	0.3	1.26	0.81	35.66	0.18	301.49
13	27	4	0.68	250	20	0.5	1.20	1.04	13.33	0.32	239.24
14	24	12	0.68	250	20	0.3	1.38	0.94	32.24	0.77	316.47
15	18	4	0.78	250	20	0.3	1.64	1.08	34.24	0.46	256.35
16	17	12	0.78	250	20	0.5	1.57	0.98	37.76	0.44	331.86
17	7	0	0.73	200	15	0.4	1.33	1.15	13.39	0.54	224.48
18	29	16	0.73	200	15	0.4	1.52	1.06	30.37	0.44	412.57
19	11	8	0.63	200	15	0.4	1.46	0.98	33.14	0.45	264.78
20	25	8	0.83	200	15	0.4	1.44	0.83	42.24	0.15	255.47
21	20	8	0.73	100	15	0.4	1.38	0.99	28.53	0.48	272.35
22	23	8	0.73	300	15	0.4	1.28	0.88	30.94	0.54	291.14
23	8	8	0.73	200	5	0.4	1.14	0.70	38.25	0.28	243.87
24	21	8	0.73	200	25	0.4	1.22	0.79	35.24	0.31	245.54
25	14	8	0.73	200	15	0.2	1.56	1.14	27.02	0.78	271.34
26	30	8	0.73	200	15	0.6	1.52	1.08	28.78	0.86	276.67
27	31	8	0.73	200	15	0.4	1.65	1.20648	26.88	0.46	302.34
28	19	8	0.73	200	15	0.4	1.58	1.09573	30.65	0.51	274.82
29	16	8	0.73	200	15	0.4	1.44	1.065024	26.04	0.48	283.57
30	26	8	0.73	200	15	0.4	1.32	1.016664	22.98	0.45	257.37
31	28	8	0.73	200	15	0.4	1.18	0.890428	24.54	0.49	285.56
32	9	8	0.73	200	15	0.4	1.24	0.94426	23.85	0.54	242.23

roughness was measured by Taylor Hobson precision machine (TAYLOR HOBSON SURTRONIC 3+) having resolution of 0.01 micrometer. Surface roughness was measured at five different positions of the work piece and average of these values was taken for the initial and final value of surface roughness. The difference between the maximum  $R_a$  value and minimum  $R_a$  value was taken as SSR. Micro hardness of the surface was measured by FISCHERSCOPE micro-hardness tester (HM 2000S).

The percentage composition of brass workpiece was shown in Table 4.

### 2.5 Determination of Number of Abrasive Particles Interacting with the Workpiece Surface

If  $N_a$  = Active number of abrasive particles interacting with the workpiece surface

$N$  = Total number of abrasive particles in full volume of media

Then, Total number of abrasive particle in full volume of media ( $N$ )<sup>21</sup> = —

where,  $V_a$ <sup>21</sup> = Volume of abrasive particle in full volume of the abrasive laden media = ———

where,  $C$  = concentration of abrasive particle in the media

$d_m$  = density of abrasive laden media

$V_m$  = Volume of abrasive laden media

$d_a$  = Density of abrasive particle

$V_o$  = Unit abrasive particle volume =  $-\pi \cdot (-)$

where,  $d$  = diameter of abrasive particle

Total number of active abrasive particle interacting with the workpiece surface ( $N_a$ )<sup>21</sup> = ———,

where,  $D$  = Diameter of workpiece

Fraction of abrasive particle involved in cutting action = —

### 2.6 Determination of Velocity of Impact for Abrasive Particle

In TACAFM process two types of forces are responsible for material removal, i.e axial force ( $F_a$ ) (direction of media flow) and radial force ( $F_c$ ) (perpendicular to the workpiece surface). Consider an abrasive particle having mass  $m$  entering the passage of the workpiece.

If  $v_a$  = velocity of abrasive particle by which it is moving in axial direction

and  $v_c$  = velocity of abrasive particle by which it is moving towards the inner surface of the workpiece

Centrifugal force developed on abrasive particle due to rotation of electrode  $F_c$ <sup>7</sup> =  $m \cdot r \cdot \omega^2 = 3\pi \cdot \mu \cdot d \cdot v_c$

So  $v_c = \frac{\text{---}}{\mu \cdot d}$

If media is passing through the restrictive passage with pressure  $P$ , then force applied ( $F$ ) will be =  $P \cdot A$  where,  $A$  = Area of the restrictive passage = —.

If abrasive particle takes time  $t$  for passing the restrictive passage then ( $F$ ) = —

Therefore = —

Therefore the resultant velocity of abrasive particle by which it is impacting over the workpiece surface will be the resultant of axial and radial velocity.

$v_R =$

## 3 Results and Discussion

Analysis of Variance technique was used to identify the responses of the selected model. All the significant parameters were identified and the interaction effect of the parameters on the measured responses was studied by the response surface graphs. The regression equation for scatter of surface roughness, micro-hardness and % improvement in  $R_a$  was presented in Table 5.

Figure 4 shows a magnificent acceptability of the regression model. Each observed value was comparable to the predicted value obtained from the model. The results of the ANOVA analysis were presented in Tables (6, 7 and 8) respectively. The model F-value of 57.59, 41.79 and 21.63 for scatter of surface roughness, micro hardness and % improvement in  $R_a$  respectively, with its Probability > F value less than 0.0001 shows model as significant. There was a 0.01% chance for having a large F-value due to noise. Value of P less than the 0.0500 represented model terms as significant.

For scatter of roughness the terms A, B, AC, AD, BC, BE, CD, CE, DE, B<sup>2</sup>, D<sup>2</sup>, E<sup>2</sup> were significant model terms with their percentage contribution as 1.32, 17.5, 22.63, 0.61, 3.95, 1.98, 0.88, 0.98, 13, 7.18, 7.56, 22.36 respectively. The values more than 0.1000, showed model terms as insignificant. If more number of insignificant model terms existed, then the model could be improved by the reduction process. The lack of fit values for SSR, micro hardness and %

Table 4 — Percentage proportion of basic elements in brass workpiece

Basic Elements	Zn	Cu	Pb	Sn	Ni	Fe
Proportion	36.4	59.2	3.26	1.2	0.4	0.6



Table 5 — Regression relation for scatter of surface roughness, micro- hardness and % Improved  $R_a$ .

Responses	$R^2$	Adjusted $R^2$	Regression Model
Scatter of surface roughness	0.9818	0.9648	$SSR = -3.04445 - 0.105313 * A + 19.3259 * B - 0.0179033 * C + 0.116923 * D - 7.92013 * E + 0.00056875 * A * C - 0.0009375 * A * D + 0.019 * B * C + 6.75 * B * E + 9e-05 * C * D - 0.00475 * C * E - 0.1725 * D * E - 18.8077 * B^2 - 0.00193077 * D^2 + 8.29808 * E^2$
Micro-hardness	0.9447	0.9221	$Micro-hardness = 98.8049 + 0.71692 * A + 30.275 * B + 0.519633 * C + 8.1976 * D - 150.733 * E - 0.0450344 * A * C + 22.0016 * A * E + 0.709939 * A^2 - 0.283839 * D^2$
% Improved $R_a$	0.9469	0.9031	$\% Improved R_a = 745.313 + 0.753073 * A - 1951.96 * B - 0.619535 * C + 0.079125 * D + 147.115 * E + 0.0731562 * A * D + 1.11325 * B * C - 0.0126775 * C * D - 0.377125 * C * E - 4.28125 * D * E - 0.0570378 * A^2 + 1215.96 * B^2 + 0.000420458 * C^2 + 0.112146 * D^2$

Current (A), Ampere; Duty Cycle (B), Fraction; Rotational speed of electrode (C), rpm; Extrusion Pressure (D), MPa; Abrasive Concentration (E), Fraction

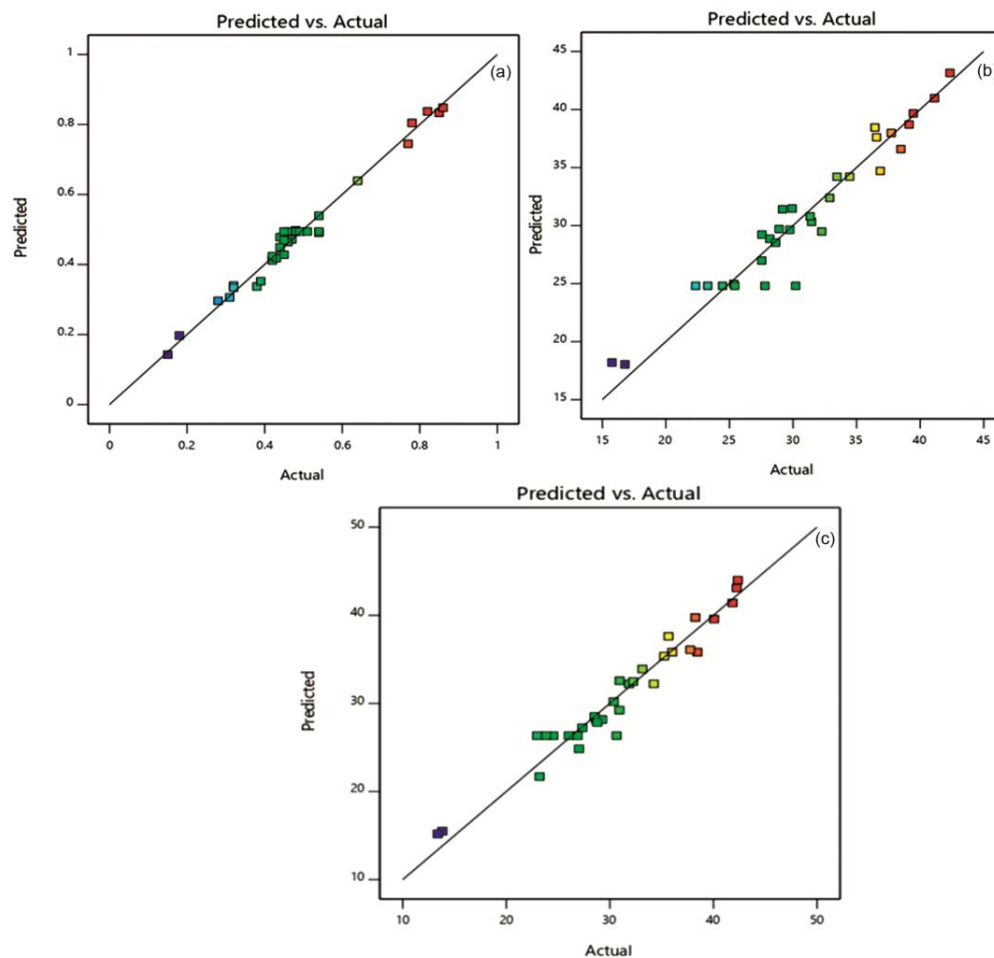


Fig.4 — (a) Predicted and actual responses for scatter of surface roughness, (b) Predicted and actual responses for micro-hardness, and (c) Predicted and actual responses for % Improved  $R_a$ .

improvement in  $R_a$  were found as 0.99, 0.18, and 0.52, respectively, showing the developed model for SSR, micro-hardness and % improvement in  $R_a$  as adequate and satisfactory.

The adequacy of the model fit can be observed by the value of  $R^2$ . The determination coefficient for scatter of surface roughness is observed as 0.9818, showing the developed model is significant for predicting the



Table 6 — ANOVA outcome for fitted RSM model for Scatter of surface roughness

Source	Sum of Squares	Degree of Freedom	Mean Square	F- value	Probability>F	
Model	0.9437	15	0.0629	57.59	< 0.0001	Significant
A	0.0121	1	0.0121	11.12	0.0042	
B	0.1601	1	0.1601	146.52	< 0.0001	
C	0.0001	1	0.0001	0.0610	0.8080	
D	0.0001	1	0.0001	0.1373	0.7158	
E	0.0028	1	0.0028	2.58	0.1279	
AC	0.2070	1	0.2070	189.51	< 0.0001	
AD	0.0056	1	0.0056	5.15	0.0374	
BC	0.0361	1	0.0361	33.05	< 0.0001	
BE	0.0182	1	0.0182	16.68	0.0009	
CD	0.0081	1	0.0081	7.41	0.0150	
CE	0.0090	1	0.0090	8.26	0.0110	
DE	0.1190	1	0.1190	108.95	< 0.0001	
B <sup>2</sup>	0.0657	1	0.0657	60.13	< 0.0001	
D <sup>2</sup>	0.0692	1	0.0692	63.37	< 0.0001	
E <sup>2</sup>	0.2046	1	0.2046	187.30	< 0.0001	
Residual	0.0175	16	0.0011			
Lack of Fit	0.0120	11	0.0011	0.9944	0.5414	Not Significant
Pure Error	0.0055	5	0.0011			
Cor Total	0.9612	31				
Std. Dev.	0.0331		R <sup>2</sup>	0.9818		
Mean	0.4850		Adjusted R <sup>2</sup>	0.9648		
C.V. %	6.81		Predicted R <sup>2</sup>	0.9006		
Press	0.0955		Adequate Precision	30.1653		

Table 7 — ANOVA outcome for fitted RSM model for micro-hardness

Source	Sum of Squares	Degree of Freedom	Mean Square	F- value	Probability>F	
Model	64135.16	9	7126.13	41.79	< 0.0001	Significant
A	54101.56	1	54101.56	317.28	< 0.0001	
B	54.99	1	54.99	0.3225	0.5759	
C	1523.70	1	1523.70	8.94	0.0068	
D	60.52	1	60.52	0.3549	0.5574	
E	153.37	1	153.37	0.8994	0.3532	
AC	1297.98	1	1297.98	7.61	0.0115	
AE	1239.22	1	1239.22	7.27	0.0132	
A <sup>2</sup>	3853.61	1	3853.61	22.60	< 0.0001	
D <sup>2</sup>	1503.88	1	1503.88	8.82	0.0071	
Residual	3751.41	22	170.52			
Lack of Fit	1437.07	17	84.53	0.1826	0.9965	Not significant
Pure Error	2314.34	5	462.87			
Cor Total	67886.57	31				
Std. Dev.	13.06		R <sup>2</sup>	0.9447		
Mean	276.36		Adjusted R <sup>2</sup>	0.9221		
C.V. %	4.73		Predicted R <sup>2</sup>	0.9143		
Press	5817.13		Adequate Precision	28.3475		

variation on scatter of surface roughness up to 98.18 percent, and the model is able to demonstrate the process. However the Predicted R<sup>2</sup> of 0.9006 has a good relation with the adjusted R<sup>2</sup> of 0.9648. The lower value (6.81) of CV % shows better accuracy and consistency

of the performed experiments. Adequate precision for the model as 30.16, was more than 4, which shows an adequate signal.

For micro hardness terms A, C, AC, AE, A<sup>2</sup>, D<sup>2</sup> were significant model terms with their percentage

Table 8 — ANOVA outcome for fitted RSM model for % Improved  $R_a$ 

Source	Sum of Squares	Degree of Freedom	Mean Square	F- value	Probability>F	
Model	1518.41	14	108.46	21.63	< 0.0001	Significant
A	337.73	1	337.73	67.37	< 0.0001	
B	126.91	1	126.91	25.32	0.0001	
C	24.75	1	24.75	4.94	0.0402	
D	28.84	1	28.84	5.75	0.0282	
E	13.40	1	13.40	2.67	0.1205	
AD	34.25	1	34.25	6.83	0.0182	
BC	123.93	1	123.93	24.72	0.0001	
CD	160.72	1	160.72	32.06	< 0.0001	
CE	56.89	1	56.89	11.35	0.0036	
DE	73.32	1	73.32	14.62	0.0014	
A <sup>2</sup>	24.60	1	24.60	4.91	0.0407	
B <sup>2</sup>	272.96	1	272.96	54.45	< 0.0001	
C <sup>2</sup>	32.64	1	32.64	6.51	0.0206	
D <sup>2</sup>	232.19	1	232.19	46.32	< 0.0001	
Residual	85.22	17	5.01			
Lack of Fit	47.14	12	3.93	0.5157	0.8389	Not significant
Pure Error	38.09	5	7.62			
Cor Total	1603.64	31				
Std. Dev.	2.24		R <sup>2</sup>	0.9469		
Mean	30.83		Adjusted R <sup>2</sup>	0.9031		
C.V. %	7.26		Predicted R <sup>2</sup>	0.8275		
Press	276.67		Adequate Precision	18.7774		

contribution as 85.17, 2.39, 2.04, 1.95, 6.06, and 2.36, respectively. The values more than 0.1000, showed model terms as insignificant. If more number of insignificant model terms existed, then the model could be improved by the reduction process. This case had 99.65 % chance of having large Lack of Fit F-value due to noise. Predicted  $R^2$  of 0.9143 was in a close relation with the Adjusted  $R^2$  of 0.9221; i.e. the difference between them was less than 0.2. The small value (6.81) of CV % shows better accuracy and consistency of the performed experiments. Adequate Precision showed the value of S/N ratio. The ratio more than 4 was advisable. The ratio of 28.347 represented an adequate signal.

For % improvement in  $R_a$  terms A, B, C, D, AD, BC, CD, CE, DE, A<sup>2</sup>, B<sup>2</sup>, C<sup>2</sup>, D<sup>2</sup> were significant model terms with their percentage contribution as 21.88, 8.22, 1.6, 1.86, 2.22, 8.03, 10.42, 3.68, 4.75, 1.59, 17.69, 2.12 and 15.05, respectively. The values more than 0.1000, showed model terms as insignificant. If more number of insignificant model terms existed, then the model could be improved by the reduction process. This case had 83.89 % chance of having large Lack of Fit F-value due to noise. Predicted  $R^2$  of 0.8275 was in a close relation with the Adjusted  $R^2$  of 0.9031; i.e. the difference between

them was less than 0.2. The small value (7.26) of CV % shows better accuracy and consistency of the performed experiments. Adequate Precision showed the value of S/N ratio. The ratio more than 4 was advisable. The ratio of 18.777 represented an adequate signal.

Figure 4 shows the plot across the actual and predicted responses. It was observed that the results between both the responses were very close for SSR, micro-hardness and % improvement in  $R_a$ . It shows the predicted model was acceptable.

The perturbation graph presented in Fig. 5 (a), shows the effect of variable process parameters on the SSR. The midpoint indicated as (value 0), in design expert shows reference point for all the parameters. The keen slope for the parameters current (A), duty cycle (B), rotational speed (C), pressure (D), abrasive concentration (E) shows that these parameters are highly dependent on scatter of surface roughness. The reason for this circumstance was discussed while demonstrating the interaction effects of parameters. Table 6, Table 7 and Table 8 shows the interactions, in which current (A), duty cycle (B), current and rotational speed (AC), current and extrusion pressure (AD), duty cycle and rotational speed (BC), duty cycle and abrasive concentration (BE), rotational

speed and extrusion pressure (CD), rotational speed and abrasive concentration (CE), extrusion pressure and abrasive concentration (DE) were the interactions for scatter of surface roughness while current (A), rotational speed (C), current and rotational speed (AC), and current and abrasive concentration (AE) were the interactions for the micro-hardness and also current (A), duty cycle (B), rotational speed (C), extrusion pressure (D), Current and extrusion pressure (AD), duty cycle and rotation (BC), rotation and extrusion pressure (CD), rotation and abrasive concentration (CE), extrusion pressure and abrasive concentration (DE) were the interactions for the % improvement in  $R_a$ . The relevant plots to these interactions were shown from Fig. 5 (b-h) for scatter

of surface roughness, from Fig. 6 (b and c) for micro-hardness and from Fig. 7(b-f) for % improvement in  $R_a$ , respectively.

Figure 5 (b) shows as the value of current and rotational speed increased, scatter of surface roughness decreased keeping other parameter as constant. Due to increase in the current value, discharge energy density also enhanced which means, for each pulse higher amount of energy could be available for melting the work surface. This developed high temperature on the surface and formed deeper craters on it<sup>22</sup>. When abrasive particles came in contact with the melted material, less amount of force was required for removing the molten/semi molten material from the surface. This provided good

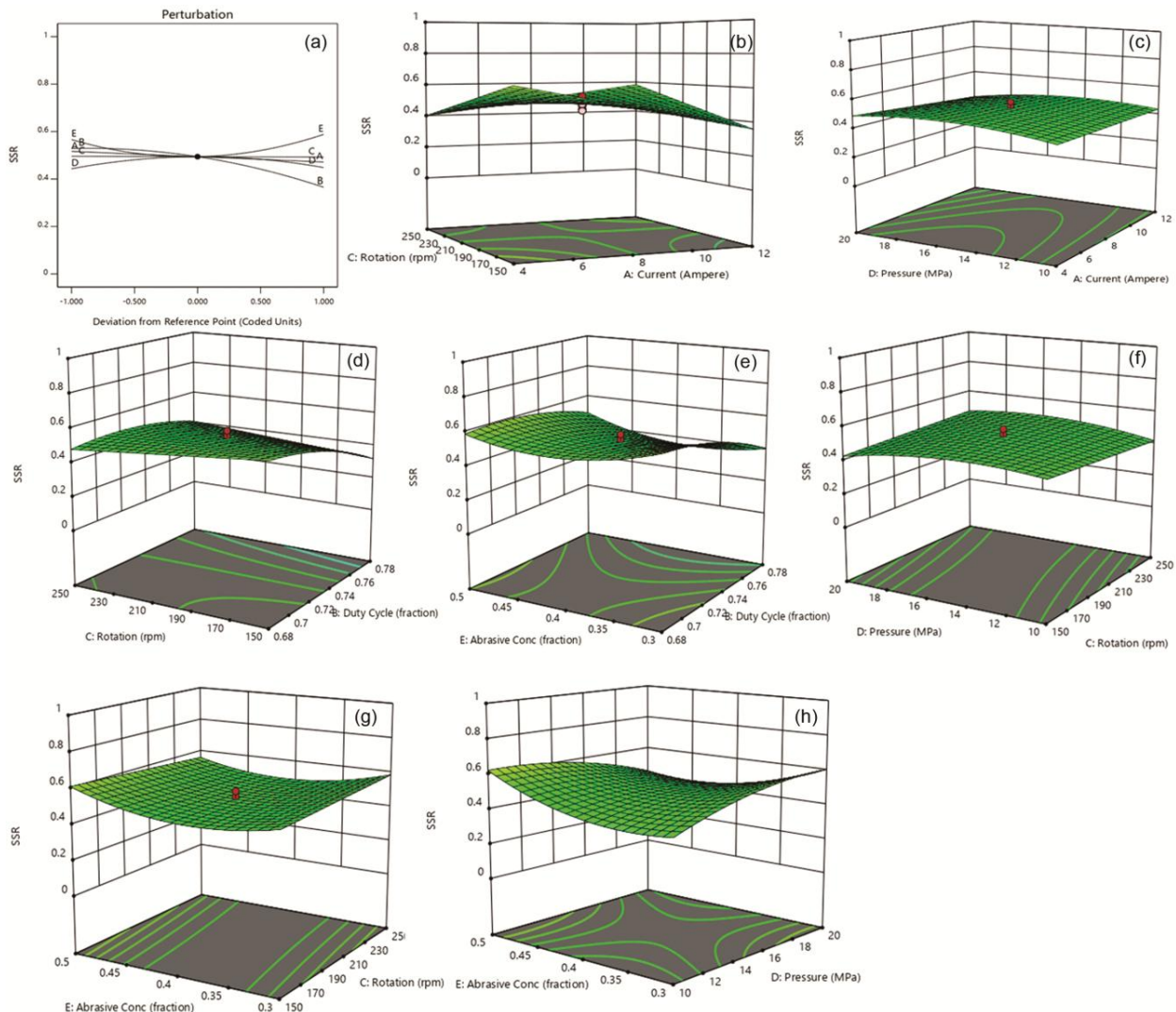


Fig. 5 — (a) Perturbation plots for scatter of surface roughness, (b - h) Response 3D surface plot showing the interactive influence of variable parameters for scatter of surface roughness.

level of finishing and decreased the scatter of roughness. As the rotational speed of electrode was increased, it developed a centrifugal force of larger magnitude in the media flow path. This may cause more number of abrasive particles contacting the workpiece surface. Therefore number of dynamic abrasive particles participating in material removal can increase that leads increase in the resultant force applied by the abrasive particles. Due to increase in the resultant force, abrasive particles can easily remove the molten/semi molten state material from the surface and produced good level of finishing with reduced scatter of roughness of the workpiece surface<sup>23</sup>. Figure 5 (c) shows as the extrusion pressure and current was increased; scatter of roughness increased. As the value of current was increased, a huge temperature was developed on the surface. A large energy pulse of EDM produces deeper crater on the surface and causes larger material removal. Therefore, large value of material removal leads to deteriorate the surface quality and increases the value of scatter of roughness. As Fig. 5(d) shows increase in scatter of surface roughness with extrusion pressure. It can be said that as the extrusion pressure was increased, the abrasive particles impacted the surface with larger force and caused deep cut on the surface. These deeper cut enhanced the material removal; larger material removal led to decrease in surface integrity. This increased the value of scatter of surface roughness. The similar trend was also observed by Walia *et al.* in 2006<sup>23</sup>. Figure 5(d) shows decrease in scatter of surface roughness with rotation and duty cycle, while keeping other parameters as constant. As duty cycle increased, the spark frequency enhanced due to availability of discharge energy for a longer duration. Larger amount of discharge energy developed high temperature on the workpiece surface and melted more material, which could be easily removed by the abrasive particles in form of micro-chips. This reduced the value of scatter of roughness. Figure 5(e) also shows initially decrease and further increase in scatter of surface roughness value with abrasive concentration. As the abrasive concentration increased, numbers of active abrasive particles participating in the abrasion process were increased. This increased the available energy for breaking the atomic bond of the material and developed new surface by displacing the atoms<sup>24</sup>. This caused increase in MR and provided better surface finish, which reduced the scatter of surface roughness.

However on further increasing the abrasive concentration, large number of abrasive particles participates in the abrasion and leads to improvement in material removal which corresponded deterioration in surface finish. This increased the scatter of surface roughness. Figure 5(f) shows that as the extrusion pressure is increased, scatter of roughness initially increased and further decreased. However there is very negligible change in the response with respect to the rotational speed keeping other parameters constant. The reason may be that at lower extrusion pressure both material removal and surface finish are low because shearing energy required by the abrasive particles is not enough to shear the peaks. Shearing strength of the sharp edged abrasive particles should be greater than the strength of material for more material removal. This increased the scatter of surface roughness, while on further increasing the extrusion pressure abrasive particles imparted a larger force on the surface and corresponded more material removal along with better surface integrity. This reduced the value of scatter of surface roughness<sup>23</sup>. Figure 5(g) shows that there was initially decrease and then further increase in scatter of roughness with increasing abrasive concentration. However there was some increase in scatter of roughness with respect to the rotation. The reason may be that on enhancing the abrasive concentration in the media, more abrasive particles interacted with the finishing surface and performed efficient cutting action. This improved the material removal and surface finish which corresponded decrease in scatter of surface roughness. On further increasing the abrasive concentration surface finish deteriorated due to more material removal. This increased the scatter of surface roughness of the surface. Figure 5(h) shows increase in scatter of roughness with the extrusion pressure and abrasive concentration while keeping other parameters as constant.

The perturbation graph presented in Fig. 6(a), shows the effect of variable parameters over the micro-hardness of surface. Figure 6(b) shows effect of current intensity and rotational speed on the micro-hardness of the material keeping other parameter constant. In TACAFM process increase in micro hardness with current occurred due to heating and better quenching. The spark generation in TACAFM process developed large amount of heat in the finishing zone. The huge amount of heat caused better quenching of material and facilitated hard carbides

and other hardened compounds formation over the surface. These hardened compounds developed over the surface increased the micro hardness of the surface which was similar with the results obtained by Gill and Kumar<sup>25</sup>.

Figure 6(b) shows as the rotational speed was increased, micro hardness of the surface also increased. In TACAFM process two forces (axial and radial) are acting during the finishing process and resultant of these two forces may be responsible for the material removal mechanism. On increasing the rotational speed, radial component of force increases which enhances the resultant force. Therefore abrasive particles were thrown on the surface with larger force due to the involvement of the centrifugal force during the media flow. It caused work hardening of the surface and increased the micro hardness of finishing surface. The similar results were also obtained by Walia *et al.* in 2008<sup>26</sup>. Fig. 6(c) shows as the current intensity was increased, micro hardness of the surface also increased. On increasing the current intensity more melting was done on the workpiece surface which caused work hardening of the surface. This work hardening increased the micro hardness of the surface. Fig. 6(c) also shows that micro hardness of the surface was enhanced due to increase in abrasive concentration. On increasing the abrasive concentration, more and more abrasive particles impacted on the workpiece surface which induced a residual stress over the surface and enhanced the micro hardness of surface.

The perturbation graph presented in Fig. 7(a), shows the effect of variable parameters over the % improvement in surface roughness. Figure 7(b) shows as the current intensity and extrusion pressure is increased keeping other parameters constant, surface finish gets better. As the current intensity is increased,

discharge energy density is also increased. This causes availability of higher energy for each pulse which corresponds rapid melting of surface due to increase of temperature over the surface. This rapid melting of surface produces deeper craters on it<sup>22</sup>, which can be easily carried away by the abrasive particles and provides good level of finish.

In conventional AFM process only axial force is responsible for the material removal which acts in the direction of media flow. But TACAFM process is a combination of Centrifugal force assisted AFM (CFAAFM) and EDM process. In the present developed process axial force and radial force both are responsible for the material removal. Radial force acts in the direction perpendicular to the surface. The resultant force subjected by the abrasive particles will be the resultant of axial and radial forces. Initially at lower pressure the resultant force is less and the abrasive particles are not able to shear the peaks of the surface. But as the pressure is increased, the resultant force also increases and removes more material which improves the surface finish. Figure 7(c) shows increase in surface finish with duty cycle and rotation keeping other parameter constant. As the duty cycle is increased, discharge energy may be applied for a longer duration. It can develop high temperature over the surface and softens more volume of material. This soft material can be easily carried by the abrasive particles with lesser amount of force that causes more material removal and provide good surface finish. As the rotational speed of the electrode is increased radial force increases which causes increase in resultant force. This may cause more availability of energy to break the atomic bond and will develop new surface with better surface integrity which was similarly observed by Walia *et al.* (2006)<sup>24</sup>. Figure 7(d) shows improvement in surface finish

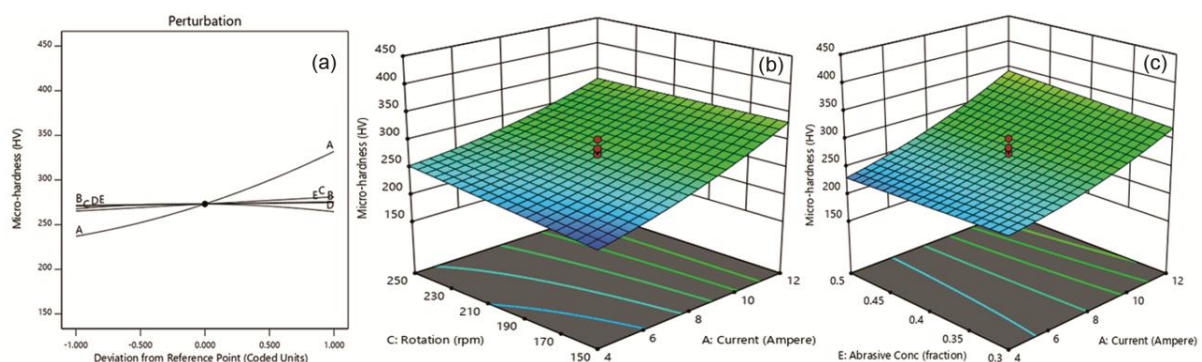


Fig. 6 — (a) Perturbation plots for micro-hardness, (b - c) Response 3D surface plot showing the interactive influence of variable parameters for micro-hardness.



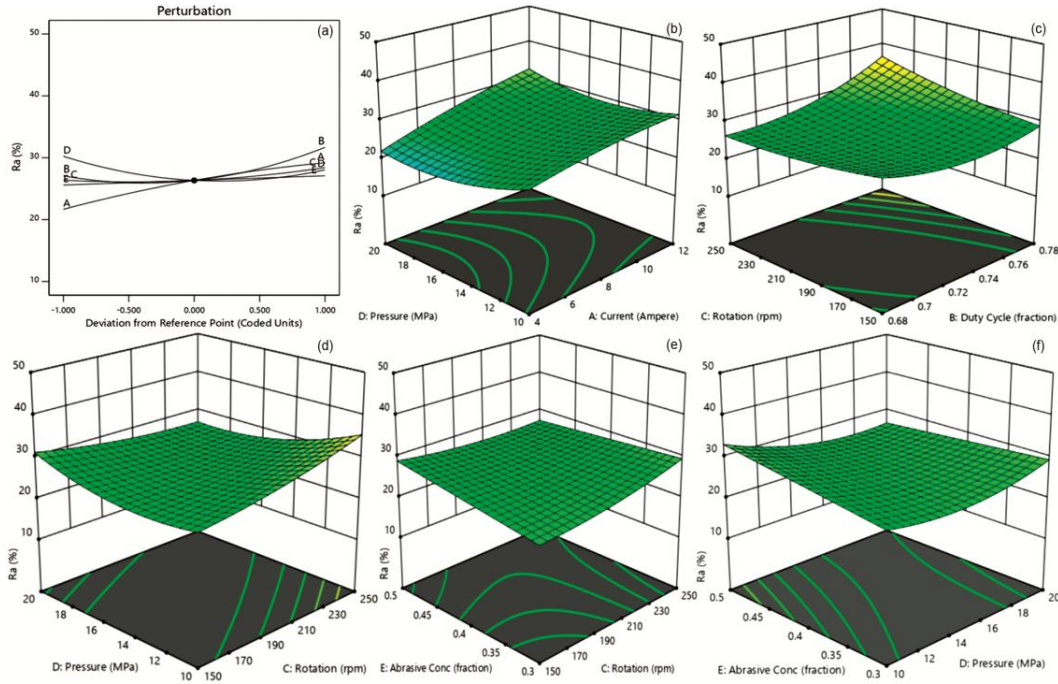


Fig. 7 — (a) Perturbation plots for % improved  $R_a$ , (b - f) Response 3D surface plot showing the interactive influence of variable parameters for % Improved  $R_a$ .

with rotation and pressure keeping other parameter constant. This is due to the increase in the resultant force by which the abrasive particles are impacting over the surface. Figure 7(e) shows as the rotation and abrasive concentration is increased, percentage improvement in surface finish also increases. On increasing the abrasive concentration, dynamic number of abrasive particles is increased and more number of abrasive particles may come in contact with the surface. This can also increase the contact region between the abrasive particles and workpiece surface. More number of abrasive particles can easily remove the peak of the surface and provide good level of surface finish. The results are comparable to the results obtained by Gorana *et al.* (2004)<sup>27</sup>. However, on further increase of abrasive concentration, dynamic abrasive particles are increased but the interference between the abrasive particles also increases leading to energy loss and reduction in rate of material removal will reduce. Figure 7(f) shows as the pressure and abrasive concentration is increased, percentage improvement in surface finish also increases. This might be due to increase in active abrasive particles.

### 3.1 Optimization

Optimization was performed to minimize scatter of surface roughness (SSR), maximize micro-hardness and maximize % improvement in  $R_a$  having the

constrained limit of five factors as shown in Table 2. The optimization was performed by using Design expert software. The desirability factor was in the range of 0 to 1, in which the smallest value shows the low desirable factor. The process variables with ultimate desirability mean it has optimum variable setting. The optimum values for the input parameters and their corresponded responses were analyzed by the software and the details were presented in Table 9. For single factor optimization, other response variables were neglected while in case of multivariable optimization, all the responses were considered and provided the equivalence significance. For the validation of optimized responses confirmation experiments were performed as shown in Table 9. It was seen from the experimental validation predicted values were very close to the experimental values.

### 3.2 Calculation of Number of Abrasive Particle Interacting with the Workpiece Surface

For the optimized condition,

$C = 0.5$ ,  $V_m = 290 \times 10^{-6} \text{ m}^3$ ,  $d_m = 1219 \text{ kg/m}^3$ ,  $d_a = 1340 \text{ kg/m}^3$ ,  $d = 33 \text{ microns}$ ,  $D = 8 \times 10^{-3} \text{ m}$ ,

Putting the above values,  $V_a = 1.32 \times 10^{-4} \text{ m}^3$ ,  $V_o = 0.188 \times 10^{-13} \text{ m}^3$

Number of abrasive particle interacting with the workpiece surface,  $N_a = 1.153 \times 10^8$

Table 9 — Single factor and multifactor optimization and comparative study of optimized outcomes and experimental facts of process variables

Optimization type	Objective	Optimized process variables					Response (predicted)	Response (Experimental)	Desirability
		Current (Amp)	Duty cycle (Fraction)	Rotational speed (rpm)	Pressure (MPa)	Abrasive concentration (Fraction)			
Single Response	To minimize SSR	11.832	0.778	150.3	11.024	0.343	0.1497 $\mu\text{m}$	0.15 $\mu\text{m}$	0.828
Single Response	To maximize Micro-hardness	12	0.78	150.746	14.506	0.5	345.963 HV	345.951 HV	0.689
Single Response	To maximize % Improved $R_a$	11.696	0.78	248.621	11.115	0.326	42.44%	39.52%	1.000
Multi response	To minimize SSR, maximize Micro-hardness and maximize % Improved $R_a$	12	0.78	150	19.944	0.5	0.142 $\mu\text{m}$ , 336.038 HV and 38.52%	0.15 $\mu\text{m}$ and 335.175 HV and 35.02%	0.8

Number of abrasive particle in full volume of media,  $N = 70.21e8$

Fraction of abrasive particle involved in material removal = — = 0.016

### 3.3 Calculation of Velocity of Impact

$m^7 = 2.6e-7$  kg,  $r = 16.5e-6$  m,  $N = 150$  rpm,  $\omega = 15.707$  rad/s

So,  $F_c = 1.0583e-9$  N

Therefore,  $v_c = 1.56e-10$  m/s

If  $t = 2e-6$  sec, then  $v_a = 6.7e-5$  m/s

So, impact velocity will be =  $6.7e-5$  m/s

### 3.4 SEM Analysis

The work pieces finished by TACAFM process were studied by scanning electron microscopy. SEM analysis was done for analyzing the microstructure of the workpiece profile at 500 X magnification. Figure 8 shows the microstructure images of the workpiece during and after the finishing in TACAFM process. Figure 8(a) shows the SEM image of surface before finishing. In TACAFM process material removal occurs due to thermal evaporation, melting and abrasion mechanism. When the spark is generated between the rotating electrode and finishing surface, oxide formation occurred due to presence of atmospheric gases in the finishing zone. The surface imperfections such as oxide layers, recast layers could be seen on the surface as revealed in Fig. 8(b). The formation of oxide layers on the surface leads to poor electrical conductivity of the workpiece and further hampers the machining process. The continuous flow of media in the gap, made nearly about an oxygen free environment around the finishing zone. The abrasive

laden media had a high dielectric resistance which corresponded negligible breakdown during the machining and further ionize on the collision of electrons with the molecules. The spark developed high temperature in the finishing zone and melted the material on the surface. Figure 8(b) shows the craters developed on the surface due to EDM mechanism. When the electrode is kept stationary, it develops spark at a particular position where surface is nearer to the electrode tip and deteriorate the surface quality. Therefore rotation of electrode is necessary to maintain the uniformity over the surface. Figure 8(b) clearly shows the molten/ semi molten material and spark craters on the surface during the electrode rotation. These soften material required less amount of force by abrasive particles to take away from the surface. Figure 8(c) shows surface structure was improved after the finishing and no cracks were observed on the surface, although some abrasive particles cutting marks were observed on the surface infrequently.

### 3.5 XRD Results

X-ray diffraction technique was used for the micro structural study of the finished surface. X-ray diffraction technique analysis of finished workpiece was performed with a software tool “X’Pert High Score”. XRD graph of finished brass workpiece was shown in Fig. 9. It can be observed from the XRD graphs that maximum peaks were identified for copper magnesium oxide and copper oxide at  $2\theta$  of 42.32 degree on (200) plane and had a cubic crystal system. Also some more peaks were identified for Copper oxide at  $2\theta$  of 36.44 degree and 52.48 degree



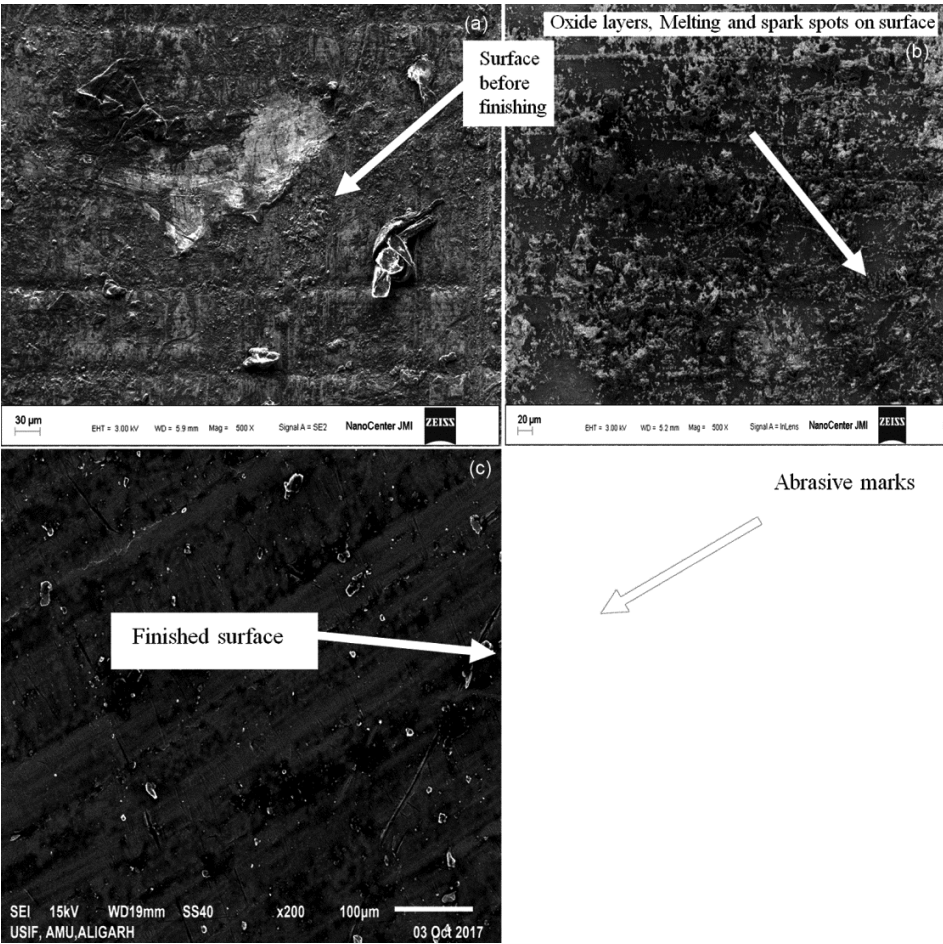


Fig. 8 — (a) SEM image of surface before finishing, (b) Oxide layers formation and Spark spots during the electrode rotation, and (c) SEM image of finished surface.

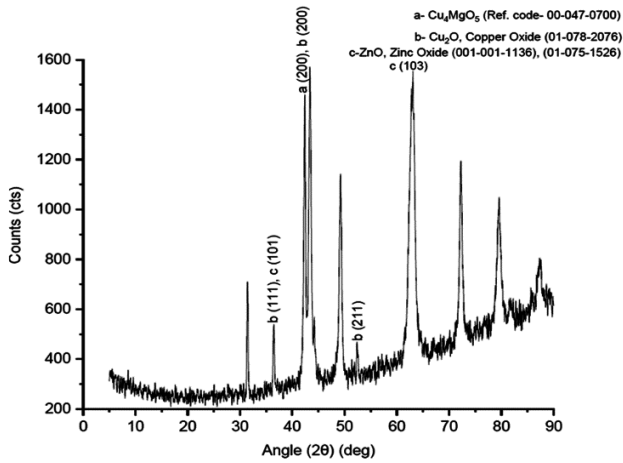


Fig. 9 — XRD results of surface produced by TACAFM process. at plane (111), (211) respectively with a cubic crystal system. Zinc Oxide peaks were identified at 2θ of 36.49 degree and 63.102 degree at plane (101), (103) respectively. Both were having hexagonal crystal

system. During the XRD interpretation some groups with lower peaks were discounted.

#### 4 Conclusion

In the present work, brass workpiece was finished using the developed Thermal additive Centrifugal AFM process. Experiments were performed for the input parameters such as current, duty cycle, rotational speed, extrusion pressure and abrasive concentration against responses scatter of surface roughness, micro-hardness of surface and % improvement in  $R_a$ . The results were obtained from the regression correlations and software tools. The conclusions drawn on the basis of the above study are:

- Duty cycle and current intensity have the major effect on finishing performance in comparison of the other input parameters.
- Duty cycle has the highest contribution of 17.5 % on the scatter of surface roughness. Scatter of

roughness decreased on increasing the duty cycle. The best scatter of roughness value obtained for the brass workpiece was 0.15  $\mu\text{m}$ .

- Current has the largest contribution of 85.17 % against the micro-hardness of surface. On increasing the current, micro-hardness of the surface was increased. The best micro-hardness of the surface was achieved as 332.58 HV.
- Current has the largest contribution of 21.88% against % improvement in  $R_a$ . On increasing the current intensity, more material removal was obtained due to more number of craters formed by the EDM mechanism on the surface. This makes easier for the abrasive particles to take away the material and provide good level of finish. The best % improvement in  $R_a$  was 39.52%.
- A fraction of 0.016 abrasive particles have participated in material removal.
- The process is capable of removing the craters and cracks on the surface at higher values of variables.

## References

- 1 Tan K L, Yeo S H, & Ong C H, *Proc Inst Mech Eng B J Eng Manuf*, 271 (2016) 2302.
- 2 Singh G, Singh A K, & Garg P, *Mater Manuf Process*, 32 (2016) 581.
- 3 Singh S, *Studies in metal finishing with magnetically assisted abrasive flow machining, Ph. D thesis, IIT Roorkee*, 2002.
- 4 Subramanian KT, & Balashanmugam N, *Int J Adv Manuf Technol*, 85 (2016) 2189.
- 5 Han S, Salvatore F, & Rech J, *J Mater Process Technol*, 267 (2018) 348.
- 6 Tzeng Hsinn Jyh, Yan Biing-Hwa, Hsu Rong-Tzong, & Lin Yan-Cherng, *Int J Adv Manuf Technol*, 34 (2007) 649.
- 7 Mittal S, Kumar V, & Kumar H, *Mater Manuf Process*, 30 (2015) 902.
- 8 Wan S, Ang Y J, Sato T, & Lim G C, *Int J Adv Manuf Technol*, 71 (2014) 1077.
- 9 Chen K Y, & Cheng K C, *Int J Adv Manuf Technol*, 74 (2014) 781.
- 10 Fu Y, Wang X, Gao H, Wei H, & Shichong L, *Int J Adv Manuf Technol*, 84 (2016) 1725.
- 11 Mali H S, & Manna A, *Int J Adv Manuf Technol*, 50 (2010) 1013.
- 12 Lv Z, Hou R, Huang C, Zhu H, & Qi H, *Int J Adv Manuf Technol*, 100 (2019) 2021.
- 13 Wang X C, Wang C C, Wang C Y, & Sun F H, *Chin. J Mech Eng*, 31 (2018) 97.
- 14 Shao Y, & Cheng K, *Int J Adv Manuf Technol*, 105 (2019) 4571.
- 15 Sankar M R, Jain V K, & J Ramkumar, *Int J Mach Tools Manuf*, 50 (2010) 637.
- 16 Sankar M R, Mondal S, Ramkumar J, & Jain V K, *Int J Adv Manuf Technol*, 42 (2009) 678.
- 17 Brar B S, Walia R S, Singh V P, & Sharma M, *J Inst Eng India Ser C94* (2013) 21.
- 18 Walia R S, Shan H S, & Kumar P K, *Int J Adv Manuf Technol*, 44 (2009) 700.
- 19 Marzban M Ali, & Hemmati Seyed Jalal, *Int J Adv Manuf Technol*, 89 (2017) 125.
- 20 Tzeng Hsinn Jyh, Yan Biing Hwa, Hsu Rong Tzong, & Lin Yan Cherng, *Int J Adv Manuf Technol*, 32 (2007) 1163.
- 21 Walia R S, Shan H S, & Kumar P K, *Int J Adv Manuf Technol*, 38 (2008) 1157.
- 22 Yan B H, Tsai H C, & Huang F Y, *Int J Mach Tools Manuf*, 45 (2005) 194.
- 23 Walia R S, Shan H S, & Kumar P, *Mater Manuf Process*, 21 (2006) 375.
- 24 Walia R S, Shan H S, & Kumar P, *Mach Sci Technol*, 10 (2006) 1.
- 25 Gill A S, & Kumar S, *Mater Manuf Process*, 31 (2016) 514.
- 26 Walia R S, Shan H S, & Kumar P, *Int J Adv Manuf Technol*, 39 (2008) 1171.
- 27 Gorana V K, Jain V K, & Lal G K, *Int J Mach Tools Manuf*, 44 (2004) 201.

# Comparative Analysis of Software Reliability Prediction Using Machine Learning and Deep Learning

1<sup>st</sup> Akshat Jindal

Department of Software Engineering  
Delhi Technological University  
New Delhi, India  
[akshatjindal\\_2k18se017@dtu.ac.in](mailto:akshatjindal_2k18se017@dtu.ac.in)

2<sup>nd</sup> Ashi Gupta

Department of Software Engineering  
Delhi Technological University  
New Delhi, India  
[ashigupta\\_2k18se040@dtu.ac.in](mailto:ashigupta_2k18se040@dtu.ac.in)

3<sup>rd</sup> Rahul

Department of Software Engineering  
Delhi Technological University  
New Delhi, India  
[rahul@dtu.ac.in](mailto:rahul@dtu.ac.in)

**Abstract**— Software Reliability is an integral part to determine Software Quality. Software is considered to be of high quality if its reliability is high. There exist many statistical models that can help in predicting Software Reliability, but it is very difficult to consider all the real-world factors and hence it makes the task of reliability prediction very difficult. Therefore, it becomes more challenging for the IT industry to predict if a software is dependable or not. Machine Learning and Deep Learning can be used for the prediction of Software Reliability by programming a model that assesses reliability by fault prediction in a more meticulous manner. Therefore, in this study the use of predefined Artificial Intelligence algorithms, mainly Artificial Neural Network (ANN), Recurrent Neural Network (RNN), Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) are intended for predicting software reliability on a time series software failure dataset and are compared on the basis of selected performance metrics. Each of the algorithm trained on software failure dataset will be used to predict the software failure time after a certain number of corrective modifications are performed on the software. Based on the result of the studies, it is discovered that LSTM produces superior outcomes in predicting the software failure trend as it can capture long and short-term trends in the software failure dataset.

**Keywords**— *Software Reliability, Deep Learning, Time Series data, comparative analysis*

## I. INTRODUCTION

The term "software reliability" refers to operational dependability. Software reliability may alternatively be defined as the probability that a software system will perform its assigned task in a given environment for a certain number of input cases, assuming that the hardware and input are both error-free. It is of crucial importance to evaluate Software Reliability for determining system dependability. However, it is difficult to accomplish reliability given the increasing complexity in software requirements.

Machine Learning and Deep Learning techniques can predict the fault rate for a given software more precisely by learning on past input data without human judgment thus leaving less room for errors and assumptions contrary to statistical methods. (Malhotra and Negi 2013) [1]. System behavior can be anticipated by utilizing Machine Learning which learns from its past and current software failure data

as it's a tool to automate data processing. Deep learning techniques include Artificial Neural Networks (ANNs), Recurrent Neural Networks (RNNs), and others that use a collection of algorithms to replicate the brain's actions. A neural network is made up of four primary components: inputs, weights, a bias or threshold, and an output. However, a deep learning model requires more data points to enhance its accuracy, although a machine learning model requires less data due to its fundamental data structure.

For this experiment, a heuristic examination of several Machine Learning and Deep Learning techniques on a univariate software failure time series data to investigate which approach can be used extensively for predicting software reliability has been proposed. Then, the metrics like Mean Absolute Error, Mean Squared Error, Median Absolute Error, and Maximum Error are used to determine their accuracies. The reason for choosing these metrics is that they can capture which technique closely represents the actual software failure dataset. For this paper, an impact on the data-driven approaches are chosen to be laid compared to the hardware/architecture-based approaches where other factors like the environment and time play a significant role and leave a major probabilistic factor.

## II. RELATED WORKS

The most prevalent Machine Learning algorithms used for Software Reliability Prediction and Modeling include Genetic Programming, Decision Trees, Support Vector Machines, and Particle Swarm Optimization. ML techniques have been proven to work better than stochastic ones due to the nature of models to learn from previous errors are thereby leading to close precision and fewer errors (Malhotra and Negi 2013) [1].

In order to facilitate the acceptance of connectionist models and their usage in software reliability models, Cai et al. (1991) [2] and Karunanithi et al. (1992) [3] undertook considerable research. G Krishan et alii. (2018) [4] compared mainly Artificial Neural Networks and SVM to conclude that geometric understanding leads to better results for SVM than NN's.

Prediction tests to oversee software reliability were utilized by Pai and Hong (2006) [5] by using SVM algorithms. Loui et alii. (2016) [6] employed a relevance vector machine for

the prediction of software dependability. Machine learning approaches such as fuzzy inference systems, cascade correlation neural networks, and decision trees are used by Kumar and Singh (2012) [7] to predict outcomes. Jaiswal and Malhotra (2016) [8] talk about predicting software reliability using ANFIS. Other methods like Bagging, GRNN, SVM, MLP, MSP, FFBPNN, CFBPNN, Lin Reg, RepTree are studied by Xingguo and Yanhua (2007) [9]

In papers like (Gokhale 1998) [10], an application was run against some software and fixed test cases to figure out a scientific model (architecture of application) in terms of criteria like branching probabilities and failure model of its components issues that could be later solved for better reliability. Future works have been shown to find a systematic way of predicting software reliability by incorporating debugging functionalities (S Trivedi 2006) [11].

However, with moving towards Machine Learning and Deep Learning techniques these tools and methods have been outperformed by the data gathered. In his research, NR Kiran (2008) [12] presented a unique soft computing-based technique that employs a non-linear ensemble trained using Back Propagation Neural Networks. A similar notion was observed by (Kumar and Jayaram 2014) [13], who used Artificial Neural Networks and Genetic Algorithms to forecast Software Reliability.

(Amin, Grunske, and Colman, 2013) [14] proposed using ARIMA modeling to solve the unrealistic assumptions, environment-dependent applicability, and questionable predictability associated with Software Reliability Growth Models (SRGMs), and showed that their method performed better due to the data-driven approach used.

### III. PROBLEM STATEMENT

The main goal of the project is to predict Software Reliability using various Machine Learning and Deep Learning algorithms and to compare and analyze the performance of the chosen algorithms. The chosen algorithms are trained on a software failure dataset to determine when the given software would malfunction based on its relationship between previous failures and modifications. The goal of modeling software dependability is to determine the likelihood of a piece of software failing in a given environment. A Theoretical justification have also been structured for the results of comparison for Machine and Deep Learning algorithms.

### IV. EXPERIMENTAL DESIGN

#### A. Dataset

- The Software Failures Dataset has been used from [6]. The dataset has a total of 101 samples. There are 2 attributes in the dataset as shown in Table 1.

- The variable  $t$  as shown in the table below is taken to be the cumulative number of corrective changes performed on the software, after each failure.

Table.1 Dataset description

Attribute	Description
$t$	The number of modifications made to the software
$Y_t$	Failure time after $t$ modification is made

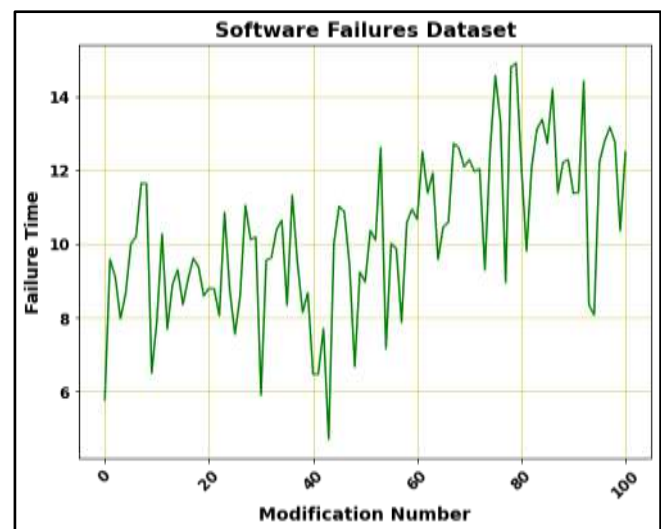


Fig.1 Dataset visualization: representing relationship between the time after which software fails to the number of modifications made in the software.

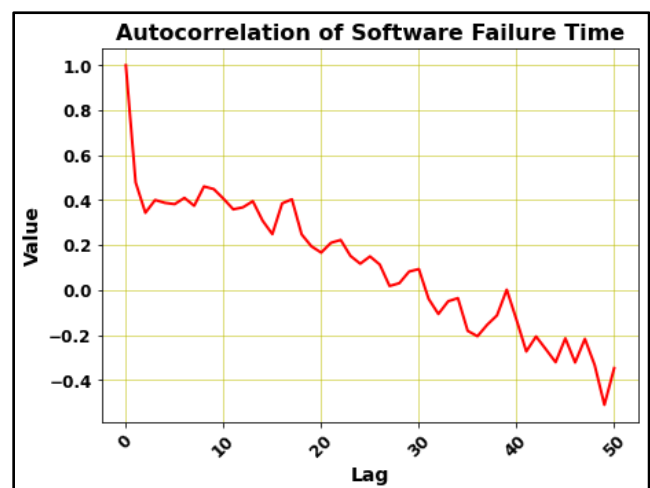


Fig.2 Dataset visualization: The Auto Correlation of Software Failure Time with Lag to measure the linear relationship between an observation at number of modification  $t$  and the observations at previous modifications

#### B. Data Transformation

- The dataset is a time series dataset. Before application of any ML/DL algorithm, it needs to be transformed into a suitable representation of the time series.

- It is transformed into an **86 x 16 shaped dataset** where each row contains the current failure time and along with it a sequence of **past 15 failure time** values.
- The dataset is then divided into training, validation and test set as shown in Table 2.

Table.2 Dataset bifurcation

Dataset Partition	Percentage	No. of Samples
Training Set	70%	61
Validation Set	10%	7
Test Set	20%	18

### C. ALGORITHMS USED

#### 1. ANN (Artificial Neural Network)

The ANN algorithm is a machine learning method that is based on the structure of the human brain and is one of the most commonly used for regression and classification issues.

Fig 3 shows it consists of three layers: an input layer, an arbitrary number of hidden levels, and an output layer.

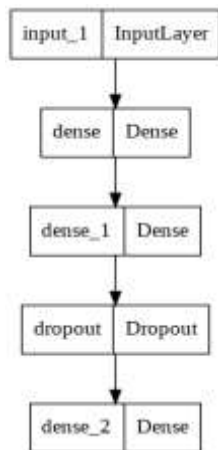


Fig. 3 ANN Model Summary

**Reason for selecting ANN:** To get a baseline prediction performance.

Training epochs: 100, Batch Size:5

#### 2. RNN (Recurrent Neural Network)

The RNN algorithm is a type of Neural Network. The variation in this algorithm is that it has a hidden state, which takes into consideration the information stored in a sequence as represented in Fig 4. In RNN, the independent activations are converted into dependent activations, which is done by using the same weights and biases for all the layers.

This reduces the complexity and helps in memorizing the previous layer outputs.

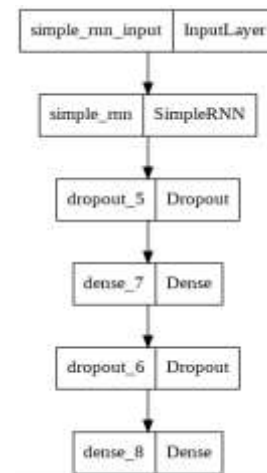


Fig. 4 RNN Model Summary

**Reason for selecting RNN:** Since the hidden state of RNN is used for remembering information about a sequence, it can be used for a time series prediction problem.

Training epochs: 100, Batch Size:5

#### 3. GRU (Gated Recurrent Unit)

GRU is a version of RNN that is designed to solve disappearing or exploding gradient issues since the model does not lose input from the current cell while also transmitting important information to the next cell. GRU decides which information should be sent to the output for prediction by employing two vectors to determine what information should be sent to the output for prediction.

GRU consists of 3 gates:

- **Update Gate:** This gate determines how much information should be handed on to future generations.
- **Reset Gate:** It determines how much past knowledge is useless and hence can be forgotten.

- **Current Memory Gate:** This gate is further incorporated into the reset gate and brings non-linearity in the input.

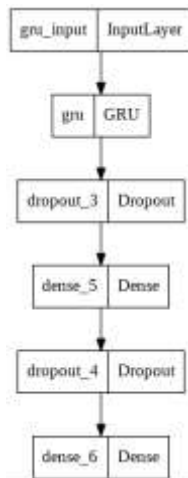


Fig. 5 GRU Model Summary

**Reason for selecting GRU:** Since GRU is a variation of RNN only, hence it can be effective in a time series prediction problem by virtue of its capability to memorize some information about a sequence.

Training epochs: 100, Batch Size:5

#### 4. LSTM (Long and Short term Memory)

LSTM is another variation of RNN. LSTM was created to address the issue of long-term dependencies while also preserving information over several timestamps of input data. To store information, it comprises a chain-like structure with four neural networks and cells called memory blocks as represented in Fig. 6. When there are long-term dependencies in a data series, the efficiency of RNN decreases.

These issues are addressed by LSTM's three gates:

- Input Gate:** It adds useful information to a cell state.
- Forget Gate:** It is used to forget long-term information that is not required anymore.
- Output Gate:** It is used for extracting useful information from the current cell.

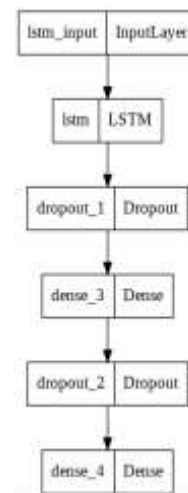


Fig 6. LSTM Model Summary

**Reason for selecting LSTM:** Since LSTM was specially developed to solve the problem of long-term dependencies in a sequence, it can be very effective in a time series prediction problem.

Training epochs: 100, Batch Size: 5

## V. RESULTS

Fig. 7, Fig. 8, Fig. 9 and Fig. 10 are the graphs plotted for the performance of the different algorithms/models on the given dataset and compared with the actual failure time values.

From the graphs the results are not very obvious therefore a more mathematical formulation of the results is needed which can be provided by performing an error analysis on the results obtained.

Error analysis also helps us to represent the algorithm's performance and compare with other algorithms. In this paper, mean absolute error, mean squared absolute error, median squared error and maximum error have been chosen as described in the following section. **The following metrics have been chosen as they help to model regression problems where the output is a real or continuous value i.e., software failure time in this case as well as serves as an L1 loss function.**



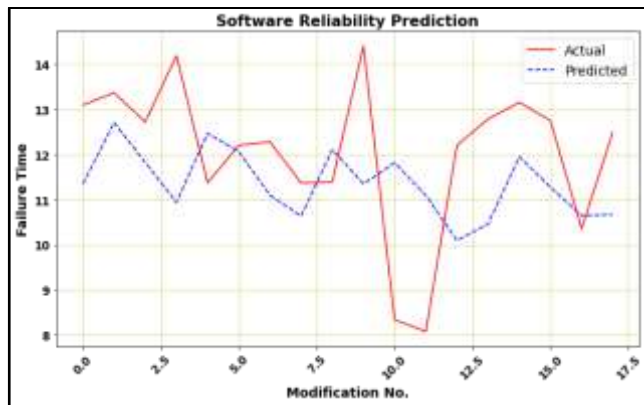


Fig 7. Prediction for ANN on Test Set

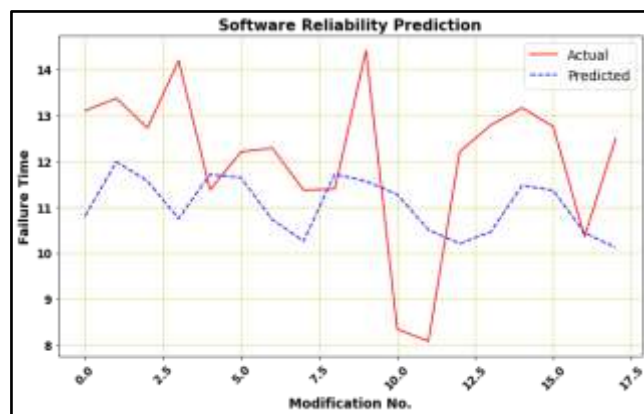


Fig 8. Prediction for RNN on Test Set

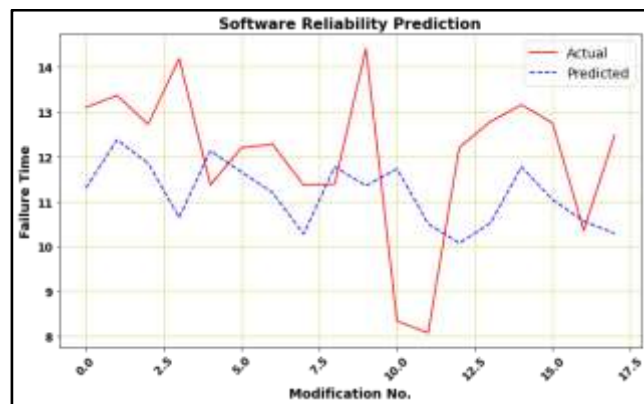


Fig 9. Prediction for GRU Test Set

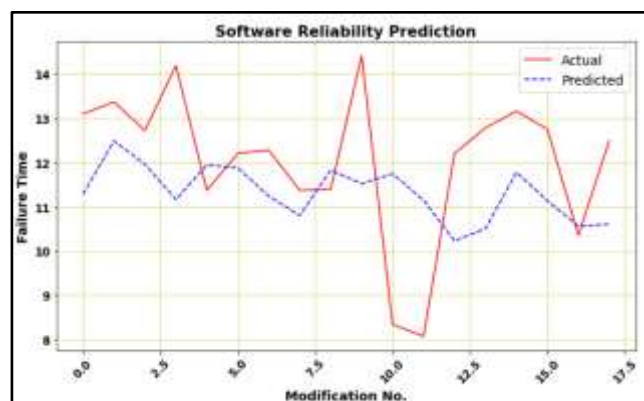


Fig 10. Prediction for LSTM on Test Set

## VI. METRICS USED FOR MODEL EVALUATION.

- A. **Mean Absolute Error:** For a given observation, the provided metric calculates the average of the absolute difference between the predicted and actual values
- B. **Mean Squared Error:** This metric is used to compute the average values for the square of difference between values predicted by the algorithm and values actually observed from the simulated domain.
- C. **Median Absolute Error:** It is the amount of the absolute difference between expected and actual values for a group of data that falls in the center.
- D. **Maximum Error:** This metric talks about the error arising from difference between absolute predicted and actual observation i.e., software failure time.

## VII. RESULTS JUSTIFICATION

- It can be inferred from Table 3, and from Fig. 11 that LSTM has the least Mean Absolute Error, Mean Squared Error, and Maximum Error amongst the chosen 4 algorithms.
- The errors in the predictions of RNN and GRU are very close to each other and are also close to the baseline errors as seen by Table 3.
- Software Reliability Prediction is a typical time series problem. It can have long and short-term trends that must be learned by any model for making decent predictions.
- Although RNN and GRU have memory units that can remember some information about a sequence, their performance efficiency gets decreased when there are long-term dependencies in the input data sequence and therefore performed poorly when compared to LSTM.
- LSTM is specially developed to capture Long and Short term trends in the input data sequence. Therefore, LSTM outperformed all the algorithms and gave relatively low errors in the reliability prediction task.

Table 3. Compiled error scores for all different models

	Mean Absolute Error	Mean Squared Error	Median Absolute Error	Maximum Error
ANN	1.6247	3.6756	<b>1.3449</b>	3.4827
RNN	1.6821	3.7328	1.6267	3.4486
GRU	1.6587	3.7310	1.5486	3.5465
<b>LSTM</b>	<b>1.5639</b>	<b>3.4735</b>	1.4992	<b>3.4095</b>



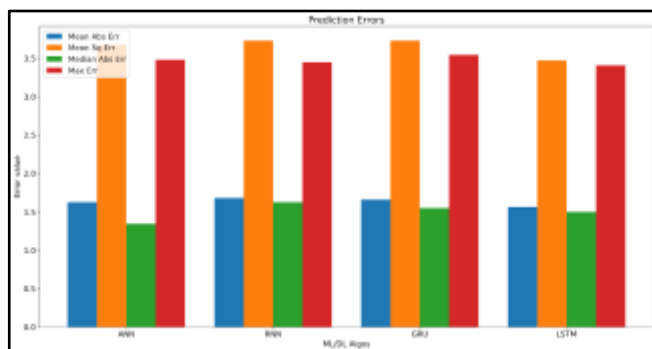


Fig 11. Graphical representation of error metrics i.e., mean absolute, mean squared, median absolute and maximum error in all 4 algorithms

## VIII. CONCLUSION

Machine Learning and Deep Learning algorithms have been applied for Software Reliability Prediction. The dataset was analyzed, visualized, and transformed to make it a time series prediction problem.

Four models have been trained for software reliability prediction and reported their performance in terms of different sorts of prediction errors.

Considering ANN as the baseline model, the performance of all of the above-described models are compared and found that LSTM outperformed all the algorithms by virtue of its ability to effectively capture Long and Short term trends in a data sequence and gave a theoretical justification of the results obtained.

## IX. REFERENCES

- [1]. Malhotra R, Negi A (2013) Reliability modeling using particle swarm optimisation. The society for reliability engineering, quality and operations management (SREQOM), India and The Division of Operation and Maintenance, Lulea University of Technology, Sweden. Int J Syst Assur Eng Manag. doi:10.1007/s13198-012-0139-0
- [2]. Cai YK, Wen YC, Zhang LM (1991) A critical review on software reliability modeling. Reliab Eng Syst Saf 32(3):357-371
- [3]. Karunanithi N, Whitley D, Malaiya Y (1992) Prediction of software reliability using connectionist models. IEEE Trans Softw Eng 18(7):563-574
- [4]. Mohan, G. & Yoshitha, N. & Lavanya, M.L.N. & Priya, A.. (2018). Assessment and Analysis of Software Reliability Using Machine Learning Techniques. International Journal of Engineering and Technology (UAE). 7. 201-205. 10.14419/ijet.v7i2.32.15567.
- [5]. Pai, F.P., Hong, C.W.: Software reliability forecasting by support vector machines with simulated annealing algorithms. J. Syst. Softw. 79, 747-755 (2006)
- [6]. Lou J, Jiang Y, Shen Q, Shen Z, Wang Z, Wang R (2016) Software reliability prediction via relevance vector regression. Neuro-computing 186:66-73
- [7]. Kumar P, Singh Y (2012) An empirical study of software reliability prediction using machine learning techniques. Int J Syst Assur Eng Manag 3(3):194-208. doi:10.1007/s13198-012-0123-8
- [8]. Arunima Jaiswal, & Ruchika Malhotra. (2016). Software Reliability Prediction Using Machine Learning Techniques. Proceedings of Fifth International Conference on Soft Computing for Problem Solving, 141-163. doi:10.1007/978-981-10-0448-3\_12
- [9]. Xingguo L, Yanhua S (2007) An early prediction method of software reliability based on support vector machine. In: Proceedings international conference on wireless communications, network- ing and mobile computing (WiCom'07), pp 6075-6078
- [10]. S. S. Gokhale, W. E. Wong, K. S. Trivedi and J. R. Horgan, "An analytical approach to architecture-based software reliability prediction," *Proceedings. IEEE International Computer Performance and Dependability Symposium. IPDS'98 (Cat. No. 98TB100248)*, 1998, pp. 13-22, doi: 10.1109/IPDS.1998.707705.
- [11]. S. S. Gokhale and K. S. Trivedi, "Analytical Models for Architecture-Based Software Reliability Prediction: A Unification Framework," in *IEEE Transactions on Reliability*, vol. 55, no. 4, pp. 578-590, Dec. 2006, doi: 10.1109/TR.2006.884587.
- [12]. N. Raj Kiran, V. Ravi, Software reliability prediction by soft computing techniques, Journal of Systems and Software, Volume 81, Issue 4, 2008, Pages 576-583, ISSN 0164-1212, <https://doi.org/10.1016/j.jss.2007.05.005>.
- [13]. M.A. Jayaram & H.S. Vijayakumar, (2014). ON APPLICATIONS OF SOFT COMPUTING ASSISTED ANALYSIS FOR SOFTWARE RELIABILITY. International Journal of Innovative Research in Advanced Engineering (IJRAE). Volume 1.
- [14]. Ayman Amin, Lars Grunske, Alan Colman, An approach to software reliability prediction based on time series modeling, Journal of Systems and Software, Volume 86, Issue 7, 2013, Pages 1923-1932, ISSN 0164-1212, <https://doi.org/10.1016/j.jss.2013.03.045>.
- [15]. Patel, Viral, Daanyaal Kapadia, Deval Ghevariya, and Shiburaj Pappu. "All India Grievance Redressal App." Journal of Information Technology and Digital World 2, no. 2: 91-99.
- [16]. Chen, Joy Iong Zong, and Joy Iong Zong. "Automatic Vehicle License Plate Detection using K-Means Clustering Algorithm and CNN." Journal of Electrical Engineering and Automation 3, no. 1 (2021): 15-23.
- [17]. Mugunthan, S. R., and T. Vijayakumar. "Design of Improved Version of Sigmoidal Function with Biases for Classification Task in ELM Domain." Journal of Soft Computing Paradigm (JSCP) 3, no. 02 (2021): 70-82.
- [18]. Manoharan, Samuel. "Study on Hermitian graph wavelets in feature detection." Journal of Soft Computing Paradigm (JSCP) 1, no. 01 (2019): 24-32
- [19]. Tripathi, Milan. "Analysis of Convolutional Neural Network based Image Classification Techniques." Journal of Innovative Image Processing (JIIP) 3, no. 02 (2021): 100-117.
- [20]. Rahul, Ayush, Divya Agarwal, Devika Vijay. "Genre Classification using Character Networks", 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021
- [21]. J.R. Horgan. "An analytical approach to architecture-based software reliability prediction", Proceedings IEEE International Computer Performance and Dependability Symposium IPDS 98 (Cat No 98TB100248) IPDS-98, 1998
- [22]. Swagat Ranjit, Shruti Shrestha, Sital Subedi, Subarna Shakya. "Comparison of algorithms in Foreign Exchange Rate Prediction", 2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS), 2018

# Design of Low Cost Bio-impedance Measuring Instrument

Rajesh Birok, Rajiv Kapoor  
Dept. of Electronics & Communication Engineering  
Delhi Technological University Delhi, India

**Abstract**—It is a well-established fact that the electrical bio-impedance of a part of the human body can provide valuable information regarding physiological parameters of the human body, if the signal is correctly detected and interpreted. Accordingly, an efficient low-cost bio-electrical impedance measuring instrument was developed, implemented, and tested in this study. Primarily, it is based upon the low-cost component-level approach so that it can be easily used by researchers and investigators in the specific domain. The measurement setup of instrument was tested on adult human subjects to obtain the impedance signal of the forearm which is under investigation in this case. However, depending on the illness or activity under examination, the instrument can be used on any other part of the body. The current injected by the instrument is within the safe limits and the gain of the biomedical instrumentation amplifier is highly reasonable. The technique is easy and user-friendly, and it does not necessitate any special training, therefore it can be effectively used to collect bio-impedance data and interpret the findings for medical diagnostics. Moreover, in this paper, several existing methods and associated approaches have been extensively explored, with in-depth coverage of their working principles, implementations, merits, and disadvantages, as well as focused on other technical aspects. Lastly, the paper also deliberates upon the present status, future challenges and scope of various other possible bio-impedance methods and techniques.

**Keywords**—Noninvasive; bio-electrical; Impedance; bio-impedance; bio-medical; instrumentation

## I. INTRODUCTION

Extensive research is going on in the field of Bio-Medical Instrumentation. Researchers and investigators in this field are striving hard to find out new ways and methods for diagnosis and measurement of health parameters for the welfare of the mankind. There are two types of techniques for measuring biomedical signals namely non-invasive and invasive techniques. Non-invasive techniques are more suitable than the invasive ones if sufficient accuracy can be achieved using them. Whether, it is animal and plant cells or tissues, these are always made up of three-dimensional arrangement of cells and tissues. Therefore, human body is a complex biological structure and system, which is also made up of billions of cells and tissues arranged in 3-D formation [1]. The biological cells and tissues of both animals and plants floats in ECF which is known as Extra-Cellular Fluids. This ECF comprises Intra-Cellular Fluids (ICF) and Cell Membranes (CM) which may be with or without cell wall. When biological cells and tissues are subjected to the external electrical stimulus they respond and produces a complex bio-electrical impedance or simply known as bio-impedance. This bio-impedance is highly frequencydependent [2], [3].

Accordingly, frequency response of bio-impedance of cells and tissues of humans is greatly affected by physiological and physiochemical composition and structure of these cells and tissues. Moreover, it also changes from person to person. As a result, learning about cell and tissue anatomy and physiology through biological cell and tissue bio-impedance analysis will be a valuable resource. Therefore, it has been found that studying complex bio-impedance of biological cell and tissues is a useful method for non-invasive physiological and pathological investigations. As we know that the bio-electrical impedance of a biological cells or tissues is dependent on the signal frequency, however, multifrequency application may also be used for non-invasive diagnostics and medical investigations, so as to determine their physiological or pathological behavior or even properties. There are numerous Non-invasive bio-impedance techniques such as BIA (Bio-Impedance Analysis), EIT (Electrical Impedance Tomography), IPG (Impedance Plethysmography), ICG (Impedance Cardiography), etc. The bio-impedance measurement technique proposed in this research paper, is a low-cost, efficient, and effective non-invasive diagnostic technique.

## II. LITERATURE SURVEY

Impedance offered by a living tissue is known as bio-impedance. The broad variability of Cole parameters makes it difficult to use bio-impedance to distinguish animal and plant tissues. [4] defines a novel electronic procedure for distinguishing fruit or vegetable tissue. This system uses a custom-built electrode pair to compute bio-impedance and Cole parameters covering a wide range of frequencies from 1 Hz to 1 MHz [4]. However, impedance of human cell and tissue consists of resistive and capacitive components [5], [6]. It's determined by injecting an alternating current in the cell or tissue and then measuring the output voltage across it. The Linear Time-Varying (LTV) bio-impedance is measured with a specified precision using stepped-sine excitations, as given in [7], but it is susceptible to temporal distortions affecting the data, which limits the device's temporal bandwidth and sets the data accuracy. Current source and voltage sensing circuit are the essential blocks in the instrument. Several authors have successfully designed current sources operating up to few hundreds of kHz [8]. Paul Annus et.al. [9] have systematically analyzed the design of a current source using transfer function approach and they have measured the load impedance using load in loop method or configuration. The voltage sensing circuit consists of amplifier, demodulator and low pass filter. The use of instrumentation amplifier for bio-impedance measurements has been analyzed by Areny and Webster [10]. The measurement technique has been used in a

number of applications such as calculating Total Body Water (TBW), calculating Intracellular Fluid (ICF) and extracellular fluid (ECF) [6] Electrical Cardiometry [11], Skin Water Content, Impedance Imaging (Tomography), Ablation Monitoring and measurement of Respiration Rate [12]. The technique also has the potential to be used in biometrics [13]. Body Composition Assessment, Transthoracic Impedance Pneumography, Electrical Impedance Tomography (EIT), and Skin Conductance are examples of bio-impedance applications that are described and analysed in [14]. [15] looked into the possibility of non-invasively tracking blood glucose levels using bio-impedance data, which would allow for more regular testing and better diabetes management and monitoring. The bio-impedance measurement is not a new technique in biomedical diagnostic techniques but research is still going on to make this technique a standard procedure for diagnosis of a particular disease. For this purpose, it is required that a large amount of data be collected for a particular disease, for a particular environmental society and analysis to be done, correlating the disease with the signal recorded. To attain this objective, we have designed a low-cost instrument using common analog signal processing blocks, which gives an accurate recording of the bio-impedance signal.

### III. BACKGROUND

#### A. Basics and Origin of Bio-impedance

Bio-impedance is a passive electrical property that describes a biological cell or tissue's ability to obstruct (oppose) the flow of electrical current through it. The reaction to electrical excitation (current or potential) applied to biological tissue is used to determine bio-impedance. The same or other electrodes applies the excitation signal and picks up the reaction in bio-impedance measurements, then charge conversion takes place from electronic to ionic charge and vice versa [7]. Simply, the ratio of voltage (V) to alternating current (I) is known as electrical impedance (Z). Since, Direct Current (DC) is quite hazardous to humans, therefore, it is never used for any experimentation on humans. In fact, Alternating Current (AC) is more preferable choice for such type of applications. The calculated or observed bio-impedance (Z) is highly influenced by the Resistive (R), Capacitive (C), and Inductive (L) parts of the cells and tissues. The bio-impedance (Z) is given by using the modulus  $|Z|$  and the phase change. Since, bio-impedance (Z) is a complex function or parameter, therefore its Resistance (R) is the real part and whereas, the Capacitance Reactance ( $X_c$ ) is the imaginary part. As the ICF, CM, and ECF are made of dissimilar materials with nonidentical electrical properties, therefore every cell and tissue components react differently to the applied AC signal. As we know that ICF and ECF are made up of ionic solution which is highly conducting in nature, thus it provides low resistance path to the applied AC signal [16]. The CM are composed of lipid bilayers which are electrically nonconducting and inserted between two layers of conducting proteins. This sandwiched structure, produces a capacitive reactance ( $X_c$ ) to the applied AC signal [17], [18]. Due to this, biological cells and tissues produces a complex bio-impedance (Z) which can be considered as overall response to an applied AC signal [2], [3]. Thus, bio-impedance (Z) is a complex function depends upon cell and tissue composition and structure, health of person and applied AC signal frequency. Moreover, it also changes with measurement direction,

from one subject to the other subject and even within the tissue itself.

The human body composition comprises, water (64%), protein (20%), fat (10%), and minerals (5%) and starch (1%). The human body mass is mainly due to O (65%), C (18%), and H (8%). The majority of muscles are made up of protein whereas, majority of bones are made up of minerals [19]. The bio-impedance is proportional to Total Body Water (TBW), which contains Intra-Cellular Water (ICW) and Extra-Cellular Water (ECW). Body water, body fat, and body muscle have different impedance values according to the amount of presence of water in these. Thus applied AC signal pass through paths that contain more water as it provides high conductivity [20]. The physiological, morphological, pathological settings and also applied AC signal frequency, all these affect and influence cells and tissues and their electrical properties [21], [22]. The biological cells and tissues may have active (endogenous) or passive (exogenous) electrical properties, depending on the type of source of applied AC signal. The bio-electric signals from the heart known as electrocardiograph (ECG), signals from the brain known as electroencephalograph (EEG), whereas electromyograph (EMG) signals from the muscles are few examples of active properties (bio-electricity) produced by ionic activities within cells and tissues (typical of nerve cells). Passive properties are generated by simulating them with an external electrical excitation source [23], [24].

The extracellular fluids surround all cells with membranes in biological tissues. The main constituents of Extra-Cellular Fluid (ECF) are fluid component of the blood known as plasma and the other one is Interstitial Fluid (IF) which surrounds all cells that are not in blood. The extracellular space is the part of a multicellular organism outside the cells, whereas intracellular space is within the organism's cells. The cell membranes separate extracellular spaces and intracellular space thus producing two electrically conducting compartments known as extracellular media and intracellular media. The resistive pathways are provided by ECF and intracellular fluids (ICF). Due to its insulating design and structure, the lipid bilayer cell membrane is very-very thin measuring approximately 6-7 nm. This lipid bilayer cell membrane is semi-permeable, due to which it has a high capacitance and which produces capacitive reactance [25], [26], [27]. Although biological cells and tissues may have inductive properties, inductance is much lower at low frequencies than resistance and reactance, so it is often overlooked [28]. Thus, biological cells or tissue's complex bio-impedance is the contributions from both frequency-dependent capacitance and conductance [21], [29], [30], [31], [32], [33]. Bio-electrical impedance often differs from one tissue to the next, as well as from one subject to the next. The complex bio-electrical impedance is affected by changes in cell and tissue composition and structure, and even health condition or status of the subject [5], [6].

#### B. Frequency Response of Bio-impedance

The anatomical, physiological, and pathological state of biological cells and tissues determine the bio-impedance frequency response. Therefore, the bio-impedance study can provide much more information related to the anatomy and physiology of a cell or tissue. Since the bio-impedance response is a variable of signal frequency, therefore bio-impedance analysis

with multifrequency inputs can give detailed cell or tissue attributes information, that can help greatly in cell or tissue specifications. Furthermore, the applied ac signal frequency also drastically affects it [2], [3]. Thus, the bio-impedance of cells or tissues and their frequency response is greatly affected not only by the composition and structure of these cells and tissues in the given physiological and physiochemical setting but also by the applied signal frequency. The few bio-impedance analysis techniques which makes use of lumped estimation of the bio-impedance values of the cell or tissue samples are BIA, IPG, and ICG etc. Bio-Electrochemical Impedance Spectroscopy (EIS) measures and analyses bio-impedance at different frequencies. Thus, EIS provides not only a lumped approximation of the cell or tissue sample's bio-impedance values at relatively higher frequency (generally 50 kHz), but also the details required to have better understanding of the many complex bio-electrical phenomena such as dielectric dispersions and relaxation.

The bio-impedance measurement can be broadly divided into two categories, namely, "single-tone" signals and "multi-tone" signals measurements. The analysis of "single-tone" signals is very straightforward, but measurements take longer, whereas use of a multi-tone signal allows for simultaneous coverage of the entire frequency spectrum. However, use of a multi-tone signal may result in an algorithm which can be more complex for analysis purpose [34]. Moreover, especially the dielectric properties of the object determine the required frequency range for bio-impedance measurements, which typically spans 3 to 4 decades in the kHz to MHz range. In general, wider range of frequencies improves fitting accuracy but at the cost of complicating measurements. Furthermore, (SNR) of measured signals also effects the fitting accuracy [35].

### C. Types of Electrode Configurations

When an alternating current is used in bio-impedance testing, the electrode displays a frequency-dependent Electrode Polarisation Impedance (EPI) at the contact point with the tissue or solution as the case may be. Any change in the electrode material that is in contact with the tissue or solution also affects the magnitude and as well as the phase of the electrode impedance. As a result, the total calculated impedance of the system is equal to the sum of the EPI and impedance of the tissue/solution [25]. Finally, the impedance depends upon type of electrode used, electrode material used, the applied signal amplitude and also on its size, shape and structure [36]. In order to measure bio-impedance, we need at least two electrodes for the electrical current to pass through the closed circuit [25]. Therefore, bio-impedance measurements are performed with two or four electrodes. In both methods, the red coloured electrodes depicted in Fig. 1 are known as input or excitation electrodes. Basically, these are the current or driving electrodes. On the other hand, blue coloured electrodes also depicted in Fig.1 are known as voltage/sensing or output electrodes used to determine output signal which is frequency dependent. It should be observed that bio-impedance measurements can be inaccurate due to factors such as movement and improper electrode placement [36]. Bio-impedance measurements are typically done with gel electrodes to minimise electrode-skin impedance. The usability of dry electrodes is studied in [37] because this type of electrode

is not suitable in many measurement environments. For bio-impedance measurement, there are two kinds of electrode configurations. According to its name the two-electrode system or setup depicted in Fig.1(a) for measurement of impedance makes use of two electrodes only. Thus, the current signal injection or current-carrying as well as voltage measurement or voltage pickup are done with the same electrodes. For unipolar measurements, electrode configuration with quite large and small electrodes can also be used [38].

The two-electrode technique suffers from contact impedance due to polarisation impedance at the electrodes' surfaces and the measured signal also covers the contact impedance due to voltage drop [25]. While, analysing the measured signal, the polarisation impedance should be considered and removed in the output signal [39], [40]. Overall, the results obtained using this method are interesting, however they do not provide an accurate signal on the electrode's surface [41]. In four electrode configuration or setup, two independent pair of electrodes are used for current injection and detecting changes in voltage or voltage measurements [39], [42]. In this system, the input that is constant amplitude current signal is injected using the outer electrodes. These electrodes are also known as current or driving electrodes depicted in red colour in Fig.1(b). The output voltage signal produced which is frequency dependent is measured at two points within the current electrode. These electrodes are also known as voltage or sensing electrodes depicted in blue colour in Fig. 1(b). In this configuration, as the distance between electrode pairs was increased, the magnitude of the measured impedance decreased [25], [43]. A number of factors influence the effect of electrode polarisation impedance which includes electrode content, size, measuring frequency, sample impedance, and so on. The main advantage of four electrode configuration over its counterpart two electrode configuration is that the voltage or sensing electrodes do not carry current, due to this it eliminates polarisation impedance. Thus, at the connecting surface of electrode with tissue or electrolyte the influence of contact impedance is greatly decreased. Therefore, use of this method is a common and popular practice to lower the effect of EPI [44]. Furthermore, four electrode configuration measurements are more sensitive and accurate.

## IV. PROPOSED METHOD

The instrument designed consists of power supply, current source, voltage sensing unit and data acquisition system (DAS). Fig.2 depicts the general block diagram of the measuring instrument. The current source circuit is used to generate a sinusoidal current of amplitude in the range of 600  $\mu$ A- 800  $\mu$ A and a frequency of 50 kHz [4]. The voltage sensing unit is used to amplify and remove the high frequency components from the signal sensed by voltage sensing electrodes. The signal is then given to the data acquisition system. The contact surface dimensions of all test electrodes are the same, and the carrier is a circular printed circuit board. To compare the electrodes' characteristics, the electrode-skin impedances are measured under a variety of signal frequencies, contact durations, contact pressures, positioning positions, and subjects. All measurements are often done with silver/silver chloride (Ag/AgCl) dry gel electrodes for contrast. [37] The instrument has four Ag/AgCl electrodes, two on the outside and two on the inside,

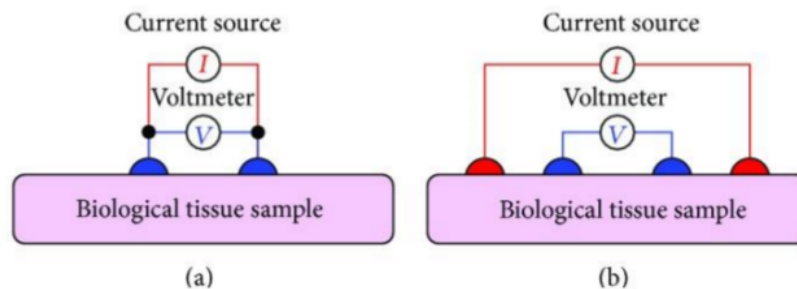


Fig. 1. Bio-impedance Measurement using: (a) Two- Electrode Method, (b) Four-electrode Method.

which are used as current electrodes and voltage electrodes, respectively. This type of arrangement is known as a tetra polar arrangement. Such an arrangement has been implemented by J. J. Wang et.al. [45] to use forearm impedance plethysmography for monitoring cardiac pumping function.

#### A. Waveform Generation

This is the first stage of the instrument and also the most critical one. All the parameters for this stage are very important as far as subject's safety and the output of the device is concerned. As this is the first stage any noise added at this stage will get amplified in the subsequent stages. A good waveform generator should have a single frequency output with very low value of total harmonic distortion and should have a very low frequency drift. A Wein bridge oscillator can be used to generate a sinusoidal waveform [46]. The oscillator is designed for a frequency of 50 kHz as this frequency is widely accepted in clinical use as a standard [47]. Among all types of voltage source generators, an oscillator produces the output with the most stable frequency as compared to other sources. Wein bridge oscillator is implemented using an op-amp, which gives a stable output with a single frequency for a particular combination of resistors and capacitors. Proper limiter circuit is also used in the circuit to keep the poles of the oscillator on the imaginary axis. One major problem with using Wein bridge oscillator is that the instrument designed cannot be used for bio-impedance spectroscopy.

A comparator, integrator and wave shaping circuit can be used to generate square, triangular and sinusoidal waveform of variable frequency [48]. The comparator and integrator circuit used together generate square and triangular waveforms. Triangular waveform is shaped using a shaping circuit to generate sinusoidal waveform. The maximum frequency achieved by this circuit using 741 op-amp is 26 kHz. This method can be used for excitation in lower frequency range but lower frequencies are not used in bio-impedance spectroscopy as electrodes get polarized at lower frequencies. It has also been observed during our laboratory experiments that there was no biomodulation obtained in the output signal when a current of frequency less than 20 kHz was injected into human body. The instrument designed uses a monolithic integrated circuit ICL8038, to produce high accuracy sine, square and triangular waveforms [49]. The frequency can be selected externally from 10 kHz to 100 kHz using a potentiometer. The output of ICL8038 is stable over a wide range of temperature and supply variations. The total harmonic distortion for ICL8038 varies

from 1% to 2% depending on the model selected.

#### B. V To I Convertor

The current driver is one of the most important sub circuit for the measurements of bio-impedance. The current driver can easily work over a fairly wide range of impedance and frequency. The main requirements of a current driver are high output impedance, short phase delay and minimal harmonic distortion. Depending on whether they are open loop or closed loop, these are categorized into two groups. The features of each design are described [50].

The voltage waveform generated using 8038 is converted into current, which is injected into human body. The current generated is within the safe limits ( $600\mu\text{A}$ -  $800\mu\text{A}$ ). The V to I converter is intended to run in the frequency range of 10-100 kHz.

The main requirement for a current source is that it should supply a constant amount of current irrespective of the impedance of the load connected to it. When the output impedance of the current source is much greater than the load impedance then the current through the load is maintained constant regardless of the load value. F. Seoane et.al have analyzed Howland circuit as V to I converter [51]. In this instrument Howland circuit with buffer is used for voltage to current conversion. Fig. 3 is the circuit diagram for voltage to current converter. The buffer stage increases the output impedance of the V to I converter. The input impedance of the circuit is around 20 k $\Omega$ . Fig. 4 depicts the simulation results for circuit of Fig.-3. The resistance {R} in the Fig. 3 represents the load to which current would be injected. The human body impedance is in the range 1 k $\Omega$  - 3 k $\Omega$ . In simulation the resistance {R} is varied from 0.1k to 5k in steps of 100  $\Omega$ . It is verified using the simulation results that the current through the load is almost constant (10  $\mu\text{A}$  variation) irrespective of the load resistance.

#### C. Instrumentation Amplifier

The voltage sensed from the section of human body is in range of millivolts. The input signal is amplified using a difference amplifier. A single op-amp can also be used as a difference amplifier but due to its low input impedance it cannot be used. The input impedance of the single op-amp difference amplifier can be increased by introducing a buffer at each of the inputs of the amplifier and instead of using a unity gain follower one can also have some gain from the first

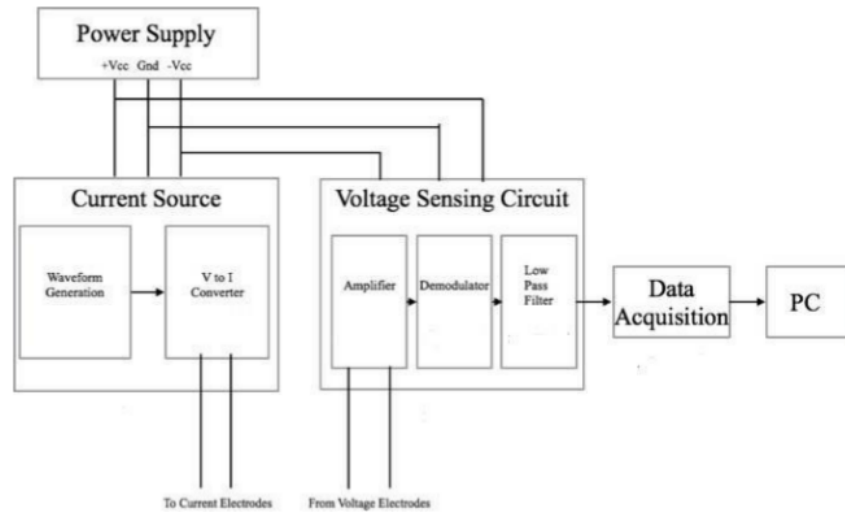


Fig. 2. General Block Diagram of Measuring Instrument.

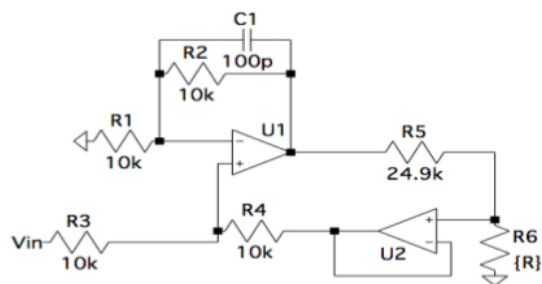


Fig. 3. Circuit Diagram for Voltage to Current Converter.

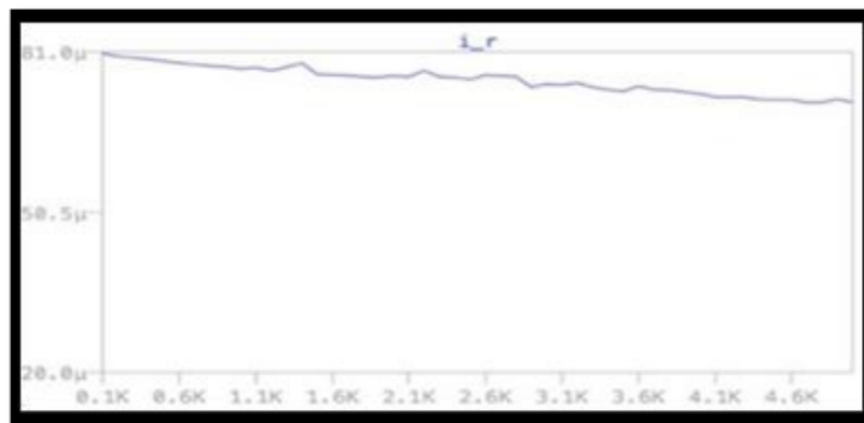


Fig. 4. Variation of Current Through {R} with Respect to Change in Value of Load Resistance, y-axis is the Current Through the Load and x-axis is the Impedance of the Load Connected to the V to I Converter.



stage. This leads to the circuit of instrumentation amplifier as shown in Fig. 5. The instrumentation amplifier has high CMRR and high input impedance and has been very effectively used in biomedical applications such as EEG and ECG.

The analysis for instrumentation amplifier has been given by Adel S. Sedra and Kenneth C. Smith [1] and its use as a biopotential amplifier has been analyzed by Nagel J.H [52]. The gain for the instrumentation amplifier is given below.

$$\text{gain} = \frac{R_4}{R_3} \left( 1 + \frac{R_2}{R_1} \right) \quad (1)$$

The signal from the electrodes is given as input to the instrumentation amplifier through its two terminals. The gain of the amplifier is 26.44 dB. The first two op-amps are used for gain and the third operational amplifier is used for common mode rejection. In the circuit designed  $R_4=R_3$  and  $R_2/R_1$  is equal to 20. The amplifier's input impedance is equal to the input impedance of the operational amplifier. Frequency response analysis of the instrumentation amplifier depicts that the phase and gain are constant in the desired frequency range. Fig. 6 depicts the frequency response of the instrumentation amplifier in which the continuous plot is the magnitude plot and dotted plot is the phase plot.

#### D. Demodulator

The signal obtained from the human body is an amplitude modulated signal, where the carrier is the current waveform that we have introduced into the human body. In this case the high frequency carrier is a sinusoidal waveform with a frequency of 50 kHz and the modulating bio-impedance signal is low frequency signal with frequency components less than 50 Hz. Webster and Tompkins [53] have suggested use of full wave rectifier for demodulation. The modulating signal can be extracted using a simple envelope detector circuit as shown in Fig. 7. Asynchronous demodulation has been used so that square and triangular waveform excitation can also be used in addition to the sinusoidal waveform.

For the first cycle diode is forward biased and it charges the capacitor to the first peak value. The charging time constant should be such that the capacitor voltage follows the input signal.

$$\tau \text{ charging} \ll \frac{1}{f_c} \quad (2)$$

where  $f_c$  is the frequency of the carrier. The charging times constant depends on source resistance  $R_s$ , forward bias diode resistance  $r_d$  and capacitance  $C_1$ .

$$\tau \text{ charging} = ((R_s + r_d + R_2) \parallel R_1) C_1 \quad (3)$$

$$R_s + r_d + R_2 \ll R_1 \quad (4)$$

$$\therefore \tau \text{ charging} = (R_s + r_d + R_2) C_1$$

When the input signal drops, diode becomes reverse bias (as capacitor is charged to a higher voltage) and the capacitor voltage remains at the initial level. During this time (when diode is reverse bias) the capacitor discharges through the resistor  $R_1$ .

$$\tau \text{ discharging} = R_1 C_1 \quad (5)$$

The discharging time constant should be large so that the capacitor discharges slowly but it should not be so large that it is unable to trace the low frequency modulating signal. The discharging time constant should follow the following relation:

$$\frac{1}{B} \gg \tau \text{ discharging} \gg \frac{1}{f_c} \quad (6)$$

Where B is the bandwidth of the low frequency bio-impedance signals. For the calculation of the values of resistances and capacitance, value of B is taken equal to 50 Hz. The diode used in the circuit is forward biased when the incoming voltage is greater than 0.7V, therefore the low voltage signals cannot be detected. Taking into account this problem related to cut off voltage of diode, a precision diode [46] is used in this circuit in place of normal diode. A precision diode is an operational amplifier with a diode in the negative feedback followed by a diode whose anode is connected at the output pin of the op-amp. The cutoff voltage for a precision diode is approximately equal to zero volts. Fig. 8 is the circuit for precision diode. The output of the demodulator is given to the low pass filter stage through a buffer. Improved output buffering and peak detector gain greater than unity is achieved with an output voltage follower. This leads to circuit of precision envelope detector in Fig. 9. The precision envelope detector is much more accurate than the simple envelope detector as the voltages below 0.7V are also detected by this circuit. Droop due to detector diode leakage can be removed through the use of bootstrapping feedback that holds detector diode bias at zero when the diode is not conducting.

#### E. Low Pass Filter

The signal from the demodulator contains high frequency components, which needs to be filtered out before the signal is given as input to the analog to digital converter of data acquisition. Here, an antialiasing filter is used as a Low Pass Filter (LPF). Accordingly, a second order Chebyshev LPF with a cutoff frequency of 40 Hz is used. Chebyshev filters differ from Butterworth filters in that they have a roll-off which is steeper and more ripple in passband (type I) or stopband (type II). Chebyshev filters have the property of minimizing the difference between real filter characteristics and the idealized one over the desired filter range, but with passband ripples. The LPF output is given to data acquisition system implemented using Arduino uno board.

#### F. Assumptions, Measurement Protocol and Data Acquisition

According to the BIA assumptions, the human body can be considered as a homogenous conductor having cylindrical dimensions (Fig. 10). In which bio-impedance is directly proportional to 'L' and inversely proportional to 'A', where 'L' is the cylinder's length and 'A' is the cylinder's base cross-sectional area (Fig. 10).

BIA formulation processes typically make the following assumptions for ease of calculation, though in practice the human body varies from these assumptions:

- The human body is considered as cylindrical.
- The cylindrical shape is defined by its height and weight.

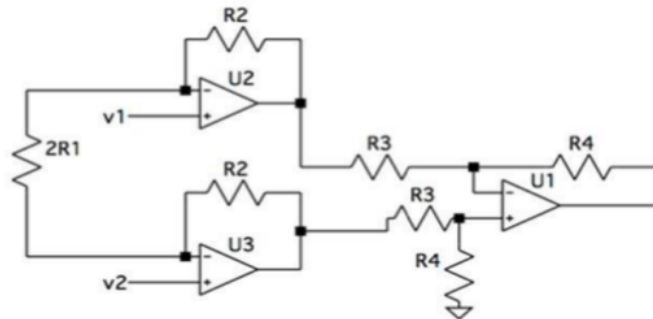


Fig. 5. Circuit Diagram for Instrumentation Amplifier.

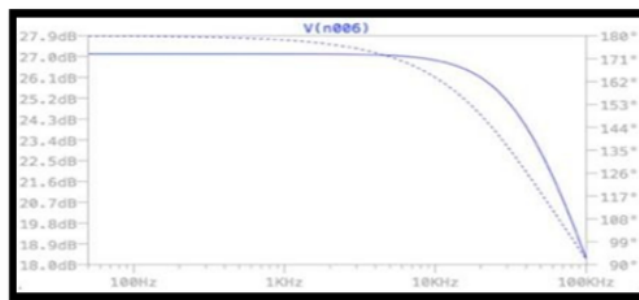


Fig. 6. Frequency Response Analysis of the Instrumentation Amplifier.

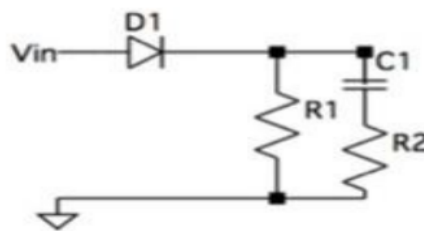


Fig. 7. Circuit Diagram of Simple Envelope Detector.

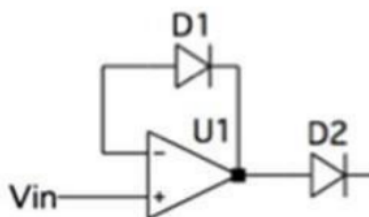


Fig. 8. Precision Detector.

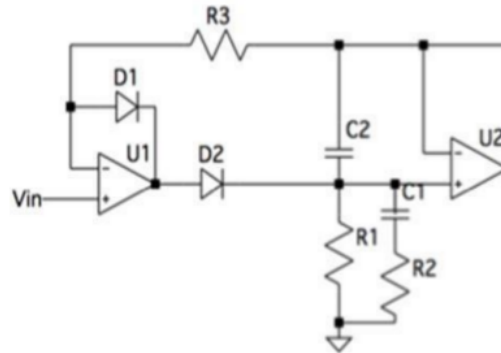


Fig. 9. Precision Envelope Detector.

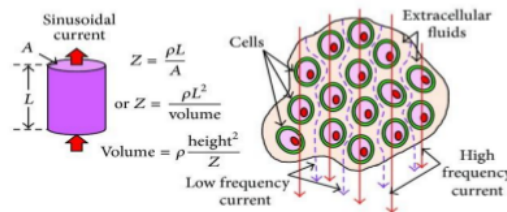


Fig. 10. The Human Body's Impedance when it is Modelled as a Homogeneous Conductor with Cylindrical Volume.

- Homogeneous and evenly distributed body composition is considered.
- The body compositions have no individual differences or variations.
- The environment parameters like temperature, other physiological parameters such as body heat or stress are assumed to be constant.

The cell membranes produce capacitive reactance due to the capacitive nature and caused by a selectively applied AC signal allows current to pass through these cell membranes using current paths largely depends upon the signal frequency (Fig. 10). The cell membranes are penetrated by current with high frequency. Thus, the current penetrates ECFs, CMs and ICFs. The low frequency current passes through ECF only as the cell membrane reactance prevents its flow through it. The BIA technique can be used to determine the (TBW) which is nothing but combination of ECW and ICW. However, it should be done at a particular frequency of the applied AC signal.

The instrument designed was tested on patients at St. Joseph hospital, Ghaziabad, India. Verbal prior approval for the study was taken telephonically from the concerned authorities. Twenty-five adult human subjects (aged 25-60 years) participated in this study. The general test setup for BIA is shown in Fig. 11. However, in our study it differed little bit for the ease of the subjects. They were made to rest in supine position for about ten to fifteen minutes. Then, Ag/AgCl electrodes were put their right and left forearms. Volunteers were told to lie down straight and remain still during measurement.

## V. RESULTS

The test setup used to validate the instrument is similar to the setup used in impedance plethysmography for the cardiac output measurement. In general, impedance is a measure of resistance. Plethysmography has become the gold standard for measuring changes in blood volume in any part of the body based on electrical impedance changes [54]. It is also being used in the diagnosis of peripheral vascular diseases. In impedance plethysmography, values of instantaneous impedance ( $Z$ ), basal impedance ( $Z_0$ ), which is an average of calculated bio-impedance values and derivative of impedance with respect to time ( $dZ/dt$ ) is used to calculate cardiac output parameters. Since the technique is a standard procedure, the values of basal impedance measured by this technique can be used to validate the instrument designed. The output of the demodulator is essentially the basal impedance of the section across which electrodes have been applied. The basal impedance values of forearm given in literature, [55] are used to validate the output signal of the instrument designed.

The output voltage signal of three sets of volunteers has been displayed in Fig. 12. The slight variations in the voltage values depict the variations in impedance ( $Z_0$ ). The basal impedance is calculated using the constant voltage level. The calculated value of basal impedance and its comparison with values of basal impedance in literature has been shown in Table I. The comparison in Table I shows that the values measured from the output of the instrument is in accordance with that of standard technique, which validates the design of the instrument.

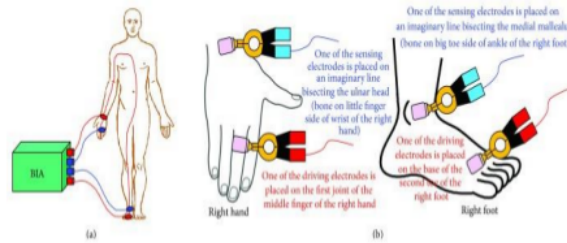


Fig. 11. General Test Setup for BIA.

TABLE I. BASAL IMPEDANCE (FOREARM)

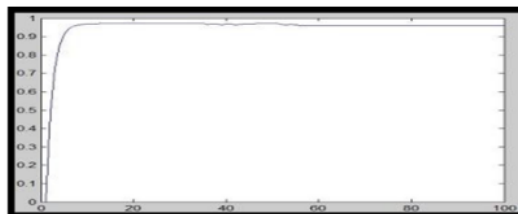
SET OF VOLUNTEERS [NOs]	MEASURED BY INSTRUMENT DESIGNED		STANDARD VALUES IN LITERATURE	
	$A_V$ . AGE	$A_V$ . $Z_0(\Omega)$	AGE GROUP	$A_V$ . $Z_0(\Omega)$
A [5]	40	58.16	36-45	$65.44 \pm 12.03$
B [8]	48	80.57	46-55	$67.50 \pm 6.38$
C [12]	57	58.77	> 55	$69.13 \pm 8.74$



(a)



(b)



(c)

Fig. 12. Output Voltage Signals Measured using the Instrument. (a) is the Output Signal of Volunteer Set-A (b) being the Output Signal of Volunteer Set-B and (c) is the Output Signal of Volunteer Set-C. The y-axis in Each Figure is the Output Voltage Amplitude Value and the x-axis Represents Number of Samples.

## VI. CONCLUSION

A simple low cost bio-impedance measuring instrument has been designed using common electronic blocks. The results have been validated and have been found to be accurate and reliable. Each block of the instrument has been tested using simulation and verified experimentally. The method and device developed fulfills required specifications and can be used in clinical examinations. Bio-impedance measurement techniques can prove very useful for first hand diagnosis. The major problem for using this technique for diagnosis is that its results are not standardized. By standardized one means that just by looking the graph or using its results one can say whether the results correspond to a normal person or a patient suffering from a disease. A large amount of analysis has to be done for a particular application on bio-impedance signals to enable the doctors to rely on results produced by bio-impedance analysis. The standardization of bio-impedance values for human body may be difficult due to variable ambient conditions, eating habits, lifestyle and anthropological background. So intensive and extensive research and in-depth analysis is required to be carried out towards normalization and standardization of the bio-impedance values.

Researchers in the field of bio-impedance have discovered that multifrequency bio-impedance analysis (BIA) is one of their main interests [56]. A group of researchers are exploring instrumentation and designing more advanced instrumentation. In the future, investigation and research can be carried out for Bluetooth-enabled wireless instrumentation for BIA techniques. Moreover, in future studies Bluetooth-enabled wireless instrumentation for Electrochemical Impedance Spectroscopy (EIS) techniques could be investigated. To improve cardiac health monitoring and to have potential transthoracic parameter assessment; ICG along with a bio-impedance tool with multi-frequency features can be used. Furthermore, in an ambulatory or long-term monitoring requirement in the Intensive Care Unit (ICU) bio-impedance based ICG can be used. However, if future research on Bio-impedance calculation and BIA can overcome these challenges, it can be implemented more appropriately than other commonly used and popular imaging methods in specific medical diagnostic applications especially related to the imaging of brain, breast, abdominal and whole body, etc.

Bio-impedance technology is becoming more prevalent as a result of an increasing number of healthcare monitoring applications. The study presented has a wide range of implications, including the ability to improve bio-impedance studies and healthcare devices that use bio-impedance technology [7]. Bio-impedance spectroscopy measures the bio-impedance of cells and tissues covering fairly good frequency range. Physiological testing and health-monitoring systems benefit greatly from this method. Devices must be compact, wearable, or even implantable for a wide variety of applications. As a result, the next generation of bio-impedance sensing systems must be designed to save energy and resources [34]. It's critical to consider the application as well as the type of cell or tissue culture to be monitored when selecting a bio-impedance measurement technique [57]. As a result, choosing the right electrode configuration is crucial. Future research should focus on the electrodes and the bio-impedance measurement method [57]. Overall, this paper describes a low-cost, bio-electrical impedance

measurement system that has been successfully developed and tested and can be used effectively and efficiently for non-invasive health monitoring. The paper also discussed some of the most important technological aspects and limitations of bio-impedance calculation and study. Finally, in this paper the theoretical dimensions, operating principles, implementations, benefits, disadvantages, and current research status, upcoming developments, and bottlenecks in bio-impedance analysis and calculation all have been covered in great detail.

## STATEMENTS & DECLARATIONS

### FUNDING

The authors hereby declare that no funds, grants, or other support in any form were received during the preparation of this manuscript.

### COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

### AUTHOR'S CONTRIBUTION

All authors contributed almost equally for material preparation, data collection and analysis to the said study related to conception and design of bio-impedance measuring instrument. Moreover, all authors read and approved the final manuscript.

### ETHICS APPROVAL

Being an observational study only, it did not require any formal ethical approval. However, verbal consent from the competent authority obtained telephonically.

### CONSENT TO PARTICIPATE

No such Informed consent was required for this observational study, however individual participants if any were informed verbally.

### CONSENT TO PUBLISH

As no figure in the manuscript depicts human participants, therefore no such consent for publication of the images of the human research participants is required.

## REFERENCES

- [1] F. H. Netter, "Atlas of Human Anatomy, Rittenhouse. 2nd. Book Distributors Inc," 1997, [Google Scholar].
- [2] J. J. Ackmann, "Complex bioelectric impedance measurement system for the frequency range from 5 Hz to 1 MHz," *Annals of Biomedical Engineering*, vol. 21, no. 2, pp. 135–146, 1993, [PubMed] [Google Scholar].
- [3] K. Cha, G. M. Chertow, J. Gonzalez, J. M. Lazarus, and D. W. Wilmore, "Multifrequency bioelectrical impedance estimates the distribution of body water," *Journal of Applied Physiology*, vol. 79, no. 4, pp. 1316–1319, 1995, [PubMed] [Google Scholar].
- [4] A. Roy *et al.*, "An experimental method of bioimpedance measurement and analysis for discriminating tissues of fruit or vegetable," *AIMS Biophysics*, vol. 7, no. 1, pp. 41–53, 2020.
- [5] S. Grimnes and G. Martinsen, "Academic Press, Bioimpedance & Bioelectricity Basics," 2021.

- [6] L. Jodal, "MSc, Medical Physicist, lecture notes on the electrical theory behind the measurement of body fluids with bioimpedance spectroscopy (BIS)," 2021.
- [7] B. Louarroudi and Sanchez, "On the correct use of stepped-sine excitations for the measurement of time-varying bioimpedance," *Physiological Measurement*, vol. 38, no. 2, pp. N73–N80, 2017.
- [8] J. W. Haslett and M. K. N. Rao, "A High Quality Controlled Current Source," in *Instrumentation and Measurement*, IEEE Transactions on, vol. 28, no. 2, pp. 132–140, 1979.
- [9] P. Annus, A. Krivoshei, M. Min, and T. Parve, "Excitation Current Source for Bioimpedance Measurement Applications: Analysis and Design," *Instrumentation and Measurement Technology Conference Proceedings*, pp. 848–863, 2008.
- [10] R. Pallas-Areny and J. G. Webster, "AC Instrumentation," 2021.
- [11] "Electrical Cardiometry" Wikipedia, 2021.
- [12] "Respiration rate measurement based on impedance pneumography Application report, SBAA1811-February 2011, Texas Instruments," 2021.
- [13] O. G. Martinsen, "Utilizing characteristic electrical properties of the epidermal skin layers to detect fake fingers in biometric fingerprint systems-A pilot study," *Biomedical Engineering*, vol. 54, pp. 891–894, 2007.
- [14] D. Naranjo-Hernández *et al.*, "Fundamentals, Recent Advances, and Future Challenges in Bioimpedance Devices for Healthcare Applications," *Journal of Sensors, Hindawi*, 2019, Article ID 9210258.
- [15] P. S. H. Jose *et al.*, "A Non-Invasive Method for Measurement of Blood Glucose using Bio Impedance Technique," in *2nd International Conference on Signal Processing and Communication (ICSPC)*, 2019, IEEE.
- [16] J. Nyboer, "Electrical impedance plethysmography; a physical and physiologic approach to peripheral vascular study," *Circulation*, vol. 2, no. 6, pp. 811–821, 1950, [PubMed] [Google Scholar].
- [17] R. W. Griffiths, M. E. Philpot, B. J. Chapman, and K. A. Munday, "Impedance cardiography: non-invasive cardiac output measurement after burn injury," *International Journal of Tissue Reactions*, vol. 3, no. 1, pp. 47–55, 1981, [PubMed] [Google Scholar].
- [18] C. J. Schuster and H. P. Schuster, "Application of impedance cardiography in critical care medicine," *Resuscitation*, vol. 11, no. 3-4, pp. 255–274, 1984, [PubMed] [Google Scholar].
- [19] "Internet Article, JAWON Medical, South Korea," 2021, <http://www.jawon.co.kr/eng/technology/body-composition/principles-of-bia.php>.
- [20] "Internet Article. Bodystat Limited, USA," 2021, <http://www.bodystat.com/science>.
- [21] D. Miklavcic, N. Pavselsj, and F. X. Hart, "Electric Properties of Tissues," *Wiley Encyclopedia of Biomedical Engineering*, 2006, [Google Scholar].
- [22] H. P. Schwan, "Electrical properties of tissue and cell suspensions: Mechanisms and models," *Proc IEEE Adv Biol Med Soc*, vol. 1, pp. 1:A70–A1, 2002, [Google Scholar].
- [23] R. Pethig and D. B. Kell, "The passive electrical properties of biological systems: Their significance in physiology, biophysics and biotechnology," *Phys Med Biol*, vol. 32, no. 933, 1987, [PubMed] [CrossRef] [Google Scholar].
- [24] U. G. Kyle, I. Bosaeus, D. L. AD, and A. D, "Bioelectrical impedance analysis part I: review of principles and methods," *Clinical Nutrition*, vol. 23, pp. 1226–43, 2004, [PubMed] [CrossRef] [Google Scholar].
- [25] S. Grimnes and Ø. G. Martinsen, "Bioimpedance & Bioelectricity Basics. Elsevier Science," 2014, 3rd ed. [Google Scholar].
- [26] K. Asami, "Characterization of heterogeneous systems by dielectric spectroscopy," *Prog Polym Sci*, vol. 27, pp. 1617–59, 2002, [CrossRef] [Google Scholar].
- [27] K. Heileman, J. Daoud, and M. Tabrizian, "Dielectric spectroscopy as a viable biosensing tool for cell and tissue characterization and analysis," *Biosens Bioelectron*, vol. 49, pp. 348–59, 2013, [PubMed] [CrossRef] [Google Scholar].
- [28] P. J. Riu and C. On, "Bioelectrical parameters of the whole human body obtained through bioelectrical impedance analysis," *Bioelectromagnetics*, vol. 25, pp. 69–71, 2004, [PubMed] [CrossRef] [Google Scholar].
- [29] C. Gabriel, S. Gabriel, and E. Corthout, "The dielectric properties of biological tissues: I. Literature survey," *Phys Med Biol*, vol. 41, pp. 2231–49, 1996, [PubMed] [CrossRef] [Google Scholar].
- [30] Ø. G. Martinsen, S. Grimnes, and H. P. Schwan, "Interface phenomena and dielectric properties of biological tissue," *Encyclopedia of Surface and Colloid Science*, vol. 20, pp. 2643–53, 2002, [Google Scholar].
- [31] D. A. Dean, T. Ramanathan, D. Machado, and R. Sundararajan, "Electrical Impedance Spectroscopy Study of Biological Tissues," *J Electrostat*, vol. 66, no. 3-4, pp. 165–77, 2008, [PMC free article] [PubMed] [CrossRef] [Google Scholar].
- [32] H. Kwon, A. L. McEwan, T. I. Oh, A. Farooq, E. J. Woo, and J. K. Seo, "A local region of interest imaging method for electrical impedance tomography with internal electrodes," *Comput Math Methods Med*, vol. 9, 2013, [PMC free article] [PubMed] [CrossRef] [Google Scholar].
- [33] J. K. Seo, T. K. Bera, H. Kwon, and R. J. Sadleir, "Effective Admittivity of Biological Tissues as a Coefficient of Elliptic PDE," *Computational and Mathematical Methods in Medicine*, vol. 2, 2013, [PMC free article] [PubMed] [CrossRef] [Google Scholar].
- [34] F. Soulier *et al.*, "Very Low Resource Digital Implementation of Bioimpedance Analysis," *Sensors (Basel), MDPI*, vol. 19, p. 3381, 2019.
- [35] J. Ojarand *et al.*, "How many frequencies to use in electrical bioimpedance measurements," in *Impedance Spectroscopy. Advanced Applications: Battery Research, Bioimpedance, System Design*. Publisher: Walter de Gruyter, 2018, pp. 161–168.
- [36] S. F. Khalil, M. S. Mohktar, and F. Ibrahim, "The Theory and Fundamentals of Bioimpedance Analysis in Clinical Status Monitoring and Diagnosis of Diseases," *A Review. Sensors*, vol. 14, 2014, [PMC free article] [PubMed] [Google Scholar].
- [37] R. Kusche *et al.*, "Dry electrodes for bioimpedance measurements - Design, characterization and comparison," *Biomedical Physics & Engineering Express, IOP*, vol. 5, no. 1, 2018.
- [38] H. Kalvøy, L. Frich, S. Grimnes, and Ø. G. Martinsen, "Impedance-based tissue discrimination for needle guidance," *Physiological Measurements*, vol. 30, 2009, [PubMed] [Google Scholar].
- [39] A. Yúfera and A. Rueda, "A Method for Bioimpedance Measure with Four- and Two-Electrode Sensor Systems," in *30th Annual International IEEE EMBS Conference Vancouver*. British Columbia, Canada, 2008, [PubMed] [CrossRef] [Google Scholar].
- [40] X. Huang, "Simulation of Microelectrode Impedance Changes Due to Cell Growth," *IEEE Sensors Journal*, vol. 4, pp. 576–83, 2004, [CrossRef] [Google Scholar].
- [41] I. Giaever, "Use of Electric Fields to Monitor the Dynamical Aspect of Cell Behavior in Tissue Cultures," *IEEE Transaction on Biomedical Engineering*, vol. 33, pp. 242–7, 1986, BME. [PubMed] [CrossRef] [Google Scholar].
- [42] T. S. Carvalho, A. L. Fonseca, A. Coutinho, B. Jotta, A. V. Pino, and M. N. Souza, "Comparison of bipolar and tetrapolar techniques in bioimpedance measurement. XXIV Congresso Brasileiro de Engenharia Biomédica – CBEB," 2014, [Google Scholar].
- [43] L. A. Geddes and R. Roeder, "Criteria for the selection of materials for implanted electrodes," *Annals of Biomedical Engineering*, vol. 31, no. 7, pp. 879–90, 2003, [PubMed] [CrossRef] [Google Scholar].
- [44] H. Kalvøy, G. K. Johnsen, Ø. G. Martinsen, and S. Grimnes, "New Method for Separation of Electrode Polarization Impedance from Measured Tissue Impedance," *The Open Biomedical Engineering Journal*, vol. 5, pp. 8–13, 2011, [PMC free article] [PubMed] [CrossRef] [Google Scholar].
- [45] J. Wang, W. Hu, T. Kao, C. Liu, and S. Lin, "Development of forearm impedance plethysmography for the minimally invasive monitoring of cardiac pumping function," *Journal of Biomedical Science and Engineering*, vol. 4, pp. 122–129, 2011, Inter instrument comparison of bioimpedance spectroscopic analyzers.
- [46] A. S. Sedra and K. C. Smith, "Microelectronic Circuits," 2021, 5th edition Oxford International Student addition.
- [47] "Bioelectrical impedance analysis in body composition measurement: National institutes of health technology assessment conference statement," 1994.
- [48] G. B. Clayton, "88 Practical Opamp Circus," 2021, 5th edition Oxford International Student addition.



- [49] "ICL 8038 datasheet," 2021.
- [50] N. Neshatvar *et al.*, "Analog Integrated Current Drivers for Bioimpedance Applications: A Review," *Sensors (Basel)*, vol. 19, no. 4, p. 756, 2019, Pub Med.
- [51] S. I. Kim and T. S. Suh, "Current Source for Wideband Electrical Bioimpedance Spectroscopy Based on a Single Operational Amplifier," in *World Congress of Medical Physics and Biomedical Engineering*. Springer-Verlag, 2006, 20006COEX Seoul, Korea. Berlin Heidelberg: Springer-Verlag., 2007.
- [52] "Biopotential Amplifiers in The Biomedical Engineering Handbook," 2000, by Joachim H. Nagel edited by Joseph D. Bronzino.
- [53] M. . Qu, Y. Zhang, J. G. Webster, and W. J. Tompkins, "Motion Artifact from Spot and Band Electrodes During Impedance Cardiography," *Biomedical Engineering*, no. 11, pp. 1029–1036, 1986.
- [54] "25 years of impedance plethysmography," 2003, BARC Newsletter No.236.
- [55] D. L. Prajapati, V. Chintan, Parmar, A. Pradnya, Gokhale, B. Hemant, C. J. S. Mehta, and Non, "Invasive Assessment of Blood Flow Index In Healthy Volunteers Using Impedance Plethysmography," *International Journal of Medical and Health Sciences*, 2013, ISSN 2277-4505.
- [56] B. T. Kanti, "Bioelectrical Impedance Methods for Non-invasive Health Monitoring: A Review," *J Med Eng*, p. 381251, 2014, PMCID: PMC4782691, PMID: 27006932.
- [57] M. Amini, J. Hisdal, and H. Kalvøy3, "Applications of Bioimpedance Measurement Techniques in Tissue Engineering," *J. Electrical Bioimpedance*, vol. 9, no. 1, pp. 142–158, 2018.

© 2022. This work is licensed under  
<https://creativecommons.org/licenses/by/4.0/> (the “License”). Notwithstanding  
the ProQuest Terms and Conditions, you may use this content in accordance  
with the terms of the License.

---

# DETECTING TROPICAL CYCLONE FROM THE BASIC OVERVIEW OF LIFE OF EXTREMELY SEVERE CYCLONIC STORM, TAUKTAE

---

**Monu Yadav**  
Department of Applied Mathematics  
Delhi Technological University  
New Delhi, India  
yadavm012@gmail.com

**Laxmi Narayan Das**  
Department of Applied Mathematics  
Delhi Technological University  
New Delhi, India  
lndas@dce.ac.in

## ABSTRACT

The natural energy in the form of sunlight hotness, create low air pressure in the atmospheric regions. In order to refill the vacated air mass the wind flows from the higher air pressure regions to lower air pressure regions. If the lower air pressure occurs in certain hot spots in the land or the coastal region the moisturized wind from the ocean or sea regions flows to the hot spots in the track of cyclone trajectories. This causes moisturized cloudy storms, severe rainfalls and wind whirls in the circulation paths. The extremely severe cyclonic storms intensified movements are tracked through the elliptic or circular whirl paths which is commonly known as cyclones eye. In this study we attempt to identify the various technical characteristics of the extremely severe cyclonic storms caused by tropical cyclones. Due to the tropical cyclone eye pattern's the storms intensity and path trajectory varies. We have compared the eye features of the tropical cyclone intensification with the life span of the extremely severe cyclonic storm TAUKTAE. The paper is organized by mentioning the data computation methods, Overview of the life of ESCS TAUKTAE, Intensity variation for regulate TAUKTAE according to Dvorak's technique, intensity estimation using ADT9.0 and comparison with SATCON and IMD best track, Description of the formation of TAUKTAE Eye, timeframes for eye scenes, sea surface temperature data comparison, and final thoughts.

*Keywords:* Cyclonic Eye; Intensity; Indian Ocean; Dvorak Method

## 1 Introduction

In addition to being violent weather manifestations, tropical cyclones (TCs) have been observed to cause devastation along the coastal regions. These conditions include storm surge, flooding, storm winds, torrential rains, thunderstorms, and lightning. In all parts of the world, TCs cause tremendous economic losses and deaths every year. Prediction of TC formation must be accurate to avoid/reduce such losses, so that preventative measures can be taken. Tropical weather is generally referred to as "Cyclone" when winds exceed or equal 34 knots (62 kmph); the term was coined by the World Meteorological Organization (WMO), 2009. The maximum sustained wind speed (MSW) is used to define tropical storms, despite the differences in names among regions. In the east Arabian Sea and adjoining Lakshadweep area, the Extremely Strong Cyclonic Storm (ESCS) "TAUKTAE" developed, and it crossed Saurashtra coast near latitude 20.8°N and longitude 71.1°E, close to northeast of Diu (about 20 km northeast of Diu) with a wind speed of 160-170 kmph gusting to 185 kmph. Through rapid intensification, the ESCS increased from 65 knots on the 16th of May to 100 knots on the 17th of May over the sea. It has affected marine life and marine activity. Over the northern Indian Ocean, it was the first cyclonic storm of the year 2021, after the Kandla cyclone in 1998, Tauktae is the most intense cyclone in the satellite era (1961-2021).

It is essential to detect and predict maximum intensity over the ocean for accurate early warning and early response to TCs, especially when looking at the basin's smaller size and the socioeconomic vulnerability of the region. Satellites are connected with computing device to detect and predict the intensity of the weather events. The numerical weather prediction models [22, 23] and dynamical-statistical models [24], and the operational forecasts [25, 26], continue to face challenges when it comes to predicting the intensity of TCs. It is also less successful at predicting rapid intensification [27].

Through the use of satellite data (INSAT-3D and Microwave images) by conducting this study, we seek to identify the different types of technical characteristics of ESCS. Due to the eye pattern of the TC, our primary objective is to determine how TC, TAUKTAE eye characteristics relate to the storm intensity. Also, different methods were used to calculate TC intensity and compare them to the best track results. In the Arabian sea, only a few studies have investigated the relationship between certain eye features and TC intensification. In the Dvorak method [4, 5], satellite imagery is used to determine the intensity of the TC based on eye's position and associated characteristics. There is no standard measure to estimate how well-formed or ragged the TC eye should be; in general, the sharper the TC eye, the greater the intensity the TC.

TC eyes have gained increasing attention for their geometric characteristics. TC eyes are usually distinguishable from each other via satellite and radar observations (e.g. circular and elliptical; [15, 27, 28, 29, 30]). Additionally, the rate of eyewall contraction is lower in large-eyed TCs when their intensification is preceded by a reduction in the eye area [15].

The geometric characteristics of the eye, described above, are important components of the inner core of TCs, respectively, and can be examined using INSAT-3D and PMW images along with the intensity of TC, TAUKTAE.

Section 2 provides an overview of the data and method used to analyze the Cyclone, as well as a brief description of the Dvorak Method. Section 3 presents the results and the discussions. Section 4 presents the Final thoughts.

## 2 Data And Methods

Information assortment for this study is basically gotten from Indian Geostationary Satellite Images and PMW Images. Following of the "TAUKTAE", ESCS is done in light of the hourly INSAT-3D and INSAT-3DR satellite information, along with Synoptic hourly information from the Bulletin and fundamental report of the "TAUKTAE", ESCS given by the RSMC, New Delhi [1], Microwave images from the NRL TC website [2] and Satellite Bulletins [3] are utilized to concentrate on the different attributes of this ESCS, "TAUKTAE".

The procedure of the review included investigation of Satellite pictures utilizing Meteorological Image Analysis Software (MS) and Dvorak's Technique [4, 5]. The MS has been utilized to concentrate on the different eye characteristics by estimating the eye temperature, eye measurement, and distance of semi-major and minor axis to assess the Eye Roundness Value (ERV) of the ESCS, TAUKTAE. The MS is a representation and investigation apparatus created at Space Application Center. It permits clients to do a top to bottom investigation of the various climate frameworks from the information of INSAT-3D and INSAT-3DR.

Based on Dvorak's intensity scale T number, tropical disturbances will undergo different phases are given in Table 1

Phases	T Number	Maximum Wind Speed	Source
low pressure area (WML)	T1.0	< 17 kts	[5]
Depression (D)	T1.5	17-27 kts	[5]
Deep Depression (DD)	T2.0	28-33 kts	[5]
Cyclonic Storm (CS)	T2.5-T3.0	34-47 kts	[5]
Severe CS (SCS)	T3.5	48-63 kts	[5]
Very Severe CS (VSCS)	T4.0-T4.5	64-89 kts	[5]
Extremely Severe CS (ESCS)	T5.0-T6.0	90-119 kts	[5]
Super CS	T6.5 -T8.0	> 120 kts	[5]

Table 1: Different phases of tropical disturbances

According to [4, 5], the intensity of ESCS, TAUKTAE using INSAT-3D/3DR visible and IR images is analyzed based on T number on a three-hourly basis. Furthermore, we analyzes the T number estimated by [4, 5] for IR based on microwave images to see how the T number estimation could be improved since the location (Centre) of TC could be better determined by the microwave images.

ESCS, TAUKTAE intensity was also estimated with Advanced Dvorak Technique(ADT) [16]. During rapid intensification & weakening, ADT (9.0) accurately represents the situation. We analyzed and compared the ADT and Manual

T Number results with that of the best track intensity of the RSMC, New Delhi. CIMSS (Cooperative Institute of Meteorological Satellite System) [17], were used to collect the ADT values during ESCS, TAUKTAE.

ESCS, TAUKTAE intensity was also estimated with SATCON [31] which is an approach for estimating intensity of the cyclone. We analyzed and compared the SATCON result with that of the best track intensity of ESCS, TAUKTAE study by the RSMC, New Delhi.

As the eye can be seen in satellite imagery, Dvorak's method [4] of estimating TC intensity can effectively be applied. Our discussion focuses on studying a variety of aspects of the "Eye" using INSAT-3D & PMW images, and how they relate to intensifying and weakening the ESCS.

Eye size are essential structural characteristics of the TC, as described in previous research, hence we examine their effect on ESCS, TAUKTAE. The increased heating in the eye is a result of strengthened eyewall convection as a TC intensifies [18, 19, 20] causing a decrease in central pressure [21]. The resulting contraction of the pressure gradient around the maximum wind creates a force that rises at and inward from the maximum wind's radius, causing the eye and eyewall to contract [21]. In this study, the size, shape, and roundness of eyes are examined. [15] report that the eye roundness value (ERV) can be calculated as follows:

$$ERV = \sqrt{(a^2 - b^2)/a^2}; \quad a, b \in R^+,$$

where a and b represent the semimajor and semiminor axes of the fitted ellipse, respectively in eye's frequently occur in the trajectory of the cyclonic storm.

Our study utilized INSAT-3D and PMW images to examine ERVs instead of radar images, which is usually the method used to trace ERVs. The accuracy of ERV determination from satellite imagery has improved recently as a result of higher resolution images and DIGI (Digital Graphic User Interface) platforms. This study utilized PMW images that have spatial resolutions of 12.5-15 km. INSAT-3D has a visible resolution of 1 km and IR resolution of 4 km, whereas DWR has a resolution of 300 m.

### 3 Results and Discussion

#### 3.1 Overview of the life of ESCS, TAUKTAE

The ESCS, TAUKTAE developed from the remnants of a Low Level circulation over the area that occurred on 13th May over the Arabian Sea, which is marked as the D at 1200 UTC (RSMC, New Delhi, 2021). It intensified into a DD at 1200 UTC on 14th May. It moved north-eastward with increase intensity and marked as CS at 1800 UTC, 14th May and was named as "TAUKTAE" later. In May 15, Tauktae intensified further, the cyclonic storm reached VSCS status later in the day. In the following morning, it became a VSCS, as it moved almost northward. The rapid intensification was observed during 16th of May morning. Early 17th of May, it intensified further into an ESCS and soon reached its peak intensity. During its progress toward the Gujarat coast, Tauktae weakened and then restrengthened before making landfall. Following landfall, as the storm moved northeastward, Tauktae gradually weakened as it moved inland. On 19th of May, Tauktae weakened into a WML over the north east.

A cyclone Tauktae intensity was the most intense during the satellite era (1961-2021) after the storm of Kandla in 1998. During its landfall, Cyclone Tauktae was of similar intensity to the Kandla cyclone of June, 1998, with sustained winds of about 160-170 kmph gusting up to 185 kmph. Although Tauktae had a higher life time maximum intensity, with gusts reaching 210 kmph during early morning to afternoon of 17th May 2021, it experienced 180-190 gusts during early morning to midafternoon.

There was adverse weather and damage caused by this rare cyclone that moved parallel to the west coast and crossed Gujarat as it moved across the nation's west coast states.

Tauktae covered a distance of approx 1880 km during its lifetime. The figure 1, one can see the observed tracking of ESCS, "TAUKTAE"

#### 3.2 Intensity variations for TAUKTAE according to Dvorak's technique

In this study, we used the visible/IR images from geostationary satellites (INSAT-3D & 3DR) as well as [5] to determine and analyze the intensity of the ESCS, TAUKTAE. As soon as TAUKTAE was conceived, the very first pattern was the "Shear Pattern". As it developed, it shifted to "Eye pattern" and remained associated until landfall.

It was shown by ESCS, "TAUKTAE" at 1300 UTC 15th of May that there was an 'Eye Pattern' with T3.5 in line with the results of the survey [9]. CDO Pattern forms before the formation of Eye, reports the TC, TAUKTAE. It also

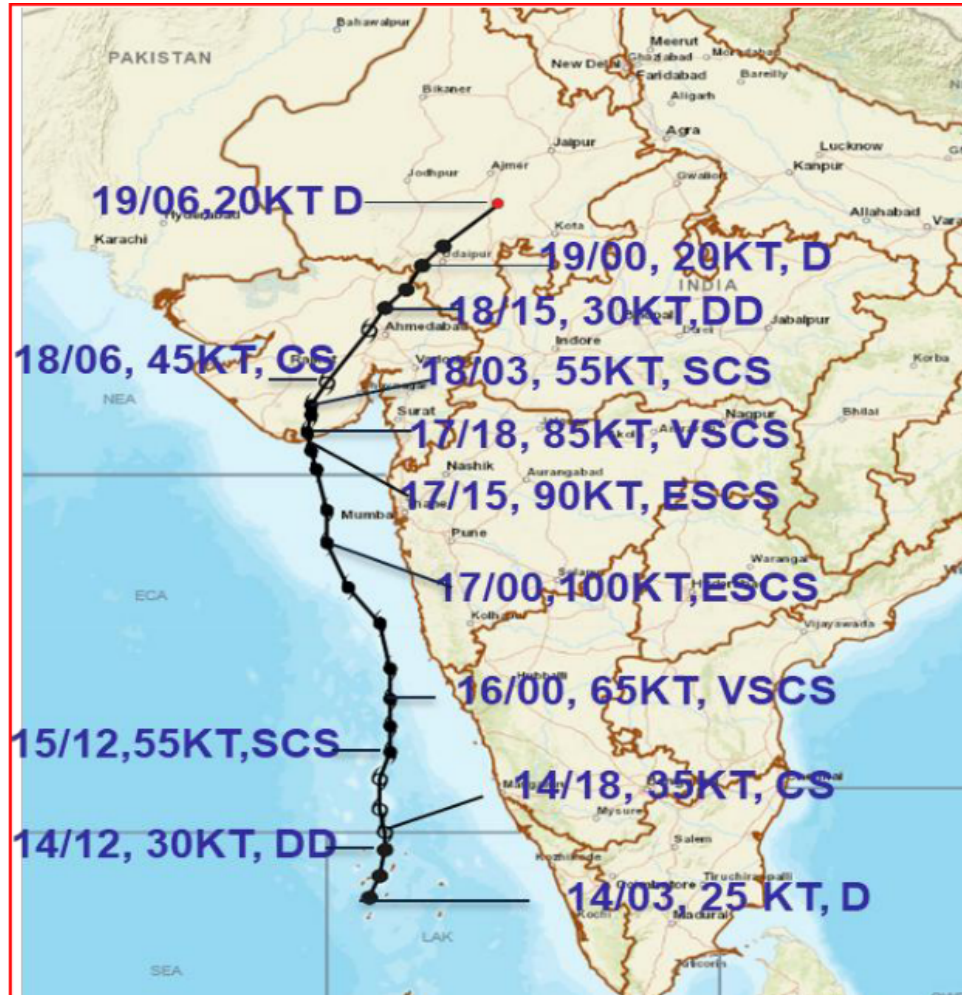


Figure 1: Observed track of ESCS, "TAUKTAE"

confirms the earlier finding by Dvorak Model development, that TCs which form eye do so within 48 hours of attaining tropical strength. The form of the eye was observed in the visible imagery of INSAT-3D on the 16th of May at 0600 UTC. In ESCS TAUKTAE, the eye grew within approximately 33h of reaching tropical storm strength on 14th May. The eye's temperature was -13 degrees C at 0900 UTC on 16th May, helping to increase its strength further to T4.5. In the following days, with the eye sharpening and further warming, the intensity increased to 'T5.0' at 2100 UTC on 16th of May and marked to 'ESCS' category. Based on [5], T5.5 was estimated as Final T.No.

By using the microwave images to determine the center and eye numbers, the intensity was computed and it affirmed the intensity computed by apply the Dvorak technique [5] to IR images. Based on microwave images obtained from NAVY NRL and CIRA web pages, eye characteristics and the Center of TC were determined to be extremely clear. The intensity of TC as calculated according to Dvorak technique ([4, 5]) was based on the accurate center determined through microwave images during the life of ESCS, TAUKTAE.

### 3.3 Based on ADT9.0, the intensity of TC, TAUKTAE

The ADT(9.0) intensity is overestimated up to T3.0 (figure 2). In comparison to IMD provides data and the Manual Dvorak T. No. the difference is about T1.0, which supports the earlier results by [10]. The ADT (Version 8.2.1) has been shown to overestimate up to T3.0 or T2.5 by them. As a result of changing the original 'Curve Band' pattern to an 'Eye' pattern, ADT captures the Best Track Intensity quite well. ADT intensity estimates were slightly underestimated in the last stage of weakening as compared to the IMD best track by approximately 0.5 T.Number. ADT's MSW was overestimated in the preliminary development phase and underestimated in the weakening phase when compared with

the best track MSW. During the initial developing phase, the Best track overestimates by about 10-30 knots. Similarly, during the weakening phase, the best track underestimates by about 05-30 knots.

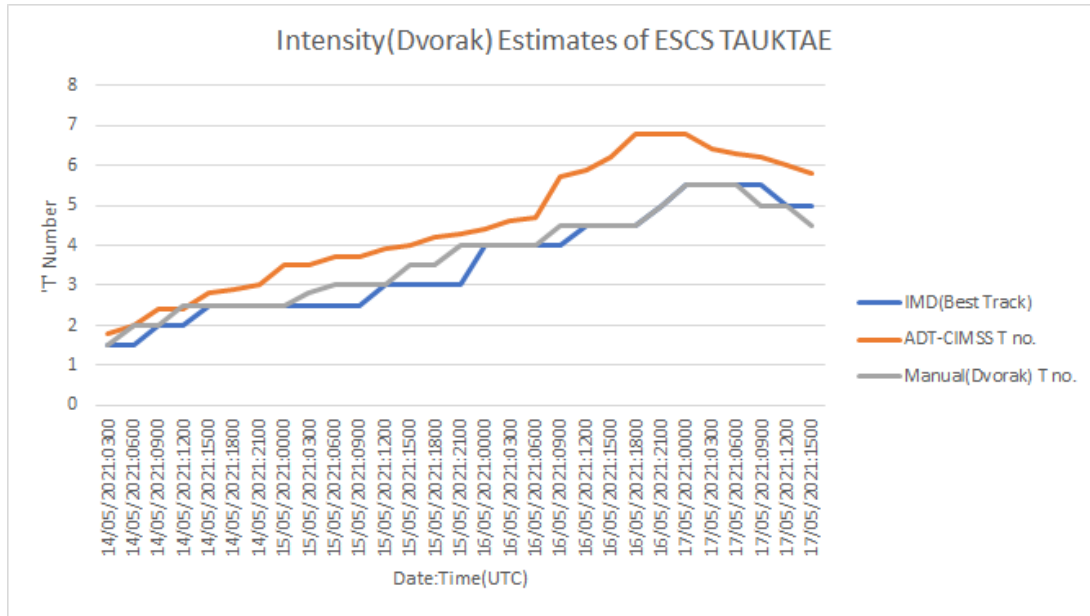


Figure 2: IMD Best Track and ADT (CIMSS) T numbers, and Manual (Dvorak) T numbers estimate intensity of ESCS-TAUKTAE

### 3.4 On the basis of SATCON, the intensity of TC, TAUKTAE

Fig. 3 illustrates the overestimation of SATCON intensity. Based on the comparison with IMD best track, it has been shown the TC, TAUKTAE was overestimated at both the beginning and the end of its lifecycle. Overestimation usually occurs during the initial stage (5-15 knots), peaking at 30 knots and weakening at 10 knots approximately.

### 3.5 Description of ESCS, TAUKTAE's eyes

#### 3.5.1 Description of the formation of an 'eye'

First appearance of the eye in a microwave image of MSG-1 SEVIRI at the 1300 UTC of the 15th May, 2021, i.e., 17 hours ahead of the eye showed up in the INSAT-3D image. This is reliable with the consequences of past exploration [6, 7]. TC's life expectancy or when the eye is obscured by cirrus was presumed to be recognized by microwave image scans. The Estimate Central Pressure(ECP) was recorded 990 hPa at 0900 UTC 15th of May resulting with MSW of 45 knots and a pressure drop of 10 hPa. Moreover, during the ECP over the Atlantic Ocean, between 987 and 997 hpa are the most common times for new eyes to form, which is in agreement with finding of [6]. In addition, the TAUKTAE's eye results are also consistent with [8] findings that the bottom of the South China Sea encounters 950 to 990 hpa as the minimum central pressure. When it first appeared in the microwave image MSG-1 SEVIRI of 15 May 1300 UTC, ESCS, TAUKTAE had an eye temperature of -12.8 degrees C.

#### 3.5.2 How the eyes are structured

Based on the INSAT-3D satellite and the Microwave satellite (PMW) images, an analysis of the eye structure was performed. The first appearance of the eye within the INSAT-3D satellite occurred at 0600 UTC on 16 May and the first appearance within Microwave on 15 May at 1300 UTC. As of 2100 UTC on 17th of May, INSAT-3D had disorganised the eye, while on 18th of May, the eye had disorganized in Microwave Satellite at 0100 UTC. Table 2 provides the structural analysis of INSAT-3D ; Table 3 provides the structural analysis of microwave stallite.



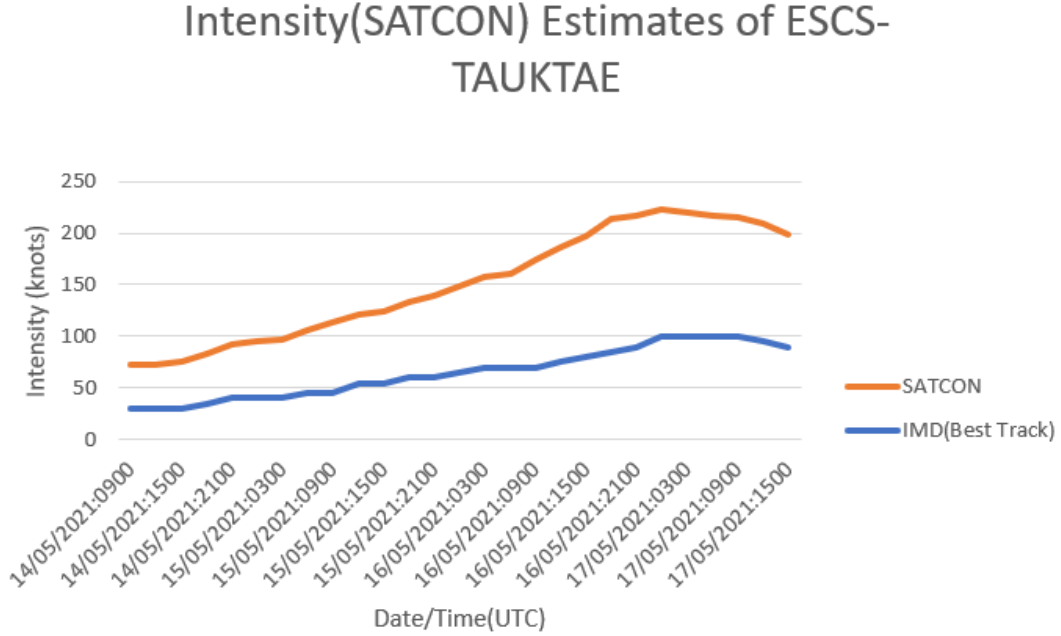


Figure 3: Intensity estimate by SATCON and IMD of ESCS-TAUKTAE

Date/Time(UTC)	Category	Eye Temp. (degree C)	Eye Structure
16/0600	VSCS	–	First Eye Pattern Seen
16/0900	VSCS	- 12.8	Become Ragged
16/1200	VSCS	- 37.0	Ragged Eye & warm
17/0300	ESCS	- 16.0	Ragged Eye
17/0600	ESCS	- 16.0	No Significant change
17/0900	ESCS	- 46.0	Warmest Ragged Eye
17/1500	ESCS	- 9.0	Large Ragged Eye
17/1800	VSCS	- 13.0	Decrease in size
17/2100	VSCS	–	Eye is Disorganised

Table 2: INSAT-3D satellite images used to analyze the structure of the TC eye.

### 3.5.3 Timeframe for ‘eye scenes’

In every three hourly INSAT-3D/3DR images and in the available PMW images, frequent eye scenes were observed over ESCS, TAUKTAE after wind speeds reached 60 knots, as can be seen from our analysis. According to [9], the frequency of eye scenes increases sharply between 50 and 70 knots and after 100 knots the frequency of eye scenes increases again. This pattern is also supported by the findings of [5]. During the weakening phase of TC, TAUKTAE, the MSW decreased to 85 knots from 1800 UTC of 17th of May, resulting in fewer eye scenes. From 16th to 17th of May (upto landfall), the ESCS displayed 13 three-hourly eye scenes for approximately 39 hours. There are far more eye scenes compared to the normal [9]. Over the course of cyclonic eye in tropical storm’s lifetime, [9] shows that a typical eye remains in contact with an eye for 30 hours, but there are many storms with more eye scenes in their lifetimes, including the SuCS, AMPHAN storm, which has 57-h eye scene, and the Hurricane ‘IOKE’ which has 98 eye scenes.

### 3.5.4 In the Eye region, SSTs are occurring

The Sea Surface Temperature(SST) over the Indian Ocean were examined taking into account the SSTs and their relationship with eye and RI. A SST was recorded in the Indian ocean during the ESCS, TAUKTAE, which reached 30 degrees Celsius [11]. Observations have found that the SSTs near the ESCS circulation are higher than 30.0 degrees Celsius. In the Atlantic Ocean, the average eye SST is 29.2 degrees Celsius, according to [6]. In eye scenes over

Date/Time(UTC)	Category	Eye Structure
15/1300	SCS	First Eye Pattern Seen
16/0800	ESCS	Ragged Eye
16/1030	ESCS	Ragged Eye with reduced size
16/1300	ESCS	Size increase
17/0800	ESCS	Circular size
17/1100	ESCS	No Significant Change
17/1300	ESCS	Circular and size increase
18/0100	ESCS	Eye is Disorganised

Table 3: Microwave satellite (PWM) images used to analyze the structure of the TC eye.

the South China Sea, [8] has shown that most SSTs exceed 29.0 degrees C near the storm circulation. In conclusion, [12, 13, 14] consistent findings suggest that the high SSTs of the Indian Ocean near TC (greater than 30 degrees C) are likely to have contributed to the RI of TC, TAUKTAE, in addition to other factors such as low wind shear. In their research, the warmth of the SST is shown to be a primary contributor to the more intense TC is fundamentally due to higher sensible and latent heat fluxes.

### 3.5.5 A characteristic of the eye is its roundness

Following [15], it was calculated and analyzed the roundness of the eye based on its ERV value. ERVs with lower TCs correspond to eyes that are rounder. The ERV of the eye usually increases with the maximum diameter, meaning that the smaller the diameter, the rounder the eye becomes [8]. An ERV of minimum 0.34 was determined at 2100 UTC on 16 th of May when TC intensified to ESCS with semi-major and semi-minor axes of 0.08 degree and 0.085 degree respectively. [15] reached a similar conclusion about ERV minimum values. According to his analysis, hurricanes with Category 5 can be identified by their higher circular eyes ( $ERV \approx 0.36$ ). Saffir-Simpson Scale classifies TC, TAUKTAE as Category 4. A mean ERV of 0.57 was found for TC, TAUKTAE, with dominant values ranging from 0.5 to 0.7. In the Atlantic Basin, [15] found that predominant values ranged from 0.5 to 0.7 and the mean ERV was 0.57. [8] agree that 0.5-0.7 ERV is the most common ERV observed by TCs over the South China Sea. TAUKTAE also shows similar results.

## 4 Final Thoughts

The results and discussions lead to the following general conclusions. A number of structural changes have occurred in the eye concerning its growth and evolution, its geometric characteristics as well as intensity of TC, TAUKTAE.

When ESCS, TAUKTAE reached tropical storm intensity, an eye had developed within 39 hours, unlike the traditional 48-hour window during which the eye develops. Before the development of the eye, TAUKTAE exhibited CDO pattern at the stage of T4.0.

A microwave image first showed the eye of ESCS, TAUKTAE, approximately 17 hours before it appeared on INSAT-3D imagery. By monitoring microwave images in real-time, we could predict based on life span data of the TAUKTAE intensification.

As with the Atlantic TCs and the South China TCs, the intensity was 45 knots with ECP of 990 hPa during the development of the eye. In contrast to the average of 30 hours of eye scenes over the Atlantic, ESCS, TAUKTAE eye scenes lasted for 39 hours.

The variability of TCs over the north Atlantic and north Pacific oceans is in accordance with ERV observations, the majority of ERVs ESCS, TAUKTAE range from 0.5 to 0.7, and the mean was 0.57 which is usually occur in the Indian Ocean.

## Declarations

### Conflict of Interest

Conflict of interest is not declared by any of the authors.

## Author Contribution

Research and manuscript preparation were done equally by each author.

## References

- [1] New Delhi RSMC. <https://rsmcnewdelhi.imd.gov.in/>.
- [2] NRL. <https://www.nrlmry.navy.mil/TC.html>.
- [3] SatelliteBulletins. <http://https://mausam.imd.gov.in/imdlatest/contents>.
- [4] Vernon F. Dvorak. Tropical cyclone intensity analysis and forecasting from satellite imagery. *Monthly Weather Review*, 103(5):420 – 430, 1975.
- [5] Vernon F. Dvorak. Tropical cyclone intensity analysis using satellite data, volume 11. US Department of Commerce, National Oceanic and Atmospheric Administration . . . , 1984.
- [6] Vigh J.L., Knaff J.A., and Schubert W.H. A climatology of hurricane eye formation. *Monthly Weather Review*, 140(5):1405 – 1426, 2012.
- [7] Velden, C., Harper, B., Wells, F., Beven, J. L., Zehr, R., Olander, T., ... & McCrone, P. (2006). The Dvorak tropical cyclone intensity estimation technique: A satellite-based method that has endured for over 30 years. *Bulletin of the American Meteorological Society*, 87(9), 1195-1210.
- [8] Xiaotong Zhu, Qingqing Li, Jinhua Yu, Dan Wu, and Kai Yao. Geometric characteristics of tropical cyclone eyes before landfall in south china based on ground-based radar observations. *Advances in Atmospheric Sciences*, 35(5):592–603, 2018.
- [9] Kenneth R. Knapp, Christopher S. Velden, and Anthony J. Wimmers. A global climatology of tropical cyclone eyes. *Monthly Weather Review*, 146(7):2089 – 2101, 2018.
- [10] Goyal S., Mohapatra M., Kumari P., Dube S.K., and Rajendra K. Validation of advanced dvorak technique (adt) over north indian ocean. *Mausam*, 68(4):689 – 698, 2017.
- [11] Eumetsat. <https://www.eumetsat.int/tropical-cyclone-tauktae>.
- [12] Miller B.I. On the maximum intensity of hurricanes. *J. Meteor.*, 15:184 – 195, 1958.
- [13] Demaria M. and Kaplan J. A statistical hurricane intensity prediction scheme (ships) for the atlantic basin. *Weather Forecasting*, 9(2):209 – 220, 1994.
- [14] Holland G.J. The maximum potential intensity of tropical cyclones. *Journal of the Atmospheric Sciences*, 54(21):2519 – 2541, 1997.
- [15] C.E. Barnes and G.M. Barnes. Eye and eyewall traits as determined with the noaa wp-3d lower-fuselage radar. *Monthly Weather Review*, 142(9):3393–3417, 2014.
- [16] Timothy L. Olander and Christopher S. Velden. The advanced dvorak technique (adt) for estimating tropical cyclone intensity: Update and new capabilities. *Weather and Forecasting*, 34(4):905 – 922, 2019.
- [17] CIMSS. <http://tropic.ssec.wisc.edu/>.
- [18] Zhang D.-L., Liu Y., and Yau M.K. A multiscale numerical study of hurricane andrew (1992). part v: Inner-core thermodynamics. *Monthly Weather Review*, 130(11):2745 – 2763, 2002.
- [19] Stern D.P. and Zhang F. How does the eye warm? part i: A potential temperature budget analysis of an idealized tropical cyclone. *Journal of the Atmospheric Sciences*, 70(1):73 – 90, 2013.
- [20] Stern D.P. and Zhang F. How does the eye warm? part ii: Sensitivity to vertical wind shear and a trajectory analysis. *Journal of the Atmospheric Sciences*, 70(7):1849 – 1873, 2013.
- [21] Willoughby H.E. Tropical cyclone eye thermodynamics. *Monthly Weather Review*, 126(12):3053 – 3067, 1998.
- [22] Krishna K. Osuri, U. C. Mohanty, A. Routray, and M. Mohapatra. The impact of satellite-derived wind data assimilation on track, intensity and structure of tropical cyclones over the north indian ocean. *International Journal of Remote Sensing*, 33(5):1627–1652, 2012.
- [23] U. C. Mohanty, Krishna K. Osuri, A. Routray, M. Mohapatra, and Sujata Pattanayak. Simulation of bay of bengal tropical cyclones with wrf model: Impact of initial and boundary conditions. *Marine Geodesy*, 33(4):294–314, 2010.
- [24] Mark DeMaria, Michelle Mainelli, Lynn K. Shay, John A. Knaff, and John Kaplan. Further improvements to the statistical hurricane intensity prediction scheme (ships). *Weather and Forecasting*, 20(4):531 – 543, 2005.

- [25] M MOHAPATRA, D P NAYAK, R P SHARMA, and B K BANDYOPADHYAY. Evaluation of official tropical cyclone track forecast over north indian ocean issued by india meteorological department. *Journal of Earth System Science*, 122(3):589–601, 2013.
- [26] M. Mohapatra, B. K. Bandyopadhyay, and D. P. Nayak. Evaluation of operational tropical cyclone intensity forecasts over north indian ocean issued by india meteorological department. *Natural Hazards*, 68(2):433–451, 2013.
- [27] Mrutyunjay Mohapatra and Monica Sharma. Cyclone warning services in india during recent years: A review. *Mausam*, 70(4):635–666, 2019.
- [28] Mitsuta, Y., & Yoshizumi, S. (1973). Periodic variations of pressure, wind and rainfall observed at Miyakojima during the second Miyakojima typhoon. *Journal of the Meteorological Society of Japan. Ser. II*, 51(6), 475-485.
- [29] Kuo H.-C., Williams R.T., and Chen J.-H. A possible mechanism for the eye rotation of typhoon herb. *Journal of the Atmospheric Sciences*, 56(11):1659 – 1673, 1999.
- [30] Oda M., Itano T., Naito G., Nakanishi M., and Tomine K. Destabilization of the symmetric vortex and formation of the elliptical eye of typhoon herb. *Journal of the Atmospheric Sciences*, 62(8):2965 – 2976, 2005.
- [31] Christopher S. Velden and Derrick Herndon. A consensus approach for estimating tropical cyclone intensity from meteorological satellites: Satcon. *Weather and Forecasting*, 35(4):1645 – 1662, 2020.

# Four-mode squeezed states in de Sitter space: A study with two field interacting quantum system

Sayantan Choudhury <sup>1</sup>, Sudhakar Panda <sup>2,3</sup>, Nilesh Pandey <sup>4</sup> and Abhishek Roy <sup>5</sup>

<sup>1</sup>*International Centre for Theoretical Sciences, Tata Institute of Fundamental Research (ICTS-TIFR), Shivakote, Bengaluru 560089, India,*

<sup>2</sup>*National Institute of Science Education and Research, Jatni, Bhubaneswar, Odisha - 752050, India,*

<sup>3</sup>*Homi Bhabha National Institute, Training School Complex, Anushakti Nagar, Mumbai - 400085, India,*

<sup>4</sup>*Department of Applied Physics, Delhi Technological University, Delhi-110042, India, and*

<sup>5</sup>*Department of Physics, Indian Institute of Technology Jodhpur, Karwar, Jodhpur - 342037, India.*

In this paper we study the application of four-mode squeezed states in the cosmological context, studying two weakly coupled scalar fields in the planar patch of the de Sitter space. We construct the four-mode squeezed state formalism and connect this concept with the Hamiltonian of the two coupled inverted harmonic oscillators having a time-dependent effective frequency in the planar patch of the de Sitter space. Further, the corresponding evolution operator for the quantum Euclidean vacuum state has been constructed, which captures its dynamics. Using the Heisenberg picture coupled differential equations describing the time evolution for all squeezing parameters (amplitude, phase and angle) have been obtained, for the weakly coupled two scalar field model. With the help of these evolutions for the coupled system, we simulate the dynamics of the squeezing parameters in terms of conformal time. From our analysis, we observe interesting dynamics, which helps us to explore various underlying physical implications of the weakly coupled two scalar field system in the planar patch of the de Sitter cosmological background.

## I. Introduction

The squeezed states in quantum optics have emerged as non-classical states of light which are a consequence of Heisenberg's uncertainty relations. When the uncertainty in one conjugate variable is below the symmetric limit, compared to the other conjugate variable, without affecting the Heisenberg's uncertainty relations, then the state obtained is called squeezed state. For review on the fundamentals of squeezed states see [1–8]. These states have been used in the field of quantum optics [9–13] for many experimental purposes. From an application perspective the squeezed states of light are being used for various applications in quantum computing [14, 15] and quantum cryptography [16] because these non-classical states of light can perform as an elementary resource in quantum information processing for the continuous variable systems. These states are also now being used in gravitational wave physics [17–19] to enhance the sensitivity of gravitational wave detectors [20–23], such as the LIGO [24] and to improve measurement techniques in quantum metrology [25–31]. For a more broad class of application of squeezed states see [32–44].

From the cosmological point of view, the use of the formalism of squeezed states was introduced in one of the early works by Grishchuk and Sidorov [45, 46] on the inflationary cosmology where they analysed the features of relic gravitons and phenomena such as particle creation and black-hole evaporation using the two-mode squeezed state formalism. They showed that the amplification of quantum fluctuations into macroscopic perturbations which occurs during cosmic inflation is a process of quantum squeezing and in the cosmological scenario

they describe primordial density perturbations, amplified by gravitational instability from the quantum vacuum fluctuations. Another important work by Andreas Albrecht et al. [47] have used the two-mode squeezed state formalism (for a single field) to understand the inflationary cosmology and the amplification process of quantum fluctuations during the inflationary epoch. Another more recent work [48] discussed on four-mode squeezed states for two quantum systems using the symplectic group theory and its Lie algebra. These works have motivated us to explore both the theoretical and numerical features of two coupled scalar fields in the planar patch of the de Sitter space by developing the formalism and constructing the corresponding four-mode squeezed operator. For more vast applications of squeezed state formalism in High energy physics and in cosmology see [41–44, 49–67].

The structure of the paper can be broadly divided into two parts, section II and section III.

In the section II, we start our analysis by considering two massive scalar fields in the planar patch of de Sitter space with  $K$  as weak coupling constant between the fields. In subsection II A, we will quantize the modes of the two coupled scalar fields. Here for simplicity we will neglect the “back action” of the scalar fields to the de Sitter background geometry. We will find the quantized position and momentum variables for the two coupled scalar fields on the planar patch of the de Sitter space and then compute the quantized version of the Hamiltonian, for our case. In subsection II B, we will briefly present the connection between the two coupled inverted quantum harmonic oscillator system and four mode squeezed state formalism constructed for our model under consideration. In subsection II C, we will construct four mode

squeezed state operator which will be useful for understanding the cosmological implications for two interacting scalar fields and also for other systems which can be explained in terms of two coupled inverted quantum harmonic oscillators.

In the section III, we will define the time evolution operator for the four mode squeezed states, which we will use to calculate the time dependent (Heisenberg picture) annihilation and creation operators for two coupled scalar fields in the planar patch of the de Sitter space. In subsection IV, we use the Heisenberg equation of motion to calculate the coupled differential equation for the mode functions of the two coupled scalar fields in the planar patch of the de Sitter space. We will also calculate the position and momentum operators in Heisenberg representation for this model. After that we give the expression for mode functions by using the set of coupled differential equation for the mode functions. We give the expression for the squeezing parameters,  $R_{1,\mathbf{k}}$ ,  $\Theta_{1,\mathbf{k}}$ ,  $\Phi_{1,\mathbf{k}}$ ,  $R_{2,\mathbf{k}}$ ,  $\Theta_{2,\mathbf{k}}$  and  $\Phi_{2,\mathbf{k}}$  which governs the evolution of the quantum state for two coupled scalar fields in the planar patch of de Sitter space.

## II. Four-mode Squeezed State Formalism for two field interacting model

In the following section, our prime objective is to construct a formalism for two scalar fields  $\phi_1$  and  $\phi_2$ , in the planar patch of de Sitter space, which are weakly interacting with each other through a coupling strength  $K$ . The action for the two scalar fields contain a usual gravitational part with  $R$  as Ricci scalar, a matter source term  $T$ , it also contains kinetic term and potential term from both the fields.

The corresponding action for two coupled interacting scalar fields [68] in the planar patch of de Sitter space can be written as:

$$S = \int d^3x dt \sqrt{-g} \left[ -\frac{R}{16\pi G} + T + \frac{1}{2} (\dot{\phi}_1^2 - m_1^2 \phi_1^2) + \frac{1}{2} (\dot{\phi}_2^2 - m_2^2 \phi_2^2) - 12KR\phi_1\phi_2 \right] \quad (1)$$

Here, the corresponding de Sitter metric in the planar patch is described by the following line element:

$$ds^2 = -dt^2 + a^2(t) d\mathbf{x}^2 \quad \text{with} \quad a(t) = \exp(Ht). \quad (2)$$

Here “ $a$ ” is the scale factor which is a function of time and  $H$  is the Hubble constant. Also here  $m_1$  is the mass of field  $\phi_1$  and  $m_2$  is the mass of field  $\phi_2$ . Due to having the homogeneity and isotropy in the spatially flat FLRW background having planar de Sitter solution both the fields are only functions of time and there is no space

dependence. For this reason there is no kinetic term appearing which involve the space derivatives. Before moving further we introduce some conditions which will help us in simplifying the calculations. We define the integral over the matter source term to be proportional to the potential function  $V(a)$  and is given by:

$$\int d^3x \sqrt{-g} T = \frac{1}{2} V(a) \quad (3)$$

We will set the Planck length  $l_p = 1$  for the rest of the computation and we also introduce some dimensionless field variables, which are given by:

$$\begin{aligned} \mu_1(t) &= \phi_1(t) a^{3/2}(t) = \exp(3Ht/2) \phi_1(t), \\ \mu_2(t) &= \phi_2(t) a^{3/2}(t) = \exp(3Ht/2) \phi_2(t). \end{aligned} \quad (4)$$

After doing all this manipulations and using the field redefinition the Lagrangian for two coupled scalar fields in the planar patch of de Sitter space can be recast in the following simplified form:

$$L = -\frac{a\dot{a}^2}{2} + \frac{V(a)}{2} + l_1 + l_2 + l_3 \quad (5)$$

Where,  $l_1$ ,  $l_2$  and  $l_3$  are given by the following expressions:

$$l_1 = \frac{\dot{\mu}_1^2}{2} - \frac{1}{2} \mu_1^2 m_1^2 - \frac{3}{2} \frac{\dot{a}}{a} \mu_1 \dot{\mu}_1 + \frac{9}{8} \left( \frac{\dot{a}}{a} \right)^2 \mu_1^2, \quad (6)$$

$$l_2 = \frac{\dot{\mu}_2^2}{2} - \frac{1}{2} \mu_2^2 m_2^2 - \frac{3}{2} \frac{\dot{a}}{a} \mu_2 \dot{\mu}_2 + \frac{9}{8} \left( \frac{\dot{a}}{a} \right)^2 \mu_2^2, \quad (7)$$

$$l_3 = K \left[ \left( \frac{\dot{a}}{a} \right)^2 \mu_1 \mu_2 - \frac{1}{2} \frac{\dot{a}}{a} (\dot{\mu}_1 \mu_2 + \mu_1 \dot{\mu}_2) \right]. \quad (8)$$

With the use of Euler Lagrange equations we get the two equation of motion each for  $\mu_1$  and  $\mu_2$  fields are given by:

$$\begin{aligned} \ddot{\mu}_1 + \left[ m_1^2 - \frac{3}{2} \frac{\ddot{a}}{a} - \frac{3}{4} \left( \frac{\dot{a}}{a} \right)^2 \right] \mu_1 \\ - \frac{K}{2} \left[ \frac{\ddot{a}}{a} + \left( \frac{\dot{a}}{a} \right)^2 \right] \mu_2 = 0, \end{aligned} \quad (9)$$

$$\begin{aligned} \ddot{\mu}_2 + \left[ m_2^2 - \frac{3}{2} \frac{\ddot{a}}{a} - \frac{3}{4} \left( \frac{\dot{a}}{a} \right)^2 \right] \mu_2 \\ - \frac{K}{2} \left[ \frac{\ddot{a}}{a} + \left( \frac{\dot{a}}{a} \right)^2 \right] \mu_1 = 0. \end{aligned} \quad (10)$$

Here we can clearly notice that the above two equations are coupled differential equation of motion due to having the interaction in the original theory.

Now we can construct the Hamiltonian for our theory from the Lagrangian given in Eq(5). The Hamiltonian for two interacting scalar fields in the planar patch of de Sitter background is given by:

$$H = -\frac{\pi^2}{2M} - \frac{V(a)}{2} + \frac{1}{2} (\pi_1^2 + m_1^2 v_1^2) + \frac{1}{2} (\pi_2^2 + m_2^2 v_2^2) \quad (11)$$

Where, we define:

$$M = a + \frac{K}{a^2} \left[ \mu_1 \mu_2 + \frac{K}{4} (\mu_1^2 + \mu_2^2) \right] \quad (12)$$

$$\pi = p_a + \frac{p_1}{2a} (3\mu_1 + k\mu_2) + \frac{p_2}{2a} (3\mu_2 + k\mu_1) \quad (13)$$

Here we introduce,  $p_a$  : Canonically conjugate momenta of scale factor  $a$ ,  $\pi_1, \pi_2$  : Canonically conjugate momenta of redefined fields  $\mu_1$  and  $\mu_2$  respectively.

### A. Quantizing the Hamiltonian

We will be quantizing the fields  $\mu_1$  and  $\mu_2$ , and will be treating gravity classically in this computation. In our analysis the back reaction from the fields is neglected. For this reason we expand  $\frac{\pi_a^2}{2M}$  in a series and retain only the terms  $\frac{\pi_a^2}{a}$  (which governs together with the potential  $V(a)$  for a which is a semi-classical behaviour of the background geometry).

With these condition defined above we get the approximated form of the Hamiltonian as follows:

$$\begin{aligned} H \sim & \frac{1}{2} (\pi_1^2 + v_1^2 m_1^2) + \frac{1}{2} (\pi_2^2 + v_2^2 m_2^2) \\ & - \frac{p_a^2}{2a^2} [3(v_1 \pi_1 + v_2 \pi_2) + K(v_1 \pi_2 + v_2 \pi_1)] \\ & + \frac{K p_a^2}{2a^4} \left( v_1 v_2 + \frac{k}{4} (v_1^2 + v_2^2) \right) \end{aligned} \quad (14)$$

where we define  $p_a \sim -a\dot{a}$ .

Now we will quantize this Hamiltonian. For this purpose we promote the fields to operators and take the Fourier decomposition. We use the following ansatz for Fourier decomposition:

$$\begin{aligned} \hat{v}_1 &= \int \frac{d^3 k}{(2\pi)^3} \hat{v}_{1,\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}, \\ \hat{\pi}_1 &= \int \frac{d^3 k}{(2\pi)^3} \hat{\pi}_{1,\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}, \\ \hat{v}_2 &= \int \frac{d^3 k}{(2\pi)^3} \hat{v}_{2,\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}, \\ \hat{\pi}_2 &= \int \frac{d^3 k}{(2\pi)^3} \hat{\pi}_{2,\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}. \end{aligned} \quad (15)$$

For our purpose we will be working in the Schrödinger picture, where the operators  $\hat{v}_{1,\mathbf{k}}, \hat{\pi}_{1,\mathbf{k}}, \hat{v}_{2,\mathbf{k}}$  and  $\hat{\pi}_{2,\mathbf{k}}$  are fixed at an initial time. We define modes and the associated canonically conjugate momenta for the two fields with initial frequency equal to  $k$  which, suitably normalized, give us:

$$\begin{aligned} \hat{v}_{1,\mathbf{k}} &= \frac{1}{\sqrt{2k_1}} (b_{1,\mathbf{k}} + b_{1,\mathbf{k}}^\dagger), & \hat{\pi}_{1,\mathbf{k}} &= i\sqrt{\frac{k_1}{2}} (b_{1,\mathbf{k}}^\dagger - b_{1,\mathbf{k}}) \\ \hat{v}_{2,\mathbf{k}} &= \frac{1}{\sqrt{2k_2}} (b_{2,\mathbf{k}} + b_{2,\mathbf{k}}^\dagger), & \hat{\pi}_{2,\mathbf{k}} &= i\sqrt{\frac{k_2}{2}} (b_{2,\mathbf{k}}^\dagger - b_{2,\mathbf{k}}) \end{aligned} \quad (16)$$

The Four-mode Hamiltonian operator after quantization for the two scalar fields interacting with each other via coupling constant  $K$  in the planar patch of de Sitter space can be written in the following simple form:

$$\begin{aligned} H(\tau) = & \left( l_1(\tau) b_{1,-\mathbf{k}} b_{1,\mathbf{k}} + l_1^*(\tau) b_{1,\mathbf{k}}^\dagger b_{1,-\mathbf{k}}^\dagger \right) \\ & + \left( l_2(\tau) b_{2,-\mathbf{k}} b_{2,\mathbf{k}} + l_2^*(\tau) b_{2,\mathbf{k}}^\dagger b_{2,-\mathbf{k}}^\dagger \right) \\ & + \left\{ \omega_1(\tau) (b_{1,-\mathbf{k}}^\dagger b_{1,\mathbf{k}} + b_{1,\mathbf{k}}^\dagger b_{1,-\mathbf{k}}) \right. \\ & + \omega_2(\tau) (b_{2,-\mathbf{k}}^\dagger b_{2,\mathbf{k}} + b_{2,\mathbf{k}}^\dagger b_{2,-\mathbf{k}}) \\ & + g_1(\tau) (b_{1,-\mathbf{k}} b_{2,\mathbf{k}} + b_{1,\mathbf{k}} b_{2,-\mathbf{k}}) \\ & + g_1^*(\tau) (b_{2,\mathbf{k}}^\dagger b_{1,-\mathbf{k}}^\dagger + b_{2,-\mathbf{k}}^\dagger b_{1,\mathbf{k}}^\dagger) \\ & + g_2(\tau) (b_{1,\mathbf{k}}^\dagger b_{2,-\mathbf{k}}^\dagger + b_{1,-\mathbf{k}}^\dagger b_{2,\mathbf{k}}^\dagger) \\ & \left. + g_2^*(\tau) (b_{2,-\mathbf{k}} b_{1,\mathbf{k}}^\dagger + b_{2,\mathbf{k}} b_{1,-\mathbf{k}}^\dagger) \right\} \end{aligned} \quad (17)$$

Where, the terms  $l_1, l_2, \omega_1, \omega_2, g_1$  and  $g_2$  are defined below:

$$\begin{aligned} l_1(\tau) &= \frac{K^2 \pi_a^2}{16a^4 m_1} + i \frac{3}{4a^2} \pi_a \\ l_2(\tau) &= \frac{K^2 \pi_a^2}{16a^4 m_2} + i \frac{3}{4a^2} \pi_a \\ \omega_1(\tau) &= \frac{1}{2} \left( m_1 + \frac{K^2 \pi_a^2}{8a^4 m_1} \right) \\ \omega_2(\tau) &= \frac{1}{2} \left( m_2 + \frac{K^2 \pi_a^2}{8a^4 m_2} \right) \\ g_1(\tau) &= \frac{K^2 \pi_a^2}{4a^4} \frac{1}{\sqrt{m_1 m_2}} \\ &+ i \frac{K^2 \pi_a^2}{4a^4} \frac{(m_1 + m_2)}{\sqrt{m_1 m_2}} = g_2(\tau) \end{aligned} \quad (18)$$

Here it is important to note that,  $\tau$  represents the conformal time which is related to the physical time  $t$  by the following expression:

$$\tau = \int \frac{dt}{a(t)} = -\frac{1}{H} \exp(-Ht). \quad (19)$$

Consequently, in terms of the conformal time coordinate the de Sitter metric in planar patch can be recast as:

$$ds^2 = a^2(\tau) (-d\tau^2 + d\mathbf{x}^2) \quad \text{where} \quad a(\tau) = -\frac{1}{H\tau}. \quad (20)$$

### B. The Two Coupled Inverted Harmonic Oscillator

The two coupled inverted quantum harmonic oscillator [69–71] can be described by the Hamiltonian  $H$  which



is the sum of the free Hamiltonian of both the inverted quantum harmonic oscillators and the interaction Hamiltonian  $H_{int}$  which depends on the type interaction between them and the coupling constant  $K$  accounts for the strength of the interaction between the two coupled inverted quantum harmonic oscillators:

$$H = \sum_{i=1}^2 H_i + KH_{int} \quad (21)$$

Here,  $H_i$  is the Hamiltonian of a free two inverted quantum harmonic oscillators:

$$\begin{aligned} H_i &= \frac{\hat{p}_i^2}{2} - \frac{\hat{q}_i^2}{2} \quad \text{where } i = 1, 2 \\ &= i \frac{\hbar}{2} \left( \hat{b}_i^2 e^{2i\frac{\pi}{4}} - \text{h.c.} \right). \end{aligned} \quad (22)$$

Here  $b_i$  corresponds to the annihilation operator for the quantum harmonic oscillator and index  $i$  runs from 1 to 2.

In the present context of discussion we will consider the construction of General squeeze Hamiltonian. Writing the free part Hamiltonian in this form facilitates the comparison with the more general squeeze Hamiltonian, which we will be considering. The setup of this two coupled inverted harmonic oscillator can be mapped to case of two coupled scalar fields in de Sitter background as indicated in the Eq(21) and Eq(17). The free part of the Hamiltonian for two coupled inverted quantum harmonic oscillator gets mapped to the terms containing  $l_1$ ,  $l_1^*$  for the first scalar field with modes  $(1, \mathbf{k}, 1, -\mathbf{k})$  and the terms containing  $l_2$ ,  $l_2^*$  for the second scalar field with modes  $(2, \mathbf{k}, 2, -\mathbf{k})$ . The free terms with  $\omega_1$  and  $\omega_2$  mimics the role of rotation operator which is absent in Eq(21) but it can be introduced explicitly in it. The interaction Hamiltonian  $H_{int}$  corresponds to the terms containing  $g_1$ ,  $g_1^*$ ,  $g_2$  and  $g_2^*$ .

### C. Four mode Squeezed State operator

Let us consider a state  $|\phi\rangle_{in}$  which is the initial reference vacuum state of the two scalar fields. The final out state  $|\phi\rangle_{out}$  can be obtained by applying the operator given in Eq (23) on the initial reference vacuum state of the two scalar fields.

$$|\phi\rangle_{out} = S_1^{(1)}(r_1, \theta_1) S_1^{(2)}(r_2, \theta_2) S(r_1, r_2, \theta_1, \theta_2) \mathcal{R}(\phi_1, \phi_2) |\phi\rangle_{in} \quad (23)$$

Here, the initial vacuum state is  $|\phi\rangle_{in}$  and we have

$$\mathcal{R}(\phi_1, \phi_2) |0, 0\rangle = e^{i(\phi_1 + \phi_2)} |0, 0\rangle \quad (24)$$

Here, we can see that the contribution of the total rotational operator on the vacuum state just introduces an

overall phase factor and we neglect it for further calculations of the total squeezed operator for the two scalar fields in de Sitter background space.

We write the most general Squeezed state in the present context, which is defined as:

$$|\Psi_{sq}\rangle = S_1^{(1)}(r_1, \theta_1) S_1^{(2)}(r_2, \theta_2) S(r_1, r_2, \theta_1, \theta_2) |0, 0\rangle \quad (25)$$

Here it is important to note that

$$\underbrace{S_1^{(1)}(r_1, \theta_1) S_1^{(2)}(r_2, \theta_2)}_{\text{with interaction 1 and 2}} \neq \underbrace{S_1^{(1)}(r_1, \theta_1) S_1^{(2)}(r_2, \theta_2)}_{\text{without interaction 1 and 2}} \quad (26)$$

Following are the important points to be noted for operators  $S_1^{(1)}(r_1, \theta_1)$ ,  $S_1^{(2)}(r_2, \theta_2)$  and  $S(r_1, r_2, \theta_1, \theta_2)$ :

- The contribution from the gravitational part from the  $a(t)$  as well as the interaction part of the fields appear in  $S(r_1, r_2, \theta_1, \theta_2)$
- The Contribution of  $\mu_1$  and  $\mu_2$  with interaction appears in  $S_1^{(1)}(r_1, \theta_1)$ ,  $S_1^{(2)}(r_2, \theta_2)$ , which makes them different from the contribution obtained from the case when there is no interaction at all.

### D. Technical details of the Constructions

Here we will give the calculation of the total squeezed operator for two scalar fields having interaction with each other in de Sitter background space. First we consider the operator  $S_1^{(1)}(r_1, \theta_1)$  and write it in the form of Eq(27) which is the product of usual squeezed operator for first field and the interaction term given in Eq(28).

$$\begin{aligned} S_1^{(1)}(r_1, \theta_1) &= \exp \left[ \frac{r_1}{2} \left( e^{-2i\theta_1} b_1^2 - e^{2i\theta_1} b_1^{\dagger 2} \right) \right. \\ &\quad \left. + \frac{r_1}{2} \left( e^{-2i\theta_1} (b_1 b_2 + b_1^\dagger b_2^\dagger) - e^{2i\theta_1} (b_2^\dagger b_1^\dagger + b_2 b_1) \right) \right] \quad (27) \\ &= S_1^{(\text{non-int.})}(r_1, \theta_1) \underbrace{S_1^{(1 \text{ int. with } 2)}(r_1, \theta_1)}_{\text{New Contribution}} \end{aligned}$$

where,  $S_1^{(1 \text{ int. with } 2)}(r_1, \theta_1)$  is the part of the operator  $S_1^{(1)}(r_1, \theta_1)$  containing interaction, which is given by:

$$\begin{aligned} S_1^{(1 \text{ int. with } 2)}(r_1, \theta_1) &= \exp \left[ \frac{r_1}{2} \left\{ e^{-2i\theta_1} (b_1 b_2 + b_1^\dagger b_2^\dagger) \right. \right. \\ &\quad \left. \left. - e^{2i\theta_1} (b_2^\dagger b_1^\dagger + b_2 b_1) \right\} + \dots \right] \quad (28) \end{aligned}$$

Similarly we can write for operator  $S_1^{(2)}(r_2, \theta_2)$ , which is also the product of usual squeezed operator for second field and the term coming from the interaction.

$$\begin{aligned} S_1^{(2)}(r_2, \theta_2) &= \\ &= S_1^{(\text{non-int.})}(r_2, \theta_2) \underbrace{S_1^{(2 \text{ int. with } 1)}(r_2, \theta_2)}_{\text{New Contribution}} \quad (29) \end{aligned}$$

where, Eq(30) represents the interaction part of the operator  $S_1^{(2)}(r_2, \theta_2)$  coming from second field.

$$S_1^{(1 \text{ int. with } 2)}(r_2, \theta_2) = \exp \left[ \frac{r_2}{2} \left\{ e^{-2i\theta_2} (b_1 b_2 + b_1^\dagger b_2^\dagger) - e^{2i\theta_1} (b_2^\dagger b_1^\dagger + b_2 b_1) \right\} + \dots \right] \quad (30)$$

Now, the full squeezed state operator is the product of the usual squeezed operators from both the fields with the operators with interaction between 1 and 2 fields, a third additional term is also present denoted as  $S(r_1, r_2, \theta_1, \theta_2)$ . This is the contribution which is appearing due to commutators between  $b_1$ ,  $b_1^\dagger$ , and  $b_2$ ,  $b_2^\dagger$ .

$$S_{\text{Full}} = \underbrace{S_1(r_1, \theta_1) S_1(r_2, \theta_2)}_{\text{Without interaction}} \times \underbrace{S_1^{\text{int.}}(r_1, \theta_1) S_1^{\text{int.}}(r_2, \theta_2)}_{\text{Interaction between 1 and 2}} \times \underbrace{S(r_1, \theta_1, r_2, \theta_2)}_{\text{Additional terms}} \quad (31)$$

In order to construct the additional term  $S(r_1, r_2, \theta_1, \theta_2)$  of the total squeezed operator we will use the Baker–Campbell–Hausdorff formula. Let us consider  $S_1^{(1)}$  and  $S_2^{(1)}$  given in Eq (32) and Eq (33).

$$S_1^{(1)}(r_1, \theta_1) = \exp \left[ \frac{r_1}{2} \left( e^{-2i\theta_1} b_1^2 - e^{2i\theta_1} b_1^{\dagger 2} \right) + \frac{r_1}{2} \left\{ e^{-2i\theta_1} (b_1 b_2 + b_1^\dagger b_2^\dagger) - e^{2i\theta_1} (b_2^\dagger b_1^\dagger + b_2 b_1) \right\} \right] \quad (32)$$

and

$$S_1^{(2)}(r_1, \theta_2) = \exp \left[ \frac{r_2}{2} \left( e^{-2i\theta_2} b_1^2 - e^{2i\theta_2} b_1^{\dagger 2} \right) + \frac{r_2}{2} \left\{ e^{-2i\theta_2} (b_1 b_2 + b_1^\dagger b_2^\dagger) - e^{2i\theta_2} (b_2^\dagger b_1^\dagger + b_2 b_1) \right\} \right] \quad (33)$$

Now we call,

$$\begin{aligned} \frac{r_1}{2} \left( e^{-2i\theta_1} b_1^2 - e^{2i\theta_1} b_1^{\dagger 2} \right) &= \alpha_1 \\ \frac{r_1}{2} \left\{ e^{-2i\theta_1} (b_1 b_2 + b_1^\dagger b_2^\dagger) - e^{2i\theta_1} (b_2^\dagger b_1^\dagger + b_2 b_1) \right\} &= \alpha_2 \end{aligned} \quad (34)$$

Similarly we have,

$$\begin{aligned} \frac{r_2}{2} \left( e^{-2i\theta_2} b_1^2 - e^{2i\theta_2} b_2^{\dagger 2} \right) &= \beta_1 \\ \frac{r_2}{2} \left\{ e^{-2i\theta_2} (b_1 b_2 + b_1^\dagger b_2^\dagger) - e^{2i\theta_2} (b_2^\dagger b_1^\dagger + b_2 b_1) \right\} &= \beta_2 \end{aligned} \quad (35)$$

Such that

$$\begin{aligned} \alpha_1 + \alpha_2 &= A_1 \\ \beta_1 + \beta_2 &= A_2 \end{aligned} \quad (36)$$

Now we apply the Baker–Campbell–Hausdorff formula between  $S_1^{(1)}$  and  $S_2^{(1)}$ , where we have defined the terms in the exponential as  $A_1$  and  $A_2$ .

$$\begin{aligned} e^{A_1} \cdot e^{A_2} &= e^{\{A_1 + A_2 + \frac{1}{2}[A_1, A_2] + \frac{1}{12}([A_1, [A_1, A_2]] - [A_2, [A_1, A_2]]) + \dots\}} \\ &= \underbrace{e^{A_1}}_{S_1^1} \underbrace{e^{A_2}}_{S_1^2} \underbrace{e^{f(A_1, A_2)}}_{S \text{ New Contribution}} \end{aligned} \quad (37)$$

We will only consider the BCH-expansion upto  $[A_1, [A_1, A_2]]$  and  $[A_2, [A_1, A_2]]$ .

$$\begin{aligned} f(A_1, A_2) &= \frac{1}{2} [A_1, A_2] + \frac{1}{12} ([A_1, [A_1, A_2]] - [A_2, [A_1, A_2]]) \\ &\quad + \dots \end{aligned} \quad (38)$$

First we consider the first commutator in the Eq(38)

$$\begin{aligned} [A_1, A_2] &= [\alpha_1 + \alpha_2, \beta_1 + \beta_2] \\ &= [\alpha_1, \beta_1] + [\alpha_2, \beta_2] + [\alpha_2, \beta_1] + [\alpha_1, \beta_2] \end{aligned} \quad (39)$$

Where,

$$\begin{aligned} [\alpha_1, \beta_1] &= 0 \\ [\alpha_2, \beta_2] &= 0 \\ [\alpha_2, \beta_1] &= \frac{r_1 r_2}{2} \left\{ \left( e^{-2i(\theta_1 + \theta_2)} - e^{-2i(\theta_1 - \theta_2)} \right) b_1 b_2^\dagger + \left( e^{2i(\theta_1 - \theta_2)} - e^{-2i(\theta_1 + \theta_2)} \right) b_1^\dagger b_2 \right\} \\ [\alpha_1, \beta_2] &= \frac{r_1 r_2}{2} \left\{ \left( e^{-2i(\theta_1 + \theta_2)} - e^{-2i(\theta_1 - \theta_2)} \right) b_1 b_2^\dagger + \left( e^{2i(\theta_1 + \theta_2)} - e^{2i(\theta_1 - \theta_2)} \right) b_1^\dagger b_2 \right\} \end{aligned} \quad (40)$$

So,

$$\begin{aligned} [A_1, A_2] &= [\alpha_1, \beta_2] + [\alpha_2, \beta_1] \\ &= \frac{r_1 r_2}{2} \left\{ f_1(\theta_1, \theta_2) b_1 b_2^\dagger + f_2(\theta_1, \theta_2) b_1^\dagger b_2 \right\} \end{aligned} \quad (41)$$

Where, we have defined  $f_1(\theta_1, \theta_2)$  and  $f_2(\theta_1, \theta_2)$  as:

$$\begin{aligned} f_1(\theta_1, \theta_2) &= 2 \cos(2(\theta_1 + \theta_2)) - 2e^{-2i(\theta_1 - \theta_2)} \\ &= f_1^{\text{Real}}(\theta_1 + \theta_2) + i f_1^{\text{Im}}(\theta_1, \theta_2) \end{aligned} \quad (42)$$

with the real and the imaginary part being,

$$\begin{aligned} f_1^{\text{Real}}(\theta_1, \theta_2) &= 2(\cos(2(\theta_1 + \theta_2)) - \cos(2(\theta_1 - \theta_2))) \\ f_1^{\text{Im}}(\theta_1, \theta_2) &= 2 \sin(2(\theta_1 - \theta_2)) \end{aligned} \quad (43)$$

and

$$\begin{aligned} f_2(\theta_1, \theta_2) &= e^{2i(\theta_1+\theta_2)} - e^{-2i(\theta_1+\theta_2)} \\ &= 2i \sin(2(\theta_1 + \theta_2)) \\ &= f_2^{\text{Real}}(\theta_1, \theta_2) + i f_2^{\text{Im}}(\theta_1, \theta_2) \end{aligned} \quad (44)$$

where,

$$\begin{aligned} f_2^{\text{Real}}(\theta_1, \theta_2) &= 0 \\ f_2^{\text{Im}}(\theta_1, \theta_2) &= 2 \sin(2(\theta_1 + \theta_2)) \end{aligned} \quad (45)$$

So, the first commutator becomes as follows:

$$\begin{aligned} [A_1, A_2] &= r_1 r_2 [\{\cos(2(\theta_1 + \theta_2)) - \cos(2(\theta_1 - \theta_2)) \\ &\quad + i \sin(2(\theta_1 + \theta_2))\} b_1 b_2^\dagger \\ &\quad + \{i \sin(2(\theta_1 - \theta_2))\} b_1^\dagger b_2] \end{aligned} \quad (46)$$

We can express it as follows:

$$\begin{aligned} [A_1, A_2] &= \left[ f_1(r_1, r_2, \theta_1, \theta_2) b_1 b_2^\dagger \right. \\ &\quad \left. + f_2(r_1, r_2, \theta_1, \theta_2) b_1^\dagger b_2 \right] \end{aligned} \quad (47)$$

Where, we define  $f_1(r_1, r_2, \theta_1, \theta_2)$  and  $f_2(r_1, r_2, \theta_1, \theta_2)$  as:

$$\begin{aligned} f_1(r_1, r_2, \theta_1, \theta_2) &= r_1 r_2 [\{\cos(2(\theta_1 + \theta_2)) - \cos(2(\theta_1 - \theta_2)) \\ &\quad + i \sin(2(\theta_1 - \theta_2))\}] \\ f_2(r_1, r_2, \theta_1, \theta_2) &= r_1 r_2 i [\sin(2(\theta_1 + \theta_2))] \end{aligned} \quad (48)$$

Let us denote the following,

$$\begin{aligned} P &= f_1(r_1, r_2, \theta_1, \theta_2) b_1 b_2^\dagger \\ Q &= f_2(r_1, r_2, \theta_1, \theta_2) b_1^\dagger b_2 \end{aligned} \quad (49)$$

So,  $[A_1, A_2] = P + Q$ . We will now compute the second commutator in the expansion which is  $[A_1, [A_1, A_2]]$ . The first part of this commutator is  $[\alpha_1, [A_1, A_2]]$ , which becomes:

$$\begin{aligned} [\alpha_1, [A_1, A_2]] &= [\alpha_1, P + Q] \\ &= (\tilde{f}_2 b_1 b_2 + \tilde{f}_1 b_1^\dagger b_2^\dagger) \end{aligned} \quad (50)$$

Where

$$\begin{aligned} \tilde{f}_2 &= r_1 f_2 e^{-2i\theta_1} \\ \tilde{f}_1 &= r_1 f_1 e^{2i\theta_1} \end{aligned} \quad (51)$$

Now we will compute the second part of the commutator  $[A_1, [A_1, A_2]]$  which is given as  $[\alpha_2, [A_1, A_2]]$ , Here, first we write  $\alpha_2$  as,

$$\alpha_2 = \Theta_1 - \Theta_2 \quad (52)$$

Where,

$$\begin{aligned} \Theta_1 &= \frac{r_1}{2} e^{-2i\theta_1} (b_1 b_2 + b_1^\dagger b_2^\dagger) \\ \Theta_2 &= \frac{r_1}{2} e^{2i\theta_1} (b_2^\dagger b_1^\dagger + b_2 b_1) \end{aligned} \quad (53)$$

Using this we have,

$$\begin{aligned} [\alpha_2, [A_1, A_2]] &= [\Theta_1 - \Theta_2, P + Q] \end{aligned} \quad (54)$$

where the commutators are as follows:

$$\begin{aligned} [\Theta_1, P] &= \frac{r_1 f_1}{2} e^{-2i\theta_1} (b_1 b_1 - b_2^\dagger b_2^\dagger) \\ [\Theta_1, Q] &= \frac{r_1 f_2}{2} e^{-2i\theta_1} (b_2 b_2 - b_1^\dagger b_1^\dagger) \\ [\Theta_2, P] &= \frac{r_1 f_1}{2} e^{2i\theta_1} (b_1 b_1 - b_2^\dagger b_2^\dagger) \\ [\Theta_2, Q] &= \frac{r_1 f_2}{2} e^{2i\theta_1} (b_2 b_2 - b_1^\dagger b_1^\dagger) \end{aligned} \quad (55)$$

The second part of the commutator  $[A_1, [A_1, A_2]]$  becomes:

$$\begin{aligned} [\alpha_2, [A_1, A_2]] &= \frac{r_1}{2} \left[ e^{2i\theta_1} \left\{ f_2 (b_1^\dagger b_1^\dagger - b_2 b_2) - f_1 (b_1 b_1 - b_2^\dagger b_2^\dagger) \right\} \right. \\ &\quad \left. + e^{-2i\theta_1} \left\{ f_1 (b_1 b_1 - b_2^\dagger b_2^\dagger) + f_2 (b_2 b_2 - b_1^\dagger b_1^\dagger) \right\} \right] \end{aligned} \quad (56)$$

For the third and the last commutator of our truncated BCH-expansion, we need to calculate  $[A_2, [A_1, A_2]]$ , and previously we had defined,  $A_2 = \beta_1 + \beta_2$  (36)

Where, we define  $\beta_1$  and  $\beta_2$  as:

$$\begin{aligned} \beta_1 &= M_1 - M_2 \\ \beta_2 &= N_1 - N_2 \end{aligned} \quad (57)$$

and hence, from Eq(35) and (57),  $M_1$  and  $M_2$  becomes:

$$\begin{aligned} M_1 &= \frac{r_2}{2} e^{-2i\theta_2} b_2^2, \quad N_1 = \frac{r_2}{2} (e^{-2i\theta_2} (b_1 b_2 + b_1^\dagger b_2^\dagger)) \\ M_2 &= \frac{r_2}{2} e^{2i\theta_2} b_2^{\dagger 2}, \quad N_2 = \frac{r_2}{2} (e^{2i\theta_2} (b_2^\dagger b_1^\dagger + b_2 b_1)) \end{aligned} \quad (58)$$

With the following commutation relation in Eq(59), we can write the first part of the  $[A_2, [A_1, A_2]]$  commutator.

$$\begin{aligned} [M_2, P] &= 0 \quad [M_1, P] = r_2 f_1 e^{-2i\theta_2} b_1 b_2 \\ [M_1, Q] &= 0 \quad [M_2, Q] = -r_2 f_2 e^{2i\theta_2} b_1^\dagger b_2^\dagger \end{aligned} \quad (59)$$

Which is given as follows:

$$\begin{aligned} [\beta_1, [A_1, A_2]] &= [\beta_1, P] + [\beta_1, Q] \\ &= r_2 (f_1 e^{-2i\theta_2} b_1 b_2 + f_2 e^{2i\theta_2} b_1^\dagger b_2^\dagger) \end{aligned} \quad (60)$$

Similarly we calculate the second term of the commutator  $[A_2, [A_1, A_2]]$  and use the following commutation

results given in the Eq(61) below,

$$\begin{aligned} [N_1, P] &= \frac{r_2 f_1}{2} e^{-2i\theta_2} (b_1 b_1 - b_2^\dagger b_2^\dagger) \\ [N_1, Q] &= \frac{r_2 f_2}{2} e^{-2i\theta_2} (b_2 b_2 - b_1^\dagger b_1^\dagger) \\ [N_2, P] &= \frac{r_2 f_1}{2} e^{2i\theta_2} (b_1 b_1 - b_2^\dagger b_2^\dagger) \\ [N_2, Q] &= \frac{r_2 f_2}{2} e^{2i\theta_2} (b_2 b_2 - b_1^\dagger b_1^\dagger) \end{aligned} \quad (61)$$

Finally we get the second term of the  $[A_2, [A_1, A_2]]$  commutator:

$$\begin{aligned} [\beta_2, [A_1, A_2]] &= [\beta_2, P + Q] \\ &= [N_1, P] - [N_2, P] + [N_1, Q] - [N_2, Q] \\ &= i r_2 \sin(2\theta_2) \left[ f_2 (b_1^\dagger b_1^\dagger - b_2 b_2) \right. \\ &\quad \left. - f_1 (b_1 b_1 - b_2^\dagger b_2^\dagger) \right] \end{aligned} \quad (62)$$

After calculating the terms in Eq(38), now we rewrite the combination of  $[A_1, [A_1, A_2]]$  and  $[A_2, [A_1, A_2]]$  commutators in terms of  $X_1, X_2, X_3$  and  $X_4$ .

$$\begin{aligned} [A_1, [A_1, A_2]] - [A_2, [A_1, A_2]] \\ = X_1 b_1 b_2 + X_2 b_1^\dagger b_2^\dagger + X_3 (b_1^\dagger b_1^\dagger - b_2 b_2) \\ + X_4 (b_1 b_1 - b_2^\dagger b_2^\dagger) \end{aligned} \quad (63)$$

Where,  $X_1, X_2, X_3$  and  $X_4$  are appended below:

$$\begin{aligned} X_1 &= (r_1 f_2 e^{-2i\theta_1} - r_2 f_1 e^{-2i\theta_2}) \\ X_2 &= (r_1 f_1 e^{2i\theta_1} - r_2 f_2 e^{2i\theta_2}) \\ X_3 &= i f_2 (r_1 \sin(2\theta_1) - r_2 \sin(2\theta_2)) \\ X_4 &= i f_1 (r_2 \sin(2\theta_2) - r_1 \sin(2\theta_1)) \end{aligned} \quad (64)$$

Note: Here the ratio of  $X_3$  and  $X_4$  satisfies the following relation given in Eq(65) :

$$\frac{X_3}{X_4} = -\frac{f_2}{f_1} \quad (65)$$

To get the full squeezed operator for our case whose form is mentioned in Eq(31), we multiply the usual free squeezed operators from both the fields with the BCH expansion given in Eq(37). The total squeezed operator for two scalar fields in de Sitter background space after applying the Baker-Campbell-Hausdorff will become:

$$S_{\text{Full}}(r_1, \theta_1, r_2, \theta_2) = e^{(M_1 + M_2 + M_3)} \quad (66)$$

Where,  $\exp M_1$ , and  $\exp M_2$  are the usual free squeezed operators given as:

$$\begin{aligned} \exp M_1 &= \underbrace{\exp \left[ \frac{r_1}{2} (e^{-2i\theta_1} b_1^2 - e^{2i\theta_1} b_1^{\dagger 2}) \right]}_{\text{for "1".}} \\ \exp M_2 &= \underbrace{\exp \left[ \frac{r_2}{2} (e^{-2i\theta_2} b_2^2 - e^{2i\theta_2} b_2^{\dagger 2}) \right]}_{\text{for "2".}} \end{aligned} \quad (67)$$

and  $\exp M_3$  is the operator term which is coming due to the interaction between both the fields but we will introduce the coupling constant  $K$  as an overall factor in the exponential which will help us in doing our further analysis perturbatively in the limit  $K \ll 1$ .

$$\exp\{K M_3\} = \exp\{K(B_1 + B_2 + B_3 + B_4 + B_5 + B_6)\} \quad (68)$$

We will introduce the coupling constant  $K$  as an overall factor in the exponential in the above equation Where the various  $B$  terms in the exponential of  $M_3$  are:

$$\begin{aligned} B_1 &= \frac{r_1}{2} \left[ e^{-2i\theta_1} (b_1 b_2 + b_1^\dagger b_2^\dagger) - e^{2i\theta_1} (b_2^\dagger b_1^\dagger + b_2 b_1) \right] \\ B_2 &= \frac{r_2}{2} \left[ e^{-2i\theta_2} (b_1 b_2 + b_1^\dagger b_2^\dagger) - e^{2i\theta_2} (b_2^\dagger b_1^\dagger + b_2 b_1) \right] \\ B_3 &= [f_1(r_1, r_2, \theta_1, \theta_2) b_1 b_2^\dagger + f_2(r_1, r_2, \theta_1, \theta_2) b_1^\dagger b_2] \\ B_4 &= \frac{1}{12} \left[ (X_1(r_1, r_2, \theta_1, \theta_2) b_1 b_2 + X_2(r_1, r_2, \theta_1, \theta_2) b_1^\dagger b_2^\dagger) \right] \\ B_5 &= X_3(r_1, r_2, \theta_1, \theta_2) (b_1^\dagger b_1^\dagger - b_2 b_2) \\ B_6 &= X_4(r_1, r_2, \theta_1, \theta_2) (b_1 b_1 - b_2^\dagger b_2^\dagger) \end{aligned} \quad (69)$$

From the next section we will change the notations for squeezing parameters and replace  $r_1 \rightarrow R_1$ ,  $\phi_1 \rightarrow \Phi_1$ ,  $\theta_1 \rightarrow \Theta_1$  and similarly for  $r_2$ ,  $\phi_2, \theta_2$  parameters.

### III. Calculation for unitary evolution

To understand the dynamics of the two scalar fields with interaction in de Sitter background space, we construct the most generic evolution operator  $\mathcal{U}(\eta_1, \eta_2)$  which is the product of total squeezed operator and the total rotational operator for both the fields. The total rotational operator is defined as follows:

$$\mathcal{R}_{\text{Total}}(\Phi_1, \Phi_2) = \mathcal{R}_1(\Phi_1) \mathcal{R}_2(\Phi_2) \quad (70)$$

$$\begin{aligned} \mathcal{U}(\eta_1, \eta_2) &= \\ &S_1(R_1, \Theta_1) S_2(R_2, \Theta_2) S_{12}(R_1, R_2, \Theta_1, \Theta_2) \times \\ &\quad \mathcal{R}_1(\Phi_1) \mathcal{R}_2(\Phi_2) \end{aligned} \quad (71)$$

In the Heisenberg picture representation the operators can be  $\hat{v}_1(\mathbf{x}, \eta)$ ,  $\hat{\pi}_1(\mathbf{x}, \eta)$ ,  $\hat{v}_2(\mathbf{x}, \eta)$  and  $\hat{\pi}_2(\mathbf{x}, \eta)$  can be written as follows

$$\begin{aligned}
\hat{v}_1(\mathbf{x}, \eta) &= \mathcal{U}^\dagger(\eta, \eta_0) \hat{v}_1(\mathbf{x}, \eta_0) \mathcal{U}(\eta, \eta_0) \\
&= \int \frac{d^3k}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x}} \left( u_{1,\mathbf{k}}^*(\eta) b_{1,\mathbf{k}} + u_{1,-\mathbf{k}}(\eta) b_{1,-\mathbf{k}}^\dagger \right) \\
\hat{\pi}_1(\mathbf{x}, \eta) &= \mathcal{U}^\dagger(\eta, \eta_0) \hat{\pi}_1(\mathbf{x}, \eta_0) \mathcal{U}(\eta, \eta_0) \\
&= \int \frac{d^3k}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x}} \left( w_{1,\mathbf{k}}^*(\eta) b_{1,\mathbf{k}} + w_{1,-\mathbf{k}}(\eta) b_{1,-\mathbf{k}}^\dagger \right) \\
\hat{v}_2(\mathbf{x}, \eta) &= \mathcal{U}^\dagger(\eta, \eta_0) \hat{v}_2(\mathbf{x}, \eta_0) \mathcal{U}(\eta, \eta_0) \\
&= \int \frac{d^3k}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x}} \left( u_{2,\mathbf{k}}^*(\eta) b_{2,\mathbf{k}} + u_{2,-\mathbf{k}}(\eta) b_{2,-\mathbf{k}}^\dagger \right) \\
\hat{\pi}_2(\mathbf{x}, \eta) &= \mathcal{U}^\dagger(\eta, \eta_0) \hat{\pi}_2(\mathbf{x}, \eta_0) \mathcal{U}(\eta, \eta_0) \\
&= \int \frac{d^3k}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{x}} \left( w_{2,\mathbf{k}}^*(\eta) b_{2,\mathbf{k}} + w_{2,-\mathbf{k}}(\eta) b_{2,-\mathbf{k}}^\dagger \right)
\end{aligned} \tag{72}$$

In Schrödinger representation the position and momentum operators  $v_{1,\mathbf{k}}$ ,  $\pi_{1,\mathbf{k}}$  and  $v_{2,\mathbf{k}}$  for both the scalar fields in terms of annihilation and creation operator can be written in the following manner:

$$\begin{aligned}
v_{1,\mathbf{k}}(\eta) &= \frac{1}{\sqrt{2k_1}} \left( b_{1,\mathbf{k}}(\eta) + b_{1,-\mathbf{k}}^\dagger(\eta) \right) \\
\pi_{1,\mathbf{k}}(\eta) &= i\sqrt{\frac{k_1}{2}} \left( b_{1,\mathbf{k}}(\eta) - b_{1,-\mathbf{k}}^\dagger(\eta) \right) \\
v_{2,\mathbf{k}}(\eta) &= \frac{1}{\sqrt{2k_2}} \left( b_{2,\mathbf{k}}(\eta) + b_{2,-\mathbf{k}}^\dagger(\eta) \right) \\
\pi_{2,\mathbf{k}}(\eta) &= i\sqrt{\frac{k_2}{2}} \left( b_{2,\mathbf{k}}(\eta) - b_{2,-\mathbf{k}}^\dagger(\eta) \right)
\end{aligned} \tag{73}$$

Where  $b_{1,\mathbf{k}}(\eta)$ ,  $b_{1,-\mathbf{k}}(\eta)$ ,  $b_{2,\mathbf{k}}(\eta)$  and  $b_{2,-\mathbf{k}}(\eta)$  are the annihilation and creation operators in time dependent Heisenberg representation for the two coupled scalar fields in de Sitter background space and using the factorized representation of the unitary time evolution operator introduced in the Eq(71), the expression for the annihilation operator for field 1 can be written at any arbitrary time scale as:

$$\begin{aligned}
b_1(\eta) &\equiv \mathcal{U}^\dagger(\eta, \eta_0) b_2 \mathcal{U}(\eta, \eta_0) \\
&= \mathcal{R}_1^\dagger(\Phi_1) \mathcal{R}_2^\dagger(\Phi_2) S_{12}^\dagger(R_1, R_2, \Theta_1, \Theta_2) S_2^\dagger(R_2, \Theta_2) S_1^\dagger(R_1, \Theta_1) b_1 S_1(R_1, \Theta_1) S_2(R_2, \Theta_2) S_{12}(R_1, R_2, \Theta_1, \Theta_2) \\
&\times \mathcal{R}_1(\Phi_1) \mathcal{R}_2(\Phi_2) \\
&= \mathcal{R}_1^\dagger(\Phi_1) \mathcal{R}_2^\dagger(\Phi_2) ((\cosh R_1) S_{12}^\dagger b_1 S_{12} - (\sinh R_1) e^{2\Theta_1} S_{12}^\dagger b_1^\dagger S_{12}) \mathcal{R}_1(\Phi_1) \mathcal{R}_2(\Phi_2)
\end{aligned} \tag{74}$$

Using the Baker–Campbell–Hausdorff formula upto linear order in coupling constant  $K$ , the annihilation operator  $b_1$  becomes (75). The total expression is the sum of the free part which is the usual time dependent expression for the annihilation operator on applying the

squeezed operator (without interaction term) and there are extra terms with factor  $K$ , coming due to the interaction between both the fields. For the creation operator  $b_1^\dagger$  of field 1, one can take the conjugate of Eq(74) or (75).

$$\begin{aligned}
b_1(\eta) &= \left( \cosh R_1 e^{-i\Phi_1} b_1 - e^{i(\Phi_1+2\Theta_1)} \sinh R_1 b_1^\dagger \right) + K \left\{ \left( -\cosh R_1 (i(R_1 \sin 2\Theta_1 + R_2 \sin 2\Theta_2)) - \frac{X_2}{12} \right. \right. \\
&\quad \left. \left. - e^{2i\Theta_1} (\sinh R_1) (f_1) \right) e^{i\Phi_2} b_2^\dagger + \left[ (\cosh R_1) f_2 + e^{2i\Theta_1} (i(R_1 \sin 2\Theta_1 + R_2 \sin 2\Theta_2)) - \frac{X_1}{12} \right] (\sinh R_1) e^{-i\Phi_2} b_2 \right. \\
&\quad \left. + \left[ \frac{\cosh R_1}{2} (R_1 \sin 2\Theta_1 + R_2 \sin 2\Theta_2)^2 - \sinh R_1 e^{2i\Theta_1} \left( \frac{X_4}{6} \right) \right] e^{-i\Phi_1} b_1 + \left[ \cosh R_1 \left( \frac{X_3}{6} \right) - \frac{e^{2i\Theta_1}}{2} \sinh R_1 \times \right. \right. \\
&\quad \left. \left. (R_1 \sin 2\Theta_1 + R_2 \sin 2\Theta_2)^2 \right] e^{i\Phi_1} b_1^\dagger \right\} + \dots
\end{aligned} \tag{75}$$

Now for operator  $b_2$ , we will apply the same factorized representation of the unitary time evolution operator mentioned in the Eq(71), the expression for the annihila-

tion operator for field 2 can be written at any arbitrary time scale as Eq(76). For the creation operator  $b_2^\dagger$  of field 2, one can take the conjugate of Eq(76) or (77).

$$\begin{aligned}
b_2(\eta) &\equiv \mathcal{U}^\dagger(\eta, \eta_0) b_2 \mathcal{U}(\eta, \eta_0) \\
&= \mathcal{R}_1^\dagger(\Phi_1) \mathcal{R}_2^\dagger(\Phi_2) S_{12}^\dagger(R_1, R_2, \Theta_1, \Theta_2) S_2^\dagger(R_2, \Theta_2) S_1^\dagger(R_1, \Theta_1) b_2 S_1(R_1, \Theta_1) S_2(R_2, \Theta_2) S_{12}(R_1, R_2, \Theta_1, \Theta_2) \times \\
&\quad \mathcal{R}_1(\Phi_1) \mathcal{R}_2(\Phi_2) = \mathcal{R}_1^\dagger(\Phi_1) \mathcal{R}_2^\dagger(\Phi_2) ((\cosh R_1) S_{12}^\dagger b_2 S_{12} - (\sinh R_1) e^{2\Theta_1} S_{12}^\dagger b_2^\dagger S_{12}) \mathcal{R}_1(\Phi_1) \mathcal{R}_2(\Phi_2)
\end{aligned} \tag{76}$$

Using the same Baker–Campbell–Hausdorff formula upto linear order in coupling constant  $K$ , for the annihilation operator  $b_2$  becomes (77). Here also we see that the total expression is sum of the free part which is the usual time dependent expression for the annihilation operator

on applying the squeezed operator and the terms with factor  $K$ , coming due the presence of  $S(R_1, \Theta_1, R_2, \Theta_2)$  operator which is accounting for the interaction between the two fields.

$$\begin{aligned}
b_2(\eta) &= \left( \cosh R_2 e^{-i\Phi_2} b_2 - e^{i(\Phi_2+2\Theta_2)} \sinh R_2 b_2^\dagger \right) + K \left( \left\{ -i \cosh R_2 (R_1 \sin 2\Theta_1 + R_2 \sin 2\Theta_2) - \cosh R_2 \left( \frac{X_2}{12} \right) \right. \right. \\
&\quad \left. \left. - e^{2i\Theta_2} \sinh R_2 (f_2) \right\} e^{i\Phi_1} b_1^\dagger + \left( f_1 \cosh R_2 + e^{2i\Theta_2} \sinh R_2 (-i(R_1 \sin 2\Theta_1 + R_2 \sin 2\Theta_2) + \frac{X_1}{12}) \right) e^{-i\Phi_1} b_1 + \right. \\
&\quad \left. \left\{ (R_1 \sin 2\Theta_1 + R_2 \sin 2\Theta_2)^2 \frac{\cosh R_2}{2} + e^{2i\Theta_2} \sinh R_2 \left( \frac{X_3}{6} \right) \right\} e^{-i\Phi_2} b_2 + \left\{ -(R_1 \sin 2\Theta_1 + R_2 \sin 2\Theta_2)^2 \sinh R_2 \times \right. \right. \\
&\quad \left. \left. \left( \frac{e^{2i\Theta_2}}{2} \right) - \frac{X_4}{6} \cosh R_2 \right\} e^{i\Phi_2} b_2^\dagger \right) + \dots
\end{aligned} \tag{77}$$

#### IV. Evolution equations

After getting the time evolution of operators  $b_1$  and  $b_2$ , we compute the expressions for the differential equation of mode functions  $u_{1,\mathbf{k}}$ ,  $u_{2,\mathbf{k}}$ ,  $w_{1,\mathbf{k}}$  and  $w_{2,\mathbf{k}}$  for the two coupled scalar fields in de Sitter background space because it is crucial step for getting the evolution equation in terms of differential equation for squeeze factor  $R_{1,\mathbf{k}}, R_{2,\mathbf{k}}$ , squeeze phase  $\Theta_{1,\mathbf{k}}, \Theta_{2,\mathbf{k}}$  and squeeze angle  $\Phi_{1,\mathbf{k}}, \Phi_{2,\mathbf{k}}$ . For calculating the differential equation we apply the Heisenberg equation of motion Eq(78) for different position and momentum operators. Using the Heisenberg equation of motion we will find that the mode functions  $u_{1,\mathbf{k}}$ ,  $u_{2,\mathbf{k}}$ ,  $w_{1,\mathbf{k}}$  and  $w_{2,\mathbf{k}}$  satisfy the Hamilton equations Eq(79). This set of four differential equation are the classical equation of motion. There are 2 pairs one for each scalar fields the two pairs are symmetric with each other (only differ in indexing variable) but they are also coupled differential equation equation with coupling constant  $K$ . These equations tells us the dynamics of the position and momentum variables of the classical two interacting scalar field theory given by the action in Eq(1).

$$\frac{d}{dt} \mathcal{O}_i = i [\mathcal{H}, \mathcal{O}_i] \tag{78}$$

Where  $\mathcal{O}_i = \{v_1, \pi_1, v_2, \pi_2\}$ .

$$\begin{aligned}
u'_{1,\mathbf{k}} &= w_{1,\mathbf{k}} - \left( \frac{P_a^2}{2a^2} \right) (3u_{1,\mathbf{k}} + Ku_{2,\mathbf{k}}) . \\
u'_{2,\mathbf{k}} &= w_{2,\mathbf{k}} - \left( \frac{Pa^2}{2a^2} \right) (3u_{2,\mathbf{k}} + Ku_{1,\mathbf{k}}) . \\
w'_{1,\mathbf{k}} &= - \left( m_1^2 + \frac{K^2 P_a^2}{4a^4} \right) u_{1,\mathbf{k}} - \left( \frac{P_a^2}{2a^2} \right) (3w_{1,\mathbf{k}} + Kw_{2,\mathbf{k}}) \\
&\quad + \left( \frac{K P_a^2}{2a^4} \right) u_{2,\mathbf{k}} . \\
w'_{2,\mathbf{k}} &= - \left( m_2^2 + \frac{K^2 P_a^2}{4a^4} \right) u_{2,\mathbf{k}} - \left( \frac{P_a^2}{2a^2} \right) (3w_{2,\mathbf{k}} + Kw_{1,\mathbf{k}}) \\
&\quad + \left( \frac{K P_a^2}{2a^4} \right) u_{1,\mathbf{k}} .
\end{aligned} \tag{79}$$

We can notice that if we switch off the interaction strength between the two fields by setting  $K = 0$ , we will get two independent sectors of scalar fields which are symmetric and decoupled with each other. Now having these mode function equations are very useful because one can calculate the equation of motion for  $R$ ,  $\Theta$  and  $\Phi$  for both the interacting fields.

Before moving further we will give a summary of various calculations done till now. We have quantized the Hamiltonian of the two coupled scalar fields in de Sitter background space Eq(17). We have shown how the quantized Hamiltonian can be mapped to the two-mode coupled harmonic oscillator. We calculated the four-mode squeezed operator for this quantized Hamiltonian and using the four-mode squeezed operator and the total

rotation operator we have defined the evolution operator, which have  $R_{1,\mathbf{k}}$ ,  $\Theta_{1,\mathbf{k}}$ ,  $\Phi_{1,\mathbf{k}}$ ,  $R_{2,\mathbf{k}}$ ,  $\Theta_{2,\mathbf{k}}$  and  $\Phi_{2,\mathbf{k}}$  as functional variables and this evolution operator governs the properties of the vacuum state for the present cosmological setup. Using the expansion of the position and momentum operators in terms of annihilation and creation operators in Heisenberg's representation we calculated the coupled differential equation for the mode functions. From Eq(72) we get the time dependent position and momentum operators where, we have used the Heisenberg representation of  $b_{1,\mathbf{k}}$  and  $b_{2,\mathbf{k}}$ .

Here, Eq(80) consists of time dependent expression for the operators  $\hat{v}_{1,\mathbf{k}}$  and  $\hat{\pi}_{1,\mathbf{k}}$

$$\begin{aligned}
\hat{v}_{1,\mathbf{k}}(\eta) &= \frac{1}{\sqrt{2k}} \left[ b_{1,\mathbf{k}} \left( \cosh R_{1,\mathbf{k}} e^{-i\Phi_{1,\mathbf{k}}} - \sinh R_{1,\mathbf{k}} e^{-i(\Phi_{1,\mathbf{k}}+2\Theta_{1,\mathbf{k}})} + K e^{i\Phi_{1,\mathbf{k}}} \left( -\sinh R_{1,\mathbf{k}} e^{2i\Theta_{1,\mathbf{k}}} \frac{X_4}{6} + \cosh R_{1,\mathbf{k}} \frac{X_3^\dagger}{6} \right) \right) \right. \\
&\quad \left. + b_{1,-\mathbf{k}}^\dagger \left( \cosh R_{1,\mathbf{k}} e^{i\Phi_{1,\mathbf{k}}} - \sinh R_{1,\mathbf{k}} e^{i(\Phi_{1,\mathbf{k}}+2\Theta_{1,\mathbf{k}})} + K e^{i\Phi_{1,\mathbf{k}}} \left( -\sinh R_{1,\mathbf{k}} e^{2i\Theta_{1,\mathbf{k}}} \frac{X_4^\dagger}{6} + \cosh R_{1,\mathbf{k}} \frac{X_3}{6} \right) \right) \right] \\
&\quad + \dots \\
\hat{\pi}_{1,\vec{k}}(\eta) &= -i\sqrt{\frac{k}{2}} \left[ b_{1,\mathbf{k}} \left( \cosh R_{1,\mathbf{k}} e^{-i\Phi_{1,\mathbf{k}}} + \sinh R_{1,\mathbf{k}} e^{-i(\Phi_{1,\mathbf{k}}+2\Theta_{1,\mathbf{k}})} - K e^{i\Phi_{1,\mathbf{k}}} \left( \sinh R_{1,\mathbf{k}} e^{2i\Theta_{1,\mathbf{k}}} \frac{X_4}{6} + \cosh R_{1,\mathbf{k}} \frac{X_3^\dagger}{6} \right) \right) \right. \\
&\quad \left. + b_{1,-\mathbf{k}}^\dagger \left( \cosh R_{1,\mathbf{k}} e^{i\Phi_{1,\mathbf{k}}} + \sinh R_{1,\mathbf{k}} e^{i(\Phi_{1,\mathbf{k}}+2\Theta_{1,\mathbf{k}})} + K e^{i\Phi_{1,\mathbf{k}}} \left( -\sinh R_{1,\mathbf{k}} e^{2i\Theta_{1,\mathbf{k}}} \frac{X_4^\dagger}{6} - \cosh R_{1,\mathbf{k}} \frac{X_3}{6} \right) \right) \right] \\
&\quad + \dots
\end{aligned} \tag{80}$$

Here, Eq(81) consists of time dependent expression for the operators  $\hat{v}_{1,\mathbf{k}}$  and  $\hat{\pi}_{1,\mathbf{k}}$  and the “...” in Eq((80), (81)) represents the contributions from other

field which we will neglect because in Eq((72),(73)) the position and momentum operators do not contain the creation and annihilation operators of other fields.

$$\begin{aligned}
\hat{v}_{2,\mathbf{k}}(\eta) &= \frac{1}{\sqrt{2k}} \left[ b_{2,\mathbf{k}} \left( \cosh R_{2,\mathbf{k}} e^{-i\Phi_{2,\mathbf{k}}} - \sinh R_{2,\mathbf{k}} e^{-i(\Phi_{2,\mathbf{k}}+2\Theta_{2,\mathbf{k}})} + K e^{i\Phi_{2,\mathbf{k}}} \left( \sinh R_{2,\mathbf{k}} e^{2i\Theta_{2,\mathbf{k}}} \frac{X_4}{6} - \cosh R_{2,\mathbf{k}} \frac{X_3^\dagger}{6} \right) \right) \right. \\
&\quad \left. + b_{1,-\mathbf{k}}^\dagger \left( \cosh R_{2,\mathbf{k}} e^{i\Phi_{2,\mathbf{k}}} - \sinh R_{2,\mathbf{k}} e^{i(\Phi_{2,\mathbf{k}}+2\Theta_{2,\mathbf{k}})} + K e^{i\Phi_{2,\mathbf{k}}} \left( \sinh R_{2,\mathbf{k}} e^{2i\Theta_{2,\mathbf{k}}} \frac{X_4^\dagger}{6} - \cosh R_{2,\mathbf{k}} \frac{X_3}{6} \right) \right) \right] \\
&\quad + \dots \\
\hat{\pi}_{1,\vec{k}}(\eta) &= -i\sqrt{\frac{k}{2}} \left[ b_{2,\mathbf{k}} \left( \cosh R_{2,\mathbf{k}} e^{-i\Phi_{2,\mathbf{k}}} + \sinh R_{2,\mathbf{k}} e^{-i(\Phi_{2,\mathbf{k}}+2\Theta_{2,\mathbf{k}})} + K e^{i\Phi_{2,\mathbf{k}}} \left( \sinh R_{2,\mathbf{k}} e^{2i\Theta_{2,\mathbf{k}}} \frac{X_4}{6} + \cosh R_{2,\mathbf{k}} \frac{X_3^\dagger}{6} \right) \right) \right. \\
&\quad \left. + b_{1,-\mathbf{k}}^\dagger \left( \cosh R_{2,\mathbf{k}} e^{i\Phi_{2,\mathbf{k}}} + \sinh R_{2,\mathbf{k}} e^{i(\Phi_{2,\mathbf{k}}+2\Theta_{2,\mathbf{k}})} + K e^{i\Phi_{2,\mathbf{k}}} \left( \sinh R_{2,\mathbf{k}} e^{2i\Theta_{2,\mathbf{k}}} \frac{X_4^\dagger}{6} + \cosh R_{2,\mathbf{k}} \frac{X_3}{6} \right) \right) \right] \\
&\quad + \dots
\end{aligned} \tag{81}$$

On comparing the time dependent form of operators for position and momentum for both the fields given in

Eq(80),(81) with (72) we can identify the mode functions to be



$$\begin{aligned}
u_{1,\mathbf{k}}(\eta) &= \frac{1}{\sqrt{2k}} \left[ \cosh R_{1,\mathbf{k}} e^{i\Phi_{1,\mathbf{k}}} - \sinh R_{1,\mathbf{k}} e^{i(\Phi_{1,\mathbf{k}}+2\Theta_{1,\mathbf{k}})} + K e^{i\Phi_{1,\mathbf{k}}} \left( -\sinh R_{1,\mathbf{k}} e^{2i\Theta_{1,\mathbf{k}}} \frac{X_4^\dagger}{6} + \cosh R_{1,\mathbf{k}} \frac{X_3}{6} \right) \right] \\
w_{1,\mathbf{k}}(\eta) &= i\sqrt{\frac{k}{2}} \left[ \cosh R_{1,\mathbf{k}} e^{i\Phi_{1,\mathbf{k}}} + \sinh R_{1,\mathbf{k}} e^{i(\Phi_{1,\mathbf{k}}+2\Theta_{1,\mathbf{k}})} + K e^{i\Phi_{1,\mathbf{k}}} \left( -\sinh R_{1,\mathbf{k}} e^{2i\Theta_{1,\mathbf{k}}} \frac{X_4^\dagger}{6} - \cosh R_{1,\mathbf{k}} \frac{X_3}{6} \right) \right] \\
u_{2,\mathbf{k}}(\eta) &= \frac{1}{\sqrt{2k}} \left[ \cosh R_{2,\mathbf{k}} e^{i\Phi_{2,\mathbf{k}}} - \sinh R_{2,\mathbf{k}} e^{i(\Phi_{2,\mathbf{k}}+2\Theta_{2,\mathbf{k}})} + K e^{i\Phi_{2,\mathbf{k}}} \left( \sinh R_{2,\mathbf{k}} e^{2i\Theta_{2,\mathbf{k}}} \frac{X_4^\dagger}{6} - \cosh R_{2,\mathbf{k}} \frac{X_3}{6} \right) \right] \\
w_{2,\mathbf{k}}(\eta) &= i\sqrt{\frac{k}{2}} \left[ \cosh R_{2,\mathbf{k}} e^{i\Phi_{2,\mathbf{k}}} + \sinh R_{2,\mathbf{k}} e^{i(\Phi_{2,\mathbf{k}}+2\Theta_{2,\mathbf{k}})} + K e^{i\Phi_{2,\mathbf{k}}} \left( \sinh R_{2,\mathbf{k}} e^{2i\Theta_{2,\mathbf{k}}} \frac{X_4^\dagger}{6} + \cosh R_{2,\mathbf{k}} \frac{X_3}{6} \right) \right]
\end{aligned} \tag{82}$$

and these above equations define the transformation between the variables which are in the Schrödinger representation with the mode functions in the Heisenberg representation. It is now a matter of algebra to show that Hamilton's equations for the mode functions (79) give the equations of motion for

$$\begin{aligned}
R'_{1,\mathbf{k}} &= \lambda_{1,\mathbf{k}} \cos 2(\varphi_{1,\mathbf{k}} - \Theta_{1,\mathbf{k}}) + KY_1 \\
\Theta'_{1,\mathbf{k}} &= -\Omega_{1,\mathbf{k}} + \frac{\lambda_{1,\mathbf{k}}}{2} (\tanh R_{1,\mathbf{k}} + \coth R_{1,\mathbf{k}}) \sin 2(\varphi_{1,\mathbf{k}} - \Theta_{1,\mathbf{k}}) \\
&\quad + KY_2 \\
\Phi'_{1,\mathbf{k}} &= \Omega_{1,\mathbf{k}} - \lambda_{1,\mathbf{k}} \tanh R_{1,\mathbf{k}} \sin 2(\varphi_{1,\mathbf{k}} - \Theta_{1,\mathbf{k}}) + KY_3 \\
R'_{2,\mathbf{k}} &= \lambda_{2,\mathbf{k}} \cos 2(\varphi_{2,\mathbf{k}} - \Theta_{2,\mathbf{k}}) + KY_4 \\
\Theta'_{2,\mathbf{k}} &= -\Omega_{2,\mathbf{k}} + \frac{\lambda_{2,\mathbf{k}}}{2} (\tanh R_{2,\mathbf{k}} + \coth R_{2,\mathbf{k}}) \sin 2(\varphi_{2,\mathbf{k}} - \Theta_{2,\mathbf{k}}) \\
&\quad + KY_5 \\
\Phi'_{2,\mathbf{k}} &= \Omega_{2,\mathbf{k}} - \lambda_{2,\mathbf{k}} \tanh R_{2,\mathbf{k}} \sin 2(\varphi_{2,\mathbf{k}} - \Theta_{2,\mathbf{k}}) + KY_6
\end{aligned} \tag{83}$$

Where  $Y_i$ 's are the contributions coming from the interaction between the two scalar fields. Here we have considered the effects of perturbation terms only upto  $\mathcal{O}(R^2)$ .  $Y_i$ 's are given in the appendix. VII A.

and

effects the evolution operator for the four mode squeezed states in de Sitter background space. Hence, these equations can be used for numerical studies of the two coupled scalar fields in de Sitter space.

$$\Omega_{i,\mathbf{k}} = \frac{k}{2} (1 + c_{si}^2) \tag{84}$$

$$\lambda_{i,\mathbf{k}} = \left[ \left( \frac{k}{2} (1 - c_{si}^2) \right)^2 + \left( \frac{3P_a^2}{2a^2} \right)^2 \right]^{\frac{1}{2}} \tag{85}$$

$$\varphi_{i,\mathbf{k}} = -\frac{\pi}{2} + \frac{1}{2} \arctan \left( \frac{ka^2}{3P_a^2} (1 - c_{si}^2) \right). \tag{86}$$

The index  $i$  runs from 1 and 2. Here the  $c_{si}$  is the effective sound speed for two fields and is given by

$$k^2 c_{si}^2 = \left[ (m_i)^2 + \left( \frac{K^2 P_a^2}{4a^4} \right) \right] \tag{87}$$

Note: We will not consider term with  $K^2$  because we are considering terms which are first order in  $K$ .

The differential equations given in Eq[(83)] are very useful because these equations govern the dynamics of the variables:  $R_{1,\mathbf{k}}$ ,  $\Theta_{1,\mathbf{k}}$ ,  $\Phi_{1,\mathbf{k}}$ ,  $R_{2,\mathbf{k}}$ ,  $\Theta_{2,\mathbf{k}}$  and  $\Phi_{2,\mathbf{k}}$  which

## V. Numerical Analysis

In this Section we present the numerical analysis for two coupled scalar fields which are weakly interacting with coupling strength  $K \ll 1$  and we will do the numerical analysis in the limit where the interaction terms are highly linear in nature. We will study the behaviour of squeezing parameter with respect to the conformal time for this particular setup. We give the differential equation for squeezed parameters of the two coupled scalar fields in de Sitter background space in eq (88). The interacting parts in these set of six equations contain only terms which are linear in nature. We have chosen this limit in order to understand the numerical behaviour of squeezing parameters in a compact manner. For numerical analysis we set the momentum  $p_a = 0.1$ , The Hubble parameter is set to 0.1 in natural units, we take  $k = 0.0001 \text{Mpc}^{-1}$  and  $a = -\frac{1}{H\eta}$ .

$$\begin{aligned}
R'_{1,\mathbf{k}}(\eta) &= \lambda_{1,\mathbf{k}}(\eta) \cos[2(\varphi_{1,\mathbf{k}}(\eta) + \Theta_{1,\mathbf{k}}(\eta))] + \frac{K}{6k} \left( i(2kP(\eta)R_{1,\mathbf{k}}(\eta) + (-2kP(\eta) + 3iS(\eta))R_{2,\mathbf{k}}(\eta) - 3S(\eta)(i + \Phi_{1,\mathbf{k}}(\eta) - \Phi_{2,\mathbf{k}}(\eta))) \right) \\
R'_{2,\mathbf{k}}(\eta) &= \lambda_{2,\mathbf{k}}(\eta) \cos[2(\varphi_{2,\mathbf{k}}(\eta) + \Theta_{2,\mathbf{k}}(\eta))] + K \left( \frac{1}{3} \left( P(\eta) - P(\eta)R_{1,\mathbf{k}}(\eta) + P(\eta)R_{2,\mathbf{k}}(\eta) + \frac{3S(\eta)\Theta_{2,\mathbf{k}}(\eta)}{k} \right) \right) \\
\Theta'_{1,\mathbf{k}}(\eta) &= -\Omega_{1,\mathbf{k}} - (\lambda_{1,\mathbf{k}} \sin[2(\varphi_{1,\mathbf{k}} + \Theta_{1,\mathbf{k}}(\eta))]) \coth[2R_{1,\mathbf{k}}(\eta)] + K \left( \frac{1}{12kR_{1,\mathbf{k}}(\eta)} (3S(\eta) + (-2ikP(\eta) - 3S(\eta))R_{2,\mathbf{k}}(\eta) - 4kP(\eta)\Theta_{1,\mathbf{k}}(\eta) - 2kP(\eta)\Phi_{1,\mathbf{k}}(\eta) - 3iS(\eta)\Phi_{1,\mathbf{k}}(\eta) + 2kP(\eta)\Phi_{2,\mathbf{k}}(\eta) + 3iS(\eta)\Phi_{2,\mathbf{k}}(\eta) + R_{1,\mathbf{k}}(\eta)(2ikP(\eta) - 3S(\eta) - 4kP(\eta)\Theta_{1,\mathbf{k}}(\eta) + 6iS(\eta)\Theta_{1,\mathbf{k}}(\eta) + R_{2,\mathbf{k}}(\eta)(3S(\eta) + 8kP(\eta)\Theta_{1,\mathbf{k}}(\eta) + 3iS(\eta)\Phi_{1,\mathbf{k}}(\eta) - 3iS(\eta)\Phi_{2,\mathbf{k}}(\eta))) \right) \\
\Theta'_{2,\mathbf{k}}(\eta) &= -\Omega_{2,\mathbf{k}} - (\lambda_{2,\mathbf{k}} \sin[2(\varphi_{2,\mathbf{k}} + \Theta_{2,\mathbf{k}}(\eta))]) \coth[2R_{2,\mathbf{k}}(\eta)] + K \left( \frac{1}{12kR_{2,\mathbf{k}}(\eta)} (3S(\eta) + R_{1,\mathbf{k}}(\eta)(-2ikP(\eta) - 3S(\eta) + 3S(\eta)R_{2,\mathbf{k}}(\eta)) - 4kP(\eta)\Theta_{2,\mathbf{k}}(\eta) + 2kP(\eta)\Phi_{1,\mathbf{k}}(\eta) + 3iS(\eta)\Phi_{1,\mathbf{k}}(\eta) - 2kP(\eta)\Phi_{2,\mathbf{k}}(\eta) - 3iS(\eta)\Phi_{2,\mathbf{k}}(\eta) + R_{2,\mathbf{k}}(\eta)(2ikP(\eta) - 3S(\eta) - 4kP(\eta)\Theta_{2,\mathbf{k}}(\eta) + 6iS(\eta)\Theta_{2,\mathbf{k}}(\eta) - 3iS(\eta)\Phi_{1,\mathbf{k}}(\eta) + 3iS(\eta)\Phi_{2,\mathbf{k}}(\eta))) \right) \\
\Phi'_{1,\mathbf{k}}(\eta) &= \Omega_{1,\mathbf{k}} - \lambda_{1,\mathbf{k}} \tanh[R_{1,\mathbf{k}}(\eta)] \sin[2(\varphi_{1,\mathbf{k}} + \Theta_{1,\mathbf{k}}(\eta))] \\
&\quad + K \left( \frac{i(2kP(\eta)R_{1,\mathbf{k}}(\eta) + (-2kP(\eta) + 3iS(\eta))R_{2,\mathbf{k}}(\eta) - 3S(\eta)(i + \Phi_{1,\mathbf{k}}(\eta) - \Phi_{2,\mathbf{k}}(\eta)))}{6k} \right) \\
\Phi'_{2,\mathbf{k}}(\eta) &= \Omega_{2,\mathbf{k}} - \lambda_{2,\mathbf{k}} \tanh[R_{2,\mathbf{k}}(\eta)] \sin[2(\varphi_{2,\mathbf{k}} + \Theta_{2,\mathbf{k}}(\eta))] \\
&\quad - K \left( \frac{i((2kP(\eta) - 3iS(\eta))R_{1,\mathbf{k}}(\eta) - 2kP(\eta)R_{2,\mathbf{k}}(\eta) + 3S(\eta)(i - \Phi_{1,\mathbf{k}}(\eta) + \Phi_{2,\mathbf{k}}(\eta)))}{6k} \right)
\end{aligned} \tag{88}$$

In fig (1), we have plotted three graphs which shows the behaviour of squeezing parameters with respect to the  $\eta$  for the two coupled scalar fields Fig (1(a)) represents behaviour of the squeezing amplitude  $R_{1,\mathbf{k}}$  for first field and  $R_{2,\mathbf{k}}$  for second field with respect to the conformal time  $\eta$ . We observe that the squeezing amplitude at early times for the two fields are different which is due to the reason that the contribution from the interacting terms in both the field is different. We see that the  $R_{1,\mathbf{k}}$  dominates over  $R_{2,\mathbf{k}}$  at early times and behaviour for both the squeezing amplitude is similar i.e, both the squeezing amplitudes decrease as we move towards the more negative value of the conformal time  $\eta$  and at  $\eta \approx 0$  which is the present time of the universe the distinguishability between the squeezing amplitude for the two fields vanishes.

In fig (1(b)), we have plotted the squeezing angle or rotation angle,  $\Theta_{1,\mathbf{k}}$  for first field and  $\Theta_{2,\mathbf{k}}$  for second field. The squeezing angle for the two fields at early time or more negative value of conformal time shows that squeezing angle of second field is dominating over the first field but as we move towards the right of the graph the field one dominates and decreases and then the second field dominates. But when we move toward the present time  $\eta = 0$  we see that the these interchanging dominance effects vanishes and squeezing angles for both fields saturates to a particular constant value.

The behaviour for the squeezing phase parameters with

respect to the conformal time  $\eta$  has been shown in fig (1(c)). Here squeezing phase  $\Phi_{1,\mathbf{k}}$  is for first field and  $\Phi_{2,\mathbf{k}}$  is for second field. We observe that its nature is similar to that of squeezing amplitude parameter shown in fig(1(a)), but there is one difference which is the dominance of squeezing phase of second field over the first field, which was another way around in case of squeezing amplitude parameter. We see that in case of squeezing phase the difference in the two fields vanishes and attains a constant value.

In fig (2) we again plot the squeezing parameters but with a different value of effective sound speed parameter  $c_{s1}$  and  $c_{s2}$ . We set the effective sound speed parameter for both fields to its lowest bound which is 0.024. Here the coupling strength  $K = 0.1$ . We observe a similar behaviour for all the three squeezing parameters but one new observation which we can see in fig (2(a)) and fig(2(c)) is that the squeezing amplitude and squeezing phase tends to zero as we move towards more negative value of conformal time. All the three graphs in fig (2) gets shifted towards right or near the origin as compared to the graphs in fig (1).

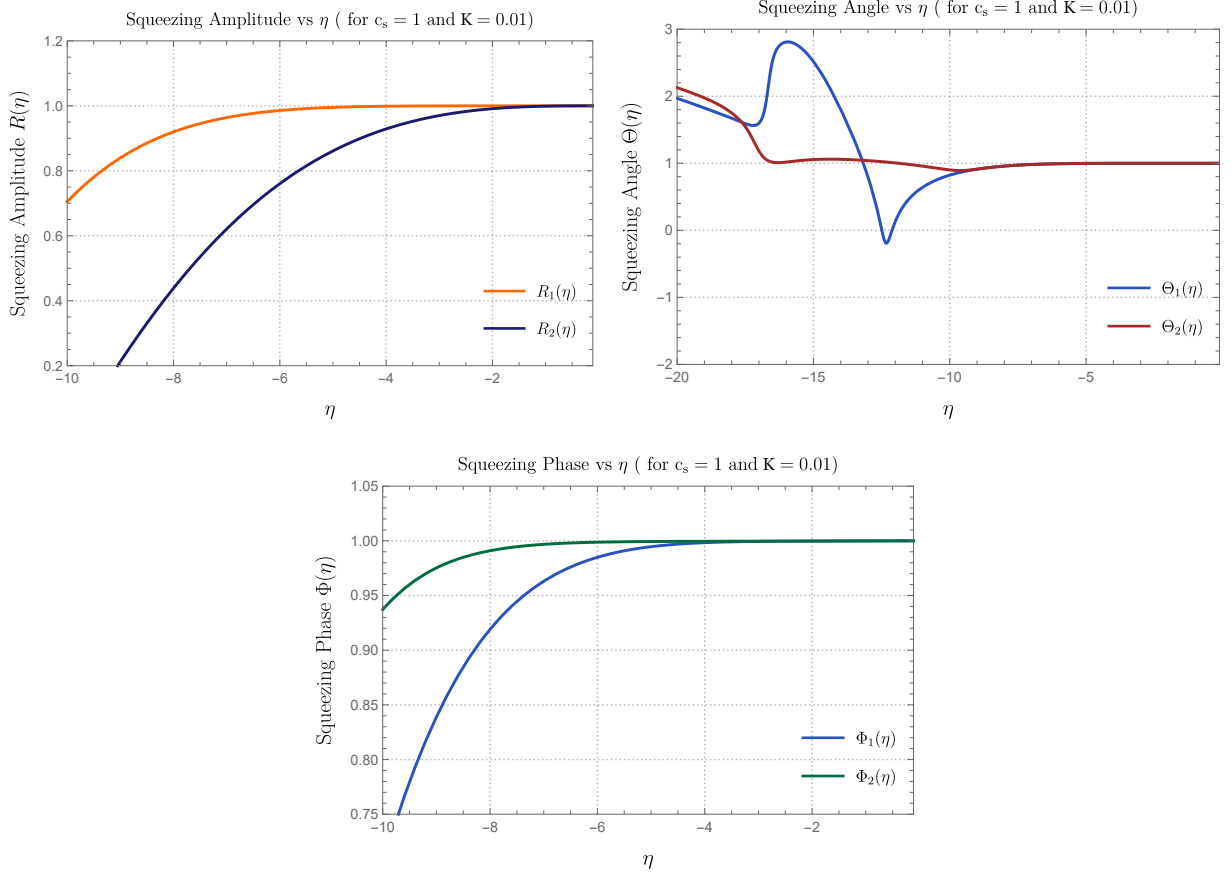


FIG. 1. Behaviour of the squeezing parameters, such as squeezing amplitude in fig (2(a)), squeezing angle in fig (2(b)), squeezing phase in fig (2(c)) with respect to the conformal time  $\eta$  for two coupled scalar fields with coupling strength  $K = 0.01$  and the effective sound speed parameter  $c_{s1}$  and  $c_{s2}$  for both the fields is equal to 1.

## VI. Conclusion

This work dealt with the mathematical formalism of four mode squeezed state in cosmology. From the cosmological perspective we have encountered two scalar fields, where the metric of the background space was set to de Sitter. The motivation for choosing this particular metric was the following setup of the action which we have chosen can be helpful to understand the dynamics of the two scalar fields in FRW cosmological universe. The analysis of our present work can be helpful in the limit where the two scalar fields are weakly coupled  $K \ll 1$  and the perturbation effects was taken only upto linear order  $\mathcal{O}(R)$ . This work could be summarised in following points:

- We have quantized the modes of the two coupled scalar fields  $\mu_1$  and  $\mu_2$ . We have also calculated the position and momentum variables for the same in de Sitter background space and obtained the quantized Hamiltonian  $H$ . We made connection between the two coupled inverted quantum harmonic oscillator system and four mode squeezed state for-

malism.

- We have given a detailed calculation for constructing the four mode squeezed state operator which is useful for understanding the cosmological action for the two interacting scalar fields and also for other systems which can be explained in terms of two coupled inverted quantum harmonic oscillators.
- The time evolution operator for the four mode squeezed states is also given, which we have used to calculate the time dependent (Heisenberg picture) annihilation and creation operators for two coupled scalar fields in de Sitter background space. Using the Heisenberg equation of motion, we have calculated the coupled differential equation for the mode functions of the two coupled scalar fields in de Sitter space.
- We presented the expression for  $R_{1,\mathbf{k}}$ ,  $\Theta_{1,\mathbf{k}}$ ,  $\Phi_{1,\mathbf{k}}$ ,  $R_{2,\mathbf{k}}$ ,  $\Theta_{2,\mathbf{k}}$  and  $\Phi_{2,\mathbf{k}}$  which are the parameters of the evolution operator for four mode squeezed state

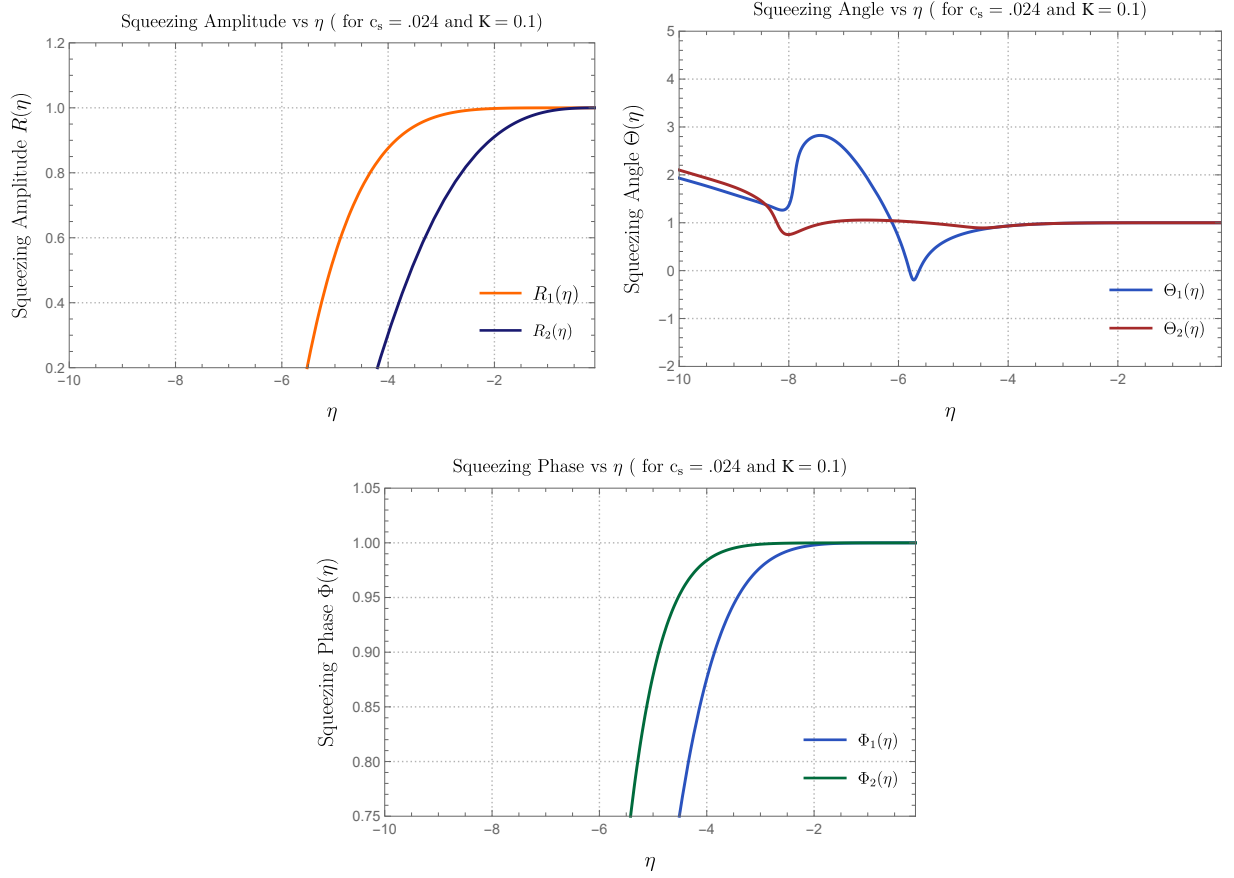


FIG. 2. Behaviour of the squeezing parameters, such as squeezing amplitude in fig (1(a)), squeezing angle in fig (1(b)), squeezing phase in fig (1(c)) with respect to the conformal time  $\eta$  for two coupled scalar fields with coupling strength  $K = 0.1$  and the effective sound speed parameter  $c_{s1}$  and  $c_{s2}$  for both the fields is equal to 0.024.

and which governs the evolution of the state for two coupled scalar fields in de Sitter metric.

- We conclude the work by analysing the behaviour of the squeezing parameters for the two coupled scalar fields. We found during early conformal time, the distinction between the squeezing parameters were noticeable. As we move to current time,  $\eta \approx 0$ , we find that the parameters converge and become indistinguishable.

With this tools in hand it would be interesting to compute quantum information quantities such as entanglement entropy, quantum discord, circuit complexity and many more quantum information theoretic measures for two coupled scalar fields in de Sitter space. This will going to help us to know about the feature and the behaviour of the long range quantum correlations for the system under consideration. Earlier the formalism was not developed for the four-mode squeezed states for which these crucial aspects was not studied in the pre-

vious works. Now since the formalism is developed and we know how to handle the system numerically, it would be really good to study the mentioned aspects in near future.

**Acknowledgement:** The Visiting Post Doctoral research fellowship of SC is supported by the J. C. Bose National Fellowship of Director, Professor Rajesh Gopakumar, ICTS, TIFR, Bengaluru. The research of SP is supported by the J. C. Bose National Fellowship. SC also would like to thank ICTS, TIFR, Bengaluru for providing the work friendly environment. SC also would like to thank all the members of our newly formed virtual international non-profit consortium Quantum Structures of the Space-Time & Matter (QASTM) for for elaborative discussions. AR and NP would like to thank the members of the QASTM Forum for useful discussions. Last but not least, we would like to acknowledge our debt to the people belonging to the various part of the world for their generous and steady support for research in natural sciences.

## VII. Appendix

### A. Interacting part of differential equations

$Y_1 =$

$$\begin{aligned} & (- (A_3 B_2 E_1 - A_2 B_3 E_1 - A_3 B_1 E_2 + A_1 B_3 E_2 + A_2 B_1 E_3 - A_1 B_2 E_3) (C_6 D_5 F_4 - C_5 D_6 F_4 - C_6 D_4 F_5 + C_4 D_6 F_5 \\ & + C_5 D_4 F_6 - C_4 D_5 F_6) (B_2 E_1 a_0 - B_1 E_2 a_0 - A_2 E_1 b_0 + A_1 E_2 b_0 + A_2 B_1 e_0 - A_1 B_2 e_0) + (x_{21} A_2 B_1 - x_{21} A_1 B_2 - x_0 A_2 E_1 \\ & + x_3 B_2 E_1 + x_0 A_1 E_2 - x_3 B_1 E_2) (A_3 B_2 E_1 - A_2 B_3 E_1 - A_3 B_1 E_2 + A_1 B_3 E_2 + A_2 B_1 E_3 - A_1 B_2 E_3) (D_6 F_5 C_0 - D_5 F_6 C_0 \\ & - C_6 F_5 D_0 + C_5 F_6 D_0 + C_6 D_5 F_0 - C_5 D_6 F_0) R_1^2 + (C_6 D_5 F_4 - C_5 D_6 F_4 - C_6 D_4 F_5 + C_4 D_6 F_5 + C_5 D_4 F_6 - C_4 D_5 F_6) \\ & (B_3 E_2 A_0 - B_2 E_3 A_0 - A_3 E_2 B_0 + A_2 E_3 B_0 + A_3 B_2 E_0 - A_2 B_3 E_0) R_2 (-x_7 A_2 E_1 + x_7 A_1 E_2 + (-x_{18} A_2 B_1 + x_{18} A_1 B_2 \\ & - x_6 A_2 E_1 + x_2 B_2 E_1 + x_6 A_1 E_2 - x_2 B_1 E_2) R_2) R_1 (x_5 (A_2 E_1 - A_1 E_2) (C_6 D_5 F_4 - C_5 D_6 F_4 - C_6 D_4 F_5 + C_4 D_6 F_5 + C_5 D_4 F_6 \\ & - C_4 D_5 F_6) (-B_3 E_2 A_0 + B_2 E_3 A_0 + A_3 E_2 B_0 - A_2 E_3 B_0 - A_3 B_2 E_0 + A_2 B_3 E_0) + (-x_{19} (A_2 B_1 - A_1 B_2) (C_6 D_5 F_4 \\ & - C_5 D_6 F_4 - C_6 D_4 F_5 + C_4 D_6 F_5 + C_5 D_4 F_6 - C_4 D_5 F_6) (-B_3 E_2 A_0 + B_2 E_3 A_0 + A_3 E_2 B_0 - A_2 E_3 B_0 - A_3 B_2 E_0 \\ & + A_2 B_3 E_0) - x_1 (B_2 E_1 - B_1 E_2) (C_6 D_5 F_4 - C_5 D_6 F_4 - C_6 D_4 F_5 + C_4 D_6 F_5 + C_5 D_4 F_6 - C_4 D_5 F_6) (-B_3 E_2 A_0 + B_2 E_3 A_0 \\ & + A_3 E_2 B_0 - A_2 E_3 B_0 - A_3 B_2 E_0 + A_2 B_3 E_0) - (x_{20} A_2 B_1 - x_{20} A_1 B_2 + x_8 A_2 E_1 - x_4 B_2 E_1 - x_8 A_1 E_2 + x_4 B_1 E_2) \\ & (A_3 B_2 E_1 - A_2 B_3 E_1 - A_3 B_1 E_2 + A_1 B_3 E_2 + A_2 B_1 E_3 - A_1 B_2 E_3) (D_6 F_5 C_0 - D_5 F_6 C_0 - C_6 F_5 D_0 + C_5 F_6 D_0 + C_6 D_5 F_0 \\ & - C_5 D_6 F_0) R_2) ) / ((A_3 B_2 E_1 - A_2 B_3 E_1 - A_3 B_1 E_2 + A_1 B_3 E_2 + A_2 B_1 E_3 - A_1 B_2 E_3)^2 (C_6 D_5 F_4 - C_5 D_6 F_4 - C_6 D_4 F_5 \\ & + C_4 D_6 F_5 + C_5 D_4 F_6 - C_4 D_5 F_6)) \end{aligned}$$

$Y_4 =$

$$\begin{aligned} & ((-C_6 D_5 F_4 + C_5 D_6 F_4 + C_6 D_4 F_5 - C_4 D_6 F_5 - C_5 D_4 F_6 + C_4 D_5 F_6) (B_3 ((-A_2 E_1 + A_1 E_2) (D_6 F_5 - D_5 F_6) c_0 + \\ & C_6 d_1 F_5 (E_2 A_0 - A_2 E_0) + C_5 d_1 F_6 (-E_2 A_0 + A_2 E_0) + C_6 (A_2 E_1 - A_1 E_2) (F_5 d_0 - D_5 f_0) - C_5 (A_2 E_1 - A_1 E_2) \\ & (F_6 d_0 - D_6 f_0)) + B_2 ((A_3 E_1 - A_1 E_3) (D_6 F_5 - D_5 F_6) c_0 + C_5 d_1 F_6 (E_3 A_0 - A_3 E_0) + C_6 d_1 F_5 (-E_3 A_0 + A_3 E_0) - \\ & C_6 (A_3 E_1 - A_1 E_3) (F_5 d_0 - D_5 f_0) + C_5 (A_3 E_1 - A_1 E_3) (F_6 d_0 - D_6 f_0)) + (A_3 E_2 - A_2 E_3) (B_1 (-D_6 F_5 + D_5 F_6) c_0 \\ & - C_6 (d_1 F_5 B_0 - B_1 F_5 d_0 + B_1 D_5 f_0) + C_5 (d_1 F_6 B_0 - B_1 F_6 d_0 + B_1 D_6 f_0))) + (A_3 B_2 E_1 - A_2 B_3 E_1 - A_3 B_1 E_2 \\ & + A_1 B_3 E_2 + A_2 B_1 E_3 - A_1 B_2 E_3) (x_{25} C_6 D_5 - x_{25} C_5 D_6 - x_{17} C_6 F_5 + x_{12} D_6 F_5 + x_{17} C_5 F_6 - x_{12} D_5 F_6) (D_6 F_5 C_0 \\ & - D_5 F_6 C_0 - C_6 F_5 D_0 + C_5 F_6 D_0 + C_6 D_5 F_0 - C_5 D_6 F_0) R_1^2 + (x_{23} (C_6 D_5 - C_5 D_6) (C_6 D_5 F_4 - C_5 D_6 F_4 - C_6 D_4 F_5 \\ & + C_4 D_6 F_5 + C_5 D_4 F_6 - C_4 D_5 F_6) (B_3 E_2 A_0 - B_2 E_3 A_0 - A_3 E_2 B_0 + A_2 E_3 B_0 + A_3 B_2 E_0 - A_2 B_3 E_0) + x_{11} (D_6 F_5 - D_5 F_6) \\ & (C_6 D_5 F_4 - C_5 D_6 F_4 - C_6 D_4 F_5 + C_4 D_6 F_5 + C_5 D_4 F_6 - C_4 D_5 F_6) (B_3 E_2 A_0 - B_2 E_3 A_0 - A_3 E_2 B_0 + A_2 E_3 B_0 + A_3 B_2 E_0 \\ & - A_2 B_3 E_0) - (A_3 B_2 E_1 - A_2 B_3 E_1 - A_3 B_1 E_2 + A_1 B_3 E_2 + A_2 B_1 E_3 - A_1 B_2 E_3) (x_{26} C_6 D_5 - x_{26} C_5 D_6 + x_{16} C_6 F_5 \\ & - x_{13} D_6 F_5 - x_{16} C_5 F_6 + x_{13} D_5 F_6) (D_6 F_5 C_0 - D_5 F_6 C_0 - C_6 F_5 D_0 + C_5 F_6 D_0 + C_6 D_5 F_0 - C_5 D_6 F_0) R_1 R_2 \\ & + (x_{24} C_6 D_5 - x_{24} C_5 D_6 + x_{10} D_6 F_5 - x_{10} D_5 F_6) (C_6 D_5 F_4 - C_5 D_6 F_4 - C_6 D_4 F_5 + C_4 D_6 F_5 + C_5 D_4 F_6 - C_4 D_5 F_6) \\ & (B_3 E_2 A_0 - B_2 E_3 A_0 - A_3 E_2 B_0 + A_2 E_3 B_0 + A_3 B_2 E_0 - A_2 B_3 E_0) R_2^2) / ((A_3 B_2 E_1 - A_2 B_3 E_1 - A_3 B_1 E_2 \\ & + A_1 B_3 E_2 + A_2 B_1 E_3 - A_1 B_2 E_3) (C_6 D_5 F_4 - C_5 D_6 F_4 - C_6 D_4 F_5 + C_4 D_6 F_5 + C_5 D_4 F_6 - C_4 D_5 F_6)^2) \end{aligned}$$

$Y_2 =$

$$\begin{aligned} & ((A_3 B_2 E_1 - A_2 B_3 E_1 - A_3 B_1 E_2 + A_1 B_3 E_2 + A_2 B_1 E_3 - A_1 B_2 E_3) (B_3 E_1 a_0 - B_1 E_3 a_0 - A_3 E_1 b_0 + A_1 E_3 b_0 \\ & + A_3 B_1 e_0 - A_1 B_3 e_0) + ((x_{21} A_3 B_1 - x_{21} A_1 B_3 - x_0 A_3 E_1 + x_3 B_3 E_1 + x_0 A_1 E_3 - x_3 B_1 E_3) (-A_3 B_2 E_1 + A_2 B_3 E_1 \\ & + A_3 B_1 E_2 - A_1 B_3 E_2 - A_2 B_1 E_3 + A_1 B_2 E_3) (D_6 F_5 C_0 - D_5 F_6 C_0 - C_6 F_5 D_0 + C_5 F_6 D_0 + C_6 D_5 F_0 - C_5 D_6 F_0) R_1^2) / \\ & (C_6 D_5 F_4 - C_5 D_6 F_4 - C_6 D_4 F_5 + C_4 D_6 F_5 + C_5 D_4 F_6 - C_4 D_5 F_6) + x_7 (A_3 E_1 - A_1 E_3) (B_3 E_2 A_0 - B_2 E_3 A_0 - A_3 E_2 B_0 \\ & + A_2 E_3 B_0 + A_3 B_2 E_0 - A_2 B_3 E_0) R_2 + (x_{18} A_3 B_1 - x_{18} A_1 B_3 + x_6 A_3 E_1 - x_2 B_3 E_1 - x_6 A_1 E_3 + x_2 B_1 E_3) (B_3 E_2 A_0 - B_2 E_3 A_0 \\ & - A_3 E_2 B_0 + A_2 E_3 B_0 + A_3 B_2 E_0 - A_2 B_3 E_0) R_2^2 + R_1 (x_5 (A_3 E_1 - A_1 E_3) (B_3 E_2 A_0 - B_2 E_3 A_0 - A_3 E_2 B_0 + A_2 E_3 B_0 \\ & + A_3 B_2 E_0 - A_2 B_3 E_0) + (-x_1 (B_3 E_1 - B_1 E_3) (B_3 E_2 A_0 - B_2 E_3 A_0 - A_3 E_2 B_0 + A_2 E_3 B_0 + A_3 B_2 E_0 - A_2 B_3 E_0) \\ & + x_{19} (A_3 B_1 - A_1 B_3) (-B_3 E_2 A_0 + B_2 E_3 A_0 + A_3 E_2 B_0 - A_2 E_3 B_0 - A_3 B_2 E_0 + A_2 B_3 E_0) - ((x_{20} A_3 B_1 - x_{20} A_1 B_3 \\ & + x_8 A_3 E_1 - x_4 B_3 E_1 - x_8 A_1 E_3 + x_4 B_1 E_3) (-A_3 B_2 E_1 + A_2 B_3 E_1 + A_3 B_1 E_2 - A_1 B_3 E_2 - A_2 B_1 E_3 + A_1 B_2 E_3) \\ & (D_6 F_5 C_0 - D_5 F_6 C_0 - C_6 F_5 D_0 + C_5 F_6 D_0 + C_6 D_5 F_0 - C_5 D_6 F_0) / (C_6 D_5 F_4 - C_5 D_6 F_4 - C_6 D_4 F_5 + C_4 D_6 F_5 \\ & + C_5 D_4 F_6 - C_4 D_5 F_6) R_2) ) / (A_3 B_2 E_1 - A_2 B_3 E_1 - A_3 B_1 E_2 + A_1 B_3 E_2 + A_2 B_1 E_3 - A_1 B_2 E_3)^2 \end{aligned}$$

$Y_5 =$

$$\begin{aligned} & ((C_6 D_5 F_4 - C_5 D_6 F_4 - C_6 D_4 F_5 + C_4 D_6 F_5 + C_5 D_4 F_6 - C_4 D_5 F_6) (B_3 ((-A_2 E_1 + A_1 E_2) (D_6 F_4 - D_4 F_6) c_0 + C_6 d_1 F_4 \\ & (E_2 A_0 - A_2 E_0) + C_4 d_1 F_6 (-E_2 A_0 + A_2 E_0) + C_6 (A_2 E_1 - A_1 E_2) (F_4 d_0 - D_4 f_0) - C_4 (A_2 E_1 - A_1 E_2) (F_6 d_0 - D_6 f_0)) \\ & + B_2 ((A_3 E_1 - A_1 E_3) (D_6 F_4 - D_4 F_6) c_0 + C_4 d_1 F_6 (E_3 A_0 - A_3 E_0) + C_6 d_1 F_4 (-E_3 A_0 + A_3 E_0) - C_6 (A_3 E_1 - A_1 E_3) \\ & (F_4 d_0 - D_4 f_0) + C_4 (A_3 E_1 - A_1 E_3) (F_6 d_0 - D_6 f_0)) + (A_3 E_2 - A_2 E_3) (B_1 (-D_6 F_4 + D_4 F_6) c_0 - C_6 (d_1 F_4 B_0 - B_1 F_4 d_0 \\ & + B_1 D_4 f_0) + C_4 (d_1 F_6 B_0 - B_1 F_6 d_0 + B_1 D_6 f_0))) + (A_3 B_2 E_1 - A_2 B_3 E_1 - A_3 B_1 E_2 + A_1 B_3 E_2 + A_2 B_1 E_3 - A_1 B_2 E_3) \\ & (x_{25} C_6 D_4 - x_{25} C_4 D_6 - x_{17} C_6 F_4 + x_{12} D_6 F_4 + x_{17} C_4 F_6 - x_{12} D_4 F_6) (-D_6 F_5 C_6 + D_5 F_6 C_6 + C_6 F_5 D_0 - C_5 F_6 D_0 \\ & - C_6 D_5 F_6 + C_5 D_6 F_6) R_1^2 + (x_{23} (C_6 D_4 - C_4 D_6) (-C_6 D_5 F_4 + C_5 D_6 F_4 + C_6 D_4 F_5 - C_4 D_6 F_5 - C_5 D_4 F_6 + C_4 D_5 F_6) \end{aligned}$$

$$(B_3E_2A_0 - B_2E_3A_0 - A_3E_2B_0 + A_2E_3B_0 + A_3B_2E_0 - A_2B_3E_0) + x_{11}(D_6F_4 - D_4F_6)(-C_6D_5F_4 + C_5D_6F_4 + C_6D_4F_5 - C_4D_6F_5 - C_5D_4F_6 + C_4D_5F_6)(B_3E_2A_0 - B_2E_3A_0 - A_3E_2B_0 + A_2E_3B_0 + A_3B_2E_0 - A_2B_3E_0) - (A_3B_2E_1 - A_2B_3E_1 - A_3B_1E_2 + A_1B_3E_2 + A_2B_1E_3 - A_1B_2E_3)(x_{26}C_6D_4 - x_{26}C_4D_6 + x_{16}C_6F_4 - x_{13}D_6F_4 - x_{16}C_4F_6 + x_{13}D_4F_6)(-D_6F_5C_0 + D_5F_6C_0 + C_6F_5D_0 - C_5F_6D_0 - C_6D_5F_0 + C_5D_6F_0)R_1R_2 + (x_{24}C_6D_4 - x_{24}C_4D_6 + x_{16}D_6F_4 - x_{10}D_4F_6)(-C_6D_5F_4 + C_5D_6F_4 + C_6D_4F_5 - C_4D_6F_5 - C_5D_4F_6 + C_4D_5F_6)(B_3E_2A_0 - B_2E_3A_0 - A_3E_2B_0 + A_2E_3B_0 + A_3B_2E_0 - A_2B_3E_0)R_2^2)/((A_3B_2E_1 - A_2B_3E_1 - A_3B_1E_2 + A_1B_3E_2 + A_2B_1E_3 - A_1B_2E_3)(C_6D_5F_4 - C_5D_6F_4 - C_6D_4F_5 + C_4D_6F_5 + C_5D_4F_6 - C_4D_5F_6)^2)$$

$$Y_3 =$$

$$(-(A_3B_2E_1 - A_2B_3E_1 - A_3B_1E_2 + A_1B_3E_2 + A_2B_1E_3 - A_1B_2E_3)(C_6D_5F_4 - C_5D_6F_4 - C_6D_4F_5 + C_4D_6F_5 + C_5D_4F_6 - C_4D_5F_6)(B_2E_1a_0 - B_1E_2a_0 - A_2E_1b_0 + A_1E_2b_0 + A_2B_1e_0 - A_1B_2e_0) + (x_{21}A_2B_1 - x_{21}A_1B_2 - x_9A_2E_1 + x_3B_2E_1 + x_9A_1E_2 - x_3B_1E_2)(A_3B_2E_1 - A_2B_3E_1 - A_3B_1E_2 + A_1B_3E_2 + A_2B_1E_3 - A_1B_2E_3)(D_6F_5C_0 - D_5F_6C_0 - C_6F_5D_0 + C_5F_6D_0 + C_6D_5F_0 - C_5D_6F_0)R_1^2 + (C_6D_5F_4 - C_5D_6F_4 - C_6D_4F_5 + C_4D_6F_5 + C_5D_4F_6 - C_4D_5F_6)(B_3E_2A_0 - B_2E_3A_0 - A_3E_2B_0 + A_2E_3B_0 + A_3B_2E_0 - A_2B_3E_0)R_2(-x_7A_2E_1 + x_7A_1E_2 + (-x_{18}A_2B_1 + x_{18}A_1B_2 - x_6A_2E_1 + x_2B_2E_1 + x_6A_1E_2 - x_2B_1E_2)R_2) + R_1(x_5(A_2E_1 - A_1E_2)(C_6D_5F_4 - C_5D_6F_4 - C_6D_4F_5 + C_4D_6F_5 + C_5D_4F_6 - C_4D_5F_6)(-B_3E_2A_0 + B_2E_3A_0 + A_3E_2B_0 - A_2E_3B_0 - A_3B_2E_0 + A_2B_3E_0) + (-x_{19}(A_2B_1 - A_1B_2)(C_6D_5F_4 - C_5D_6F_4 - C_6D_4F_5 + C_4D_6F_5 + C_5D_4F_6 - C_4D_5F_6)(-B_3E_2A_0 + B_2E_3A_0 + A_3E_2B_0 - A_2E_3B_0 - A_3B_2E_0 + A_2B_3E_0) - x_1(B_2E_1 - B_1E_2)(C_6D_5F_4 - C_5D_6F_4 - C_6D_4F_5 + C_4D_6F_5 + C_5D_4F_6 - C_4D_5F_6)(-B_3E_2A_0 + B_2E_3A_0 + A_3E_2B_0 - A_2E_3B_0 - A_3B_2E_0 + A_2B_3E_0) - (x_{20}A_2B_1 - x_{20}A_1B_2 + x_8A_2E_1 - x_4B_2E_1 - x_8A_1E_2 + x_4B_1E_2)(A_3B_2E_1 - A_2B_3E_1 - A_3B_1E_2 + A_1B_3E_2 + A_2B_1E_3 - A_1B_2E_3)(D_6F_5C_0 - D_5F_6C_0 - C_6F_5D_0 + C_5F_6D_0 + C_6D_5F_0 - C_5D_6F_0)R_2))/((A_3B_2E_1 - A_2B_3E_1 - A_3B_1E_2 + A_1B_3E_2 + A_2B_1E_3 - A_1B_2E_3)^2(C_6D_5F_4 - C_5D_6F_4 - C_6D_4F_5 + C_4D_6F_5 + C_5D_4F_6 - C_4D_5F_6)^2)$$

$$Y_6 =$$

$$((-C_6D_5F_4 + C_5D_6F_4 + C_6D_4F_5 - C_4D_6F_5 - C_5D_4F_6 + C_4D_5F_6)(B_3((-A_2E_1 + A_1E_2)(D_5F_4 - D_4F_5)c_0 + C_5d_1F_4(E_2A_0 - A_2E_0) + C_4d_1F_5(-E_2A_0 + A_2E_0) + C_5(A_2E_1 - A_1E_2)(F_4d_0 - D_4f_0) - C_4(A_2E_1 - A_1E_2)(F_5d_0 - D_5f_0)) + B_2((A_3E_1 - A_1E_3)(D_5F_4 - D_4F_5)c_0 + C_4d_1F_5(E_3A_0 - A_3E_0) + C_5d_1F_4(-E_3A_0 + A_3E_0) - C_5(A_3E_1 - A_1E_3)(F_4d_0 - D_4f_0) + C_4(A_3E_1 - A_1E_3)(F_5d_0 - D_5f_0))(A_3E_2 - A_2E_3)(B_1(-D_5F_4 + D_4F_5)c_0 - C_5(d_1F_4B_0 - B_1F_4d_0 + B_1D_4f_0) + C_4(d_1F_5B_0 - B_1F_5d_0 + B_1D_5f_0))) - (A_3B_2E_1 - A_2B_3E_1 - A_3B_1E_2 + A_1B_3E_2 + A_2B_1E_3 - A_1B_2E_3)(x_{25}C_5D_4 - x_{25}C_4D_5 - x_{17}C_5F_4 + x_{12}D_5F_4 + x_{17}C_4F_5 - x_{12}D_4F_5)(-D_6F_5C_0 + D_5F_6C_0 + C_6F_5D_0 - C_5F_6D_0 - C_6D_5F_0 + C_5D_6F_0)R_1^2 + (x_{23}(C_5D_4 - C_4D_5)(C_6D_5F_4 - C_5D_6F_4 - C_6D_4F_5 + C_4D_6F_5 + C_5D_4F_6 - C_4D_5F_6)(B_3E_2A_0 - B_2E_3A_0 - A_3E_2B_0 + A_2E_3B_0 + A_3B_2E_0 - A_2B_3E_0) + x_{11}(D_5F_4 - D_4F_5)(C_6D_5F_4 - C_5D_6F_4 - C_6D_4F_5 + C_4D_6F_5 + C_5D_4F_6 - C_4D_5F_6)(B_3E_2A_0 - B_2E_3A_0 - A_3E_2B_0 + A_2E_3B_0 + A_3B_2E_0 - A_2B_3E_0) + (A_3B_2E_1 - A_2B_3E_1 - A_3B_1E_2 + A_1B_3E_2 + A_2B_1E_3 - A_1B_2E_3)(x_{26}C_5D_4 - x_{26}C_4D_5 + x_{16}C_5F_4 - x_{13}D_5F_4 - x_{16}C_4F_5 + x_{13}D_4F_5) \cdot (-D_6F_5C_0 + D_5F_6C_0 + C_6F_5D_0 - C_5F_6D_0 - C_6D_5F_0 + C_5D_6F_0)R_1R_2 + (x_{24}C_5D_4 - x_{24}C_4D_5 + x_{10}D_5F_4 - x_{10}D_4F_5)(C_6D_5F_4 - C_5D_6F_4 - C_6D_4F_5 + C_4D_6F_5 + C_5D_4F_6 - C_4D_5F_6)(B_3E_2A_0 - B_2E_3A_0 - A_3E_2B_0 + A_2E_3B_0 + A_3B_2E_0 - A_2B_3E_0)R_2^2)/((A_3B_2E_1 - A_2B_3E_1 - A_3B_1E_2 + A_1B_3E_2 + A_2B_1E_3 - A_1B_2E_3)(C_6D_5F_4 - C_5D_6F_4 - C_6D_4F_5 + C_4D_6F_5 + C_5D_4F_6 - C_4D_5F_6)^2)$$

## B. Coefficients

Here are the constants which appear in Y's where  $S = \frac{P_a}{3a^2}$  and  $P = \frac{3P_a^2}{2a^2}$  terms are given as:

$$\begin{aligned} A_0 &= \frac{1}{\sqrt{2k}} \cosh [R_{1,k}(\eta)] (P \cos [\Phi_{1,k}(\eta)] + k \sin [\Phi_{1,k}(\eta)]) + (-P \cos [2\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta)] + k \sin [2\Theta_{1,k}(\eta) \\ &\quad + \Phi_{1,k}(\eta)]) \sinh [R_{1,k}(\eta)] \\ a_0 &= \frac{P}{3\sqrt{2}\sqrt{k}} (\cos [\Phi_{2,k}(\eta)] \cosh [R_{2,k}(\eta)] - \cos [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)] \sinh [R_{2,k}(\eta)]) \\ A_1 &= \frac{1}{\sqrt{2}\sqrt{k}} (-\cos [2\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta)] \cosh [R_{1,k}(\eta)] + \cos [\Phi_{1,k}(\eta)] \sinh [R_{1,k}(\eta)]) \\ A_2 &= \frac{1}{\sqrt{k}} \left( \sqrt{2} \sin [2\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta)] \sinh [R_{1,k}(\eta)] \right) \\ A_3 &= \frac{1}{\sqrt{2}\sqrt{k}} (-\cosh [R_{1,k}(\eta)] \sin [\Phi_{1,k}(\eta)] + \sin [2\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta)] \sinh [R_{1,k}(\eta)]) \end{aligned}$$

$$\begin{aligned}
B_0 &= -\frac{1}{\sqrt{2}\sqrt{k}} (\cosh [R_{1,k}(\eta)] (-k \cos [\Phi_{1,k}(\eta)] + P \sin [\Phi_{1,k}(\eta)]) - (k \cos [2\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta)] \\
&\quad + P \sin [2\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta)]) \sinh [R_{1,k}(\eta)]) \\
b_0 &= \frac{1}{3\sqrt{2}\sqrt{k}} (P (\cosh [R_{2,k}(\eta)] \sin [\Phi_{2,k}(\eta)] - \sin [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)] \sinh [R_{2,k}(\eta)])) \\
B_1 &= \frac{1}{\sqrt{2}\sqrt{k}} (-\cosh [R_{1,k}(\eta)] \sin [2\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta)] + \sin [\Phi_{1,k}(\eta)] \sinh [R_{1,k}(\eta)]) \\
B_2 &= \frac{-1}{\sqrt{k}} \left( \sqrt{2} \cos [2\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta)] \sinh [R_{1,k}(\eta)] \right) \\
B_3 &= \frac{1}{\sqrt{2}\sqrt{k}} (\cos [\Phi_{1,k}(\eta)] \cosh [R_{1,k}(\eta)] - \cos [2\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta)] \sinh [R_{1,k}(\eta)]) \\
C_0 &= \frac{1}{\sqrt{2}\sqrt{k}} (\cosh [R_{2,k}(\eta)] (P \cos [\Phi_{2,k}(\eta)] + k \sin [\Phi_{2,k}(\eta)]) + (-P \cos [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)] \\
&\quad + k \sin [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)]) \sinh [R_{2,k}(\eta)]) \\
c_0 &= \frac{P}{3\sqrt{2}\sqrt{k}} (\cos [\Phi_{1,k}(\eta)] \cosh [R_{1,k}(\eta)] - \cos [2\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta)] \sinh [R_{1,k}(\eta)]) \\
C_4 &= \frac{1}{\sqrt{2}\sqrt{k}} (-\cos [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)] \cosh [R_{2,k}(\eta)] + \cos [\Phi_{2,k}(\eta)] \sinh [R_{2,k}(\eta)]) \\
C_5 &= \frac{1}{\sqrt{k}} \left( \sqrt{2} \sin [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)] \sinh [R_{2,k}(\eta)] \right) \\
C_6 &= \frac{1}{\sqrt{2}\sqrt{k}} (-\cosh [R_{2,k}(\eta)] \sin [\Phi_{2,k}(\eta)] + \sin [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)] \sinh [R_{2,k}(\eta)]) \\
D_0 &= \frac{1}{\sqrt{2}\sqrt{k}} (\cosh [R_{2,k}(\eta)] (-k \cos [\Phi_{2,k}(\eta)] + P \sin [\Phi_{2,k}(\eta)]) - (k \cos [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)] \\
&\quad + P \sin [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)]) \sinh [R_{2,k}(\eta)]) \\
d_0 &= \frac{1}{3\sqrt{2}\sqrt{k}} (P (\cosh [R_{1,k}(\eta)] \sin [\Phi_{1,k}(\eta)] - \sin [2\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta)] \sinh [R_{1,k}(\eta)])) \\
D_4 &= \frac{1}{\sqrt{2}\sqrt{k}} (-\cosh [R_{2,k}(\eta)] \sin [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)] + \sin [\Phi_{2,k}(\eta)] \sinh [R_{2,k}(\eta)]) \\
D_5 &= \frac{-1}{\sqrt{k}} \left( \sqrt{2} \cos [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)] \sinh [R_{2,k}(\eta)] \right) \\
D_6 &= \frac{1}{\sqrt{2}\sqrt{k}} (\cos [\Phi_{2,k}(\eta)] \cosh [R_{2,k}(\eta)] - \cos [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)] \sinh [R_{2,k}(\eta)]) \\
E_0 &= \frac{e^{i\Phi_{1,k}(\eta)}}{\sqrt{2}} \sqrt{k} \left( \cosh [R_{1,k}(\eta)] (-iP + kc_{s1}^2) - e^{2i\Theta_{1,k}(\eta)} \sinh [R_{1,k}(\eta)] (iP + kc_{s1}^2) \right) \\
e_0 &= \frac{e^{i\Phi_{2,k}(\eta)}}{3\sqrt{2}\sqrt{k}} \left( (-ikP + 3S) \cosh [R_{2,k}(\eta)] + e^{2i\Theta_{2,k}(\eta)} i(-kP + 3iS) \sinh [R_{2,k}(\eta)] \right) \\
E_1 &= \frac{e^{i\Phi_{1,k}(\eta)}}{\sqrt{2}} i\sqrt{k} \left( e^{2i\Theta_{1,k}(\eta)} \cosh [R_{1,k}(\eta)] + \sinh [R_{1,k}(\eta)] \right) \\
E_2 &= -\sqrt{2} e^{i(2\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta))} \sqrt{k} \sinh [R_{1,k}(\eta)] \\
E_3 &= -\frac{1}{\sqrt{2}} e^{i(\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta))} \sqrt{k} (\cosh [R_{1,k}(\eta)] + i\Theta_{1,k}(\eta)) + \sinh [R_{1,k}(\eta) - i\Theta_{1,k}(\eta)] \\
f_0 &= \frac{e^{i\Phi_{1,k}(\eta)}}{3\sqrt{2}\sqrt{k}} \left( (-ikP + 3S) \cosh [R_{1,k}(\eta)] + e^{2i\Theta_{1,k}(\eta)} i(-kP + 3iS) \sinh [R_{1,k}(\eta)] \right) \\
F_0 &= \frac{-e^{i\Phi_{2,k}(\eta)}}{\sqrt{2}} \sqrt{k} \left( e^{2i\Theta_{2,k}(\eta)} \sinh [R_{2,k}(\eta)] (iP + kc_{s2}^2) + i \cosh [R_{2,k}(\eta)] (P + ikc_{s2}^2) \right) \\
F_4 &= \frac{e^{i\Phi_{2,k}(\eta)}}{\sqrt{2}} i\sqrt{k} \left( e^{2i\Theta_{2,k}(\eta)} \cosh [R_{2,k}(\eta)] + \sinh [R_{2,k}(\eta)] \right)
\end{aligned}$$



$$\begin{aligned}
F_5 &= -\sqrt{2}e^{i(2\Theta_{2,k}(\eta)+\Phi_{2,k}(\eta))}\sqrt{k}\sinh[R_{2,k}(\eta)] \\
F_6 &= -\frac{1}{\sqrt{2}}e^{i\Phi_{2,k}(\eta)}\sqrt{k}\left(\cosh[R_{2,k}(\eta)]+e^{2i\Theta_{2,k}(\eta)}\sinh[R_{2,k}(\eta)]\right) \\
x_1 &= \frac{1}{6\sqrt{2k}}\sin[2\Theta_{1,k}(\eta)](-2\cos[\Phi_{1,k}(\eta)]\cosh[R_{1,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))] \\
&\quad +(\sin[2\Theta_{2,k}(\eta)-\Phi_{1,k}(\eta)]+2\sin[4\Theta_{1,k}(\eta)-2\Theta_{2,k}(\eta)+\Phi_{1,k}(\eta)]-\sin[4\Theta_{1,k}(\eta)+2\Theta_{2,k}(\eta) \\
&\quad +\Phi_{1,k}(\eta)])\sinh[R_{1,k}(\eta)]) \\
x_2 &= \frac{1}{12\sqrt{2k}}\sin[2\Theta_{2,k}(\eta)](2\cos[\Phi_{1,k}(\eta)]\cosh[R_{1,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))] \\
&\quad +(\sin[2\Theta_{2,k}(\eta)-\Phi_{1,k}(\eta)]+2\sin[4\Theta_{1,k}(\eta)-2\Theta_{2,k}(\eta)+\Phi_{1,k}(\eta)] \\
&\quad -\sin[4\Theta_{1,k}(\eta)+2\Theta_{2,k}(\eta)+\Phi_{1,k}(\eta)])\sinh[R_{1,k}(\eta)]) \\
x_3 &= -\frac{1}{12\sqrt{2}\sqrt{k}}\sin[2\Theta_{1,k}(\eta)](2\cos[\Phi_{1,k}(\eta)]\cosh[R_{1,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))] \\
&\quad +(\sin[2\Theta_{2,k}(\eta)-\Phi_{1,k}(\eta)]+2\sin[4\Theta_{1,k}(\eta)-2\Theta_{2,k}(\eta)+\Phi_{1,k}(\eta)]-\sin[4\Theta_{1,k}(\eta)+2\Theta_{2,k}(\eta) \\
&\quad +\Phi_{1,k}(\eta)])\sinh[R_{1,k}(\eta)]) \\
x_4 &= \frac{1}{6\sqrt{2k}}\sin[2\Theta_{2,k}(\eta)](2\cos[\Phi_{1,k}(\eta)]\cosh[R_{1,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))] \\
&\quad +(\sin[2\Theta_{2,k}(\eta)-\Phi_{1,k}(\eta)]+2\sin[4\Theta_{1,k}(\eta)-2\Theta_{2,k}(\eta)+\Phi_{1,k}(\eta)]-\sin[4\Theta_{1,k}(\eta)+2\Theta_{2,k}(\eta) \\
&\quad +\Phi_{1,k}(\eta)])\sinh[R_{1,k}(\eta)]) \\
x_6 &= \frac{1}{6\sqrt{2}\sqrt{k}}(\cosh[R_{1,k}(\eta)]\sin[2\Theta_{2,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))]\sin[\Phi_{1,k}(\eta)]) \\
x_7 &= (\cos[2\Theta_{2,k}(\eta)-\Phi_{1,k}(\eta)]-2\cos[4\Theta_{1,k}(\eta)-2\Theta_{2,k}(\eta)+\Phi_{1,k}(\eta)]+\cos[4\Theta_{1,k}(\eta)+2\Theta_{2,k}(\eta) \\
&\quad +\Phi_{1,k}(\eta)])\sin[2\Theta_{2,k}(\eta)] \\
x_8 &= \frac{1}{6\sqrt{2}\sqrt{k}}\sin[2\Theta_{2,k}(\eta)](\cosh[R_{1,k}(\eta)](-\cosh[i(2(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))+\Phi_{1,k}(\eta))]) \\
&\quad +\cosh[2i(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))-i\Phi_{1,k}(\eta)])+(\cos[2\Theta_{2,k}(\eta)-\Phi_{1,k}(\eta)]-2\cos[4\Theta_{1,k}(\eta)-2\Theta_{2,k}(\eta)+\Phi_{1,k}(\eta)] \\
&\quad +\cos[4\Theta_{1,k}(\eta)+2\Theta_{2,k}(\eta)+\Phi_{1,k}(\eta)])\sinh[R_{1,k}(\eta)]) \\
x_9 &= \frac{1}{12\sqrt{2}\sqrt{k}}\sin[2\Theta_{1,k}(\eta)](\cosh[R_{1,k}(\eta)](\cosh[i(2(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))+\Phi_{1,k}(\eta))]-\cosh[2i(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta) \\
&\quad -i\Phi_{1,k}(\eta))])-(\cos[2\Theta_{2,k}(\eta)-\Phi_{1,k}(\eta)]-2\cos[4\Theta_{1,k}(\eta)-2\Theta_{2,k}(\eta)+\Phi_{1,k}(\eta)] \\
&\quad +\cos[4\Theta_{1,k}(\eta)+2\Theta_{2,k}(\eta)+\Phi_{1,k}(\eta)])\sinh[R_{1,k}(\eta)]) \\
x_{10} &= \frac{1}{6\sqrt{2}(\sqrt{k})}\sin[2\Theta_{2,k}(\eta)](\cosh[R_{2,k}(\eta)](\cos[\Phi_{2,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)-\Theta_{2,k}(\eta))] \\
&\quad -2\sin[2\Theta_{1,k}(\eta)]\sin[2\Theta_{2,k}(\eta)]\sin[\Phi_{2,k}(\eta)])+\cos[2\Theta_{2,k}(\eta)+\Phi_{2,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))]\sinh[R_{2,k}(\eta)]) \\
x_{11} &= \frac{-1}{3\sqrt{2}(\sqrt{k})}\sin[2\Theta_{1,k}(\eta)](\cosh[R_{2,k}(\eta)](\cos[\Phi_{2,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)-\Theta_{2,k}(\eta))] \\
&\quad -2\sin[2\Theta_{1,k}(\eta)]\sin[2\Theta_{2,k}(\eta)]\sin[\Phi_{2,k}(\eta)])+\cos[2\Theta_{2,k}(\eta)+\Phi_{2,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))]\sinh[R_{2,k}(\eta)]) \\
x_{12} &= -\frac{1}{6\sqrt{2}\sqrt{k}}\sin[2\Theta_{1,k}(\eta)](\cosh[R_{2,k}(\eta)](\cos[\Phi_{2,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)-\Theta_{2,k}(\eta))] \\
&\quad -2\sin[2\Theta_{1,k}(\eta)]\sin[2\Theta_{2,k}(\eta)]\sin[\Phi_{2,k}(\eta)])+\cos[2\Theta_{2,k}(\eta)+\Phi_{2,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))]\sinh[R_{2,k}(\eta)]) \\
x_{13} &= \frac{1}{3\sqrt{2}\sqrt{k}}\sin[2\Theta_{2,k}(\eta)](\cosh[R_{2,k}(\eta)](\cos[\Phi_{2,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)-\Theta_{2,k}(\eta))] \\
&\quad -2\sin[2\Theta_{1,k}(\eta)]\sin[2\Theta_{2,k}(\eta)]\sin[\Phi_{2,k}(\eta)])+\cos[2\Theta_{2,k}(\eta)+\Phi_{2,k}(\eta)]\sin[2(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))]\sinh[R_{2,k}(\eta)]) \\
x_{14} &= -\frac{1}{3\sqrt{2k}}\sin[2\Theta_{1,k}(\eta)](\cosh[R_{2,k}(\eta)](2\cos[\Phi_{2,k}(\eta)]\sin[2\Theta_{1,k}(\eta)]\sin[2\Theta_{2,k}(\eta)]+\sin[2(\Theta_{1,k}(\eta) \\
&\quad -\Theta_{2,k}(\eta))]\sin[\Phi_{2,k}(\eta)]+\sin[2(\Theta_{1,k}(\eta)+\Theta_{2,k}(\eta))]\sin[2\Theta_{2,k}(\eta)+\Phi_{2,k}(\eta)]\sinh[R_{2,k}(\eta)])
\end{aligned}$$

$$\begin{aligned}
x_{15} &= \frac{1}{6\sqrt{2}\sqrt{k}} \sin [2\Theta_{2,k}(\eta)] (\cosh [R_{2,k}(\eta)] (2 \cos [\Phi_{2,k}(\eta)] \sin [2\Theta_{1,k}(\eta)] \sin [2\Theta_{2,k}(\eta)] \\
&\quad + \sin [2(\Theta_{1,k}(\eta) - \Theta_{2,k}(\eta))] \sin [\Phi_{2,k}(\eta)] + \sin [2(\Theta_{1,k}(\eta) + \Theta_{2,k}(\eta))] \sin [2\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta)] \sinh [R_{2,k}(\eta)]) \\
x_{16} &= \frac{1}{12\sqrt{2}} i e^{i(\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta))} \sqrt{k} \sin [2\Theta_{1,k}(\eta)] \left( e^{-i\Theta_{1,k}(\eta)} \sin [2\Theta_{2,k}(\eta)] (2 \cos [2\Theta_{1,k}(\eta)] \cosh [R_{1,k}(\eta)] \right. \\
&\quad \left. - (1 - 3e^{4i\Theta_{1,k}(\eta)}) \sinh [R_{1,k}(\eta)] \right) + 2 \cos [2\Theta_{2,k}(\eta)] \sin [2\Theta_{1,k}(\eta)] (\cosh [R_{1,k}(\eta) - i\Theta_{1,k}(\eta)] \\
&\quad - \sinh [R_{1,k}(\eta) + i\Theta_{1,k}(\eta)]) \\
x_{17} &= -\frac{1}{6\sqrt{2}} i e^{i(\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta))} \sqrt{k} \sin [2\Theta_{2,k}(\eta)] \left( e^{-i\Theta_{1,k}(\eta)} \sin [2\Theta_{2,k}(\eta)] (2 \cos [2\Theta_{1,k}(\eta)] \cosh [R_{1,k}(\eta)] \right. \\
&\quad \left. - (1 - 3e^{4i\Theta_{1,k}(\eta)}) \sinh [R_{1,k}(\eta)] \right) + 2 \cos [2\Theta_{2,k}(\eta)] \sin [2\Theta_{1,k}(\eta)] (\cosh [R_{1,k}(\eta) - i\Theta_{1,k}(\eta)] \\
&\quad - \sinh [R_{1,k}(\eta) + i\Theta_{1,k}(\eta)]) \\
x_{18} &= \frac{1}{6\sqrt{2}} e^{i\Phi_{1,k}(\eta)} i \sqrt{k} \sin [2\Theta_{1,k}(\eta)] \left( \sin [2\Theta_{2,k}(\eta)] (2 \cos [2\Theta_{1,k}(\eta)] \cosh [R_{1,k}(\eta)] - (1 - 3e^{4i\Theta_{1,k}(\eta)}) \sinh [R_{1,k}(\eta)] \right) \\
&\quad + 2e^{i\Theta_{1,k}(\eta)} \cos [2\Theta_{2,k}(\eta)] \sin [2\Theta_{1,k}(\eta)] (\cosh [R_{1,k}(\eta) - i\Theta_{1,k}(\eta)] - \sinh [R_{1,k}(\eta) + i\Theta_{1,k}(\eta)]) \\
x_{19} &= \frac{1}{12\sqrt{2}} e^{i\Phi_{1,k}(\eta)} i \sqrt{k} \sin [2\Theta_{2,k}(\eta)] \left( \sin [2\Theta_{2,k}(\eta)] (2 \cos [2\Theta_{1,k}(\eta)] \cosh [R_{1,k}(\eta)] - (1 - 3e^{4i\Theta_{1,k}(\eta)}) \sinh [R_{1,k}(\eta)] \right) \\
&\quad + 2e^{i\Theta_{1,k}(\eta)} \cos [2\Theta_{2,k}(\eta)] \sin [2\Theta_{1,k}(\eta)] (\cosh [R_{1,k}(\eta) - i\Theta_{1,k}(\eta)] - \sinh [R_{1,k}(\eta) + i\Theta_{1,k}(\eta)]) \\
x_{20} &= \frac{1}{6\sqrt{2}} i e^{i(\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta))} \sqrt{k} \sin [2\Theta_{2,k}(\eta)] \left( e^{-i\Theta_{1,k}(\eta)} \sin [2\Theta_{2,k}(\eta)] (2 \cos [2\Theta_{1,k}(\eta)] \cosh [R_{1,k}(\eta)] \right. \\
&\quad \left. - (1 - 3e^{4i\Theta_{1,k}(\eta)}) \sinh [R_{1,k}(\eta)] \right) + 2 \cos [2\Theta_{2,k}(\eta)] \sin [2\Theta_{1,k}(\eta)] (\cosh [R_{1,k}(\eta) - i\Theta_{1,k}(\eta)] \\
&\quad - \sinh [R_{1,k}(\eta) + i\Theta_{1,k}(\eta)]) \\
x_{21} &= \frac{1}{12\sqrt{2}} i e^{i(\Theta_{1,k}(\eta) + \Phi_{1,k}(\eta))} \sqrt{k} \sin [2\Theta_{1,k}(\eta)] \left( e^{-i\Theta_{1,k}(\eta)} \sin [2\Theta_{2,k}(\eta)] (2 \cos [2\Theta_{1,k}(\eta)] \cosh [R_{1,k}(\eta)] \right. \\
&\quad \left. - (1 - 3e^{4i\Theta_{1,k}(\eta)}) \sinh [R_{1,k}(\eta)] \right) + 2 \cos [2\Theta_{2,k}(\eta)] \sin [2\Theta_{1,k}(\eta)] (\cosh [R_{1,k}(\eta) - i\Theta_{1,k}(\eta)] \\
&\quad - \sinh [R_{1,k}(\eta) + i\Theta_{1,k}(\eta)]) \\
x_{23} &= \frac{1}{6\sqrt{2}} e^{i(\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta))} \sqrt{k} \sin [2\Theta_{1,k}(\eta)] \\
&\quad (\cosh [R_{2,k}(\eta) - 2i\Theta_{1,k}(\eta) - 3i\Theta_{2,k}(\eta)] - \cosh [R_{2,k}(\eta) + 2i\Theta_{1,k}(\eta) - i\Theta_{2,k}(\eta)] \\
&\quad - \cosh [R_{2,k}(\eta) - 2i\Theta_{1,k}(\eta) + i\Theta_{2,k}(\eta)] + \cosh [R_{2,k}(\eta) - i(2\Theta_{1,k}(\eta) + \Theta_{2,k}(\eta))] \\
&\quad + 2i \cosh [R_{2,k}(\eta)] \sin [2\Theta_{1,k}(\eta) - \Theta_{2,k}(\eta)] - 2i \cosh [R_{2,k}(\eta) + 2i(\Theta_{1,k}(\eta) + \Theta_{2,k}(\eta))] \sin [\Theta_{2,k}(\eta)]) \\
x_{24} &= \frac{1}{12\sqrt{2}} e^{i(\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta))} \sqrt{k} \\
&\quad (-\cosh [R_{2,k}(\eta) - 2i\Theta_{1,k}(\eta) - 3i\Theta_{2,k}(\eta)] + \cosh [R_{2,k}(\eta) + 2i\Theta_{1,k}(\eta) - i\Theta_{2,k}(\eta)] \\
&\quad + \cosh [R_{2,k}(\eta) - 2i\Theta_{1,k}(\eta) + i\Theta_{2,k}(\eta)] - \cosh [R_{2,k}(\eta) - i(2\Theta_{1,k}(\eta) + \Theta_{2,k}(\eta))] \\
&\quad - 2i \cosh [R_{2,k}(\eta)] \sin [2\Theta_{1,k}(\eta) - \Theta_{2,k}(\eta)] + 2i \cosh [R_{2,k}(\eta) \\
&\quad + 2i(\Theta_{1,k}(\eta) + \Theta_{2,k}(\eta))] \sin [\Theta_{2,k}(\eta)] \sin [2\Theta_{2,k}(\eta)] \\
x_{25} &= \frac{1}{12\sqrt{2}} e^{-R_{2,k}(\eta) + i(\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta))} \sqrt{k} \sin [2\Theta_{1,k}(\eta)] (-\cos [2\Theta_{1,k}(\eta) - \Theta_{2,k}(\eta)] + \cos [2\Theta_{1,k}(\eta) + 3\Theta_{2,k}(\eta)] \\
&\quad + 2i \cos [\Theta_{2,k}(\eta)] \sin [2\Theta_{1,k}(\eta)] - 2e^{2R_{2,k}(\eta)} (2i \cos [\Theta_{2,k}(\eta)] \cos [2\Theta_{1,k}(\eta) + \Theta_{2,k}(\eta)] \\
&\quad + \sin [2\Theta_{1,k}(\eta)] \sin [\Theta_{2,k}(\eta)]) \\
x_{26} &= \frac{1}{6\sqrt{2}} e^{-R_{2,k}(\eta) + i(\Theta_{2,k}(\eta) + \Phi_{2,k}(\eta))} \sqrt{k} (-\cos [2\Theta_{1,k}(\eta) - \Theta_{2,k}(\eta)] + \cos [2\Theta_{1,k}(\eta) + 3\Theta_{2,k}(\eta)] \\
&\quad + 2i \cos [\Theta_{2,k}(\eta)] \sin [2\Theta_{1,k}(\eta)] - 2e^{2R_{2,k}(\eta)} (2i \cos [\Theta_{2,k}(\eta)] \cos [2\Theta_{1,k}(\eta) + \Theta_{2,k}(\eta)] \\
&\quad + \sin [2\Theta_{1,k}(\eta)] \sin [\Theta_{2,k}(\eta)] \sin [2\Theta_{2,k}(\eta)]
\end{aligned}$$

**Corresponding author address:**

E-mail: sayantan.choudhury@icts.res.in,  
sayanphysicsisi@gmail.com

- 
- [1] B. L. Schumaker and C. M. Caves, “New formalism for two-photon quantum optics. ii. mathematical foundation and compact notation,” *Phys. Rev. A* **31** (May, 1985) 3093–3111.
- [2] M. Teich, “Squeezed states of light (A),” *Journal of the Optical Society of America A* (01, 1987) 10.
- [3] C. Fabre, G. Grynberg, and A. Aspect, *Introduction to Quantum Optics: From the Semi-classical Approach to Quantized Light*. 09, 2010.
- [4] R. Loudon and P. Knight, “Squeezed light,” *Journal of Modern Optics* **34** no. 6-7, (1987) 709–759.
- [5] U. Andersen, T. Gehring, C. Marquardt, and G. Leuchs, “30 years of squeezed light generation,” *Physica Scripta* **91** (11, 2015) .
- [6] O. Rosas-Ortiz, “Coherent and squeezed states: Introductory review of basic notions, properties, and generalizations,” *Integrability, Supersymmetry and Coherent States* (2019) .
- [7] B. L. Schumaker and C. M. Caves, “New formalism for two-photon quantum optics. 2. Mathematical foundation and compact notation,” *Phys. Rev. A* **31** (1985) 3093–3111.
- [8] A. Garcia-Chung, “Squeeze operator: a classical view,” [arXiv:2003.04257 \[math-ph\]](https://arxiv.org/abs/2003.04257).
- [9] C. M. Caves and B. L. Schumaker, “New formalism for two-photon quantum optics. i. quadrature phases and squeezed states,” *Phys. Rev. A* **31** (May, 1985) 3068–3092. <https://link.aps.org/doi/10.1103/PhysRevA.31.3068>.
- [10] S. Zubairy, “Quantum squeezing,” *Journal of Optics B: Quantum and Semiclassical Optics* **7** no. 5, (May, 2005) 156–156.
- [11] A. Furusawa, “Applications of squeezed states of light - quantum teleportation and related quantum information processing,” in *Conference on Coherence and Quantum Optics*, p. CWA2. Optical Society of America, 2007.
- [12] V. Dodonov, “‘nonclassical’ states in quantum optics: A ‘squeezed’ review of the first 75 years,” *Journal of Optics B: Quantum and Semiclassical Optics* **4** (01, 2002) R1.
- [13] K. Zelaya, S. Dey, and V. Hussin, “Generalized squeezed states,” *Physics Letters A* **382** no. 47, (2018) 3369–3375.
- [14] N. C. Menicucci, P. van Loock, M. Gu, C. Weedbrook, T. C. Ralph, and M. A. Nielsen, “Universal quantum computation with continuous-variable cluster states,” *Phys. Rev. Lett.* **97** (Sep, 2006) 110501.
- [15] F. R. Cardoso, D. Z. Rossatto, G. P. L. M. Fernandes, G. Higgins, and C. J. Villas-Boas, “Superposition of two-mode squeezed states for quantum information processing and quantum sensing,” *Phys. Rev. A* **103** (Jun, 2021) 062405.
- [16] M. Hillery, “Quantum cryptography with squeezed states,” *Phys. Rev. A* **61** (Jan, 2000) 022309.
- [17] H. Vahlbruch, S. Chelkowski, B. Hage, A. Franzen, K. Danzmann, and R. Schnabel, “Coherent control of vacuum squeezing in the gravitational-wave detection band,” *Phys. Rev. Lett.* **97** (Jul, 2006) 011101.
- [18] S. S. Y. Chua, B. J. J. Slagmolen, D. A. Shaddock, and D. E. McClelland, “Quantum squeezed light in gravitational-wave detectors,” *Classical and Quantum Gravity* **31** no. 18, (Sep, 2014) 183001.
- [19] S. Choudhury, A. Mazumdar, and S. Pal, “Low & High scale MSSM inflation, gravitational waves and constraints from Planck,” *JCAP* **07** (2013) 041, [arXiv:1305.6398 \[hep-ph\]](https://arxiv.org/abs/1305.6398).
- [20] C. Darsow-Fromm, J. Gurs, R. Schnabel, and S. Steinlechner, “Squeezed light at 2128nm for future gravitational-wave observatories,” *Opt. Lett.* **46** no. 23, (Dec, 2021) 5850–5853.
- [21] R. Schnabel, N. Mavalvala, D. McClelland, and P. K. Lam, “Quantum metrology for gravitational wave astronomy,” *Nature communications* **1** (11, 2010) 121.
- [22] L. Barsotti, J. Harms, and R. Schnabel, “Squeezed vacuum states of light for gravitational wave detectors,” *Rept. Prog. Phys.* **82** no. 1, (2019) 016905.
- [23] R. Schnabel, “Squeezed states of light and their applications in laser interferometers,” *Phys. Rept.* **684** (2017) 1–51, [arXiv:1611.03986 \[quant-ph\]](https://arxiv.org/abs/1611.03986).
- [24] J. Aasi, J. Abadie, B. Abbott, R. Abbott, T. Abbott, M. Abernathy, C. Adams, T. Adams, P. Addesso, C. Affeldt, O. Aguiar, P. Ajith, B. Allen, E. Ceron, D. Amariutei, S. Anderson, W. Anderson, K. Arai, and J. Zweizig, “Enhanced sensitivity of the ligo gravitational wave detector by using squeezed states of light,” *Nature Photonics* **7** (07, 2013) 613.
- [25] H. Vahlbruch, S. Chelkowski, K. Danzmann, and R. Schnabel, “Quantum engineering of squeezed states for quantum communication and metrology,” *New Journal of Physics* **9** no. 10, (Oct, 2007) 371–371.
- [26] P. M. Anisimov, G. M. Raterman, A. Chiruvelli, W. N. Plick, S. D. Huver, H. Lee, and J. P. Dowling, “Quantum metrology with two-mode squeezed vacuum: Parity detection beats the heisenberg limit,” *Phys. Rev. Lett.* **104** (Mar, 2010) 103602.
- [27] V. Giovannetti, S. Lloyd, and L. Maccone, “Advances in quantum metrology,” *Phys. Rev. Lett.* **96** (02, 2011) .
- [28] J. P. Dowling, “Quantum optical metrology – the lowdown on high-noon states,” *Contemporary Physics* **49** no. 2, (2008) 125–143.
- [29] C. Xu, L. Zhang, S. Huang, T. Ma, F. Liu, H. Yonezawa, Y. Zhang, and M. Xiao, “Sensing and tracking enhanced by quantum squeezing,” *Photon. Res.* **7** no. 6, (Jun, 2019) A14–A26.
- [30] M. Riedel, P. Böhi, Y. Li, T. Haensch, A. Sinatra, and P. Treutlein, “Atom-chip-based generation of entanglement for quantum metrology,” *Nature* **464** (03, 2010) 1170–3.
- [31] S. Choudhury, S. Chowdhury, N. Gupta, and A. Swain, “QMetrology from QCosmology: Study with Entangled Two Qubit Open Quantum System in De Sitter Space,” *SciPost Phys. Core* **4** (2021) 006, [arXiv:2005.13555 \[hep-th\]](https://arxiv.org/abs/2005.13555).
- [32] C. M. Caves, “Quantum-mechanical noise in an interferometer,” *Phys. Rev. D* **23** no. 8, (Apr., 1981) 1693–1708.
- [33] S. L. Braunstein and P. van Loock, “Quantum information with continuous variables,” *Rev. Mod. Phys.* **77** (Jun, 2005) 513–577.
- [34] H. Yonezawa and A. Furusawa, “Continuous-variable

- quantum information processing with squeezed states of light,” *Optics and Spectroscopy* **108** (02, 2010) 288–296.
- [35] S. L. Braunstein and P. Van Loock, “Quantum information with continuous variables,” *Reviews of modern physics* **77** no. 2, (2005) 513.
- [36] H. Vahlbruch, M. Mehmet, K. Danzmann, and R. Schnabel, “Detection of 15 db squeezed states of light and their application for the absolute calibration of photoelectric quantum efficiency,” *Phys. Rev. Lett.* **117** (Sep, 2016) 110801.
- [37] Y. Yamamoto and H. A. Haus, “Preparation, measurement and information capacity of optical quantum states,” *Rev. Mod. Phys.* **58** (Oct, 1986) 1001–1020.
- [38] K. Adhikari, S. Choudhury, H. N. Pandya, and R. Srivastava, “PGW Circuit Complexity,” [arXiv:2108.10334 \[gr-qc\]](#).
- [39] K. Ando and V. Vennin, “Bipartite temporal Bell inequalities for two-mode squeezed states,” *Phys. Rev. A* **102** no. 5, (2020) 052213, [arXiv:2007.00458 \[quant-ph\]](#).
- [40] J. Martin, A. Micheli, and V. Vennin, “Discord and Decoherence,” [arXiv:2112.05037 \[quant-ph\]](#).
- [41] P. Bhargava, S. Choudhury, S. Chowdhury, A. Mishara, S. P. Selvam, S. Panda, and G. D. Pasquino, “Quantum aspects of chaos and complexity from bouncing cosmology: A study with two-mode single field squeezed state formalism,” *SciPost Phys. Core* **4** (2021) 026, [arXiv:2009.03893 \[hep-th\]](#).
- [42] S. Choudhury, S. Chowdhury, N. Gupta, A. Mishara, S. P. Selvam, S. Panda, G. D. Pasquino, C. Singha, and A. Swain, “Circuit Complexity From Cosmological Islands,” *Symmetry* **13** (2021) 1301, [arXiv:2012.10234 \[hep-th\]](#).
- [43] K. Adhikari, S. Choudhury, S. Chowdhury, K. Shirish, and A. Swain, “Circuit complexity as a novel probe of quantum entanglement: A study with black hole gas in arbitrary dimensions,” *Phys. Rev. D* **104** no. 6, (2021) 065002, [arXiv:2104.13940 \[hep-th\]](#).
- [44] S. Choudhury, A. Mukherjee, N. Pandey, and A. Roy, “Causality Constraint on Circuit Complexity from *COSMOEFT*,” [arXiv:2111.11468 \[hep-th\]](#).
- [45] L. P. Grishchuk and Y. V. Sidorov, “Squeezed quantum states in theory of cosmological perturbations,” in *5th Seminar on Quantum Gravity*. 1990.
- [46] L. P. Grishchuk, “Quantum Mechanics of the Primordial Cosmological Perturbations,” in *a talk given at the Sixth Marcel Grossmann Meeting*. Kyoto, 1991.
- [47] A. Albrecht, P. Ferreira, M. Joyce, and T. Prokopec, “Inflation and squeezed quantum states,” *Phys. Rev. D* **50** (1994) 4807–4820, [arXiv:astro-ph/9303001](#).
- [48] T. Colas, J. Grain, and V. Vennin, “Four-mode squeezed states: two-field quantum systems and the symplectic group  $Sp(4, \mathbb{R})$ ,” [arXiv:2104.14942 \[quant-ph\]](#).
- [49] K. Hasebe, “ $Sp(4; \mathbb{R})$  Squeezing for Bloch Four-Hyperboloid via The Non-Compact Hopf Map,” *J. Phys. A* **53** no. 5, (2020) 055303, [arXiv:1904.12259 \[quant-ph\]](#).
- [50] S. Choudhury and S. Pal, “Fourth level MSSM inflation from new flat directions,” *JCAP* **04** (2012) 018, [arXiv:1111.3441 \[hep-ph\]](#).
- [51] S. Choudhury and S. Pal, “DBI Galileon inflation in background SUGRA,” *Nucl. Phys. B* **874** (2013) 85–114, [arXiv:1208.4433 \[hep-th\]](#).
- [52] S. Choudhury and S. Pal, “Brane inflation in background supergravity,” *Phys. Rev. D* **85** (2012) 043529, [arXiv:1102.4206 \[hep-th\]](#).
- [53] S. Choudhury, T. Chakraborty, and S. Pal, “Higgs inflation from new Kähler potential,” *Nucl. Phys. B* **880** (2014) 155–174, [arXiv:1305.0981 \[hep-th\]](#).
- [54] S. Choudhury and S. Panda, “COSMOS- $e'$ -GTachyon from string theory,” *Eur. Phys. J. C* **76** no. 5, (2016) 278, [arXiv:1511.05734 \[hep-th\]](#).
- [55] A. Bhattacharyya, S. Das, S. S. Haque, and B. Underwood, “Rise of cosmological complexity: Saturation of growth and chaos,” *Phys. Rev. Res.* **2** no. 3, (2020) 033273, [arXiv:2005.10854 \[hep-th\]](#).
- [56] S. Choudhury, “COSMOS- $e'$ - soft Higgsotic attractors,” *Eur. Phys. J. C* **77** no. 7, (2017) 469, [arXiv:1703.01750 \[hep-th\]](#).
- [57] S. Akhtar, S. Choudhury, S. Chowdhury, D. Goswami, S. Panda, and A. Swain, “Open Quantum Entanglement: A study of two atomic system in static patch of de Sitter space,” *Eur. Phys. J. C* **80** no. 8, (2020) 748, [arXiv:1908.09929 \[hep-th\]](#).
- [58] S. Choudhury, S. Panda, and R. Singh, “Bell violation in primordial cosmology,” *Universe* **3** no. 1, (2017) 13, [arXiv:1612.09445 \[hep-th\]](#).
- [59] S. Choudhury, S. Panda, and R. Singh, “Bell violation in the Sky,” *Eur. Phys. J. C* **77** no. 2, (2017) 60, [arXiv:1607.00237 \[hep-th\]](#).
- [60] A. Bhattacharyya, S. Das, S. Shajidul Haque, and B. Underwood, “Cosmological Complexity,” *Phys. Rev. D* **101** no. 10, (2020) 106020, [arXiv:2001.08664 \[hep-th\]](#).
- [61] S. Choudhury and S. Panda, “Entangled de Sitter from stringy axionic Bell pair I: an analysis using Bunch–Davies vacuum,” *Eur. Phys. J. C* **78** no. 1, (2018) 52, [arXiv:1708.02265 \[hep-th\]](#).
- [62] M. B. Einhorn and F. Larsen, “Squeezed states in the de Sitter vacuum,” *Phys. Rev. D* **68** (2003) 064002, [arXiv:hep-th/0305056](#).
- [63] S. Choudhury and S. Panda, “Quantum entanglement in de Sitter space from stringy axion: An analysis using  $\alpha$  vacua,” *Nucl. Phys. B* **943** (2019) 114606, [arXiv:1712.08299 \[hep-th\]](#).
- [64] D. Baumann and L. McAllister, *Inflation and String Theory*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 5, 2015. [arXiv:1404.2601 \[hep-th\]](#).
- [65] J. Grain and V. Vennin, “Canonical transformations and squeezing formalism in cosmology,” *JCAP* **02** (2020) 022, [arXiv:1910.01916 \[astro-ph.CO\]](#).
- [66] L. Grishchuk, H. A. Haus, and K. Bergman, “Generation of squeezed radiation from vacuum in the cosmos and the laboratory,” *Phys. Rev. D* **46** (1992) 1440–1449.
- [67] J. Martin and V. Vennin, “Real-space entanglement in the Cosmic Microwave Background,” [arXiv:2106.15100 \[gr-qc\]](#).
- [68] L. P. Grishchuk and Y. V. Sidorov, “Squeezed quantum states of relic gravitons and primordial density fluctuations,” *Phys. Rev. D* **42** (Nov, 1990) 3413–3421.
- [69] S. Tarzi, “The inverted harmonic oscillator: some statistical properties,” *Journal of Physics A: Mathematical and General* **21** no. 14, (Jul, 1988) 3105–3111.
- [70] C. Yuce, A. Kilic, and A. Coruh, “Inverted oscillator,”

- Physica Scripta* **74** no. 1, (2006) 114.
- [71] V. Subramanyan, S. S. Hegde, S. Vishveshwara, and B. Bradlyn, “Physics of the inverted harmonic oscillator: From the lowest landau level to event

horizons,” *Annals of Physics* **435** (2021) 168470. Special issue on Philip W. Anderson.

# Geopolymer Concrete: A Material for Sustainable Development in Indian Construction Industries

Manvendra Verma <sup>1</sup>, Nirendra Dev <sup>1</sup>, Ibadur Rahman <sup>2</sup>, Mayank Nigam <sup>3</sup>, Mohd. Ahmed <sup>4,\*</sup> and Javed Mallick <sup>4</sup>

<sup>1</sup> Department of Civil Engineering, Delhi Technological University, Delhi 110042, India; manvendraverma\_2k16phdce05@dtu.ac.in (M.V.); nirendradev@dce.ac.in (N.D.)

<sup>2</sup> Department of Civil Engineering, Jamia Millia Islamia, Delhi 110020, India; irahman1@jmi.ac.in

<sup>3</sup> Department of Civil Engineering, GLA University, Mathura 281406, India; mayank.nigam@gla.ac.in

<sup>4</sup> Civil Engineering Department, College of Engineering, King Khalid University, Abha 61421, Saudi Arabia; jmallick@kku.edu.sa

\* Correspondence: mall@kku.edu.sa

**Abstract:** Geopolymer concrete (GPC) is a new material in the construction industry, with different chemical compositions and reactions involved in a binding material. The pozzolanic materials (industrial waste like fly ash, ground granulated blast furnace slag (GGBFS), and rice husk ash), which contain high silica and alumina, work as binding materials in the mix. Geopolymer concrete is economical, low energy consumption, thermally stable, easily workable, eco-friendly, cementless, and durable. GPC reduces carbon footprints by using industrial solid waste like slag, fly ash, and rice husk ash. Around one tonne of carbon dioxide emissions produced one tonne of cement that directly polluted the environment and increased the world's temperature by increasing greenhouse gas production. For sustainable construction, GPC reduces the use of cement and finds the alternative of cement for the material's binding property. So, the geopolymer concrete is an alternative to Portland cement concrete and it is a potential material having large commercial value and for sustainable development in Indian construction industries. The comprehensive survey of the literature shows that geopolymer concrete is a perfect alternative to Portland cement concrete because it has better physical, mechanical, and durable properties. Geopolymer concrete is highly resistant to acid, sulphate, and salt attack. Geopolymer concrete plays a vital role in the construction industry through its use in bridge construction, high-rise buildings, highways, tunnels, dams, and hydraulic structures, because of its high performance. It can be concluded from the review that sustainable development is achieved by employing geopolymers in Indian construction industries, because it results in lower CO<sub>2</sub> emissions, optimum utilization of natural resources, utilization of waste materials, is more cost-effective in long life infrastructure construction, and, socially, in financial benefits and employment generation.

**Keywords:** geopolymer concrete; fly ash; GGBFS; compressive strength; mechanical properties; durability properties and bond strength



**Citation:** Verma, M.; Dev, N.; Rahman, I.; Nigam, M.; Ahmed, M.; Mallick, J. Geopolymer Concrete: A Material for Sustainable Development in Indian Construction Industries. *Crystals* **2022**, *12*, 514. <https://doi.org/10.3390/cryst12040514>

Academic Editor: Dawei Wang

Received: 10 March 2022

Accepted: 29 March 2022

Published: 7 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

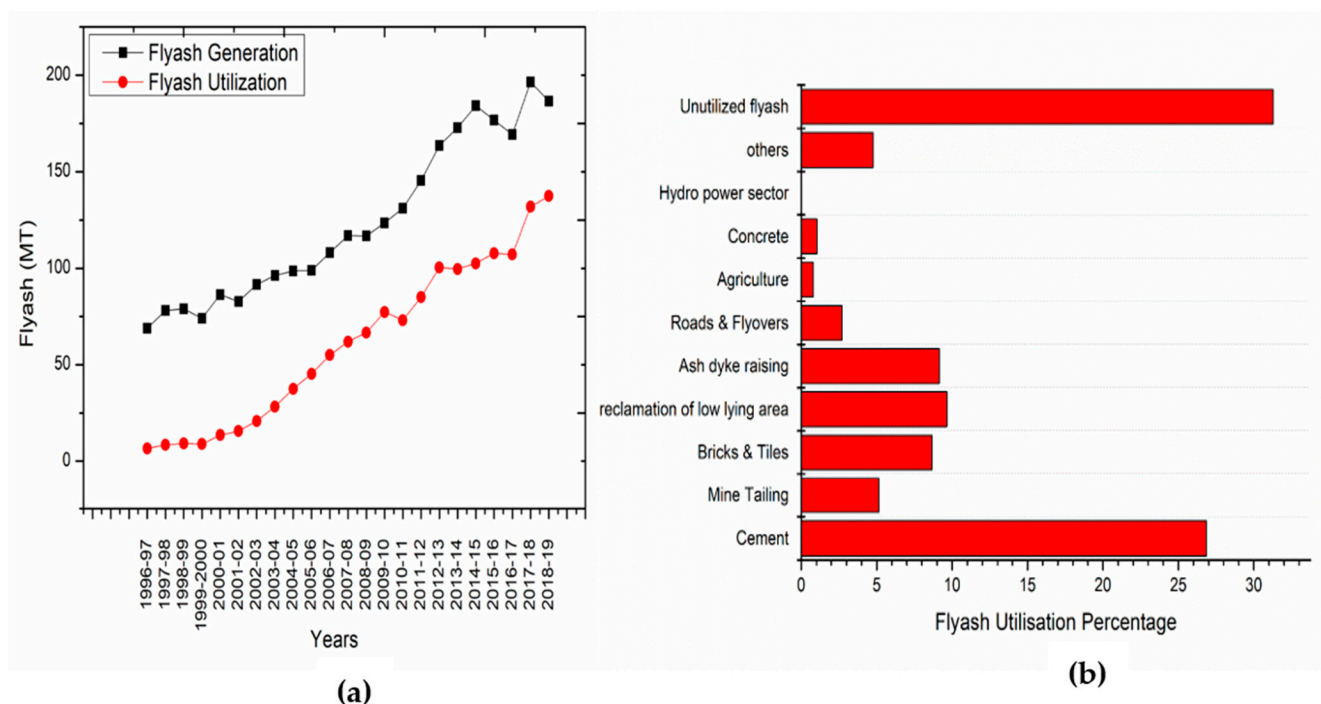
After water, concrete is the world's second most abundant resource. Portland cement was traditionally used as a binding material in concrete. In 2018, India manufactured approximately 502 million tonnes of cement. Portland cement is not environmentally sustainable, and there are many environmental concerns involved with its manufacturing. Cement processing accounts for about 8% of greenhouse gas emissions from carbon dioxide [1]. Cement development depletes renewable resources and pollutes the climate, endangering our future [2]. We must pursue sustainable development because the earth's natural resources are limited [3].

On the other hand, there is an excess of agricultural solid waste on the planet, such as fly ash, slag, and rice husk ash. In India, thermal power plants generate 200 million tonnes



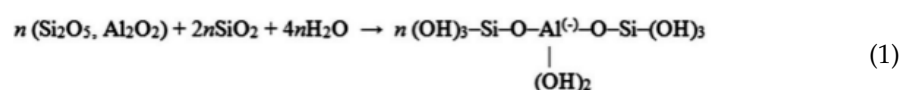
of fly ash per year, while steel manufacturing plants produce 20 million tonnes of slag per year. Fly ash and slag are wastes from that sector; they deposit industrial solid wastes on cultivated land and contaminate the atmosphere due to the restricted mode of industrial solid wastes [4].

The geopolymer was introduced by the Davidovits in 1978 and is made by the alkaline solution's activation (consists of sodium or potassium silicate and sodium or potassium hydroxide) of the high alumina silica-rich materials. It is like ceramic composites that make the bond between alumina and silica. Geopolymer-based concrete using fly ash has a high potential for the construction industry to replace ordinary Portland Cement (OPC)-based concrete with comparable structural properties [5]. Figure 1a describes the fly ash generation and utilisation from 1996 to 2019 and data from the mineral book 2019. Figure 1b shows the description of the fly ash used in various departments in 2019 with their percentage utilisation [6].

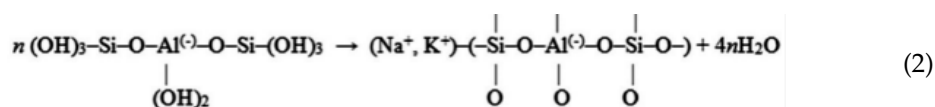


**Figure 1.** (a) Fly ash generation and utilisation in India 1996–2019; (b) Utilisation of the fly ash in various in India 2019 [6].

Geopolymerisation is a reaction in which the formation of the alumino-silicate gel structure uses silica and alumina content from the pozzolan binding materials like fly ash and slag activated by an alkaline solution ( $\text{NaOH}$  and  $\text{Na}_2\text{SiO}_3$ ). Figure 2a shows the mechanism of geopolymerisation, describing the constituents and their effect on the end products. The microstructure of the crosslinked geopolymer shows higher stability against external environmental conditions. In the geopolymerisation process, alkali metal plays an essential role in the reaction rate and end products. It is an inorganic polymer that shows an amorphous nature at high temperatures. Heat-curing increases the geopolymerisation rate compared to ambient-temperature curing. Equations (1) and (2) show the geopolymerisation reaction process from starting to the specimen hardening [7]. The end products of the geopolymerisation show a three-dimensional structure.







The alkali metals like Na and K are primarily used in the activator solution to activate the pozzolanic binding material, and they form the zeolitic structure during the geopolymerisation reaction. The  $\text{SiO}_2/\text{Al}_2\text{O}_3$  ratio plays a vital role in forming the microstructure at the curing temperature. Figure 2b shows the curing temperature's effects on the  $\text{SiO}_2/\text{Al}_2\text{O}_3$  ratio forming the end product's zeolitic structure. It also shows the gel product's amorphous nature at elevated temperatures by forming a poly-sialate-siloxo  $(\text{Si-O-Al-O-Si-O})_n$ . The content present in the mix design does not involve the end product of the geopolymerisation reaction.

### 1.1. Research Significance

Geopolymer concrete is economical, low energy consumption, thermally stable, easily workable, eco-friendly, cementless, and durable. GPC reduces carbon footprints by using industrial solid waste like slag, fly ash, and rice husk ash. Around one tonne of carbon dioxide emissions are needed to produce one tonne of cement; these emissions directly pollute the environment and increase the world's temperature by increasing greenhouse gas production. For sustainable construction, GPC reduces the use of cement and presents an alternative to cement for the material's binding property. GPC has better physical, mechanical, and durable properties than Portland cement concrete and it is highly resistant to acid, sulphate, and salt attack. Thus, geopolymer concrete is an alternative to Portland cement concrete and is a material with a potentially large commercial value for sustainable development in the construction industry. The study presents the state-of-art review of the GPC production, its performance (environmental, socio-economic) and applications as a material that can be useful for sustainable development in construction industry. This review is an effort to join and share the scientific and technical information, and to bridge the knowledge gap in GPC utilization as a material for sustainable development in construction industry.

### 1.2. Definition of Abbreviations Used in the Manuscript Script

Geopolymer concrete (GPC); plain cement concrete (PCC); ordinary portland cement (OPC); ground granulated blast furnace slag (GGBFS); air-cooled slag (ACS); Alkali-Activated Fly ash/Slag (AFS); Alkali-Silica Reaction (ASR); Biomass Fly ash (BFA); Bottom Ash (BA); alkali-activated fly ash (AAFA); fibre-reinforced plastic (FRP); GFPR (glass fibre reinforced polymer); fly ash-slag geopolymer concrete (FSGC); high-range water-reducing admixture (HRWRA); high volume fly ash (HVFA); high volume slag (HVS); light weight aggregate (LWA); lightweight expanded clay aggregate (LECA); microencapsulated phase change materials (MPCM); recycled concrete aggregates (RCA); self-compacting geopolymer concrete (SCGC); naphthalene formaldehyde (SNF); polycarboxylate ether (PCE); super plastically formed (SPF); high performance ash (HFA); rice husk ash (RHA); C-S-H (Calcium Silicate Hydrate (C-S-H); calcium aluminosilicate hydrate (C-A-S-H); calcium-sodium aluminosilicate hydrate [C-N-A-S-H]; energy dispersive X-ray analysis (EDXA).

### 1.3. Research Methodology

The main aim of this research is to review GPC as a material for sustainable development in the construction industry of developing countries, i.e., material that satisfies environmental (high performance, durability, and eco-friendly applications) and socio-economic (society acceptance, commercial, and cost) aspects of sustainability. The research methodology is based on the processing of the secondary data, i.e., published research and books. The secondary data are taken from the databases of scientific articles, including Scopus, ScienceDirect, and Google Scholar. It is focused on the keyword "geopolymer

concrete”, and is supported by other key words; for example, “mechanical properties” and “durability properties”.

## 2. GPC Manufacturing Studies

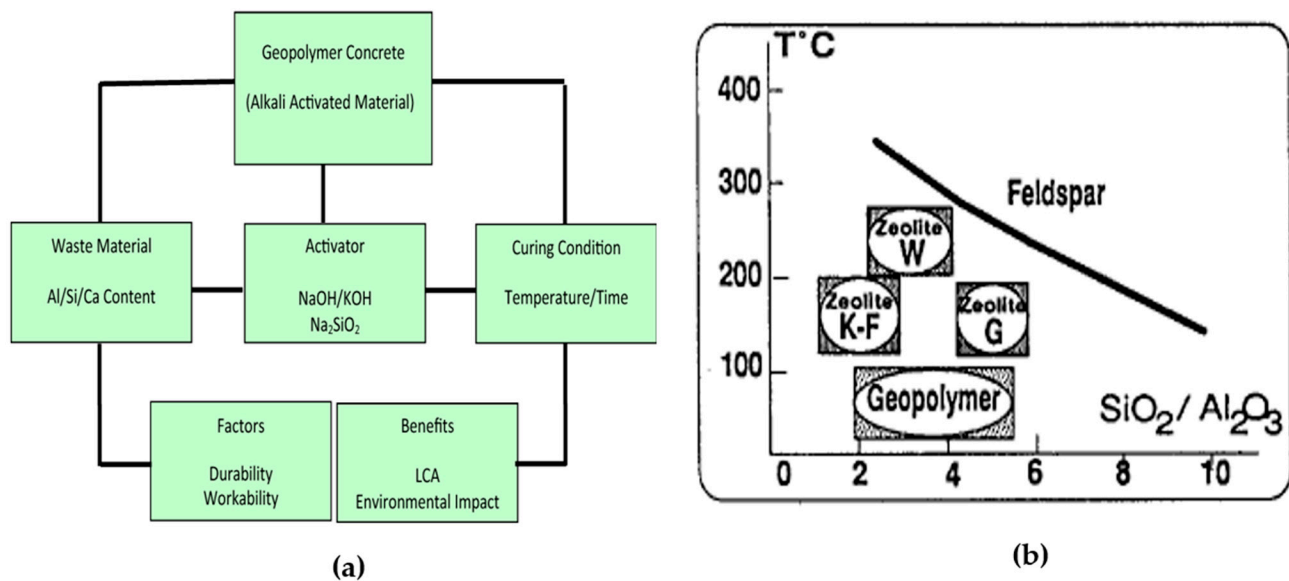
Geopolymer mortar and concrete require water and superplasticiser to increase workability because the sodium silicate and NaOH concentration increment degrades the fresh mortar’s and concrete’s flowability and workability; usage of superplasticizer also harms the strength of the geopolymer mortar [8]. The slag and sodium silicate dosage increment in the GPC mix reduces the porosity and simultaneously reduces the mesopore volume, and creates a denser matrix, due to the more reactive products. The autogenous shrinkage of geopolymer mortar develops due to the self-desiccation in the solid-state but chemical shrinkage in the fresh state. The AFS mortar shows higher drying shrinkage than the OPC specimens due to the mixed samples’ high capillary stress [9]. The AFS mortar has higher compressive strength, flexural strength, and lower water absorption than the OPC mortar samples. The hydration product of the AFS paste is mostly amorphous, and the higher fly ash content shows less length change [10].

The water to binder ratio increment in the AFS mortar increases the flow and workability of the mortar. A fine aggregate to binder ratio above 2.5 results in the instantaneous degradation in the flow of mortar. The GGBFS-based mortar shows a higher strength than the fly ash-based mortar, and the compressive strength of the mortar increases with the increase of the fine aggregate to binder ratio up to 2.5. Beyond 2.5, the strength drastically degrades [11]. The geopolymer mortar strength increases with the rise in temperature, but it declines beyond the 800 °C temperature. The geopolymer mortar compressive strength depends on the bonding between the binder and aggregates, and the increment aggregates in the mortar reduce the geopolymerization reaction [12]. The thickness of the geopolymer paste is the main factor influencing the mortar’s heat-resistant property, and changes in aggregate and the mass ratio of paste to fine aggregate affect the paste’s thickness. If it is too thick, the water vapour can collapse the structure when released, and if it is too thin, the paste cannot bond aggregate together.

High calcium fly ash usage in the mix provides a high-strength geopolymer mortar up to 86 MPa at 28 days [12]. The heat evolution during the geopolymerization is first due to sodium silicate and gelation, and second due to bulk hydration of GGBFS, responsible for the strength development observed in the calorimeter. After the reaction, the products are amorphous and uniformly distributed throughout space in the matrix [13].

### 2.1. Fly Ash Based GPC

Fly ash rich in silica and alumina constituents shows pozzolanic characteristics and binding properties. The fly ash is used in the GPC, which provides better workability and higher solid to liquid ratios than the other pozzolanic materials like slag, metakaolin, and rice husk ash due to its small particle size, and spherical and porous properties [14]. In the GPC, the bottom ash usage increases the strength and workability of the increasing surface for reactivity and collapses the pores in the matrix due to their higher fineness [15]. The increase in ultrafine fly ash dosage decreases the setting time, but up to 15% dosage, ultrafine fly ash usage and the setting time are longer for the fly ash-based GPC. The use of ultrafine fly ash increases by 10%, reducing the matrix’s porosity of the GPC [16]. The fineness of fly ash is the deciding factor for the amount of fly ash required for the mix design of GPC, because increasing the fineness of fly ash decreases the quantity requirement for mix design [17]. The fineness of fly ash plays a vital role by reducing the porosity and water absorption capacity of both concrete GPC and OPC concrete [18]. The fineness of the fly ash increases the workability, density, and strength of the GPC. It slightly affects the alkalinity of the GPC but is similar to the OPC concrete [19]. Figure 3a shows the effect of the fineness of the fly ash on the density of the concrete mix. The GGBF/fly ash ratio is 20/80, which is the optimum point of mechanical strength [20].



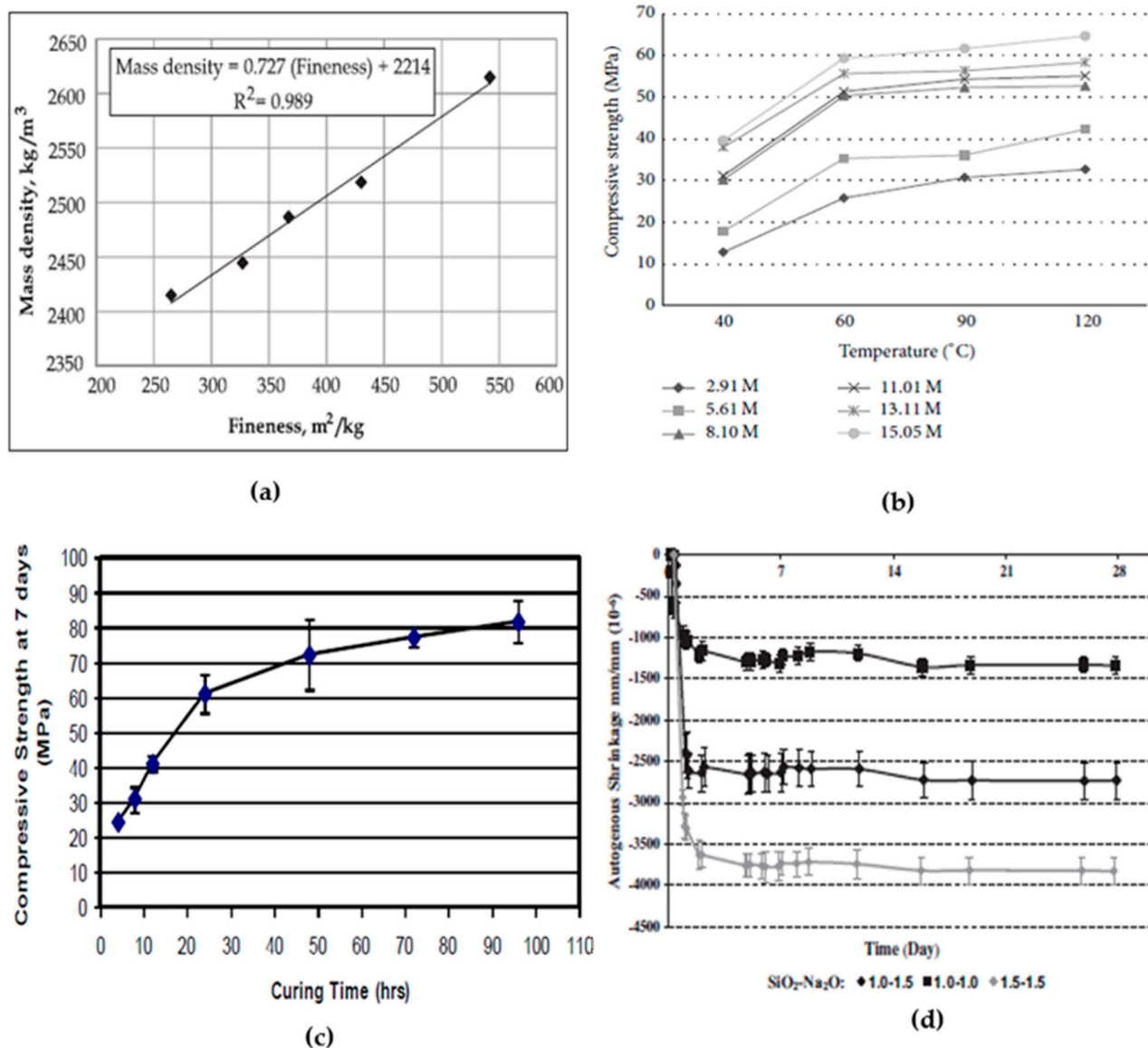
**Figure 2.** (a) Mechanism of Geopolymerisation; (b) crystallisation temperature ranges for K<sub>2</sub>O-Al<sub>2</sub>O<sub>3</sub>-SiO<sub>2</sub> system [21].

## 2.2. GGBFS Based GPC

In the GPC, increasing the dosage of GGBFS decreases the mix's workability due to the angular particles of samples compared to the spherical particles of fly ash. It decreases the ambient temperature curing setting time due to the reaction of calcium present in the slag. The slag's usage with fly ash in the GPC presents a similar compressive strength to ambient-cured specimens to the water-cured OPC concrete. The specimen strength gains slow down after 28 days but strengthen up to 180 days [22]. In the GPC, the strength increases with the dosage of GGBFS in the mix design and eliminates heating conditions for curing [23]. The mix design's slag/fly ash ratio increases, showing less mass loss at the elevated temperature [24]. The heat-cured fly ash-based GPC gives a higher strength, but this could be achieved by replacing some fly ash with the GGBFS at ambient-curing conditions [25].

## 2.3. Effect of Molar Ratios of Alkaline Solution

The compressive strength of GPC increases with an increase in the molarity of sodium hydroxide, but the workability declines [26–31]. The GPC uses sodium hydroxide to increase the mechanical properties compared to the OPC concrete, which results in its higher strength and performance [26,32]. Instead of the sodium silicate, KOH generates a much more heterogeneous structure, with higher porosity and lower strength developed after hydration [13]. Sodium hydroxide with sodium silicate is used as an alkaline activator of the binder in a 1:1 ratio in the mix design to produce GPC strength of 40 MPa and above, and verify the potential of fly ash to replace the cement [33]. The fly ash/slag-based paste setting time reduces as the slag's dosage, sodium silicate, and sodium hydroxide molarity increase. The initial time is 55 min, and the final time is 160 min for the fly ash/slag-based paste at a room temperature of 17 °C [34]. The compressive strength of 30 MPa was quickly achieved with the range of molarity of 9.5–14 M of NaOH at room temperature in 28 days [35]. The geopolymer matrix developed in GGBFS is similar to the matrix developed in the absence of GGBFS. If the NaOH concentration is low, the calcium content of the GGBFs participates in the development of amorphous CSH gel and the formation of calcium-based geopolymer. If the concentration of NaOH is high, the precipitation of calcium hydroxide develops [36].



**Figure 3.** (a) Effect of fly ash fineness on mass density; (b) effect of degree of heating for different concentrations of NaOH solution on compressive strength at solution-to-fly ash ratio of 0.40 [27]; (c) effect of curing time on compressive strength [30]; (d) autogenous shrinkage of AAFA pastes with different SiO<sub>2</sub> and Na<sub>2</sub>O content, cured at 40 °C (corrugated tube method, measurement started after the final set, the autogenous shrinkage was the mean value of three replicates for each specimen).

The addition of sodium silicate extensively increases the hybrid geopolymer system's early strength development [37]. The interface between aggregate and geopolymeric paste is not apparent after the high soluble silicate dosage in the salt-free geopolymeric mortar. In the geopolymerisation process, soluble silicates are very useful in degrading alkali saturation in the GPC pore solution even after using a high alkali concentrated activating solution and promoting the higher inter-particle bonding binder well as the aggregate surface. Geopolymerisation leads to reliable geopolymer results like binders, mortars, and concrete [38]. The ratio of sodium silicate to sodium hydroxide (by mass) increases with the increasing GPC compressive strength. If the alkaline solution content increases in the mix, then the setting time and workability increase with the reduction of the compressive strength. The alkaline ratio increased from 1.5 to 2.5, and the alkaline solution increased from 35% to 45% of the mass of binder, and the optimum point was found at 2.5 alkaline ratio and at 40% alkaline solution in the mix [39]. The Na<sub>2</sub>CO<sub>3</sub> activator is

the cheapest activator in terms of the relative cost per ton and strength of geopolymer mortars [40]. The yield point of the compressive strength of mortar and paste found at the molar ratio of  $\text{Na}_2\text{O}/\text{SiO}_2$  is 0.40 and shows the paste's denser morphology. If the  $\text{Na}_2\text{O}/\text{SiO}_2$  molar ratio increases, then the consistency of the paste decreases. It forms an alumino-silicate gel with the fly ash content and is responsible for the higher workability and the mechanical properties of the GPC samples in the hardened state [41]. The increased ratio of  $\text{Na}_2\text{SiO}_3/\text{NaOH}$  decreases the fine aggregate requirement in the mix design and increases the extra water requirement in the mix [17]. Figure 3b shows the relationship between compressive strength and curing temperature with various molarities of NaOH.

#### 2.4. Effect of Calcination Temperature

The geopolymer manufactured from calcinated clays shows better mechanical properties than a geopolymer obtained from un-calcinated clays [42,43]. Calcination conditions (calcination temperature, heating rate, calcination time, atmosphere, and cooling rate), to achieve dehydroxylation of a given clay mineral, affect the pozzolanic reactivity, hydration kinetics, and consequently, strength and durability of geopolymer concrete [44,45]. The geopolymer derived from clay minerals calcinated at temperature of 900 °C shows better characteristics, such as compressive strength, water absorption, and stability in the aggressive environment of chloride ions, than the clay minerals calcinated at lower temperature than 900 °C [46].

#### 2.5. Effect of Particle Size Fraction and Types of Aggregates

In the GPC, the aggregate size has a critical effect on the elevated temperature; the spalling of GPC reduces with the increase in the maximum aggregate size of the GPC mix design. The size of the fracture process zone increases with the aggregate size by shielding the crack tip. The maximum aggregate size increase in mix design provides higher resistance to fire by improving stability [47]. The GPC has better stability against elevated temperatures and better compatibility between the aggregates and pastes [48]. The geopolymer aggregate was developed from the mine tailings of fly ash and is used as a lightweight aggregate in mortars and concrete. The geopolymer aggregate shows better mechanical properties than the LECAs in the mix design solid state, whereas the rheological properties are the same. The geopolymer aggregate shows a better crushing value and excellent bonding due to its rough surface [49]. The geopolymer sand dosage in the mix design shows better mechanical properties up to 40% replacement of natural sand, but beyond 40%, the mix's mechanical properties decrease with the increasing content of geopolymer sand [50]. The concrete strength increases with the dosage of m-sand in the mix by replacing natural sand up to 20%, but beyond 20%, strength slightly decreases compared to the nominal mix [51]. The GPC uses 100% m-sand to replace natural sand, which shows adequate strength, and it is an economical, environment-friendly material that reduces carbon footprints by up to 80% [52]. In the GPC, the dosage of the RCA in the mix design by the partial replacement of the natural sand shows an adverse effect on the compressive strength, splitting strength, sorptivity, chloride ion penetration, water absorption, and volume of voids, but the properties are better than the OPC concrete [53]. The ductility of the GPC structural element is shown to increase with the partial replacement of natural sand in GPC design mix by waste aggregate of lime stone [54].

#### 2.6. Effect of the Alkali Metal Activator

Geopolymer concrete is made by the geopolymerisation reaction among the binding constituents of the concrete, in which alkali metal cations play an important role in every stage of the reaction of geopolymerisation. The type of alkali metal cation is also an essential factor for the geopolymerisation reaction in all the reaction stages. It also affects the speed of the setting time of the paste and the condensation reaction of the GPC. The alkali metal cation works as a template for the reaction in directing and controlling the molecules' reaction and final structure formation. The size of the alkali metal cation also plays a role



in forming the structure of the geopolymer. The potassium cation (K) is responsible for a higher condensation than the sodium (Na) cation in the same condition. The potassium cation shows a higher compressive strength due to the higher surface area, produces a higher amorphous structure, and is less resistant to HCl attack. The alkali metal depends on the source material (Si and Al content) for the reaction because the different source materials directly affect the end products' physical and chemical properties [55].

The GGBFS-based geopolymer directly depends on the chemical composition of reactions. Increasing the sodium hydroxide (NH) and sodium silicate (NS) content in the composition increases the compressive strength, setting time and fire-resistant characteristics of the end products [56]. The geopolymer product's thermal stability with sodium-containing activators is less than activated potassium, due to functional changes in the microstructure. The sodium-activated geopolymer shows deterioration of strength sharply at 800 °C due to increases in the average pore size. The materials prepared with potassium silicate and fly ash show better thermal stability than the sodium-containing activator used. In the sodium activator, the amorphous structure is replaced by the crystalline Na-feldspar, but in the potassium activator used, the materials remain amorphous up to 1200 °C. The geopolymer material produced using fly ash and sodium or potassium silicate shows very high shrinkage and significant changes in the compressive strength on increasing the temperature range of 800–1200 °C [57].

In the geopolymer materials, a sodium-containing activator shows thermal stability up to 500 °C, which is more stable than the regular concrete mix [58]. The geopolymer's thermal shrinkage mostly depends on the Si/Al ratio and the alkali used in the activator. The low Si/Al ratio used in the mix shows better thermal stability than the higher Si/Al ratio and shows densification at elevated temperatures [59]. The alkali activator content also plays an effective role in finding the compressive strength of the GPC. Increasing the alkali content increases the compressive strength due to the solubility of the alumino-silicate with the increase in Na<sub>2</sub>O concentration, but beyond 10% it does not show a significant increase in strength. The content of Na<sub>2</sub>O ranges between 6–15% by mass. Sodium and silica content at a higher level show higher autogenous shrinkage than the cement matrix, but the autogenous shrinkage mechanism of the geopolymer matrix is different from the cement paste. In the geopolymer, the reorganisation and geopolymerisation of the structure develop a finer pore size distribution that generates autogenous shrinkage instead of self-desiccation in the cement matrix [60].

### 2.7. Effect of Ratio of Activator Liquid to Fly Ash/Slag

If the activator to binder ratio decreases from 0.40 to 0.35, then the fresh mix's workability decreases, but extra water enhances the workability and strength are reduced. The strength reached up to 51 MPa at 28 days of GPC, containing 20% slag and 80% fly ash as a binder and cured at 20 °C [22]. In the fly ash-slag based GPC, the slag/fly ash plays an essential role in the mechanism of the geopolymerisation for strength development. If the slag content in the mix design is over 50% of the binder, this demonstrates the calcium silicate hydrate gel's primary reaction with the Na and Al, leading to the formation of the C–N–A–S–H bond. If increasing the fly ash content in the binder, then the form of N–C–A–S–H type gel as a primary reaction bound the water tightly in the composition and got a higher degree of crosslinking compared to the composition formed of C–A–S–H type gel in the slag binder. This study provides critical evidence of the different mechanisms of the geopolymerisation based on the slag/fly ash ratio content as a binder in the mix design. If the small dosage of the slag in the fly ash is in the mix, the reactions slow down the formation of the C–A–S–H binding gel and the formation of the hybrid C–N–A–S–H gel due to the overtime release of Si and Al in the reactions. Increasing the dosage of fly ash in the mix promotes the zeolite development in the hybrid gels after 28 days of higher temperature curing with a lower activator concentration. After the long curing period, the gel microstructure formation predicted the mechanical capability and durability performance of the GPC samples [61]. If the 15% ACS replaces the GGBFS in GPC, it leads

to the thermal stability of the GGBFS-based GPC up to 1000 °C and shows the amorphous nature to resist the heat treatments used in the formation of refractory bricks. Increasing the dosage of ACS by up to 25% decreases the thermal stability of the geopolymer. If the GGBFS is replaced by silica fume by 10%, it increases the strength at a high level and shows thermal stability up to 500 °C, whereas the ACS dosage shows thermal stability up to 800 °C [58]. If the dosage of slag and sodium silicate is increased, it increases the autogenous shrinkage and chemical shrinkage, and if the water binder ratio decreases, drying shrinkage decreases and autogenous shrinkage increases in the AFS mortar. The GPC's slag content above 70% in the mix reflects the rapid setting and cracks generated due to the autogenous shrinkage and shows the hydration products' denser matrix. The geopolymer matrix is not affected by using superplasticisers in the mix design [62]. If the GGBFS, OPC, and CH in the fly ash increase the concrete's compressive strength, it is applicable when no extra water is added to the alkaline solution [63].

With the increasing GGBF dosage, the GPC's workability and setting time would decrease. When GGBFS dosage was up to 30% of binder, the design mix got 55 MPa of GPC and 63 MPa of geopolymer mortar at 28 days [39]. In the GGBFS/fly ash-based GPC, the formation of C-S-H gel at 27 °C by activating GGBFS shows minimal interaction between the fly ash and GGBFs due to different reactions occurring. The mix design strength was developed by the C-S-H gel formation products [64]. The presence of the CSH gels and geopolymeric gels in the matrix enhances the system's strength with the usual quantity of GGBFS. The geopolymer matrix's CSH gel formation worked as a micro-aggregate and resulted in excellent mechanical strength [36].

In the GPC, the fly ash to activator ratio plays an essential factor regarding strength and fire-resistant matrix and silicate to hydroxide, binder age and curing period show negligible effects on the early strength of the GPC. The optimum point of strength and fire-resistance is found at  $\text{Na}_2\text{SiO}_3/\text{KOH} = 2.5$  and  $\text{FA}/\text{activator} = 3.0$  [65,66]. In the GPC, water in the fresh concrete mix and removed through evaporation causes a crack [67]. The workability of the fresh concrete increases with the increment of water dosage in the mortar mix. In the GPC, when the molarity of sodium hydroxide is above 10 M, it shows an increment of workability with fewer effects on strength by increasing the water content in the mix. The strength of the GPC decreases with the increase of the water to geopolymer solids ratio by mass. This leads to the loose microstructure of the geopolymer paste in the concrete [68].

## 2.8. ITZ (Interfacial Transition Zone)

The reaction mechanism of ITZ is slightly different from the matrix reaction. In the ITZ reaction, a tremendous number of voids initially appear with the water content at ITZ, but the hydration products fill the voids after the hydration process. The difference between the microstructure of the matrix and ITZ is challenging to identify after the reaction. EDAX results describe K/Al and Si/Al content in the ITZ being found at a higher level than in the bulk matrix; well-developed crystalline is not present in the ITZ, and forms a sponge-like amorphous gel [69]. The ITZ properties of sodium silicate activated mortar are good with very low porosity at the interface, and thermal activation provides early strength even when the reaction slows down to give a large age span before setting time. A high concentration of the alkaline solution is required for the strong ITZ bond between the siliceous aggregate and fly ash [70]. According to SEM and Nano-indentation analysis, there is no ITZ bond between the old cement paste and the geopolymer matrix [71]. The superplasticizer used in the mix design improved the microstructure of the ITZ of concrete. The ITZ thickness directly affects the concrete's compressive strength, and it is affected by the superplasticizer dosage in the mix. If the ITZ thickness decreases, then the compressive strength of the mix increases. The superplasticizer's lower dosage in the mix develops the loose and porous ITZ between the aggregate and binder and decreases the mixed concrete's performance by reducing the compressive strength. When the super-plasticiser's dosage is high, the mix develops a dense ITZ and increases the concrete's performance [72]. Soluble silicate in



the mix as an activator liquid plays an essential factor in developing the ITZ between the aggregate and paste in the GPC. If the soluble silicate quantity is meagre, the mix shows the weak compressive strength of the paste mortar and concrete compared to the high dosage of silicate soluble. The chloride present in the mix shows debonding between aggregate and paste by crystallising paste at ITZ. LWA used as aggregate in GPC shows excellent bonding at ITZ due to its porous and rough aggregate surface. The bonding zone deteriorates at a temperature of 800 °C due to the dehydration of microstructural water, and swelling of the unreacted silicate content in the matrix creates microcracks between the paste and LWA [73]. The liquid-to-binder ratio plays a vital role in the ITZ of GPC [74].

### 2.9. Effect of Curing Conditions

The curing temperature plays a vital role in the setting and hardening of the GPC. The GPC samples cured at higher temperatures obtain concrete in 4 h at ambient temperature, but at an ambient temperature below 10 °C, the setting of fresh concrete takes up to 4 days without losing quality. The GPC specimens cured at the higher temperature reached mechanical strength within 1 day, but the hardening of the fresh concrete at ambient temperature increased quality strength after 28 days compared to strength gain in 1 day. Curing time also plays an essential role in gaining strength with temperature [75]. The compressive strength of the GPC increases with the increasing curing temperature from 30 °C to 90 °C [26,30]. If the curing at the ambient temperature is impossible due to the delayed time of fly ash-based GPC, the temperature for curing is favourable for gaining higher strength, and the long curing time enhances the geopolymerisation process. GPC samples cured at higher temperatures for a long time lead to microcavities developing in the microstructure, creating cracks in the sample due to water evaporation from the matrix [76]. The GPC flexural strength is higher than the OPC concrete of the same compressive strength [63]. In the natural pozzolanic-based GPC, the compressive strength increases with increasing time and temperature. The GPC applications are cured in both conditions; atmospheric pressure up to 100 °C and autoclave curing above 100 °C. It enhances the mixed samples' compressive strength by eliminating the micro-cracks in the samples [77]. If the curing time increases from 6 h to 96 h, it increases the compressive strength of the GPC, but beyond 48 h of curing, the compressive strength of the samples is not significant [30]. Pre-curing at an ambient temperature above 95% humidity at room temperature before heat curing is beneficial for strength development [68]. Temperature curing is responsible when the specimens cured for 1 h at elevated temperatures do not reach remarkable strength. The longer curing time is responsible for the strength development at early periods, by accelerating the reaction rates. The curing age enhances the polymerisation process from 4 h to 96 h and increases the compressive strength, but the strength development of the samples is achieved at 24 h of curing, so there is no need to cure beyond 24 h [78].

The geopolymerisation reaction increases with the increase in curing temperature and develops the early strength of the GPC. The GPC samples cured in the fog show the high absorption of moisture. The end products of GPC had a mostly open microstructure when steam cured [58]. In the GPC, the oven-cured specimens reached 90% strength in 3 days and 28 days of compressive strength, but the ambient-cured samples reached up to 82% strength in 28 days. The ambient-cured specimens' ultimate strength was higher than the oven-cured samples because the strength rate beyond 7 days is not significant [79].

The elastic modulus of GPC is directly affected by curing temperature. The MOE of specimens increases with the curing's increasing temperature up to a limit and related to the water to binder ratio. The evaporation of the water from the matrix during the temperature curing reduces the elastic modulus of the GPC specimens [80]. Figure 3c shows the effect of curing time on the strength of mixed samples.

### 2.10. Effect of Calcium Content

The calcium content in the slag and OPC is used in the GPC to form the CSH gel with the geopolymeric gel at low alkaline conditions and improve the mix's compressive

strength. The lower calcium content available to less CSH gel resulted in the mix's lower overall strength. The calcium content plays a minor role in strength improvement in the high alkaline conditions and forms the CSH gels' precipitation [81]. The high calcium BA's mechanical strength depends on the fineness of the raw sample of BA and the water content in the mix design [82]. The calcium content present in the mix's slag is essential for both early and more prolonged age. If the reaction rate is slow and low strength development confirmed, the low calcium fly ash is used as a binder with the lower concentration of alkali activator used without heat curing. The C-S-H/C-A-S-H precipitation formation initiated the strength development of the GPC fresh concrete, but in the fly ash GPC, the concrete hardens by forming alumina-silicate precipitation. The free calcium content from slag and fly ash dissolution increases the gel's formation and later develops the hardened concrete's strength [83]. The higher calcium in the mix emphasises the C-A-S-H type gel end product with a chain structure. A high amount of calcium content is attributed to high strength GPC [26,31].

#### *2.11. Effect of Superplasticizer Addition*

A superplasticizer is used to increase concrete strength by lowering the water content of the mix design. The addition of a superplasticizer to the binder mass of up to 2% improves the workability of fresh GPC while having little effect on the strength of hardened GPC [30,84]. The use of superplasticizers is detrimental to the GPC's high-temperature performance [85]. The superplasticizer based on SNF is very effective for GPC [86]. The naphthalene-based superplasticizer is very useful, improving the workability of a slump by about 136%, while the PCE-based superplasticizer increases the workability by around 145%, but it affects the strength of the product blend specimens, reducing the strength by about 29%. In certain cases, SNF-based superplasticizer has no negative impact on concrete strength [87].

The addition of citric acid and sucrose to fly ash-based GPC is an excellent alternative chemical admixture for improving the paste's rheological properties while increasing its weight. Sucrose acts as a retarder in the GPC mixture, while citric acid speeds up the hardening process. The mixture influenced the compressive strength after the sucrose was added, affecting the paste's porosity. Sucrose-added samples resulted in a relationship between compressive strength and porosity, directly attached to compressive strength. Sucrose could be used as a chemical admixture as a retarder in the GPC [88]. The polycarboxylate-based superplasticizer has a greater retarding effect on the fly ash-slag-based GPC without affecting the paste's heat of hydration and has greater workability than the SNF-based superplasticizer. Increasing the content of PCE-based superplasticizer above 2% has little impact on the growing strength before 7 days, but it can have a detrimental effect on the strength of the GPC. The superplasticizer improved the workability and mechanical strength of SCGC and the microstructure of bonding between the paste and aggregate at GPC's ITZ. The microstructure of ITZ differs due to changes in thickness caused by variations in superplasticizer material in the blend, which also influences the compressive strength of GPC. By decreasing ITZ thickness, increased superplasticizer use improves the compressive strength of SCGC and the engineered quality microstructure. Superplasticizer material greater than 2% was inadequate to produce desirable workability with resistance to segregation.

In contrast, 6% and 7% of superplasticizer contents have the necessary workability properties within the EFNARC limits [89], with 7% producing the essential intensity at all ages and increasing the microstructural properties [71]. Because of its intrinsic resilience in alkaline media, the conventional superplasticizer is commonly used as an additive to the binder in OPC concrete to increase strength, but it degrades the hardened specimen strength in GPC blends. The use of a superplasticizer at high temperatures has a detrimental impact on the GPC mixed specimens. At higher temperatures, the SNF-based superplasticizer had no impact on strength. The high molar NaOH content in the GPC mix configuration creates the helpful SNF-based superplasticizer [5].

### 2.12. Effect of Handling Time

The fresh GPC is easily workable for up to 120 min without any strength deterioration [30]. The workability of the fresh GPC increases with increased hand mixing time up to 30 min [28]. If the mixing time is increased, it drastically retards the setting time of the fresh concrete.

### 2.13. Effect of Silicate and Alumina

In the SEM analysis, if the Si/Al ratio  $\leq 1.40$  is present in the matrix, it shows a clustered dense microstructure with large interconnected pores, and if Si/Al  $\geq 1.65$ , then a homogeneous microstructure with tiny pores is present. The matrix gel's microstructure increases with the increase of the silicon content, available when the ratio is  $1.40 \leq \text{Si/Al} \leq 1.65$ . The geopolymer microstructure is affected by the absorption of nitrogen and results in the volume expansion of the matrix. The larger gel volume is responsible for higher compressive load and increases the young modulus when the microstructure of the gel is homogenous at the ratio of Si/Al is 1.65. Therefore, the young modulus depends on both compressive strength and homogeneity in the microstructures of the gel. The mixed specimens' ultimate strength is reduced beyond the Si/Al = 1.90 due to the unreacted silica present in the matrix [90]. The thermal shrinkage increases as the mixed content's Si/Al ratio increases due to dehydration, dehydroxylation, sintering, and resilience [59]. The silica and alumina content play a vital role in the reaction of geopolymerisation. Silica content presents the amorphous end products in the reaction, contributing to the higher compressive strength of the mix design through the denser matrix development. The mechanical properties also increase with silica content and achieve a maximum strength of 65 MPa [91].

In the GPC, the  $\text{SiO}_2/\text{Al}_2\text{O}_3$  and  $\text{SiO}_2/\text{Fe}_2\text{O}_3$  ratios increase with the curing temperature, increasing the mechanical properties of the GPC. They also reduce the water absorption capacity compared to the OPC concrete. The CaO content present in the mix does not affect the geopolymer matrix's reaction [92]. The molar ratio of  $\text{SiO}_2/\text{Al}_2\text{O}_3$  increases to 3.4–3.8, which is highly responsible for the high strength gain at a later age [93]. In the fly ash-based GPC, the Si:Al ratio  $\geq 5$  present in the fly ash activated with sodium silicate shows low to moderate strength at ambient curing, but after the heat, the curing shows excellent dimensional stability and high compressive strength.

When the Si:Al ratio is less than 2, it shows high compressive strength but poor dimensional stability and reduces the strength after heating [94]. In the geopolymerisation reaction, increases in alumina and silica content accelerate the range of 3.20–3.70. The mix alumina content increases neither show any zeolitic phase development nor show the mix samples strength development [95]. The alumina content present in the mix is highly responsible for the setting time of the mix. Increasing the Si/Al ratio leads to a longer setting time. In addition, increases in Al content decrease the strength of the concrete [94]. Figure 3d shows the effect of the  $\text{SiO}_2/\text{Na}_2\text{O}$  ratio on the autogenous shrinkage of the geopolymer paste.

## 3. Durability and Other Related Aspect Studies

Geopolymers show a great potential in construction industry for durable structures subjected to extreme environment and for sustainable infrastructure development [96,97]. In the durability studies, the long-term strength of the GPC deteriorates with time in aggressive environmental conditions [98]. Acid attack, seawater conditions, sulphate attack, carbonation of concrete, chloride penetration, alkali-aggregate reactions, and free-thaw conditions were included in the durability studies.

### 3.1. Effect of Sulphate Attack

The magnesium sulphate deteriorates the GPC at a very high level in the calcium-rich geopolymer formed in the end products. It breaks the CSH bond and forms the Mg-SH by replacing the calcium present in the structure. The magnesium-formed structure expanded the volume that created the crack formation in the GPC. At the same time, the sodium

sulphate does not deteriorate the GPC at a very high level. The magnesium sulphate reduces the mechanical properties of the GPC-mixed specimens [99]. The high-calcium BA geopolymer mortar shows excellent resistant properties against sodium sulphate [83]. The Fly ash/GGBFS-based GPC shows a 33% deterioration in mechanical strength and a 0.04% expansion after immersion in magnesium sulphate for 360 days, but the OPC concrete deteriorates to 48% mechanical strength and 0.8 expansion in concrete in the same conditions. In the  $\text{Na}_2\text{SO}_4$ , exposure to OPC concrete shows a deterioration of strength and expansion of 30% and 0.412%, respectively, but in the GPC, the strength is increased in the same condition [100]. The clay-fly ash-based GPC is less affected by the sulphate attack on the GPC than the OPC concrete because the clay/fly ash-based GPC contains significantly less calcium in the mix [101]. BFA-based GPC is significantly less susceptible to sulphate exposure after 18 months. The OPC concrete reduces up to 20%, but BFA-based GPC deteriorates up to 4% in strength in the same exposure condition as sodium sulphate [102].

### 3.2. Effect of Acid Attack

The acid attack on the concrete decreases the concrete's performance and strength by reducing the specimens' mass loss in the acidic conditions below the 6.5 pH of the concrete. The sulphuric acid immersion in the exposure condition for 28 days shows the weakening of the concrete, and mass loss increased with the GPC matrix weakening. The loss reduction of concrete increases with the increase in acid content. The GPC shows better stability against acidic conditions than OPC concrete due to less calcium content present in the GPC [103]. In the slag-based geopolymer mortar, the addition of nano-silica increases the micro-structure of cement paste (formation of additional calcium silicate hydrate (CSH) gels in the system) and strength properties [104]. If the pozzolanic content of geopolymer mortar is more than 50%, it shows better durable properties than conventional OPC concrete. It is less affected by acid attacks and chloride penetration in concrete. Therefore, alkali-activated slag/fly ash cement are highly useful in acidic or seawater conditions [105].

### 3.3. Effect of Sea Water

The GPC concrete shows better properties against seawater conditions by reducing the concrete's sulphate and chloride penetration [106]. The fly ash-based GPC shows high compressive strength, tensile and flexural strength, low elasticity, water absorption, drying shrinkage, and sorptivity in seawater conditions. The fly ash-based GPC achieved a strength of 55 MPa after 28 days, outperforming the OPC concrete, and was less susceptible to seawater in the same conditions [107].

### 3.4. Effect on Carbonation

The carbonation reaction rate of the GPC depends on the contents of the mixed design present in the concrete. The fly ash/GGBFS based GPC shows weak resistance to the carbonation reaction due to the pozzolanic binder's activation by the mix's sodium silicate. The carbonation reaction increases the concrete's permeability, which is very hazardous for the concrete's durability [108].

### 3.5. Effect of Alkali-Silica Reaction and Leaching

The RCA dosage increases in conventional concrete reduce the concrete's strength and mechanical properties and lead to leaching in the concrete, but in the GPC, the RCA dosage does not affect the strength at a minimal level and reduces the leaching in the concrete [109]. The ASR susceptibility of fly ash-based GPC is lower than that of OPC concrete [110,111]. The non-wood biomass ash-based GPC shows excellent properties against acidic conditions compared to the OPC concrete because the OPC concrete shows 9% mass loss in 28 days under sulphuric acid conditions, while the biomass ash-based GPC shows less than 2% mass loss in the same conditions [112].

### 3.6. Effect of Elevated Temperature

The aggregate size in the mix design of the concrete plays a vital role under high temperatures. If the maximum aggregate size is less than 10 mm in the mix designs, it shows the explosive spalling of the concrete specimens under high temperatures in both types of concrete: GPC and OPC concrete. The spalling of concrete is prevented using a maximum aggregate size of more than equal to 14 mm in the concrete's design mix. The concrete's spalling is explained by the fracture process zone's size ( $l_p$ ), which varies with the aggregate size. The aggregate size is larger than the  $l_p$ , also long and healthy because of the crack-tip shielding. The GPC is chemically stable under elevated temperatures, whereas the OPC concrete chemically decomposes and dehydrates under the same conditions and decreases in the evaporation water content, decreasing the spalling probability of the concrete [47]. The geopolymer mortars exhibited better performance at elevated temperatures in comparison to control cement mortar mixture [113]. The Si/Al ratio plays a vital role under the elevated temperature; the strength increases with the Si/Al ratio in the exposure of 800 °C of the mixed samples. The heat-cured specimens above 80 °C show higher stability against the elevated temperature. However, the ambient-cured specimens show lower stability in the same conditions, and potassium-based geopolymer shows higher stability than sodium-based geopolymer in high-temperature conditions [66,67]. The GPC has better stability against elevated temperatures than the OPC concrete and is more porous than the OPC concrete analysed by the sorptivity test. It reduces the risk of the spalling of concrete under high-temperature conditions [114].

The FSGC shows similar trends to those of Portland cement concrete in weight loss under elevated temperatures of 600 °C [115]. The GPC shows a higher degree of transient creep and the OPC pastes below the 250 °C temperature. When the temperature ranges between 250 °C and 550 °C, the geopolymer does not show the transient creep increase while the OPC paste shows the higher transient creep, and the geopolymer increases the elastic modulus. OPC concrete shows a minor change in the elastic modulus [116]. The powder slag of ferrosilicon (FSS), an industrial waste, may be blended for manufacturing light weight GPC to conserve its strength properties under elevated temperature [117]. The electric arc furnace steel slag (EAFSS), barite and ilmenite heavy aggregates-based heavyweight geopolymer concrete is very effective in radiation shielding at high temperatures [118]. The dosage of the MPCM shows the stimulating effects on the thermal performance of the PCC and GPC. The number of microcapsules affects the thermal conductivity and latent heat of concrete. Microcapsule dosage increases concrete porosity and has more substantial effects on GPC than PCC [119]. The geopolymer matrix's thermal conductivity is higher than the OPC pastes and shows that the geopolymer paste's specific heat is less than the PC pastes [120]. The pore size distribution plays a vital role in the FA/M-based GPC under elevated temperatures [121].

### 3.7. Effect on the Bond Strength

The GPC specimens explain a similar cracking pattern to the OPC concrete in the pull-out load test, and both fail in a brittle manner due to the splitting of concrete along with the bonding of concrete and bars. The bond strength increases with the concrete strength and concrete cover in both types of concrete. GPC shows higher bond strength than OPC concrete due to the higher splitting strength of GPC for the same compressive strength [122,123]. The GPC beams with lap-spliced reinforcement show similar failure behaviour to the OPC concrete beams. For both types of concrete, the reinforcement is based on Australian Standards and the ACI code. The bond strengths of the beam-ends specimens have lower strength compared to the direct pull-out tests. The bar size plays a vital role in the bond strength; the bond strength increases with the reduction of the bar size [124]. The bond strength of the sand-coated GFRP-reinforced GPC compared to OPC concrete shows higher failure loads than the OPC concrete [125].

The ultimate strength and crack load increases with the increase of the fibre concentration in the mix designs, and it also reduces the cracking rate in the beam [126]. The



sand-coated GFPR bars are a perfect alternative to the internal reinforcement of GPC structures [127]. The elastic behaviour of the GPC under reinforced beams is similar to the under-reinforced OPC concrete beams. GPC specimens show a more brittle flexural strength than OPC concrete specimens [122,128]. The compressive strength of the GPC increases by 5% with the 1% dosage of the 1% steel fibre in the mix [129]. The load capacity of fly ash/GLSS-based geopolymer concrete column increases with the increment of concrete strength, reduction in the load eccentricity, and increase in the design's longitudinal-reinforcement ratio [130]. The column-designed specimens show a similar failure to the design code AS3600 [131].

#### 4. Geopolymer Material Applications

The GPC materials have found number of applications in infrastructure development and different fields [132]. The potential applications of geopolymers in sustainable construction are given in [133,134]. Studies have been conducted on the barriers and challenges in the effective application of GPC and reinforced geopolymer composites [135,136]. The geopolymers have been successfully used for soil stabilization in transportation and geotechnical engineering [137,138]. The geopolymers are most suitable for marine engineering construction due to their corrosive resistance and excellent aggressive environment performances [139]. The geopolymers can be used to construct light weight [140] and blast-resistant structures [141]. The geopolymers are used in buildings for floor heating [142] and for energy saving applications, i.e., to reduce HVAC demand of buildings [143]. GEO-based materials can be a sound choice in construction industries in place of OPC for sustainable development due to their thermal energy storage capacity [144], fire retardant capabilities [145], and electrical and self-sensing characteristics [146]. The geopolymer is cost-effective due to its stable performance against elevated temperatures and is used to replace epoxy resins in structural retrofitting with FRP. It is also used as a cost-efficient lining of the trenches to rehabilitate sewage pipelines [147,148]. The geopolymers have proven to be sound materials for structural retrofitting and rehabilitation of heritage buildings [149,150], for building preservation [151,152], and for damaged structural elements [153].

Geopolymers as aggregates in concrete show better strength and stability compared to natural aggregates [154]. The geopolymers meet the desired requirements as a material for 3D printing construction [155,156]. The 3D-printed GPC can be potentially used in a wide range of structural applications in construction industry due to its sustainable processing [157,158]. Table 1 describes the geopolymer applications as per the Si/Al content ratio in the geopolymer mix [159].

**Table 1.** Geopolymer materials applications [159].

Si/Al Ratio	Applications
1	<ul style="list-style-type: none"> <li>➤ Bricks</li> <li>➤ Ceramics</li> </ul>
2	<ul style="list-style-type: none"> <li>➤ Low CO<sub>2</sub> cements and concretes</li> <li>➤ Radioactive and toxic waste encapsulation</li> </ul>
3	<ul style="list-style-type: none"> <li>➤ Foundry equipment</li> <li>➤ Thermal insulation materials, 200–1000 °C</li> <li>➤ Tooling for titanium processing</li> </ul>
>3	<ul style="list-style-type: none"> <li>➤ Sealant for industry, 200–600 °C</li> <li>➤ Tooling for aeronautics SPF aluminium</li> </ul>
20–35	<ul style="list-style-type: none"> <li>➤ Fire-resistant and heat resistant fibre composites</li> </ul>

## 5. Sustainability

A review of the sustainability of geopolymer concrete is found in [160,161]. In terms of energy used in the production of concrete constituents, sustainability was increased by using alternative materials and reducing energy consumption [162]. In the production of GPC, industrial solid waste is used as a binder, activated by alkaline chemicals for concrete production [154]. Cement, aggregate (fine and coarse), and water are used to produce the PCC. The GPC production uses fly ash, GGBFS, NaOH, sodium silicate, aggregate (fine and coarse), and water. In India, cement production's mostly dry process generates the energy of 4.2 MJ/kg [163]. The fly ash and GGBFS are solid industrial wastes produced by the thermal power plant and steel plant. The fly ash is directly used in the mix, but the GGBFS requires grinding before use in the mix. NaOH production's embodied energy is 20.5 MJ/kg, and sodium silicate production is 5.37 MJ/kg [164]. There is zero embodied energy in fly ash and water. The coarse aggregate and fine aggregates' embodied energy are 0.22 MJ/kg and 0.02 MJ/kg, respectively. The superplasticizer's embodied energy in SNF-based is 12.6 MJ/kg. The total embodied energy of the OPC concrete is 1897.86 MJ/m<sup>3</sup>, whereas the embodied energy of the GPC is 1749.21 MJ/m<sup>3</sup>. The GPC's embodied energy is less compared to the OPC concrete. Table 2 describes the total constituents present in both the mix design and their embodied energy.

**Table 2.** GPC and OPC concrete constituents details.

	Embodied Energy (MJ/kg)	OPC Concrete		Geopolymer Concrete	
		Mix Content (kg/m <sup>3</sup> )	Embodied Energy Content (MJ/kg)	Mix Content (kg/m <sup>3</sup> )	Embodied Energy Content (MJ/kg)
OPC	4.2	370	1554	0.0	0.0
Fly ash	0.0	0.0	0.0	303.75	0.0
GGBFS	0.31	0.0	0.0	101.25	31.38
NaOH	20.5	0.0	0.0	40.5	830.25
Na <sub>2</sub> SiO <sub>3</sub>	5.37	0.0	0.0	101.25	543.71
Fine Aggregate	0.02	683	13.66	683	13.66
Coarse Aggregate	0.22	1289	283.58	1269	279.18
Water	0.0	148	0.0	40.5	0.0
Superplasticizer	12.6	3.7	46.62	4.05	51.03
Total		2493.7	1897.86	2543.7	1749.21

The geopolymer materials promote a circular economy process and sustainable development [165,166]. The use of sustainable material, such as geopolymers, their application and economical production techniques in the construction industry create employment and increase energy efficiency [167]. The economic and commercial manufacturing aspects (social), an important component of sustainability of the GPC, has not received much attention. The cost studies of GPC were undertaken by Mathew et al. [168]. The GPC is more cost-effective than OPC. Weil et al. [169] have studied the cost comparison and cost drivers for OPC concrete and GPC. The cement is the main cost driver in cement concrete while the activators are the cost driver in GPC. The cost-effectiveness of GPC manufacturing is demonstrated by Youssef et al. [170]. Van Deventer et al. [171] have discussed the factors affecting the acceptance of GPC technology in industry. They pointed out that the regulations and supply chain issues barrier must be removed to achieve widespread acceptance of GPC. Shamsaei et al. [172] have reviewed the studies devoted to the obstacles related to commercialization of GPC. They pointed out that economic analysis and social/national standards attitude towards the application of GPC requires further research. The 3D-printed geopolymer concrete is cheaper as compared to 3D-printed cement concrete with the same properties [173]. It can be concluded that sustainable development is achieved by employing geopolymers in Indian construction industries because it results in lower CO<sub>2</sub> emissions, optimum utilization of natural resources, utilization of waste



materials, economic and long life infrastructure construction, and societal income and employment generation.

## 6. Conclusions

The comprehensive survey of the journals and papers and energy analysis of the OPC concrete and GPC concrete concluded as follows:

- Geopolymers are a perfect alternative to OPC concrete in concrete performance in terms of strength, durable properties, and sustainability.
- GPC reduces carbon footprints by using industrial solid waste like fly ash and slag and reducing cultivated land for dumping.
- GPC has reduced the cost of concrete by the use of industrial waste in the concrete production.
- The GPC shows better mechanical properties compared to OPC concrete.
- The GPC shows better durable properties, for all extreme environmental conditions i.e., acid attack, seawater conditions, sulphate attack, carbonation of concrete, chloride penetration, alkali-aggregate reactions and elevated temperature, than OPC concrete.
- The embodied energy of the GPC is less compared to the OPC concrete for the same compressive strength.
- The GPC materials have found a number of applications in infrastructure development and other various fields, and have become a proven material for the sustainable development in the construction industry.
- The application of GPC and economical production techniques in the construction industry create employment and increase energy efficiency.

## 7. Recommendations and Future Scope

This comprehensive study of the geopolymerisation process shows the development of strength by bonding the raw materials in the reaction process. The comprehensive literature survey shows that the GPC has better strength and durability properties than OPC concrete. It also utilises industrial solid waste and reduces land dumping costs. The following aspects of GPC require further research:

- The requirement to implement GPC on major structural projects such as roads, bridges, buildings, hospitals, and other structures.
- More research on the other non-ferrous solid waste used in the GPC and analysis of such material's capacity to work as a binder in the GPC.
- Further research on the hazards associated with some hazardous material components of GPC for its use on a large scale in the construction sector, and research on ways to reduce the hazard vulnerability in production/application of GPC.
- Further research work on long term effect on strength and durability characteristics.
- Further study on social/national standards attitude towards the application of GPC.
- Further research work to make geopolymer 3D printing process a viable construction approach.
- Further research work to draw relationships between composition, structure, and strength characteristics.

**Author Contributions:** Conceptualization, M.V. and N.D.; methodology, M.N.; validation, I.R. and M.A.; formal analysis, M.V. and M.N.; investigation, M.V.; resources, I.R. and J.M.; data curation, M.V.; writing—original draft preparation, M.V.; writing—review and editing, I.R. and M.A.; supervision, N.D.; project administration, M.A.; funding acquisition, J.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through Research Group (small) Project under grant number [R.G.P1/90/1442].

**Institutional Review Board Statement:** This study did not involve humans or animals.

**Informed Consent Statement:** As a corresponding author or on behalf of all research paper authors, I consent to participate. All author of the research paper is consent to the publication.

**Data Availability Statement:** The available data had been used.

**Acknowledgments:** The authors acknowledge the Dean of the Faculty of Engineering for his valuable support and help.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

- Shobeiri, V.; Bennet, B.; Xie, T.; Visintin, P. A comprehensive assessment of the global warming potential of geopolymers concrete. *J. Clean. Prod.* **2021**, *297*, 126669. [\[CrossRef\]](#)
- Imtiaz, L.; Rehman, S.K.U.; Memon, A.S.; Khizar, K.M.; Faisal, J.M. A Review of Recent Developments and Advances in Eco-Friendly Geopolymer Concrete. *Appl. Sci.* **2020**, *10*, 7838. [\[CrossRef\]](#)
- Qaidi, S.M.A.; Tayeh, B.A.; Zeyad, A.M.; de Azevedo, A.R.G.; Ahmed, H.U.; Emad, W. Recycling of mine tailings for the geopolymers production: A systematic review. *Case Stud. Constr. Mater.* **2022**, *16*, e00933. [\[CrossRef\]](#)
- Qaidi, S.M.A.; Tayeh, B.A.; Isleem, H.F.; de Azevedo, A.R.G.; Ahmed, H.U.; Emad, W. Sustainable utilization of red mud waste (bauxite reidue) and slag for the production of geopolymer composites: A review. *Case Stud. Constr. Mater.* **2022**, *16*, e00994. [\[CrossRef\]](#)
- Albitar, M.; Visintin, P.; Ali, M.S.M.; Drechsler, M. Assessing Behaviour of Fresh and Hardened Geopolymer Concrete Mixed with Class-F Fly Ash. *KSCE J. Civ. Eng.* **2015**, *19*, 1445–1455. [\[CrossRef\]](#)
- Bhavan, I. *Indian Minerals Yearbook 2019*; (Advance Release), Slag-Iron and Steel; Government of India, Ministry of Mines, Indian Bureau of Mines, ENVIS-Centre on Environmental Problems of Mining: Dhanbad, India, 2019.
- Davidovits, J. Geopolymers: Inorganic polymeric new materials. *J. Mater. Educ.* **1994**, *16*, 91–138. [\[CrossRef\]](#)
- Wongkvanklom, A.; Posi, P.; Kampala, A.; Kaewngao, T.; Chindaprasirt, P. Beneficial utilization of recycled asphaltic concrete aggregate in high calcium fly ash geopolymer concrete. *Case Stud. Constr. Mater.* **2021**, *15*, e00615. [\[CrossRef\]](#)
- Lee, N.K.; Jang, J.G.; Lee, H.K. Cement and Concrete Composites Shrinkage characteristics of alkali-activated fly ash/slag paste and mortar at early ages. *Cem. Concr. Compos.* **2014**, *53*, 239–248. [\[CrossRef\]](#)
- Chi, M. Effects of the alkaline solution/binder ratio and curing condition on the mechanical properties of alkali-activated fly ash mortars. *Sci. Eng. Compos. Mater.* **2017**, *24*, 773–782. [\[CrossRef\]](#)
- Huseien, G.F.; Ismail, M.; Tahir, M.; Mirza, J.; Hussein, A.; Khalid, N.H.; Sarbini, N.N. Effect of binder to fine aggregate content on performance of sustainable alkali activated mortars incorporating solid waste materials. *Chem. Eng. Trans.* **2018**, *63*, 667–672. [\[CrossRef\]](#)
- Luukkonen, T.; Abdollahnejad, Z.; Yliniemi, J.; Kinnunen, P.; Illikainen, M. One-part alkali-activated materials: A review. *Cem. Concr. Res.* **2017**, *103*, 21–34. [\[CrossRef\]](#)
- Xing, J.; Zhao, Y.; Qiu, J.; Sun, X. Microstructural and Mechanical Properties of Alkali Activated Materials from Two Types of Blast Furnace Slags. *Materials* **2019**, *12*, 2089. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhang, P.; Gao, Z.; Wang, J.; Guo, J.; Hu, S.; Ling, Y. Properties of fresh and hardened fly ash/slag based geopolymer concrete: A review. *J. Clean. Prod.* **2020**, *270*, 122389. [\[CrossRef\]](#)
- Xie, T.; Ozbakkaloglu, T. Behavior of low-calcium fly and bottom ash-based geopolymer concrete cured at ambient temperature. *Ceram. Int.* **2015**, *41*, 5945–5958. [\[CrossRef\]](#)
- Deb, P.S.; Sarker, P.K. Effects of Ultrafine Fly Ash on Setting, Strength, and Porosity of Geopolymers Cured at Room Temperature. *J. Civ. Eng.* **2016**, *29*, 1–5. [\[CrossRef\]](#)
- Kawade, U.R.; Hussain, M.; Shirule, P.P.A. Enhancement of compressive strength of geopolymer concrete by varying ratio of Na<sub>2</sub>SiO<sub>3</sub>/NaOH and by varying molarity of NaOH. *Int. J. Latest Res. Eng. Technol.* **2016**, *2*, 47–50.
- Luhar, S.; Khandelwal, U. A Study on Water Absorption and Sorptivity of Geopolymer Concrete. *Int. J. Civ. Eng.* **2015**, *2*, 1–10. [\[CrossRef\]](#)
- Patankar, S.V.; Ghugal, Y.M.; Jamkar, S.S. Mix design of fly ash based geopolymer concrete. *Adv. Struct. Eng. Mater.* **2015**, *3*, 1619–1633. [\[CrossRef\]](#)
- Verma, M.; Dev, N. Effect of ground granulated blast furnace slag and fly ash ratio and the curing conditions on the mechanical properties of geopolymer concrete. *Struct. Concr.* **2021**, *21*, 1–15. [\[CrossRef\]](#)
- Davidovits, J.; Quentin, S. Geopolymers Inorganic polymeric new materials. *J. Therm. Anal.* **1991**, *37*, 1633–1656. [\[CrossRef\]](#)
- Deb, P.S.; Nath, P.; Sarker, P.K. The effects of ground granulated blast-furnace slag blending with fly ash and activator content on the workability and strength properties of geopolymer concrete cured at ambient temperature. *J. Mater.* **2014**, *62*, 32–39. [\[CrossRef\]](#)
- Krishnan, L.; Karthikeyan, S.; Nathiya, S.; Suganya, K. Geopolymer Concrete an Eco-Friendly Construction Material. *Int. J. Res. Eng. Technol.* **2014**, *3*, 164–167. [\[CrossRef\]](#)

24. Gao, X.; Yu, Q.L.; Brouwers, H.J.H.H. Reaction kinetics, gel character and strength of ambient temperature cured alkali activated slag–fly ash blends. *Constr. Build. Mater.* **2015**, *80*, 105–115. [CrossRef]
25. Verma, M.; Dev, N. Review on the effect of different parameters on behavior of Geopolymer Concrete. *Int. J. Innov. Res. Sci. Eng. Technol.* **2017**, *6*, 11276–11281. [CrossRef]
26. Nagalia, G.; Park, Y.; Abolmaali, A.; Aswath, P. Compressive Strength and Microstructural Properties of Fly Ash-Based Geopolymer Concrete. *J. Mater. Civ. Eng.* **2016**, *18*, 040161441–11. [CrossRef]
27. Patankar, S.V.; Ghugal, Y.M.; Jamkar, S.S. Effect of Concentration of Sodium Hydroxide and Degree of Heat Curing on Fly Ash-Based Geopolymer Mortar. *Indian J. Mater. Sci.* **2014**, *2*, 1–6. [CrossRef]
28. Bidwe, S.S.; Hamane, A.A. Effect of different molarities of Sodium Hydroxide solution on the Strength of Geopolymer concrete. *Am. J. Eng. Res.* **2015**, *4*, 139–145. Available online: [https://www.ajer.org/papers/v4\(03\)/S04301390145.pdf](https://www.ajer.org/papers/v4(03)/S04301390145.pdf) (accessed on 9 March 2022).
29. Kabir, S.M.A.A.; Alengaram, U.J.; Jumaat, M.Z.; Sharmin, A.; Islam, A. Influence of Molarity and Chemical Composition on the Development of Compressive Strength in POFA Based Geopolymer Mortar. *Adv. Mater. Sci. Eng.* **2015**, *2*, 1–15. [CrossRef]
30. Aldawsari, S.; Kampmann, R.; Harnisch, J.; Rohde, C. Setting Time, Microstructure, and Durability Properties of Low Calcium Fly Ash/Slag Geopolymer: A Review. *Materials* **2022**, *15*, 876. [CrossRef]
31. Topark-Ngarm, P.; Chindaprasirt, P.; Sata, V. Setting Time, Strength, and Bond of High-Calcium Fly Ash Geopolymer Concrete. *J. Mater. Civ. Eng.* **2015**, *27*, 04014198. [CrossRef]
32. Verma, M.; Dev, N. Sodium hydroxide effect on the mechanical properties of flyash-slag based geopolymer concrete. *Struct. Concr.* **2020**, *22*, E368–E379. [CrossRef]
33. Ryu, G.S.; Lee, Y.B.; Koh, K.T.; Chung, Y.S. The mechanical properties of fly ash-based geopolymer concrete with alkaline activators. *Constr. Build. Mater.* **2013**, *47*, 409–418. [CrossRef]
34. Lee, N.K.; Lee, H.K. Setting and mechanical properties of alkali-activated fly ash/slag concrete manufactured at room temperature. *Constr. Build. Mater.* **2013**, *47*, 1201–1209. [CrossRef]
35. Somna, K.; Jaturapitakkul, C.; Kajitvichyanukul, P.; Chindaprasirt, P. NaOH-activated ground fly ash geopolymer cured at ambient temperature. *Fuel* **2011**, *90*, 2118–2124. [CrossRef]
36. Bella, N.; Gudiel, E.; Soriano, L.; Font, A.; Borrachero, M.V.; Paya, J.; Monzó, J.M. Formulation of Alkali-Activated Slag Binder Destined for Use in Developing Countries. *Appl. Sci.* **2020**, *10*, 9088. [CrossRef]
37. Ban, C.C.; Ken, P.W.; Ramli, M. Effect of Sodium Silicate and Curing Regime on Properties of Load Bearing Geopolymer Mortar Block. *J. Mater. Civ. Eng.* **2017**, *29*, 1–9. [CrossRef]
38. Amran, Y.H.; Alyousef, R.; Alabduljabbar, H.; El-Zeadani, M. Clean production and properties of geopolymer concrete; A review. *J. Clean. Prod.* **2020**, *251*, 119679. [CrossRef]
39. Nath, P.; Sarker, P.K. Effect of GGBFS on setting, workability and early strength properties of fly ash geopolymer concrete cured in ambient condition. *Constr. Build. Mater.* **2014**, *66*, 163–171. [CrossRef]
40. Lv, Q.; Yu, J.; Ji, F.; Gu, L.; Chen, Y.; Shan, X. Mechanical Property and Microstructure of Fly Ash-Based Geopolymer Activated by Sodium Silicate. *KSCE J. Civ. Eng.* **2021**, *25*, 1765. [CrossRef]
41. De Vargas, A.S.; Molin, D.C.C.D.; Vilela, A.C.F.; da Silva, F.J.; Pavão, B.; Veit, H. The effects of Na<sub>2</sub>O/SiO<sub>2</sub> molar ratio, curing temperature and age on compressive strength, morphology and microstructure of alkali-activated fly ash-based geopolymers. *Cem. Concr. Compos.* **2011**, *33*, 653–660. [CrossRef]
42. Tchadjie, L.N.; Ekelu, S.O. Enhancing the reactivity of aluminosilicate materials toward geopolymer synthesis. *J. Mater. Sci.* **2018**, *53*, 4709–4733. [CrossRef]
43. Elimbi, A.; Tchakoute, H.K.; Njopwouo, D. Effects of calcination temperature of kaolinite clays on the properties of geopolymer cements. *Constr. Build. Mater.* **2011**, *25*, 2805–2812. [CrossRef]
44. Freire, C.B.; Dos Santos, B.L.D.; De Miranda, I.L.F.; Rodrigues, M.A.; Lameiras, F.S. Influence of the Kaolinite Calcination Conditions on the Compressive Strength of Geopolymer. *KnE Eng.* **2020**, *5*, 1–10.
45. Boakye, K.; Khorami, M.; Ganjian, E.; Saidani, M. A review of the Effect of Calcination Temperature on the Properties of Calcined Clay Concrete. *Int. J. Eng. Technol. Inf.* **2021**, *2*, 72–74. [CrossRef]
46. Ghani, U.; Hussain, S.; Amin, N.; Imtiaz, M.; Khan, S.A. Role of calcination on geopolymerization of lateritic clay by alkali treatment. *J. Saudi Chem. Soc.* **2021**, *25*, 101198. [CrossRef]
47. Pan, Z.; Sanjayan, J.G.; Kong, D.L.Y. Effect of aggregate size on spalling of geopolymer and Portland cement concretes subjected to elevated temperatures. *Constr. Build. Mater.* **2012**, *36*, 365–372. [CrossRef]
48. Gluth, G.J.G.; Rickard, W.D.A.; Werner, S.; Pirsawetz, S. Acoustic emission and microstructural changes in fly ash geopolymer concretes exposed to simulated fire. *Mater. Struct.* **2016**, *49*, 5243–5254. [CrossRef]
49. Yliniemi, Paiva; Ferreira; Tiainen; Illikainen. Development and incorporation of lightweight waste-based geopolymer aggregates in mortar and concrete. *Constr. Build. Mater.* **2017**, *131*, 784–792. [CrossRef]
50. Sreenivasulu, C.; Ramakrishnaiah, A.; Guru Jawahar, J. Mechanical Properties of Geopolymer Concrete Using Granite Slurry as Sand Replacement. *Int. J. Adv. Eng. Technol.* **2015**, *8*, 83–91.
51. Mahaboob, S.; Reddy, C.H.B.; Vasugi, K. Strength behaviour of geopolymer concrete replacing fine aggregates by M- sand and E-waste. *Int. J. Eng. Trends Technol.* **2016**, *40*, 401–407. [CrossRef]

52. Saravanan, S.; Elavenil, S. Strength properties of geopolymer concrete using mssand by assessing their mechanical characteristics. *J. Eng. Appl. Sci.* **2018**, *13*, 4028–4041.
53. Shaikh, F.A. Mechanical and durability properties of fly ash geopolymer concrete containing recycled coarse aggregates. *Int. J. Sustain. Built Environ.* **2016**, *5*, 277–287. [\[CrossRef\]](#)
54. Taher, S.M.S.; Saadullah, S.T.; Haido, J.H.; Tayeh, B.A. Behavior of geopolymer concrete deep beams containing waste aggregate of glass and limestone as a partial replacement of natural sand. *Case Stud. Constr. Mater.* **2021**, *15*, e00744. [\[CrossRef\]](#)
55. Matsuda, A.; Maruyama, I.; Meawad, A.; Pareek, S.; Araki, Y. Reaction, Phases, and Microstructure of Fly Ash-Based Alkali-Activated Materials. *J. Adv. Concr. Technol.* **2019**, *17*, 93–101. [\[CrossRef\]](#)
56. Abdel-Ghani, N.T.; Elsayed, H.A.; AbdelMoied, S. Geopolymer synthesis by the alkali-activation of blastfurnace steel slag and its fire-resistance. *HBRC Journal* **2018**, *14*, 159–164. [\[CrossRef\]](#)
57. Hosan, A.; Haque, S.; Shaikh, F. Comparative Study of Sodium and Potassium Based Fly Ash Geopolymer at Elevated Temperatures. In Proceedings of the International Conference on Performance-based and Life-Cycle Structural Engineering, Brisbane, Australia, 9–11 December 2015. [\[CrossRef\]](#)
58. Khater, H.M. Studying the effect of thermal and acid exposure on alkali-activated slag geopolymer. *Adv. Cem. Res.* **2014**, *29*, 1–9. [\[CrossRef\]](#)
59. Lu, Y.; Cui, N.; Xian, Y.; Liu, J.; Xing, C.; Xie, N.; Wang, D. Microstructure Evolution Mechanism of Geopolymers with Exposure to High-Temperature Environment. *Crystals* **2021**, *11*, 1062. [\[CrossRef\]](#)
60. Ma, Y.; Ye, G. The shrinkage of alkali activated fly ash. *Cem. Concr. Res.* **2015**, *68*, 75–82. [\[CrossRef\]](#)
61. Ismail, I.; Bernal, S.A.; Provis, J.L.; San Nicolas, R.; Hamdan, S.; Van Deventer, J.S.J. Modification of phase evolution in alkali-activated blast furnace slag by the incorporation of fly ash. *Cem. Concr. Compos.* **2014**, *45*, 125–135. [\[CrossRef\]](#)
62. Jang, J.G.; Lee, N.K.; Lee, H.K. Fresh and hardened properties of alkali-activated fly ash/slag pastes with superplasticizers. *Constr. Build. Mater.* **2014**, *50*, 169–176. [\[CrossRef\]](#)
63. Nath, P.; Sarker, P.K. Flexural strength and elastic modulus of ambient-cured blended low-calcium fly ash geopolymer concrete. *Constr. Build. Mater.* **2017**, *130*, 22–31. [\[CrossRef\]](#)
64. Xu, A.; Shayan, A. Effect of activator and water to binder ratios on setting and strength of geopolymer concrete. In Proceedings of the ARRB 27th Conference, Melbourne, VIC, Australia, 16–18 November 2016.
65. Zulkifly, K.; Cheng-Yong, H.; Yun-Ming, L.; Bayuaji, R.; Abdullah, M.M.A.B.; Ahmad, S.B.; Stachowiak, T.; Szmidla, J.; Gondro, J.; Jez, B.; et al. Elevated-Temperature Performance, Combustibility and Fire Propagation Index of Fly Ash-Metakaolin Blend Geopolymers with Addition of Monoaluminium Phosphate (MAP) and Aluminum Dihydrogen Triphosphate (ATP). *Materials* **2021**, *14*, 1973. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Korniejenko, K.; Figiela, B.; Ziejewska, C.; Marczyk, J.; Bazan, P.; Hebda, M.; Choińska, M.; Lin, W.-T. Fracture Behavior of Long Fiber Reinforced Geopolymer Composites at Different Operating Temperatures. *Materials* **2022**, *15*, 482. [\[CrossRef\]](#) [\[PubMed\]](#)
67. Nurrudin, M.F.; Haruna, S.; Mohammed, B.S.; Sha'aban, I.G. Methods of curing geopolymer concrete: A review. *International. J. Adv. Appl. Sci.* **2018**, *5*, 31–36. [\[CrossRef\]](#)
68. Bondar, D.; Lynsdale, C.J.; Milestone, N.B.; Hassani, N. Sulfate Resistance of Alkali Activated Pozzolans. *Int. J. Concr. Struct. Mater.* **2015**, *9*, 145–158. [\[CrossRef\]](#)
69. San Nicolas, R.; Provis, J.L. The Interfacial Transition Zone in Alkali-Activated Slag Mortars. *Front. Mater.* **2015**, *2*, 70. [\[CrossRef\]](#)
70. Alanazi, H. Study of the Interfacial Transition Zone Characteristics of Geopolymer and Conventional Concretes. *Gels* **2022**, *8*, 105. [\[CrossRef\]](#)
71. Liu, Z.; Cai, C.S.; Peng, H.; Fan, F. Experimental Study of the Geopolymeric Recycled Aggregate Concrete. *J. Mater. Civ. Eng.* **2016**, *28*, 1–9. [\[CrossRef\]](#)
72. Demie, S.; Nuruddin, M.F.; Shafiq, N.; Fadhil, M.; Shafiq, N. Effects of micro-structure characteristics of interfacial transition zone on the compressive strength of self-compacting geopolymer concrete. *Constr. Build. Mater.* **2020**, *41*, 91–98. [\[CrossRef\]](#)
73. Abdulkareem, O.A.; Al Bakri, A.M.M.; Kamarudin, H.; Nizar, I.K.; Saif, A.A. Effects of elevated temperatures on the thermal behavior and mechanical performance of fly ash geopolymer paste, mortar and lightweight concrete. *Constr. Build. Mater.* **2014**, *50*, 377–387. [\[CrossRef\]](#)
74. Verma, M.; Dev, N. Effect of Liquid to Binder Ratio and Curing Temperature on the Engineering Properties of the Geopolymer Concrete. *Silicon.* **2022**, *14*, 1743–1757. [\[CrossRef\]](#)
75. Hadi, M.N.S.; Zhang, H.; Parkinson, S. Optimum mix design of geopolymer pastes and concretes cured in ambient condition based on compressive strength, setting time and workability. *J. Build. Eng.* **2019**, *23*, 301–313. [\[CrossRef\]](#)
76. Heah, C.Y.; Kamarudin, H.; Al Bakri, A.M.M.; Binhussain, M.; Luqman, M.; Nizar, I.K.; Ruzaidi, C.M.; Liew, Y.M. Effect of Curing Profile on Kaolin-based Geopolymers. *Phys. Procedia* **2011**, *22*, 305–311. [\[CrossRef\]](#)
77. Chen, T.A. Optimum curing temperature and duration of alkali-activated glass inorganic binders. *J. Chin. Inst. Eng.* **2020**, *43*, 592. [\[CrossRef\]](#)
78. Elyamany, H.E.; Abd Elmoaty, A.M.; Elshaboury, A.M. Setting time and 7-day strength of geopolymer mortar with various binders. *Constr. Build. Mater.* **2018**, *187*, 974–983. [\[CrossRef\]](#)
79. Islam, A.; Alengaram, U.J.; Jumaat, M.Z.; Bashir, I.I.; Kabir, S.M.A.A. Engineering properties and carbon footprint of ground granulated blast-furnace slag-palm oil fuel ash-based structural geopolymer concrete. *Constr. Build. Mater.* **2015**, *101*, 503–521. [\[CrossRef\]](#)



80. Bondar, D.; Lynsdale, C.J.; Milestone, N.B.; Hassani, N.; Ramezaniapour, A.A. Engineering Properties of Alkali Activated Natural Pozzolan Concrete. *ACI Mater. J.* **2011**, *108*, 64–72.
81. Saif, M.S.; El-Hariri, M.O.R.; Sarie-Eldin, A.I.; Tayeh, B.A.; Farag, M.F. Impact of Ca<sup>+</sup> content and curing condition on durability performance of Metakaolin-based Geopolymer Mortars. *Case Stud. Constr. Mater.* **2022**, *16*, e00922. [\[CrossRef\]](#)
82. Chotetanorm, C.; Chindaprasirt, P.; Sata, V.; Rukzon, S.; Sathonsaowaphak, A. High-Calcium Bottom Ash Geopolymer: Sorptivity, Pore Size, and Resistance to Sodium Sulfate Attack. *J. Mater. Civ. Eng.* **2013**, *25*, 105–111. [\[CrossRef\]](#)
83. Puligilla, S.; Mondal, P. Role of slag in microstructural development and hardening of fly ash-slag geopolymer. *Cem. Concr. Res.* **2013**, *43*, 70–80. [\[CrossRef\]](#)
84. Verma, M.; Nigam, M. Mechanical Behaviour of Self Compacting and Self Curing Concrete. *Int. J. Innov. Res. Sci. Eng. Technol.* **2017**, *6*, 14361–14366. [\[CrossRef\]](#)
85. Maanser, A.; Benouis, A.; Ferhoune, N. Effect of high temperature on strength and mass loss of admixed concretes. *Constr. Build. Mater.* **2018**, *166*, 916–921. [\[CrossRef\]](#)
86. Verma, M.; Dev, N. Effect of SNF-Based Superplasticizer on Physical, Mechanical and Thermal Properties of the Geopolymer Concrete. *Silicon* **2022**, *14*, 1–11. [\[CrossRef\]](#)
87. Nematollahi, B.; Sanjayan, J. Effect of different superplasticizers and activator combinations on workability and strength of fly ash-based geopolymer. *Mater. Des.* **2014**, *57*, 667–672. [\[CrossRef\]](#)
88. Kusbiantoro, A.; Ibrahim, M.S.; Muthusamy, K.; Alias, A. Development of sucrose and citric acid as the natural-based admixture for fly ash-based geopolymer. *Procedia Environ. Sci.* **2013**, *17*, 596–602. [\[CrossRef\]](#)
89. European Federation for Specialist Construction Chemicals and Concrete Systems. *European Guidelines for Self-Compacting Concrete: Specification, Production and Use*; Association House: Farnham, UK, 2005.
90. Kaur, M.; Singh, J.; Kaur, M. Synthesis of fly ash based geopolymer mortar considering different concentrations and combinations of alkaline activator solution. *Ceram. Int.* **2018**, *44*, 1534–1537. [\[CrossRef\]](#)
91. Hajimohammadi, A.; van Deventer, J.S.J. Characterisation of One-Part Geopolymer Binders Made from Flyash. *Waste Biomass Valorization* **2017**, *8*, 225–233. [\[CrossRef\]](#)
92. Priyanka, R.; Selvaraj, R.; Rajesh, B. Characterization Study on Clay Based Geopolymer Concrete. *Int. J. Res. Engg. Appl. Sci.* **2015**, *5*, 49–55.
93. De Silva, P.; Sagoe-Crenstil, K.; Sirivivatnanon, V. Kinetics of geopolymerization: Role of Al<sub>2</sub>O<sub>3</sub> and SiO<sub>2</sub>. *Cem. Concr. Res.* **2007**, *37*, 512–518. [\[CrossRef\]](#)
94. Rickard, W.D.A.; Temuujin, J.; Van Riessen, A. Thermal analysis of geopolymer pastes synthesised from five fly ashes of variable composition. *J. Non. Cryst. Solids.* **2012**, *358*, 1830–1839. [\[CrossRef\]](#)
95. Chindaprasirt, P.; De Silva, P. Effect of SiO<sub>2</sub> and Al<sub>2</sub>O<sub>3</sub> on the setting and hardening of high calcium fly ash-based geopolymer systems. *J. Mater. Sci.* **2012**, *47*, 4876–4883. [\[CrossRef\]](#)
96. Wasim, M.; Ngo, T.D.; Law, D. A state-of-the-art review on the durability of geopolymer concrete for sustainable structures and infrastructure. *Constr. Build. Mater.* **2021**, *291*, 123381. [\[CrossRef\]](#)
97. Wasim, M.; Ngo, T.D.; Law, D. Durability performance of reinforced waste-based geopolymer foam concrete under exposure to various corrosive environments. *Case Stud. Constr. Mater.* **2021**, *16*, e00703. [\[CrossRef\]](#)
98. Wong, L.S. Durability Performance of Geopolymer Concrete: A Review. *Polymers* **2022**, *14*, 868. [\[CrossRef\]](#) [\[PubMed\]](#)
99. Ismail, I.; Bernal, S.A.; Provis, J.L.; Hamdan, S.; Van Deventer, J.S.J. Microstructural changes in alkali activated fly ash/slag geopolymers with sulfate exposure. *Mater. Struct. Constr.* **2013**, *46*, 361–373. [\[CrossRef\]](#)
100. Saavedra, W.G.V.; Angulo, D.E.; De Gutiérrez, R.M. Fly Ash Slag Geopolymer Concrete: Resistance to Sodium and Magnesium Sulfate Attack. *J. Mater. Civ. Eng.* **2016**, *28*, 1–9. [\[CrossRef\]](#)
101. Sukmak, P.; De Silva, P.; Horpibulsuk, S.; Chindaprasirt, P. Sulfate Resistance of Clay-Portland Cement and Clay High-Calcium Fly Ash Geopolymer. *J. Mater. Civ. Eng.* **2015**, *27*, 04014158. [\[CrossRef\]](#)
102. Bhutta, M.A.R.; Hussin, W.M.; Azreen, M.; Tahir, M.M. Sulphate Resistance of Geopolymer Concrete Prepared from Blended Waste Fuel Ash. *J. Mater. Civ. Eng.* **2014**, *26*, 1–6. [\[CrossRef\]](#)
103. Manjeeth, K.V.; Rama, J.S.K. An Experimental Investigation on the behaviour of Portland Cement Concrete and Geopolymer Concrete in acidic environment. *Int. J. Civ. Eng.* **2015**, *2*, 40–44. [\[CrossRef\]](#)
104. Hosan, A.; Shaikh, F.U.A. Compressive strength development and durability properties of high volume slag and slag-fly ash blended concretes containing nano-CaCO<sub>3</sub>. *J. Mater. Res. Technol.* **2021**, *10*, 1310–1322. [\[CrossRef\]](#)
105. Wiyono, D.; Hardjito, D.; Antoni, P.; Hardjito, D. Improving the durability of pozzolan concrete using alkaline solution and geopolymer coating. *Procedia Eng.* **2015**, *125*, 747–753. [\[CrossRef\]](#)
106. Reddy, D.V.; Edouard, J.-B.; Sobhan, K. Durability of Fly Ash-Based Geopolymer Structural Concrete in the Marine Environment. *J. Mater. Civ. Eng.* **2013**, *6*, 781–787. [\[CrossRef\]](#)
107. Olivia, M.; Nikraz, H. Properties of fly ash geopolymer concrete designed by Taguchi method. *Mater. Des.* **2012**, *36*, 191–198. [\[CrossRef\]](#)
108. Pasupathy, K.; Berndt, M.; Castel, A.; Sanjayan, J.; Pathmanathan, R. Carbonation of a blended slag-fly ash geopolymer concrete in field conditions after 8 years. *Constr. Build. Mater.* **2016**, *125*, 661–669. [\[CrossRef\]](#)
109. Sanusi, O.; Tempest, B.; Ogunro, V.O.; Gergely, J. Leaching Characteristics of Geopolymer Cement Concrete Containing Recycled Concrete Aggregates. *J. Hazard. Toxic Radioact. Waste* **2016**, *20*, 1–8. [\[CrossRef\]](#)

110. Kupwade-Patil, K.; Allouche, E.N. Impact of Alkali Silica Reaction on Fly Ash-Based Geopolymer Concrete. *J. Mater. Civ. Eng.* **2013**, *25*, 131–139. [\[CrossRef\]](#)
111. Kupwade-Patil, K.; Allouche, E.N. Examination of Chloride-Induced Corrosion in Reinforced Geopolymer Concretes. *J. Mater. Civ. Eng.* **2013**, *25*, 1465–1476. [\[CrossRef\]](#)
112. Matalkah, F.; Soroushian, P.; Balchandra, A.; Peyvandi, A. Characterization of Alkali-Activated Non-wood Biomass Ash-Based Geopolymer Concrete. *J. Mater. Civ. Eng.* **2017**, *29*, 1–9. [\[CrossRef\]](#)
113. Kaya, M.; Uysal, M.; Yilmaz, K.; Atis, C.D. Behaviour of Geopolymer Mortars after Exposure to Elevated Temperatures. *Mater. Sci.* **2018**, *24*, 428–436. [\[CrossRef\]](#)
114. Zhao, R.; Sanjayan, J.G. Geopolymer and Portland cement concretes in simulated fire. *Mag. Concr. Res.* **2011**, *63*, 163–173. [\[CrossRef\]](#)
115. Ren, W.; Xu, J.; Bai, E. Strength and Ultrasonic Characteristics of Alkali-Activated Fly Ash-Slag Geopolymer Concrete after Exposure to Elevated Temperatures. *J. Mater. Civ. Eng.* **2016**, *28*, 1–8. [\[CrossRef\]](#)
116. Pan, Z.; Sanjayan, J.G.; Collins, F. Cement and Concrete Research Effect of transient creep on compressive strength of geopolymer concrete for elevated temperature exposure. *Cem. Concr. Res.* **2014**, *56*, 182–189. [\[CrossRef\]](#)
117. Tayeh, B.A.; Hakamy, A.; Sherif, M.A.; Zeyad, A.M.; Agwa, I.S. Effect of air agent on mechanical properties and microstructure of lightweight geopolymer concrete under high temperature. *Case Stud. Constr. Mater.* **2022**, *16*, e00951. [\[CrossRef\]](#)
118. Sherif, M.A.; Zeyad, A.M.; Tayeh, B.A.; Agwa, I.S. Effect of high temperatures on mechanical, radiation attenuation and microstructure properties of heavyweight geopolymer concrete. *Struct. Eng. Mech.* **2021**, *80*, 181–199. [\[CrossRef\]](#)
119. Cao, V.D.; Pilehvar, S.; Salas-Bringas, C.; Szczotok, A.M.; Rodriguez, J.F.; Carmona, M.; Al-Manasir, N.; Kjøniksen, A.L. Microencapsulated phase change materials for enhancing the thermal performance of Portland cement concrete and geopolymer concrete for passive building applications. *Energy Convers. Manag.* **2017**, *133*, 56–66. [\[CrossRef\]](#)
120. Snell, C.; Tempest, B.; Gentry, T. Comparison of the Thermal Characteristics of Portland Cement and Geopolymer Cement Concrete Mixes. *J. Archit. Eng.* **2017**, *23*, 1–10. [\[CrossRef\]](#)
121. Rivera, O.G.; Long, W.R.; Weiss, C.A., Jr.; Moser, R.D.; Williams, B.A.; Torres-Cancel, K.; Gore, E.R.; Allison, P.G. Effect of elevated temperature on alkali-activated geopolymeric binders compared to portland cement-based binders. *Cem. Concr. Res.* **2016**, *90*, 43–51. [\[CrossRef\]](#)
122. Sarker, P.K. Bond strength of reinforcing steel embedded in fly ash-based geopolymer concrete. *Mater. Struct.* **2011**, *44*, 1021–1030. [\[CrossRef\]](#)
123. Luan, C.; Wang, Q.; Yang, F.; Zhang, K.; Utashev, N.; Dai, J.; Shi, X. Practical Prediction Models of Tensile Strength and Reinforcement-Concrete Bond Strength of Low-Calcium Fly Ash Geopolymer Concrete. *Polymers* **2021**, *13*, 875. [\[CrossRef\]](#)
124. Al-Azzawi, M.; Yu, T.; Hadi, M.N.S. Factors Affecting the Bond Strength Between the Fly Ash-based Geopolymer Concrete and Steel Reinforcement. *Structures* **2018**, *14*, 262–272. [\[CrossRef\]](#)
125. Tekle, B.H.; Khennane, A.; Kayali, O. Bond Properties of Sand-Coated GFRP Bars with Fly Ash-Based Geopolymer Concrete. *J. Compos. Constr.* **2016**, *20*, 1–13. [\[CrossRef\]](#)
126. Sing Ng, T.; Amin, A.; Foster, S.J. The behaviour of steel-fibre-reinforced geopolymer concrete beams in shear. *Mag. Concr. Res.* **2013**, *65*, 308–318. [\[CrossRef\]](#)
127. Maranan, G.; Manalo, A.; Karunasena, K.; Benmokrane, B. Bond Stress-Slip Behavior: Case of GFRP Bars in Geopolymer Concrete. *J. Mater. Civ. Eng.* **2015**, *27*, 1–9. [\[CrossRef\]](#)
128. Yost, J.R.; Radlińska, A.; Ernst, S.; Salera, M. Structural behavior of alkali activated fly ash concrete. Part 1: Mixture design, material properties and sample fabrication. *Mater. Struct.* **2012**, *46*, 435–447. [\[CrossRef\]](#)
129. Prakash, A.S.; Kumar, G.S. Experimental Study on Geopolymer Concrete using Steel Fibres. *Int. J. Eng. Trends Technol.* **2015**, *21*, 396–399. [\[CrossRef\]](#)
130. Albitar, M.; Ali, M.S.M.; Visintin, P. Experimental study on fly ash and lead smelter slag-based geopolymer concrete columns. *Constr. Build. Mater.* **2017**, *141*, 104–112. [\[CrossRef\]](#)
131. AS 3600–2009: *Australian Standard for Concrete Structures*; Standards Australia Limited: Sydney, NSW, Australia, 2001. Available online: [www.standards.org.au](http://www.standards.org.au) (accessed on 9 March 2022).
132. Srividya, T.; Kannan, R.P.R.; Sivasakthi, M.; Sujitha, A.; Jeyalakshmi, R. A state-of-the-art on development of geopolymer concrete and its field applications. *Case Stud. Constr. Mater.* **2022**, *16*, e00812. [\[CrossRef\]](#)
133. Cong, P.; Cheng, Y. Advances in geopolymer materials: A comprehensive review. *J. Traffic Transp. Eng. (Engl. Ed.)* **2021**, *8*, 283–314. [\[CrossRef\]](#)
134. Almutairi, A.L.; Tayeh, B.A.; Adesina, A.; Isleem, H.F.; Zeyad, A.M. Potential applications of geopolymer concrete in construction: A review. *Case Stud. Constr. Mater.* **2021**, *15*, e00733. [\[CrossRef\]](#)
135. Shehata, N.; Sayed, E.T.; Abdelkareem, M.A. Recent progress in environmentally friendly geopolymers: A review. *Sci. Total Environ.* **2021**, *762*, 143166. [\[CrossRef\]](#)
136. Figiela, B.; Šimonová, H.; Korniejewko, K. State of the art, challenges, and emerging trends: Geopolymer composite reinforced by dispersed steel fibers. *Rev. Adv. Mater. Sci.* **2022**, *61*, 1–15. [\[CrossRef\]](#)
137. Huang, J.; Kogbara, R.B.; Hariharan, N.; Masad, E.A.; Little, D.N. A state-of-the-art review of polymers used in soil stabilization. *Constr. Build. Mater.* **2021**, *305*, 124685. [\[CrossRef\]](#)

138. Arpajirakul, S.; Pungrasmi, W.; Likitlersuang, S. Efficiency of microbially-induced calcite precipitation in natural clays for ground improvement. *Constr. Build. Mater.* **2021**, *282*, 122722. [\[CrossRef\]](#)
139. Yang, G.; Zhao, J.; Wang, Y. Durability properties of sustainable alkali-activated cementitious materials as marine engineering material: A review. *Mater. Today Sustain.* **2022**, *17*, 100099. [\[CrossRef\]](#)
140. Shi, J.; Liu, B.; Liu, Y.; Wang, E.; He, Z.; Xu, H.; Ren, X. Preparation and characterization of lightweight aggregate foamed geopolymer concretes aerated using hydrogen peroxide. *Constr. Build. Mater.* **2020**, *256*, 119442. [\[CrossRef\]](#)
141. Meng, Q.; Wu, C.; Hao, H.; Li, J.; Wu, P.; Yang, Y.; Wang, Z. Steel fibre reinforced alkali-activated geopolymer concrete slabs subjected to natural gas explosion in buried utility tunnel. *Constr. Build. Mater.* **2020**, *246*, 118447. [\[CrossRef\]](#)
142. Abdollahnejad, Z.; Miraldo, S.; Pacheco-Torgal, F.; Aguiar, J.B. Cost-efficient one-part alkali-activated mortars with low global warming potential for floor heating systems applications. *Eur. J. Environ. Civ. Eng.* **2015**, *21*, 412–429. [\[CrossRef\]](#)
143. Sierraab, V.; Chejne, F. Energy saving evaluation of microencapsulated phase change materials embedded in building systems. *J. Energy Storage* **2022**, *49*, 104102. [\[CrossRef\]](#)
144. Rahjoo, M.; Goracci, G.; Martauz, P.; Rojas, E.; Dolado, J.S. Geopolymer Concrete Performance Study for High-Temperature Thermal Energy Storage (TES) Applications. *Sustainability* **2022**, *14*, 1937. [\[CrossRef\]](#)
145. Mohd Basri, M.S.; Mustapha, F.; Mazlan, N.; Ishak, M.R. Rice Husk Ash-Based Geopolymer Binder: Compressive Strength, Optimize Composition, FTIR Spectroscopy, Microstructural, and Potential as Fire-Retardant Material. *Polymers* **2021**, *13*, 4373. [\[CrossRef\]](#)
146. Rovnanik, P.; Kusák, I.; Bayer, P.; Schmid, P.; Fiala, L. Comparison of electrical and self-sensing properties of Portland cement and alkali-activated slag mortars. *Cem. Concr. Res.* **2019**, *118*, 84–91. [\[CrossRef\]](#)
147. Pacheco-Torgal, F.; Abdollahnejad, Z.; Miraldo, S.; Baklouti, S.; Ding, Y. An overview on the potential of geopolymers for concrete infrastructure rehabilitation. *Constr. Build. Mater.* **2012**, *36*, 1053–1058. [\[CrossRef\]](#)
148. Khan, H.A.; Castel, A.; Khan, M. Corrosion investigation of fly ash based geopolymer mortar in natural sewer environment and sulphuric acid solution. *Corros. Sci.* **2020**, *168*, 108586. [\[CrossRef\]](#)
149. Cavallaro, G.; Milioto, S.; Lazzara, G. Halloysite Nanotubes: Interfacial Properties and Applications in Cultural Heritage. *Langmuir* **2020**, *36*, 3677–3689. [\[CrossRef\]](#)
150. Giacobello, F.; Ielo, I.; Belhamdi, H.; Plutino, M.R. Geopolymers and Functionalization Strategies for the Development of Sustainable Materials in Construction Industry and Cultural Heritage Applications: A Review. *Materials* **2022**, *15*, 1725. [\[CrossRef\]](#)
151. Abulencia, A.B.; Villoria, M.B.D.; Libre, R.G.D.; Quiatchon, P.R.J.; Dollente, I.J.R.; Guades, E.J.; Promentilla, M.A.B.; Garciano, L.E.O.; Ongpeng, J.M.C. Geopolymers as sustainable material for strengthening and restoring unreinforced masonry structures: A review. *Buildings* **2021**, *11*, 532. [\[CrossRef\]](#)
152. Ielo, I.; Giacobello, F.; Sfameni, S.; Rando, G.; Galletta, M.; Trovato, V.; Rosace, G.; Plutino, M.R. Nanostructured Surface Finishing and Coatings: Functional Properties and Applications. *Materials* **2021**, *14*, 2733. [\[CrossRef\]](#)
153. Kantarc, F.; Maras, M.M. Formulation of a novel nano TiO<sub>2</sub>-modified geopolymer grout for application in damaged beam-column joints. *Constr. Build. Mater.* **2022**, *317*, 125929. [\[CrossRef\]](#)
154. Rao, S.M.; Acharya, I.P. Synthesis and Characterization of Fly Ash Geopolymer Sand. *J. Mater. Civ. Eng.* **2014**, *1*, 912–917. [\[CrossRef\]](#)
155. Kashani, A.; Ngo, T. Optimisation of mixture properties for 3D printing of geopolymer concrete. In Proceedings of the International Symposium on Automation and Robotics in Construction, Berlin, Germany, 20–25 July 2018; pp. 1–8.
156. Bagheri, A.; Cremona, C. Formulation of mix design for 3D printing of geopolymers: A machine learning approach. *Mater. Adv.* **2020**, *1*, 720–727. [\[CrossRef\]](#)
157. Ur Rehman, A.; Sglavo, V.M. 3D printing of geopolymer-based concrete for building applications. *Rapid Prototyp. J.* **2020**, *26*, 1783–1788. [\[CrossRef\]](#)
158. Voney, V.; Odaglia, P.; Brumaud, C.; Dillenburger, B.; Habert, G. From casting to 3D printing geopolymers: A proof of concept. *Cem. Concr. Res.* **2021**, *143*, 106374. [\[CrossRef\]](#)
159. Gasca-Tirado, R.; Manzano-Ramírez, A.; RiveraMuñoz, E.M.; Velázquez-Castillo, R.; Apátiga-Castro, M.; Nava, R.; Rodríguez-López, A. Ion Exchange in Geopolymers. In *New Trends in Ion Exchange Studies*; Karakuş, S., Ed.; IntechOpen: London, UK, 2018; Chapter 5. [\[CrossRef\]](#)
160. Krishna, R.S.; Mishra, J.; Zribi, M.; Adeniyi, F.; Saha, S.; Baklouti, S.; Shaikh, F.U.A.; Gökçe, H.S. A review on developments of environmentally friendly geopolymer technology. *Materialia* **2021**, *20*, 101212. [\[CrossRef\]](#)
161. Jihui, Z. Eco-friendly geopolymer materials: A review of performance improvement, potential application and sustainability assessment. *J. Clean. Prod.* **2021**, *307*, 127085.
162. Karthik, S.; Mohan, K.S.R. A Taguchi Approach for Optimizing Design Mixture of Geopolymer Concrete Incorporating Fly Ash, Ground Granulated Blast Furnace Slag and Silica Fume. *Crystals* **2021**, *11*, 1279. [\[CrossRef\]](#)
163. Reddy, B.V.V.; Jagadish, K.S. Embodied energy of common and alternative building materials and technologies. *Energy Build.* **2003**, *35*, 129–137. [\[CrossRef\]](#)
164. Anvekar, S.R.; Manjunatha, L.R.; Anvekar, S.R.; Sagari, S.; Archana, K. An Economic and Embodied Energy Comparison of Geo-polymer, Blended Cement and Traditional Concretes. *J. Civ. Eng. Technol. Res.* **2014**, *1*, 33–40.
165. Al-Hamrani, A.; Kucukvar, M.; Alnahhal, W.; Mahdi, E.; Onat, N.C. Green Concrete for a Circular Economy: A Review on Sustainability, Durability, and Structural Properties. *Materials* **2021**, *14*, 351. [\[CrossRef\]](#)



166. de Azevedo, R.; Marvila, M.T.; de Oliveira, L.B.; Ferreira, W.M.; Colorado, H.; Rainho Teixeira, S.; Vieira, C.M.F. Circular economy and durability in geopolymers ceramics pieces obtained from glass polishing waste. *Int. J. Appl. Ceram. Technol.* **2021**, *18*, 1891–1900. [[CrossRef](#)]
167. Asim, N.; Badiei, M.; Torkashvand, M.; Mohammad, M.; Alghoul, M.A.; Gasaymeh, S.S.; Sopian, K. Wastes from the petroleum industries as sustainable resource materials in construction sectors: Opportunities, limitations, and directions. *J. Clean. Prod.* **2021**, *284*, 125459. [[CrossRef](#)]
168. Mathew, B.J.; Sudhakar, M.; Natarajan, C. Strength, Economic and Sustainability Characteristics of Coal Ash –GGBS Based Geopolymer Concrete. *Int. J. Comput. Eng. Res.* **2013**, *3*, 207–212.
169. Weil, M.; Dombrowski-Daube, K.; Buchwald, A. Geopolymer binders-Part 3: Ecological and economic analyses of geopolymer concrete mixes for external structural elements. *ZKG Int.* **2011**, *64*, 76–87.
170. Youssef, N.; Lafhaj, Z.; Chapiseau, C. Economic Analysis of Geopolymer Brick Manufacturing: A French Case Study. *Sustainability* **2020**, *12*, 7403. [[CrossRef](#)]
171. van Deventer, J.S.J.; Provis, J.L.; Duxson, P. Technical and commercial progress in the adoption of geopolymer cement. *Miner. Eng.* **2012**, *29*, 89–104. [[CrossRef](#)]
172. Shamsaei, E.; Bolt, O.; Basquiroto de Souza, F.; Benhelal, E.; Sagoe-Crentsil, K.; Sanjayan, J. Pathways to Commercialisation for Brown Coal Fly Ash-Based Geopolymer Concrete in Australia. *Sustainability* **2021**, *13*, 4350. [[CrossRef](#)]
173. Munir, Q.; Kärki, T. Cost Analysis of Various Factors for Geopolymer 3D Printing of Construction Products in Factories and on Construction Sites. *Recycling* **2021**, *6*, 60. [[CrossRef](#)]

# Global, regional, and national burden of colorectal cancer and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019

GBD 2019 Colorectal Cancer Collaborators\*



## Summary

**Background** Colorectal cancer is the third leading cause of cancer deaths worldwide. Given the recent increasing trends in colorectal cancer incidence globally, up-to-date information on the colorectal cancer burden could guide screening, early detection, and treatment strategies, and help effectively allocate resources. We examined the temporal patterns of the global, regional, and national burden of colorectal cancer and its risk factors in 204 countries and territories across the past three decades.

**Methods** Estimates of incidence, mortality, and disability-adjusted life years (DALYs) for colorectal cancer were generated as a part of the Global Burden of Diseases, Injuries and Risk Factors Study (GBD) 2019 by age, sex, and geographical location for the period 1990–2019. Mortality estimates were produced using the cause of death ensemble model. We also calculated DALYs attributable to risk factors that had evidence of causation with colorectal cancer.

**Findings** Globally, between 1990 and 2019, colorectal cancer incident cases more than doubled, from 842 098 (95% uncertainty interval [UI] 810 408–868 574) to 2.17 million (2.00–2.34), and deaths increased from 518 126 (493 682–537 877) to 1.09 million (1.02–1.15). The global age-standardised incidence rate increased from 22.2 (95% UI 21.3–23.0) per 100 000 to 26.7 (24.6–28.9) per 100 000, whereas the age-standardised mortality rate decreased from 14.3 (13.5–14.9) per 100 000 to 13.7 (12.6–14.5) per 100 000 and the age-standardised DALY rate decreased from 308.5 (294.7–320.7) per 100 000 to 295.5 (275.2–313.0) per 100 000 from 1990 through 2019. Taiwan (province of China; 62.0 [48.9–80.0] per 100 000), Monaco (60.7 [48.5–73.6] per 100 000), and Andorra (56.6 [42.8–71.9] per 100 000) had the highest age-standardised incidence rates, while Greenland (31.4 [26.0–37.1] per 100 000), Brunei (30.3 [26.6–34.1] per 100 000), and Hungary (28.6 [23.6–34.0] per 100 000) had the highest age-standardised mortality rates. From 1990 through 2019, a substantial rise in incidence rates was observed in younger adults (age <50 years), particularly in high Socio-demographic Index (SDI) countries. Globally, a diet low in milk (15.6%), smoking (13.3%), a diet low in calcium (12.9%), and alcohol use (9.9%) were the main contributors to colorectal cancer DALYs in 2019.

**Interpretation** The increase in incidence rates in people younger than 50 years requires vigilance from researchers, clinicians, and policy makers and a possible reconsideration of screening guidelines. The fast-rising burden in low SDI and middle SDI countries in Asia and Africa calls for colorectal cancer prevention approaches, greater awareness, and cost-effective screening and therapeutic options in these regions.

**Funding** Bill & Melinda Gates Foundation.

**Copyright** © 2022 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

## Introduction

In 2019, colorectal cancer was the third leading cause of cancer deaths and the second leading cause of disability-adjusted life years (DALYs) for cancer worldwide.<sup>1</sup> Around 70–75% of colorectal cancer cases occur sporadically and are associated with modifiable risk factors, whereas 25–30% of cases are linked to non-modifiable risk factors such as genetic factors, a personal history of polyps or adenoma, or a family history of colorectal cancer or hereditary risk (eg, Lynch syndrome or familial adenomatous polyposis).<sup>2–5</sup> Because of the increased prevalence of modifiable risk factors such as smoking, alcohol consumption, unhealthy diets, sedentary behaviour, physical inactivity, obesity, increasing life

expectancy, increasing awareness and affordability of colorectal cancer screening, and increasing screening capacity, incident cases of colorectal cancer are growing rapidly in low-income and middle-income countries (LMICs).<sup>6,7</sup>

The UN Sustainable Development Goal (SDG) target 3.4 focuses on reduction of premature mortality from non-communicable diseases (including cancers) by a third by 2030.<sup>8</sup> Colorectal cancer can be prevented by ameliorating modifiable risk factors, and deaths can be prevented through early detection of polyps with proven screening interventions;<sup>9–11</sup> therefore, addressing the global colorectal cancer burden must serve as one of the considerations towards progress on SDG 3.4 relating to

*Lancet Gastroenterol Hepatol*  
2022

Published Online  
April 7, 2022  
[https://doi.org/10.1016/S2468-1253\(22\)00044-9](https://doi.org/10.1016/S2468-1253(22)00044-9)

See Online/Comment  
[https://doi.org/10.1016/S2468-1253\(22\)00089-9](https://doi.org/10.1016/S2468-1253(22)00089-9)

\*Collaborators are listed at the end of the Article

Correspondence to:  
Dr Rajesh Sharma, University School of Management and Entrepreneurship, Delhi Technological University, Delhi 110095, India  
[rajesh.sharma@dtu.ac.in](mailto:rajesh.sharma@dtu.ac.in)

### Research in context

#### Evidence before this study

Colorectal cancer is one of the leading causes of cancer deaths worldwide. Previously, the Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) 2017 provided estimates for colorectal cancer incidence, deaths, and disability-adjusted life years (DALYs) for the period 1990–2017. Apart from GBD 2017, the International Agency for Research on Cancer (IARC) provided estimates for colorectal cancer for 2020 under the GLOBOCAN project. The present study was done as a part of GBD 2019, which produced estimates for 302 causes of death, 369 diseases and injuries, and 87 risk factors for 204 countries and territories for 1990–2019.

#### Added value of this study

In this study, we provide age-sex-location-specific estimates of colorectal cancer incidence, deaths, and DALYs for 204 countries and territories between 1990 and 2019. GBD 2019 produced estimates with technical collaboration from WHO, which has led to the inclusion of nine more WHO member countries. 28 714 site-years of data were used to estimate colorectal cancer incidence, deaths, and DALYs in GBD 2019, 16% more than in GBD 2017. In comparison with GLOBOCAN 2020, which provided colorectal cancer estimates for 2020, we provide estimates for full time series through 1990 to 2019 for all 204 countries and territories included in GBD 2019. Apart from estimates of incident cases, deaths, and age-standardised rates, as was done in GLOBOCAN 2020, we also estimated the burden of deaths and disability quantified with DALYs. The colorectal cancer burden was also examined in the light of country-level socioeconomic

development measured by Socio-demographic Index (SDI). The contribution of the main risk factors to colorectal cancer DALYs was also examined by sex in 21 world regions.

#### Implications of all the available evidence

Incident cases of colorectal cancer doubled or more than doubled in 16 of 21 world regions, and the number of deaths doubled or more than doubled in 15 of 21 world regions in the past three decades. The age-standardised incidence and death rates (per 100 000 person-years) either remained the same or decreased in high SDI quintiles and increased in low SDI and middle SDI quintiles. Large increases in colorectal cancer incidence rates were observed in middle SDI countries, as well as in people aged 20–49 years in high SDI countries. Further research is required to understand the causes of the colorectal cancer burden in younger adults (aged <50 years) and the main risk factors, including obesity, physical inactivity, alcohol consumption, smoking, and an altered gut microbiome, that might have led to the rise in the colorectal cancer burden. The increasing incidence of colorectal cancer in people younger than 50 years in high SDI countries also necessitates reconsideration of screening recommendations to include younger age groups (ie, those aged 40–49 years). The public health interventions for colorectal cancer awareness, screening, and prevention through containment of modifiable risk factors such as alcohol, smoking, an unhealthy diet (high in processed meat and fat, and low in fruits and vegetables), and obesity are key to stemming the tide of colorectal cancer worldwide.

non-communicable diseases. In this endeavour, recent changes in the colorectal cancer burden should be tracked at the global, regional, and national levels to identify those countries making progress and those countries and regions where more work is needed.

This study aimed to investigate the global, regional, and national burden of colorectal cancer in 204 countries and territories from 1990 to 2019. We examined the age-sex-location-specific burden of colorectal cancer using estimates from the Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) 2019.<sup>1,12–14</sup> The colorectal cancer burden was examined in light of the development status of countries measured by the Socio-demographic Index (SDI). We also aimed to quantify health loss due to colorectal cancer using DALYs, which encompasses the burden of a disease due to both deaths and disabilities caused by it. Last, we also examined the risk-attributable burden of the main risk factors for colorectal cancer. Apart from GBD, the International Agency for Research on Cancer (IARC) produced cancer estimates for 2020; however, GBD estimates facilitate examination of temporal patterns at global, regional, and national levels. An assessment of recent trends at the global, regional, and national levels, and the associated risk factors for countries at different levels of

development, can help track progress, map resource requirements, and help in policy making and implementation towards prevention and tackling the growing burden of colorectal cancer.

This manuscript was produced as part of the GBD Collaborator Network and in accordance with the GBD Protocol.

### Methods

#### Overview

The GBD 2019 estimates were generated for 286 causes of death, 369 causes of non-fatal burden, and 87 risk factors.<sup>1,12–14</sup> In comparison with the 195 countries and territories included in GBD 2017, nine more countries (Cook Islands, Monaco, San Marino, Nauru, Niue, Palau, Saint Kitts and Nevis, Tokelau, and Tuvalu) were included in GBD 2019, providing full time-series estimates from 1990 to 2019 for 204 countries and territories, which were grouped under seven super-regions and 21 regions. The colorectal cancer estimates mapped to the International Classification of Diseases (ICD) codes are available in the appendix (p 14).<sup>15</sup> The GBD estimation framework and calculation of all the metrics are detailed in GBD 2019 capstone publications and the supplementary appendices published with these reports.<sup>1,12–14</sup>

See Online for appendix

## Data sources

For GBD 2019, input data from various sources such as vital registration, verbal autopsy, and cancer registries were used to generate the colorectal cancer estimates. 28714 site-years of data (22849 site-years for vital registration, 516 site-years for verbal autopsy, and 5349 site-years for cancer registries) were used to estimate the colorectal cancer burden in GBD 2019, 3962 site-years (16%) more than GBD 2017. The vital registration system records data of vital events in a person's life (birth, death, and cause of death). Verbal autopsy is a data collection method generally used in populations without a complete vital registration system; in this method, a trained interviewer uses a questionnaire to collect information about signs, symptoms, and demographic characteristics of a recently deceased person from someone familiar with the deceased. Cancer registries are a data collection system that record and manage the data relating to a person with cancer. The information on various input sources of data used in GBD 2019 can be obtained from the GBD 2019 Data Input Sources Tool.

## Mortality estimation

GBD estimation begins with mortality estimation in multiple steps. The mortality data from cancer registries might be sparse, although incidence data can be available; therefore, to maximise data availability, mortality-to-incidence ratios (MIRs) were generated from the cancer registries that contained both incidence and mortality data. In the first step, incidence and mortality data from cancer registries were processed before they were matched by cancer, age, sex, year, and location to generate crude MIRs. Final MIRs were estimated by use of spatio-temporal Gaussian Process Regression (ST-GPR) using the Healthcare Access and Quality (HAQ) index, age, and sex<sup>16</sup> as covariates. The MIR estimates from the ST-GPR model were multiplied with incidence data to generate crude mortality estimates. The final mortality estimates were produced with the Cause of Death Ensemble Model (CODEm) using crude mortality estimates from the last step and those from vital registration and verbal autopsy as inputs along with other variables taken as covariates.<sup>1,15</sup> Only those variables were chosen that have been found to have a plausible relation with death due to colorectal cancer. The list of covariates at different levels used in CODEm for colorectal cancer is presented in the appendix (pp 15–16).<sup>1,15</sup> The final mortality estimates from CODEm were then divided by the MIRs from ST-GPR to generate age-sex-location-specific estimates for incident cases.

## Non-fatal estimation

The mortality estimates generated from CODEm were combined with reference life tables to generate estimates for years of life lost (YLLs).<sup>12</sup> The 10-year prevalence was divided into four sequelae of a fixed

duration based on expected person-time spent in each of the sequelae: diagnosis and primary therapy (4·0 months), metastatic phase (9·7 months), and terminal phase (1·0 month), with the remaining duration assigned to the controlled phase.<sup>15</sup> Sequela-specific years lived with disability (YLDs) were calculated by multiplying disability weights with sequelae-specific prevalence. Total YLDs were calculated by summing the sequela-specific YLDs. The sum of YLLs and YLDs produced the DALYs estimates, with one DALY being equivalent to 1 year of healthy life lost.<sup>17</sup>

The age-specific rates of incidence, mortality, and DALYs were expressed per 100 000 person-years and calculated with GBD population estimates,<sup>12</sup> and the age-standardised rates were calculated as weighted averages of age-specific rates per 100 000 people, in which weights are the proportion of people in corresponding age groups as per the GBD world population age standard.<sup>12</sup> All GBD estimates in this Article are provided with 95% uncertainty intervals (UIs). For each computational step, 1000 draws are generated; 95% UIs are calculated by taking values at the 2·5th and 97·5th percentile from the 1000 draws, and are provided alongside the mean estimates.<sup>12</sup>

## Socio-demographic Index

The colorectal cancer burden was evaluated against country-level development measured with the SDI,<sup>12,15</sup> which is a composite indicator of three indicators: lag-distributed income per capita, average educational attainment for people aged 15 years and older, and the total fertility rate (in people aged <25 years). Each of these indicators was first rescaled on a scale of 0 (lowest) to 1 (highest) based on country-specific values. The geometric mean of these three indices provided the final value of the country-level SDI. Based on SDI values, the 204 countries and territories were categorised into five groups: low SDI (<0·45), low-middle SDI ( $\geq 0\cdot45$  and <0·61), middle SDI ( $\geq 0\cdot61$  and <0·69), high-middle SDI ( $\geq 0\cdot69$  and <0·80), and high SDI ( $\geq 0\cdot80$ ).

## Risk factors

Estimation of GBD risk factors is based on a comparative risk assessment framework and involves six steps. The first is identification of risk-outcome pairs: only those risk-outcomes that have convincing or plausible evidence, as per World Cancer Research Fund criteria,<sup>18</sup> are included in GBD risk factor estimation. The second is estimation of relative risk (RR) as a function of exposure for each risk-outcome pair. The third is distribution of exposure for each risk factor by age, sex, location, and year. The fourth is determining the theoretical minimum risk exposure level (TMREL). The fifth is estimation of the population attributable fraction (PAF) and attributable burden. The RR for each risk-outcome pair, exposure levels, and TMREL are used to model the PAF.<sup>13</sup> The PAF of a particular risk factor is multiplied by colorectal cancer

For the GBD 2019 Data Input Sources Tool see <http://ghdx.healthdata.org/gbd-2019/data-input-sources>

DALYs to generate the DALYs attributable to that risk factor. The sixth is estimating the PAF and attributable burden for the combination of risk factors.

The details of each of these steps and the underlying methodology are provided elsewhere.<sup>13</sup> In this GBD iteration, 87 risk factors were included, of which ten (alcohol use, diet high in processed meat, diet high in red meat, diet low in calcium, diet low in fibre, diet low in milk, high body-mass index [BMI], high fasting plasma glucose, low physical activity, and smoking) have a non-zero contribution to colorectal cancer deaths and DALYs. We assessed the percentage contribution of these ten risk factors to colorectal cancer DALYs in 2019. The total number of input data sources used for the exposure of different risk factors in GBD 2019 are detailed elsewhere,<sup>13</sup> of which we present the number of input data sources for the ten risk factors related to colorectal cancer (appendix p 17).

The percentage changes between 1990 and 2019 were interpreted as statistically significant if the 95% UI did not include zero. All data analysis and data visualisation in this study were done with statistical software R (version 4.1.1), Stata (version 13.1), and Python (version 3.8.8).

### Role of the funding source

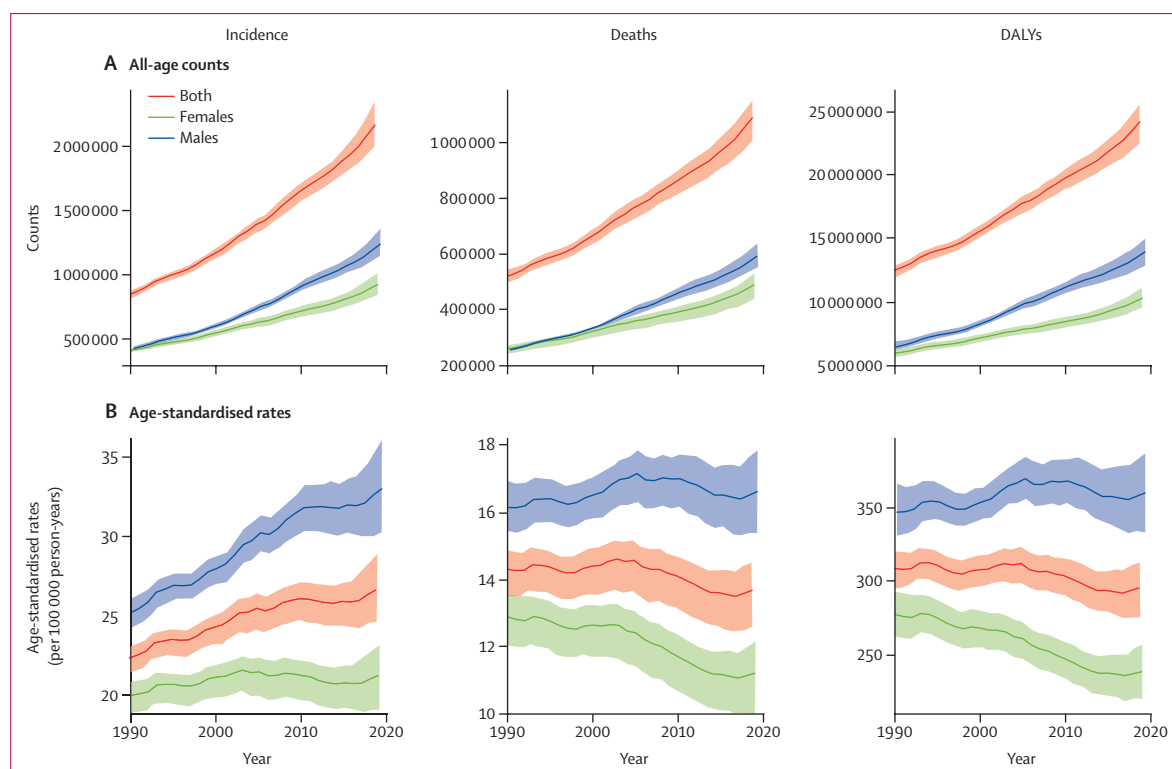
The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

## Results

### Overview of the global burden

Globally, for both sexes combined, incident cases of colorectal cancer more than doubled, from 842 098 (95% UI 810 408–868 574) in 1990 to 2·17 million (2·00–2·34) in 2019 (figure 1A). Between 1990 and 2019, deaths due to colorectal cancer increased from 518 126 (493 682–537 877) to 1·09 million (1·02–1·15), and DALYs increased from 12·4 million (11·9–12·9) to 24·3 million (22·6–25·7). By 2019, 95·6% of colorectal cancer DALYs were due to YLLs and 4·4% were due to YLDs (appendix p 5). The global age-standardised incidence rate of colorectal cancer increased from an estimated 22·2 (21·3–23·0) per 100 000 to 26·7 (24·6–28·9) per 100 000 from 1990 through 2019 (figure 1B). By contrast, the global age-standardised mortality rate decreased from 14·3 (13·5–14·9) per 100 000 to 13·7 (12·6–14·5) per 100 000, and the age-standardised DALY rate decreased from 308·5 (294·7–320·7) per 100 000 to 295·5 (275·2–313·0) per 100 000 from 1990 through 2019.

Males experienced greater increases in colorectal cancer incidence, deaths, and DALYs, than females in terms of absolute counts, and the age-standardised rates increased in males and remained similar throughout or decreased in females (figure 1). In 2019, males accounted for 57·2% (1·2 million [95% UI 1·1–1·4]) of colorectal cancer incident cases, and 54·9% (594 176 [550 959–638 031]) of deaths due to colorectal cancer. In 2019, the age-standardised



**Figure 1: Global temporal patterns of colorectal cancer burden, 1990–2019**

(A) All-age counts. (B) Age-standardised rates. Data source: Global Burden of Diseases, Injuries, and Risk Factors Study 2019. DALYs=disability-adjusted life-years.

incidence rate was 1.5-times higher in males than in females (33.1 [30.2–36.1] per 100 000 vs 21.2 [19.0–23.2] per 100 000), with a similar disparity between males and females in terms of the age-standardised mortality rate (16.6 [15.4–17.8] per 100 000 in males vs 11.2 [10.0–12.2] per 100 000 in females) and the age-standardised DALY rate (360.0 [333.1–387.8] per 100 000 in males vs 237.9 [218.7–257.1] per 100 000 in

females; figure 1). The number of incident cases increased in all SDI quintiles, whereas the age-standardised incidence rate decreased in the high SDI quintile only (appendix p 6). Similarly, deaths and DALYs increased in all SDI quintiles, but age-standardised mortality rates stagnated or decreased only in high-middle SDI and high SDI quintiles (appendix pp 7–8). Notably, the rise in all-age counts was steeper in all SDI quintiles than in the

	Incidence (95% UI)	Mortality (95% UI)	DALYs (95% UI)	Age-standardised incidence rate (95% UI)	Age-standardised mortality rate (95% UI)	Age-standardised DALY rate (95% UI)
Global	2 166 168 (1 996 298–2 342 842)	1 085 797 (1 002 795–1 149 679)	24 284 087 (22 614 920–25 723 221)	26.7 (24.6–28.9)	13.7 (12.6–14.5)	295.5 (275.2–313.0)
Andean Latin America	11 094 (8935–13 467)	5630 (4593–6791)	125 578 (101 753–151 796)	20.0 (16.1–24.2)	10.3 (8.4–12.4)	220.8 (179.0–266.3)
Australasia	23 671 (19 439–28 848)	8382 (7575–8978)	163 248 (150 872–173 959)	48.3 (39.6–59.1)	16.2 (14.8–17.3)	348.6 (324.2–370.7)
Caribbean	13 813 (11 813–15 959)	7995 (6935–9176)	172 016 (147 186–200 175)	26.7 (22.9–30.9)	15.5 (13.4–17.7)	333.3 (285.2–387.9)
Central Asia	10 949 (9999–12 008)	7467 (6822–8166)	199 841 (182 012–219 941)	15.2 (13.9–16.6)	11.2 (10.3–12.2)	256.8 (234.4–281.1)
Central Europe	84474 (74 551–95 453)	51 567 (45 636–57 749)	1 052 146 (922 923–1 184 246)	39.9 (35.2–45.1)	23.6 (20.8–26.4)	512.6 (448.7–577.9)
Central Latin America	37 542 (32 211–43 870)	22 470 (19 542–25 997)	539 638 (465 200–627 069)	15.9 (13.7–18.6)	9.7 (8.4–11.2)	223.7 (193.1–259.5)
Central sub-Saharan Africa	3957 (3015–5113)	3544 (2705–4609)	100 988 (75 749–131 447)	7.7 (5.9–10.1)	7.4 (5.7–9.9)	169.3 (129.2–220.2)
East Asia	637 096 (548 895–738 549)	275 604 (238 238–317 886)	6 712 862 (5 774 277–7 735 907)	30.9 (26.8–35.7)	14.1 (12.2–16.2)	325.2 (280.7–373.2)
Eastern Europe	106 017 (96 250–117 074)	63 476 (57 180–70 011)	1 419 105 (1 287 540–1 571 374)	31.1 (28.2–34.4)	18.3 (16.5–20.2)	423.7 (384.0–469.3)
Eastern sub-Saharan Africa	14 227 (12 130–16 886)	12 717 (10 940–15 001)	356 433 (301 931–425 606)	8.8 (7.6–10.4)	8.5 (7.4–9.9)	193.9 (166.0–229.6)
High-income Asia Pacific	196 371 (166 417–225 643)	76 929 (64 821–83 603)	1 327 823 (1 186 117–1 414 814)	44.6 (38.4–51.1)	15.3 (13.4–16.4)	323.9 (298.6–342.1)
High-income North America	260 911 (229 909–295 693)	95 664 (88 321–99 688)	1 987 109 (1 895 869–2 059 774)	42.7 (37.6–48.6)	14.9 (13.9–15.5)	339.9 (325.9–351.9)
North Africa and Middle East	60 010 (53 354–67 555)	39 147 (34 761–44 107)	1 013 634 (896 161–1 146 526)	13.9 (12.3–15.6)	9.8 (8.7–11)	218.7 (194.1–246.5)
Oceania	691 (555–855)	551 (443–682)	16 315 (12 915–20 556)	10.0 (8.2–12.1)	8.8 (7.2–10.7)	203.6 (163.6–252.5)
South Asia	113 711 (98 190–129 352)	94 846 (81 524–109 075)	2 419 098 (2 078 019–2 782 570)	8.3 (7.2–9.4)	7.3 (6.2–8.3)	165.1 (141.7–189.9)
Southeast Asia	117 010 (96 631–136 244)	82 024 (67 617–94 606)	2 142 434 (1 780 490–2 482 287)	19.3 (16–22.4)	14.4 (11.9–16.6)	334.0 (276.6–386.4)
Southern Latin America	26 866 (21 480–33 612)	17 930 (16 774–18 975)	366 436 (347 729–385 441)	32.2 (25.7–40.4)	21.2 (19.9–22.4)	447.6 (424.7–470.5)
Southern sub-Saharan Africa	7106 (6389–7882)	5922 (5329–6580)	147 780 (132 439–165 539)	13.1 (11.8–14.5)	11.5 (10.4–12.7)	250.4 (225.1–279.3)
Tropical Latin America	42 891 (40 118–44 928)	27 704 (25 668–29 090)	660 129 (625 562–687 740)	17.8 (16.6–18.6)	11.7 (10.8–12.3)	268.3 (253.7–279.8)
Western Europe	382 442 (332 800–432 448)	172 454 (155 345–181 815)	3 008 234 (2 815 060–3 152 895)	42.4 (37.1–48.3)	17.3 (15.8–18.1)	351.2 (332.0–366.8)
Western sub-Saharan Africa	15 321 (12 895–17 824)	13 773 (11 698–16 069)	353 242 (295 571–420 704)	8.7 (7.4–10.0)	8.4 (7.3–9.7)	176.1 (149.0–206.2)

Numbers in parenthesis represent 95% uncertainty intervals (UIs). DALYs=disability-adjusted life years. The age-standardised incidence rate, age-standardised mortality rate, and age-standardised DALY rate are shown per 100 000 person-years. Source: Global Burden of Diseases, Injuries, and Risk Factors Study 2019.

**Table: Region-wise colorectal cancer burden in 2019**



high SDI quintile such that the share of low, low-middle, middle, and high-middle SDI quintiles in the total colorectal cancer burden increased from 47.3% to 62.1% in terms of incident cases, from 57.1% to 69.8% in terms of deaths, and from 62.4% to 74.6% in terms of DALYs between 1990 and 2019 (appendix pp 6–8).

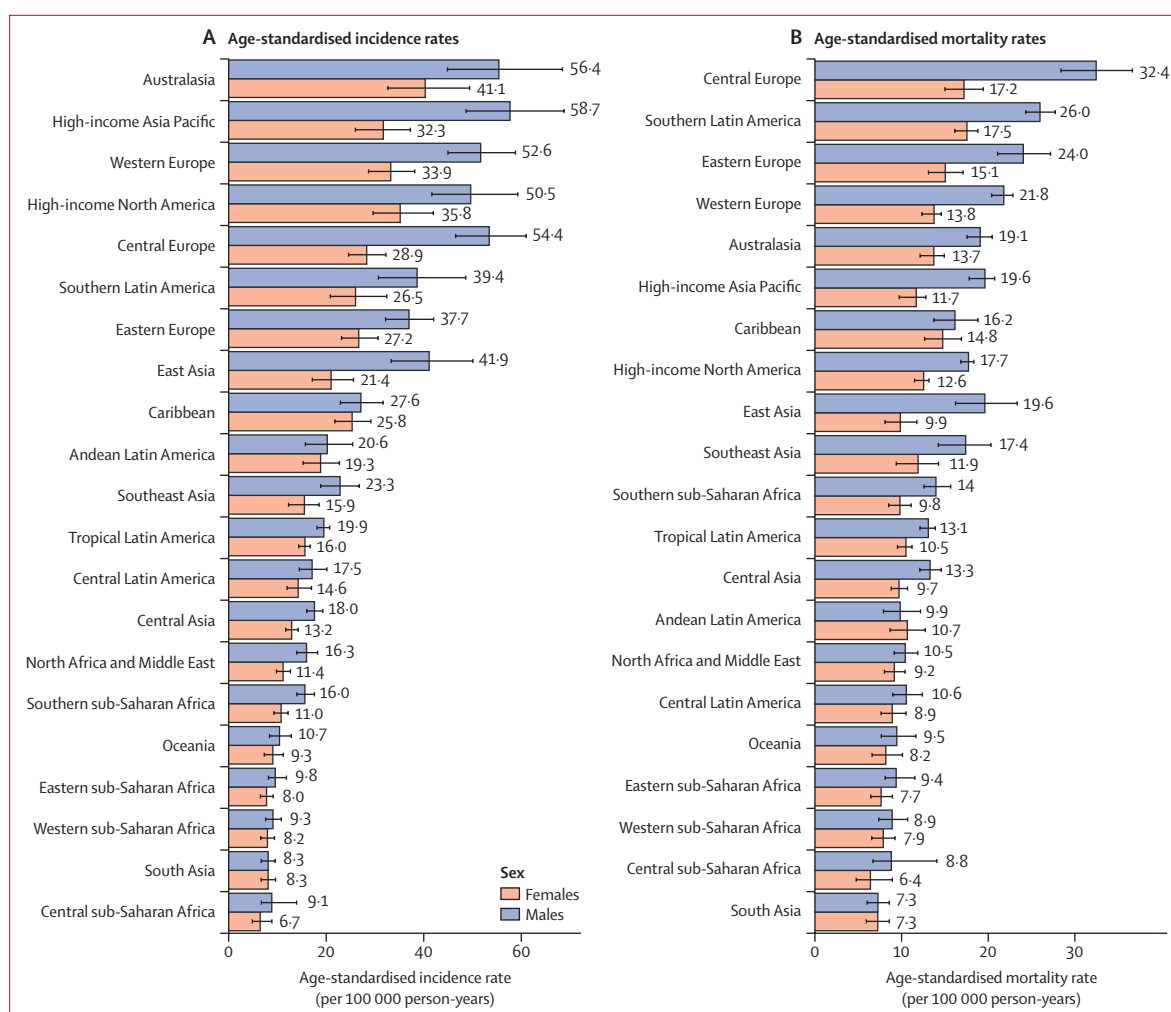
### Colorectal cancer burden by region

In 2019, east Asia was the worst-affected region, with 637 096 (95% UI 548 895–738 549) new cases, 275 604 (238 238–317 886) deaths, and 6.7 million (5.8–7.7) DALYs due to colorectal cancer (table). Australasia had the highest age-standardised incidence rate (48.3 [39.6–59.1] per 100 000) and central Europe had the highest age-standardised mortality rate (23.6 [20.8–26.4] per 100 000) across 21 regions in 2019. The age-standardised incidence rate was the lowest in central sub-Saharan Africa (7.7 [5.9–10.1] per 100 000) and south Asia (8.3 [7.2–9.4] per 100 000). South Asia also had the lowest

age-standardised mortality rate (7.3 [6.2–8.3] per 100 000). The age-standardised DALY rate was the highest in central Europe (512.6 [448.7–577.9] per 100 000) and lowest in south Asia (165.1 [141.7–189.9] per 100 000) in 2019 (table).

The preponderance of colorectal cancer in males in 2019, in terms of both age-standardised incidence rates and age-standardised mortality rates, was more apparent in developed regions (eg, Australasia, central Europe, and the high-income Asia Pacific) and differences between males and females were smaller in south Asia and regions of Africa (eg, eastern sub-Saharan Africa and western sub-Saharan Africa; figure 2).

Three patterns emerged from region-wise temporal trends of age-standardised rates in GBD regions (appendix p 9). First, in regions with already high incidence rates in 1990 (eg, Australasia and the high-income Asia Pacific), the age-standardised incidence rates either stagnated or decreased and age-standardised



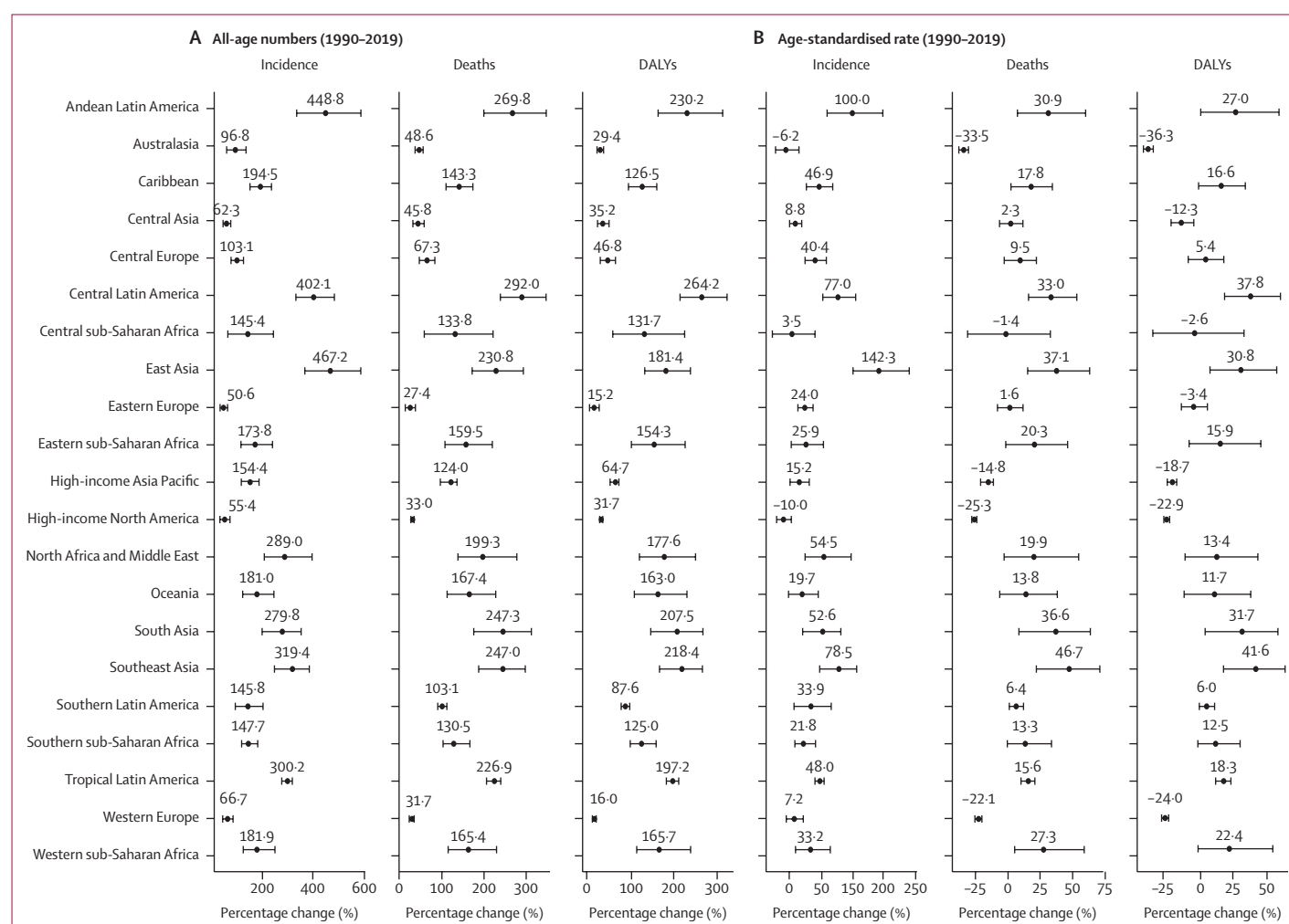
**Figure 2: Age-standardised rates of colorectal cancer in 2019, by sex and region**

(A) Age-standardised incidence rate (per 100 000 person-years). (B) Age-standardised mortality rate (per 100 000 person-years). Error bars denote 95% uncertainty intervals. Data source: Global Burden of Diseases, Injuries, and Risk Factors Study 2019.

mortality rates decreased in the past three decades. Second, concomitant rises in age-standardised incidence and mortality rates were observed in a few regions of Asia, Africa (eg, western sub-Saharan Africa and south Asia), and Oceania. Third, large increases occurred in age-standardised incidence rates, whereas a smaller increase was observed in age-standardised mortality rates in a few regions (eg, east Asia and south east Asia).

From 1990 to 2019, incident cases doubled or more than doubled in 16 of 21 regions, deaths doubled or more than doubled in 15 of 21 regions, and DALYs doubled or more than doubled in 13 of 21 regions, led by regions of Latin America and Asia (figure 3A). The changes in age-standardised rates were modest, with age-standardised incidence rates increasing by 50% or more in six GBD regions (east Asia, Andean Latin America, southeast Asia, central Latin America, North Africa and the Middle East, and south Asia; figure 3B). Age-standardised

incidence rates decreased, although these decreases were not statistically significant, in high-income North America ( $-10.0\%$  [95% UI  $-21.1$  to  $2.5$ ]) and Australasia ( $-6.2\%$  [ $-22.7$  to  $14.7$ ]) between 1990 and 2019. Age-standardised mortality rates increased the most in southeast Asia ( $46.7\%$  [ $21.8$  to  $70.0$ ]) followed by east Asia ( $37.1\%$  [ $15.1$  to  $62.3$ ]), and decreased in Australasia ( $-33.5\%$  [ $-29.8$  to  $-37.2$ ]) and high-income North America ( $-25.3\%$  [ $-23.5$  to  $-27.3$ ]) from 1990 through 2019, with a couple of other high-income regions experiencing a reduction ( $-22.1\%$  [ $-25.0$  to  $-19.6$ ] in western Europe and  $-14.8\%$  [ $-20.6$  to  $-10.9$ ] in the high-income Asia Pacific). Due to large increases in incidence and death rates, the age-standardised DALY rate increased the most in regions of Asia and Latin America, led by southeast Asia ( $41.6\%$  [ $18.2$  to  $62.7$ ]) and central Latin America ( $37.8\%$  [ $19.0$  to  $59.4$ ]) and decreased in high-income regions and regions of Europe (figure 3B).



**Figure 3: Region-wise percentage change in colorectal cancer burden, 1990–2019**

(A) All-age numbers. (B) Age-standardised rate (per 100 000). Data source: Global Burden of Diseases, Injuries, and Risk Factors Study 2019. Error bars denote 95% uncertainty intervals. DALYs=disability-adjusted life-years.

### Colorectal cancer burden by country

China, the USA, and Japan had the highest all-age incident counts for both sexes combined, with 607 900 (95% UI 521 805–708 420) new cases in China, 227 242 (197 022–261 375) new cases in the USA, and 160 211 (130 730–186 831) new cases in Japan, in 2019 (appendix pp 18–43). China (261 777 [224 403–303 318]), India (79 098 [67 137–92 723]), and the USA (84 026 [77 987–87 516]) had the highest death counts. Somalia (5·0 [3·1–9·2] per 100 000), Niger (5·6 [4·2–7·6] per 100 000), and Bangladesh (5·6 [3·9–8·0] per 100 000) had the lowest age-standardised incidence rates, whereas Taiwan (province of China; 62·0 [48·9–80·0] per 100 000), Monaco (60·7 [48·5–73·6] per 100 000), and Andorra (56·6 [42·8–71·9] per 100 000) had the highest age-standardised incidence rates (figure 4A; appendix pp 18–43). Greenland (31·4 [26·0–37·1] per 100 000), Brunei (30·3 [26·6–34·1] per 100 000), and Hungary (28·6 [23·6–34·0] per 100 000) had the highest age-standardised mortality rates, whereas Bangladesh (4·9 [3·4–7·1] per 100 000), Somalia (5·0 [3·2–9·3]), and Nepal (5·4 [3·9–7·4] per 100 000) had the lowest age-standardised mortality rates among 204 countries and territories in 2019 (figure 4B; appendix pp 18–43). Age-standardised DALY rates varied from 107·4 (74·6–152·7) per 100 000 in Bangladesh to 680·3 (555·7–812·4) per 100 000 in Greenland in 2019 (figure 4C; appendix pp 18–43).

Between 1990 and 2019, for both sexes combined, incident cases doubled or more than doubled in 157 of 204 countries and territories and deaths doubled or more than doubled in 129 of 204 countries and territories (appendix pp 44–66). Austria was the only country that reported a significant reduction in the number of colorectal cancer deaths (–20·5% [–26·4 to –14·8]) and DALYs (–29·7% [–34·4 to –25·0]) between 1990 and 2019. The changes in age-standardised rates were modest across countries in comparison with changes in absolute counts. From 1990 through 2019, the largest increase in age-standardised rates occurred in Cape Verde (age-standardised incidence rate 180·6% [121·1 to 237·3], age-standardised mortality rate 152·9% [97·3 to 204·0], and age-standardised DALY rate 114·2% [71·8 to 159·4]) and the largest reduction occurred in Austria (age-standardised incidence rate –34·1% [–46·7 to –19·1], age-standardised mortality rate –50·2% [–53·3 to –46·8], and age-standardised DALY rate –53·2% [–56·2 to –50·0]; appendix pp 44–66).

The relationship between country-level age-standardised rates of colorectal cancer and SDI in 2019 is shown in the appendix (p 10). SDI seemed to exert a positive relationship with age-standardised rates, with the slope becoming steeper in the upper end of the development spectrum for age-standardised incidence rates, whereas a positive relationship between age-standardised DALY rates and SDI seemed to taper off slightly towards high SDI countries.

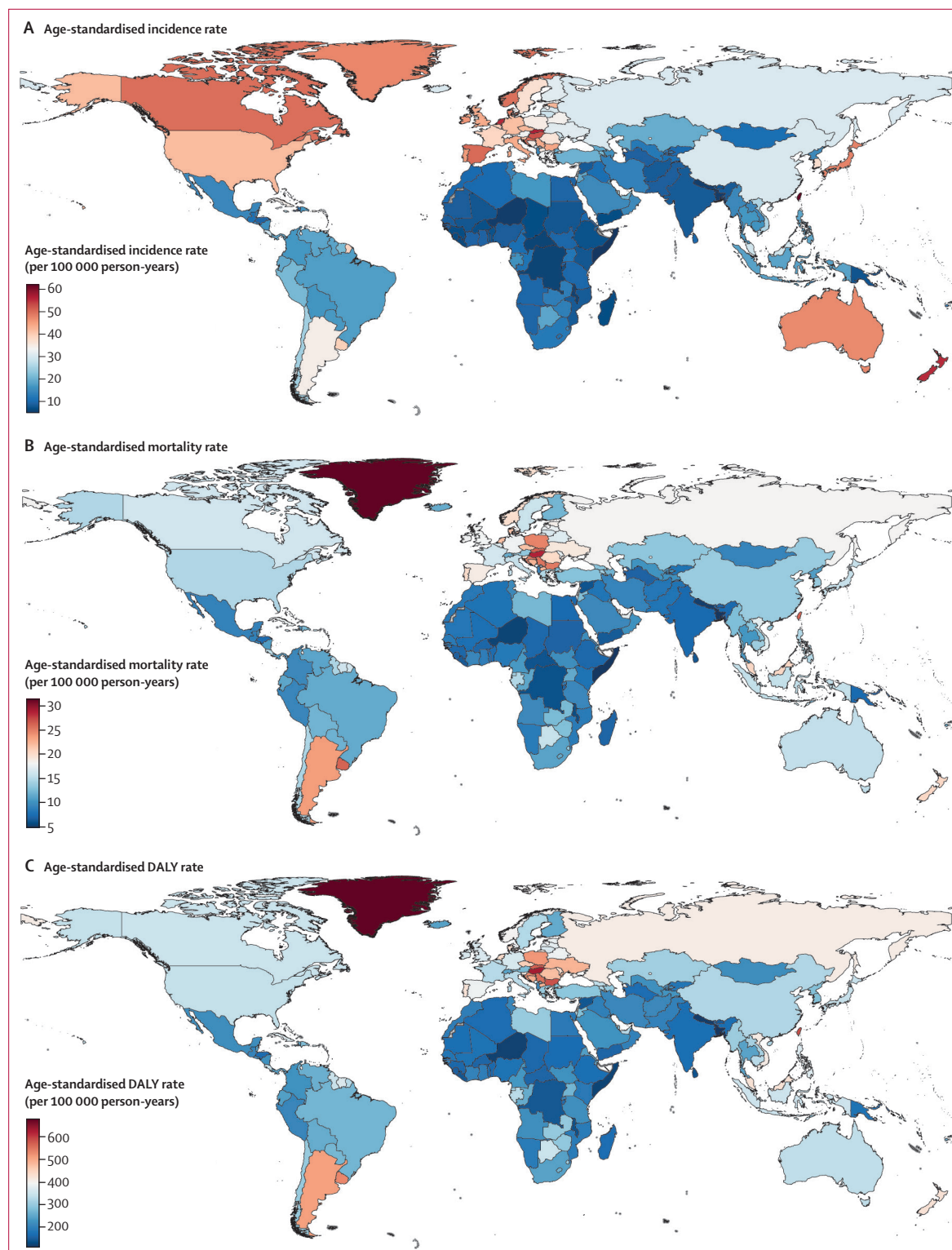
### Colorectal cancer burden by age group

Figure 5 illustrates the colorectal cancer incidence count and age-specific rates (per 100 000 person-years) in 2019. The incidence count followed a bell-shaped distribution with a peak in individuals aged 60–74 years for both males and females. Incident cases were higher in males than in females in all age groups up to age 80–84 years, with a greater number of new cases in females aged 85 years and older (figure 5A). Unlike incident cases, incidence rates continued to increase with age, increasing much faster in those aged 50–54 years and older (figure 5B). Between 1990 and 2019, all age groups experienced a rise in incident cases, with the highest increases occurring in those aged 85 years and older (appendix p 11). Furthermore, incidence rates increased in younger age groups (20–49 years) and decreased in older age groups (50–80 years) only in the high SDI quintile (appendix p 12). Percentage changes in incidence rates were higher in females compared to males only in low SDI quintiles; in all other SDI quintiles, males had higher percentage changes in incidence rates. Moving up in the development spectrum, the contrast between increases (or decreases) in younger age groups (20–49 years) and older age groups (50–74 years) became more pertinent. For instance, in the high SDI quintile, the incidence rates either remained unchanged or decreased in older age groups (50–74 years), whereas large increases were observed in younger age groups (20–49 years; appendix p 12).

### Risk factors

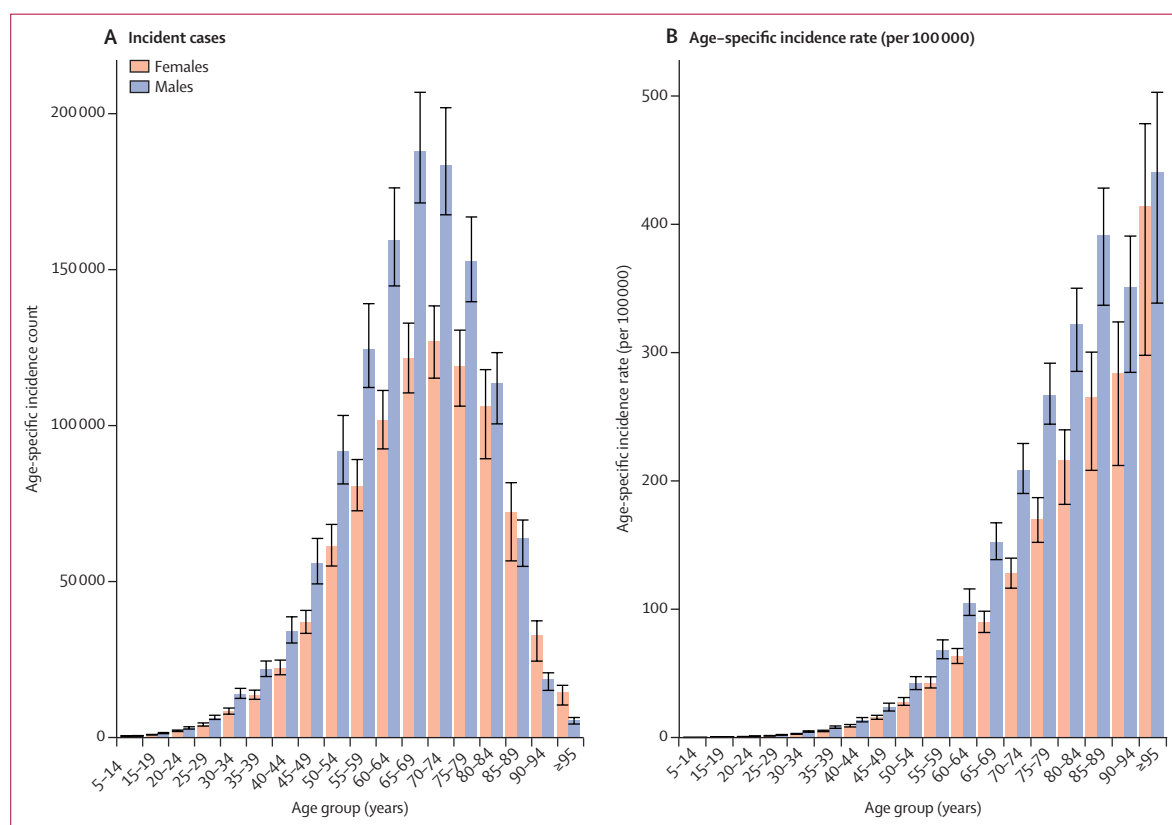
Figure 6 depicts the contribution of ten risk factors to all-age DALYs due to colorectal cancer, for both sexes combined, for 21 GBD regions in 2019. At the global level, a diet low in milk (15·6%), smoking (13·3%), a diet low in calcium (12·9%), and alcohol use (9·9%) were the main contributors to colorectal cancer DALYs, with the relative contribution of different risk factors varying as per the region's development status. In sub-Saharan Africa and Asia (barring the high-income Asia Pacific), a diet low in calcium and milk were the main risk factors, whereas smoking and alcohol use were the main risk factors in high-income regions (figure 6). At the global level, high BMI contributed only 8·3% of DALYs, with a higher contribution in comparatively high-income regions (eg, central Europe [14·0%] and high-income North America [13·8%]). The contribution of smoking to DALYs was greater than 15% in regions of Europe (eg, 18·2% in central Europe, 15·9% in western Europe, and 15·5% in eastern Europe) as well as in southern Latin America, east Asia, and high-income North America.

The severity of risk factors also varied by sex, with alcohol use contributing 13·9% of global DALYs and smoking contributing 18·9% of global DALYs in males, whereas alcohol use contributed only 4·5% of global DALYs and smoking contributed only 5·7% of global DALYs in females (appendix p 13). In females, a diet low in milk (15·5%), a diet low in calcium (12·0%), and high



**Figure 4: Geographical distribution of age-standardised rates of colorectal cancer in 2019**

(A) Age-standardised incidence rate. (B) Age-standardised mortality rate. (C) Age-standardised DALY rate. Data source: Global Burden of Diseases, Injuries, and Risk Factors Study 2019. DALY=disability-adjusted life-year.



**Figure 5: Age-specific burden of colorectal cancer in 2019**

(A) Incident cases. (B) Age-specific incidence rate (per 100 000). Error bars denote 95% uncertainty intervals. Data source: Global Burden of Diseases, Injuries, and Risk Factors Study 2019.

fasting plasma glucose (7.5%) were the major contributors to colorectal cancer DALYs. Similar differences were observed between sexes with regard to the contribution of high BMI, which contributed 11.1% of DALYs in males and 4.6% in females. The differences between sexes in the contribution of risk factors to colorectal cancer DALYs were also apparent in all GBD regions. The distribution of inadequate dietary risk factors (eg, a diet low in milk and calcium) was similar across the two sexes across regions of sub-Saharan Africa and Asia (appendix p 13).

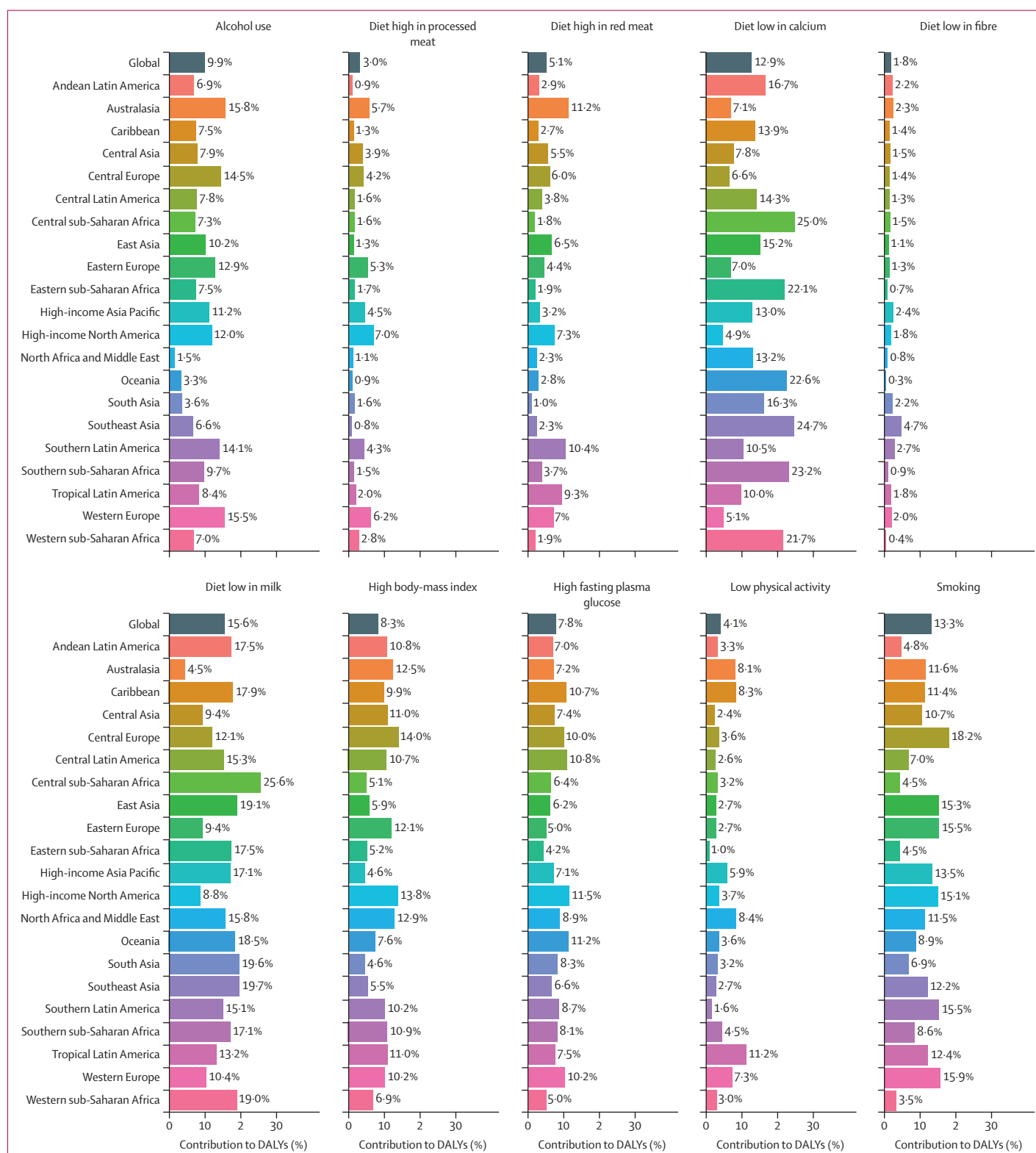
## Discussion

Between 1990 and 2019, new cases of colorectal cancer doubled or more than doubled in 157 of 204 countries and territories, and deaths due to colorectal cancer doubled or more than doubled in 129 of 204 countries and territories, and large increases were experienced in low SDI and middle SDI countries. Incidence and mortality rates mostly decreased in high SDI countries, whereas several low SDI and middle SDI countries and regions saw increases in age-standardised rates. The incident cases shared by low SDI, low-middle SDI, middle SDI, and high-middle SDI countries increased from 47.3% to 62.1%, the share of death counts rose from

57.1% to 69.8%, and the share of DALYs increased from 62.4% to 74.6% from 1990 through 2019.

We found large increases in age-standardised incidence rates in Asia (eg, 142.3% in east Asia and 78.5% in southeast Asia) and Latin America (eg, 100.0% in Andean Latin America, and 77.0% in central Latin America) between 1990 and 2019. Due to fast economic growth and rapid industrialisation, the thriving middle class in developing countries is adopting a westernised lifestyle characterised by an unhealthy diet (low in fruits and vegetables and high in red meat and processed meat), sedentary behaviour (eg, spending a long time watching television), and less physical activity, along with substance abuse (alcohol and smoking).<sup>19,20</sup> These behavioural changes have resulted in an increased incidence of lifestyle-related illnesses, including colorectal cancer.

The age-standardised mortality rate decreased in high-income regions (eg, -33.5% [95% UI -29.8 to -37.2] in Australasia and -25.3% [-23.5 to -27.3] in high-income North America) between 1990 and 2019. The reduction in incident cases and deaths, particularly among individuals older than 50 years, in high-SDI regions and countries suggests early detection of colorectal cancer due to screening, cancer registries, and technological improvements, as well as normalisation of early referral



**Figure 6: Percentage contribution of risk factors to all-age DALYs of colorectal cancer in 2019, for both sexes, globally and by regions**

Data source: Global Burden of Diseases, Injuries, and Risk Factors Study 2019. DALYs=disability-adjusted life-years.



to physicians. In most high SDI countries, 50–75 years is also the age group in which colorectal cancer screening is generally recommended for early detection of adenomatous polyps and adenomas.<sup>21</sup> In the USA, the impact of colorectal cancer screening in reduction of deaths due to colorectal cancer is notable given the considerable increases in colonoscopy screening in the 2000s.<sup>22</sup>

We found more colorectal cancer cases and deaths among males than females, with male age-standardised rates being 1.5–2.0-times higher than females in different GBD regions and the difference between males and females growing over time. In the 1960s, when the first international reports were published on colorectal cancer, the age-standardised incidence rate used to be higher for females than males, but the incidence has since risen faster for males than for females.<sup>23</sup> The predominance of males in the global colorectal cancer burden has been attributed to the high prevalence of visceral fat,<sup>24</sup> higher smoking prevalence,<sup>19</sup> and greater alcohol consumption among this demographic.<sup>25</sup> We also found a greater contribution of alcohol and smoking to colorectal cancer DALYs in males than in females. Some studies have also suggested protective effects of endogenous oestrogen against colorectal carcinogenesis in females.<sup>26</sup> Besides endogenous oestrogens, oral contraceptives might also contribute to the lowered colorectal cancer risk in females compared to males.<sup>27</sup> Besides the difference in incidence, there are also differences in mortality between males and females, with females tending to have better survival outcomes.<sup>28,29</sup> Sex apparently modulates the circadian clocks of males and females differently, with respect to the effects of chemotherapy in colorectal cancer, which might also contribute to the sex-based differences in survival and mortality.<sup>30</sup> In terms of the age-specific burden of colorectal cancer in males versus females, the findings of GBD 2019 differ from a previous study in which colorectal cancer incidence and mortality were higher in females than in males in individuals aged 65 years and older, implying that colorectal cancer is a major health threat in older females;<sup>31</sup> in the present study, however, such differences were mainly observed in individuals aged 80 years and older.

In line with previous studies,<sup>32–36</sup> we observed large increases in new cases and age-specific rates in individuals aged 20–49 years between 1990 and 2019, especially in high SDI countries. In the USA, a previous study<sup>37</sup> also noted substantial increases in colorectal cancer incidence rates in individuals aged 20–49 years from 1975 to 2010 and the incidence rate of colon cancers is expected to increase by an estimated 90%, and the incidence rate of rectal cancers is expected to increase by 140%, by 2030.<sup>37,38</sup> In most high-income countries, screening is recommended from the age of 50 years onwards; however, the recent trends in growing colorectal cancer incidence in younger adults (<50 years) have led to calls for a reconsideration of screening recommendations to

include individuals aged 40–49 years. In 2019, the American Cancer Society recommended colorectal cancer screening from the age of 45 years onwards,<sup>36</sup> which can be further modified to include younger people, especially targeting those with a high risk of colorectal cancer (eg, personal history of polyps or adenoma or family history of colorectal cancer or hereditary risk, males, smokers, and those with a high BMI).

The exact reasons for the increasing incidence of colorectal cancer in people younger than 50 years are less clear, but one possible reason could be due to the birth cohort effect, such that those born in the second half of the 20th century are increasingly exposed to potentially modifiable behavioural risk factors such as unhealthy diets, obesity, sedentary behaviour, low physical activity,<sup>18</sup> and increased smoking prevalence in early adulthood.<sup>32,33,39,40</sup> However, most of these risk factors have been implicated on the basis of evidence generated from patients with colorectal cancer aged 50 years or older, so the exact mechanisms or underlying risk factors remain less clear.<sup>40–42</sup> A previous study has attributed prolonged sedentary television viewing time to early onset of colorectal cancer, particularly rectal cancer.<sup>43</sup> Early-onset colorectal cancer has also been attributed to the rising obesity prevalence in younger adults.<sup>44–46</sup> A prospective cohort study of 94 217 women showed that dietary and lifestyle factors leading to hyperinsulinaemia are associated with an increased risk of colorectal cancer in younger women in the USA.<sup>47</sup> Binge drinking (episodic heavy alcohol consumption) is also implicated as one of the risk factors, with binge drinking being higher among adults younger than 50 years than in adults aged 50 years and older (ie, those of screening age).<sup>48</sup>

The GBD 2019 colorectal cancer estimates (2.17 million incident cases and 1.09 million deaths) are higher than GLOBOCAN estimates for 2020 (1.9 million incident cases and 935 173 deaths).<sup>49</sup> The global age-standardised rates estimated by GBD are also higher (age-standardised incidence rate 26.7 per 100 000; age-standardised mortality rate 13.7 per 100 000) than those for GLOBOCAN (age-standardised incidence rate 19.5 per 100 000 and age-standardised mortality rate 9.0 per 100 000). Both GBD and GLOBOCAN report high incidence and mortality rates in high-income regions and countries in Europe, North America, and Asia, and low incidence and mortality rates in LMICs of sub-Saharan Africa and south Asia, yet there are a few differences in the estimates, which stem from the different methodologies and data sources used. One of the main differences is that GLOBOCAN mostly estimates cancer incidence and deaths from cancer registry data, whereas GBD also models estimates from data sources such as vital registration, cancer registries, verbal autopsy, and sample registration systems. Another key difference between GBD and GLOBOCAN estimates stems from the redistribution of deaths from unknown or non-specific cancers to other cancers in GBD estimates.<sup>15</sup> The GBD estimation

framework has two key advantages over GLOBOCAN. In addition to producing incidence and mortality estimates, GBD produces estimates for YLLs, YLDs, and DALYs, which encompasses the disease burden due to both death and disability caused by the disease. Second, GLOBOCAN provides estimates for a single year (eg, 2002, 2008, 2012, 2018, and 2020) and there are some time series available for selected countries up to 2012 (the CI5Plus: Cancer Incidence in Five Continents Time Trends series); however, continuous time series of estimates of the colorectal cancer burden are not available from GLOBOCAN for sufficiently long time periods at the global and regional levels for all countries and territories. A track of temporal patterns can provide useful information; for instance, the rises in colorectal cancer incidence among younger adults (<50 years) in the past three decades, particularly in high-income regions, can allow researchers to examine the changing risk factors, set hypotheses, and help clinicians and policy makers to be more vigilant about the changing trends of colorectal cancer.

The major limitation of this study was the non-availability of data from cancer registries in many countries in Africa, the Caribbean, and Asia. Although GBD uses all available data from sources such as vital registration, verbal autopsy, and cancer registries, the accuracy of GBD estimates crucially depends upon cancer incidence and mortality data from cancer registries. Many LMICs in sub-Saharan Africa and Asia either do not have population-based cancer registries or the existing registries have insufficient coverage. Under-reporting also occurs because of poor documentation of cases, a shortage of medical facilities and adequately trained medical personnel, lack of a well established oncology centre, and as a result of misdiagnosis. Moreover, the 95% UIs reported in the study do not take into account several sources of bias, including measurement bias, selection bias due to missing data, and model specification bias. Second, the disease registration system was inadequate 30 years ago, especially in LMICs, and although it has improved in several LMICs in recent years, the estimates for the early years of the analysis (ie, the 1990s) are expected to be more biased, which is also reflected in the wider 95% UIs of the mean estimates and percentage changes. Third, there is a lag in the official reporting of cancer data for 2019 even in high SDI countries so the estimates for the most recent years are generated on the basis of recent trends and are provided with wider uncertainty intervals. GBD is an iterative estimation framework providing temporal estimates; the official data for 2019 will be updated as they become available and will be reflected in future GBD iterations.

In conclusion, colorectal cancer incident cases and deaths have more than doubled worldwide in the past three decades. Reducing the prevalence of modifiable risk factors and increasing screening uptake are key to reducing deaths from colorectal cancer, in line with achieving target 3.4 of the SDGs.<sup>8</sup> Low SDI and middle

SDI (including low-middle, middle, and high-middle SDI) countries, which together comprise close to 75% of colorectal cancer DALYs, are expected to experience further increases in colorectal cancer incidence as a result of population ageing and increasing life expectancy, improved screening and detection, and changing lifestyles. Strategies such as dietary and lifestyle modifications, early screening among high-risk individuals, access to high-quality health care, and better treatment modalities (including improved personalised therapy and access to clinical trials) are imperative to face this global challenge. Population-based cancer registries are important for monitoring colorectal cancer incidence and survival. The increase in colorectal cancer incidence in people younger than 50 years should serve as an early warning signal, necessitating greater awareness of these trends among clinicians, researchers, and policy makers, as well as more research into the risk factors and mechanisms that underpin these trends.

The data generated through GBD are an important resource for people and health-care providers, as they provide information on the effect of current treatment strategies, the effects of previous interventions, and the need for preventive measures. The findings from GBD 2019 can be used by policy makers and provide new perspectives for scientists and physicians throughout the world. These results provide comprehensive and comparable estimates that can inform efforts for equitable colorectal cancer control worldwide, with the larger goal of reducing the global burden of cancer.

#### GBD 2019 Colorectal Cancer Collaborators

Rajesh Sharma, Mohsen Abbasi-Kangevari, Rami Abd-Rabu, Hassan Abidi, Eman Abu-Gharbieh, Juan Manuel Acuna, Sangeet Adhikari, Shailesh M Advani, Muhammad Sohail Afzal, Mohamad Aghaie Meybodi, Bright Opoku Ahinkorah, Sajjad Ahmad, Ali Ahmadi, Sepideh Ahmadi, Haroon Ahmed, Luai A Ahmed, Muktar Beshir Ahmed, Hanadi Al Hamad, Fares Alahdab, Fahad Mashhour Alanezi, Turki M Alanzi, Fadwa Alhalaqi Naji Alhalaqi, Yousef Alimohamadi, Vahid Alipour, Syed Mohamed Aljunid, Motasem Alkhayat, Sami Almustanyir, Rajaa M Al-Raddadi, Saba Alvand, Nelson Alvis-Guzman, Saeed Amini, Robert Ancuceanu, Amir Anoushiravani, Ali Arash Anoushiravani, Alireza Ansari-Moghaddam, Jalal Arabloo, Armin Aryannejad, Mohammad Asghari Jafarabadi, Seyyed Shamsadin Athari, Floriane Ausloos, Marcel Ausloos, Atalel Fentahun Awedew, Mamaru Ayenew Awoke, Tegegn Mulatu Ayana, Sina Azadnajafabad, Hiva Azami, Mohammadreza Azangou-Khyavy, Amirhossein Azari Jafari, Ashish D Badiye, Sara Bagherieh, Saeed Bahadory, Atif Amin Baig, Jennifer L Baker, Maciej Banach, Amadou Barrow, Alemshet Yirga Berhie, Sima Besharat, Devidas S Bhagat, Akshaya Srikanth Bhagavathula, Neeraj Bhala, Kritika Bhattacharyya, Vijayalakshmi S Bhojaraja, Sadia Bibi, Ali Bijani, Antonio Biondi, Tone Bjørge, Belay Boda Abule Bodicha, Dejana Braithwaite, Hermann Brenner, Daniela Calina, Chao Cao, Yin Cao, Giulia Carreras, Felix Carvalho, Ester Cerin, Raja Chandra Chakinala, William C S Cho, Dinh-Toi Chu, Joao Conde, Vera Marisa Costa, Natália Cruz-Martins, Omid Dadras, Xiaochen Dai, Lalit Dandona, Rakhi Dandona, Anna Danielewicz, Feleke Mekonnen Demeke, Getu Debalkie Demissie, Rupak Desai, Deepak Dharmnitiya, Mostafa Dianatinasab, Daniel Diaz, Mojtaba Didehdar, Saeid Doaei, Linh Phuong Doan, Milad Dodangeh, Fatemeh Eghbalian, Debela Debela Ejeta, Michael Ekholuenetale, Temitope Cyrus Ekundayo, Iman El Sayed, Muhammed Elhadi,

For more on CI5Plus see <https://ci5.iarc.fr/CI5plus/Default.aspx>

- Daniel Berhanie Enyew, Tahir Eyayu, Rana Ezzeddini, Ildar Ravisovich Fakhraiyev, Umar Farooque, Hossein Farrokhpour, Farshad Farzadfar, Ali Fatehizadeh, Hamed Fattahi, Nima Fattahi, Masood Fereidoonnezad, Eduarda Fernandes, Getahun Fetensa, Irina Filip, Florian Fischer, Masoud Foroutan, Peter Andras Gaal, Mohamed M Gad, Silvano Gallus, Tushar Garg, Tamiru Getachew, Seyyed-Hadi Ghamari, Ahmad Ghashghaee, Nermin Ghith, Maryam Gholamalizadeh, Jamshid Gholizadeh Navashenaq, Abraham Tamirat Gizaw, James C Glasbey, Mahaveer Golechha, Pouya Goleij, Kebebe Bekele Gonfa, Giuseppe Gorini, Avirup Guha, Sapna Gupta, Veer Bala Gupta, Vivek Kumar Gupta, Rasool Haddadi, Nima Hafezi-Nejad, Arvin Haj-Mirzaian, Rabih Halwani, Shafiu Haque, Sanam Hariri, Ahmed I Hasaballah, Soheil Hassanipour, Simon I Hay, Claudiu Herteliu, Ramesh Holla, Mohammad-Salar Hosseini, Mehdi Hosseinzadeh, Mihaela Hostiuc, Mowafa Househ, Junjie Huang, Ayesha Humayun, Ivo Iavicoli, Olayinka Stephen Ilesanmi, Irena M Ilic, Milena D Ilic, Farhad Islami, Masao Iwagami, Mohammad Ali Jahani, Mihajlo Jakovljevic, Tahereh Javaheri, Ranil Jayawardena, Rime Jebai, Ravi Prakash Jha, Tamas Joo, Nitin Joseph, Farahnaz Joukar, Jacek Jerzy Jozwiak, Ali Kabir, Rohollah Kalhor, Ashwin Kamath, Neeti Kapoor, Ibraheem M Karaye, Amirali Karimi, Joonas H Kauppila, Asma Kazemi, Mohammad Keykhaei, Yousef Saleh Khader, Himanshu Khajuria, Rovshan Khalilov, Javad Khanali, Maryam Khayamzadeh, Mahmoud Khodadost, Hanna Kim, Min Seo Kim, Adnan Kisa, Sezer Kisa, Ali-Asghar Kolahi, Hamid Reza Koohestani, Jacek A Kopec, Rajasekaran Koteeswaran, Ai Koyanagi, Yuvaraj Krishnamoorthy, G Anil Kumar, Manoj Kumar, Vivek Kumar, Carlo La Vecchia, Faris Hasan Lami, Iván Landires, Caterina Ledda, Sang-woong Lee, Wei-Chen Lee, Yeong Yeh Lee, Elvynna Leong, Bingyu Li, Stephen S Lim, Stany W Lobo, Joana A Loureiro, Raimundas Lunevicius, Farzan Madadzadeh, Ata Mahmoodpoor, Azeem Majeed, Mohammad-Reza Malekpoor, Reza Malekzadeh, Ahmad Azam Malik, Fariborz Mansour-Ghanaei, Lorenzo Giovanni Mantovani, Miquel Martorell, Sahar Masoudi, Prashant Mathur, Jitendra Kumar Meena, Entezar Mehrabi Nasab, Walter Mendoza, Alexios-Fotios A Mentis, Tomislav Mestrovic, Junmei Miao Jonasson, Bartosz Miazgowski, Tomasz Miazgowski, Gelana Fekadu Worku Mijena, Seyyedmohammadsadeq Mirmoenei, Mohammad Mirza-Aghazadeh-Attari, Hamed Mirzaei, Sanjeev Misra, Karzan Abdulmuhsin Mohammad, Esmail Mohammadi, Saeed Mohammadi, Seyyede Momeneh Mohammadi, Abdollah Mohammadian-Hafshejani, Shafiu Mohammed, Teroj Abdulrahman Mohammed, Nagabhishek Moka, Ali H Mokdad, Zeinab Mokhtari, Mariam Molokhia, Sara Momtazmanesh, Lorenzo Monasta, Ghobad Moradi, Rahmatollah Moradzadeh, Paula Moraga, Joana Morgado-da-Costa, Sumaira Mubarik, Francesc Mulita, Mohsen Naghavi, Mukhammad David Naimzada, Hae Sung Nam, Zuhair S Natto, Biswa Prakash Nayak, Javad Nazari, Ehsan Nazemalhosseini-Mojarad, Ionut Negoii, Cuong Tat Nguyen, Son Hoang Nguyen, Nurulamin M Noor, Maryam Noori, Seyyed Mohammad Ali Noori, Virginia Nuñez-Samudio, Chimezie Igwegbe Nzoputani, Bogdan Oancea, Oluwakemi Ololade Odukoya, Ayodipupo Sikiru Oguntade, Hassan Okati-Aliabad, Andrew T Olagunju, Tinuke O Olagunju, Sokking Ong, Samuel M Ostroff, Alicia Padron-Monedero, Reza Pakzad, Adrian Pana, Anamika Pandey, Fatemeh Pashazadeh Kan, Urvish K Patel, Uttam Paudel, Renato B Pereira, Navaraj Perumalsamy, Richard G Pestell, Zahra Zahid Piracha, Richard Charles G Pollok, Akram Pourshams, Naeimeh Pourtaheri, Akila Prashant, Mohammad Rabiee, Navid Rabiee, Amir Radfar, Sima Rafiei, Mosiur Rahman, Amir Masoud Rahmani, Valid Rahmadian, Nazanin Rajai, Aashish Rajesh, Vajihah Ramezani-Doroh, Kiana Ramezanzadeh, Kamal Ranabhat, Sina Rashedi, Amirfarzan Rashidi, Mahsa Rashidi, Mohammad-Mahdi Rashidi, Mandana Rastegar, David Laith Rawaf, Salman Rawaf, Reza Rawassizadeh, Mohammad Sadegh Razeghinia, Andre M N Renzaho, Negar Rezaei, Nima Rezaei, Saeid Rezaei, Mohsen Rezaeian, Sahba Rezazadeh-Khadem, Gholamreza Roshandel, Maha Mohamed Saber-Ayad, Bahar Saberzadeh-Ardestani, Basema Saddik, Hossein Sadeghi, Umar Saeed, Maryam Sahebazzamani, Amirhossein Sahebkar, Amir Salek Farrokhi, Amir Salimi, Hamideh Salimzadeh, Pouria Samadi, Mehrnoosh Samaei, Abdallah M Samy, Juan Sanabria, Milena M Santric-Milicevic, Muhammad Arif Nadeem Saqib, Arash Sarveazad, Brijesh Sathian, Maheswar Satpathy, Ione Jayce Ceola Schneider, Mario Šekerija, Sadaf G Sepanlou, Allen Seylani, Feng Sha, Sayed Mohammad Shafiee, Zahra Shaghaghi, Saeed Shahabi, Elaheh Shaker, Maedeh Sharifian, Javad Sharifi-Rad, Sara Sheikhbahaei, Jeevan K Shetty, Reza Shirkoohi, Parnian Shobeiri, Sudeep K Siddappa Malleshappa, Diego Augusto Santos Silva, Guilherme Silva Julian, Achintya Dinesh Singh, Jasvinder A Singh, Md Shahjahan Siraj, Gholam Reza Sivandzadeh, Valentin Yuriech Skryabin, Anna Aleksandrovna Skryabina, Bogdan Socea, Marco Solmi, Mohammad Sadegh Soltani-Zangbar, Suhang Song, Viktória Szerencsés, Miklós Szócska, Rafael Tabarés-Seisdedos, Elnaz Tabibian, Majid Taheri, Yasaman Taheri Abkenar, Amir Taherkhani, Iman M Talaat, Ker-Kan Tan, Abdelghani Tbakhi, Bekele Tesfaye, Amir Tiyyuri, Daniel Nigusse Tollosa, Mathilde Touvier, Bach Xuan Tran, Biruk Shalmeno Tusa, Irfan Ullah, Saif Ullah, Marco Vacante, Sahel Valadan Tahbaz, Massimiliano Veroux, Bay Vo, Theo Vos, Cong Wang, Ronny Westerman, Melat Woldemariam, Seyed Hossein Yahyazadeh Jabbari, Lin Yang, Fereshteh Yazdanpanah, Chuanhua Yu, Deniz Yuce, Ismaeel Yunusa, Vesna Zadnik, Mazyar Zahir, Iman Zare, Zhi-Jiang Zhang, and Mohammad Zoladl.
- Affiliations**  
University School of Management and Entrepreneurship (R Sharma PhD), Delhi Technological University, New Delhi, India; Social Determinants of Health Research Center (M Abbasi-Kangevari MD, M Azangou-Khyavy MD, S Ghamari MD, J Khanali MD, A Kolahi MD, M Rashidi MD, A Salimi MD), Department of Epidemiology (A Ahmadi PhD), School of Advanced Technologies in Medicine (S Ahmadi PhD), Department of Community Nutrition (S Doaei PhD), Cancer Research Center (M Gholamalizadeh PhD), Department of Pharmacology (A Haj-Mirzaian MD, K Ramezanzadeh PharmD), Obesity Research Center (A Haj-Mirzaian MD), Department of Health & Community Medicine (A Kolahi MD), Gastroenterology and Liver Diseases Research Center (E Nazemalhosseini-Mojarad PhD), Medical Genetics Genomic Research Center (H Sadeghi PhD), Phytochemistry Research Center (J Sharifi-Rad PhD, Y Taheri Abkenar PharmD), Medical Ethics and Law Research Center (M Taheri PhD), Shahid Beheshti University of Medical Sciences, Tehran, Iran (M Khayamzadeh MD); Evidence-Based Practice Center (R Abd-Rabu MD), Mayo Clinic, Rochester, MN, USA; Laboratory Technology Sciences Department (H Abidi PhD), Department of Nursing (M Zoladi PhD), Yasuj University of Medical Sciences, Yasuj, Iran; Clinical Sciences Department (E Abu-Gharbieh PhD, Prof R Halwani PhD, M M Saber-Ayad MD, Prof I M Talaat PhD), College of Medicine (Prof R Halwani PhD), Sharjah Institute for Medical Research (B Saddik PhD), University of Sharjah, Sharjah, United Arab Emirates; Department of Epidemiology and Population Health (Prof J M Acuna MD), Khalifa University, Abu Dhabi, United Arab Emirates; FIU Robert Stempel College of Public Health & Social Work (Prof J M Acuna MD), Department of Epidemiology (R Jebai MPH), Florida International University, Miami, FL, USA; Biodesign Center for Environmental Health Engineering (S Adhikari MS), Arizona State University, Tempe, AZ, USA; Terasaki Institute for Biomedical Innovation, Los Angeles, CA, USA (S M Advani PhD); School of Medicine (S M Advani PhD), Georgetown University, Washington DC, USA; Department of Life Sciences (M S Afzal PhD), School of Sciences (M N Saqib PhD), University of Management and Technology, Lahore, Pakistan; Department of Medicine (M Aghaie Meybodi MD), Rutgers University, Newark, NJ, USA; The Australian Centre for Public and Population Health Research (ACPPHR) (B O Ahinkorah MPH), University of Technology Sydney, Sydney, NSW, Australia; Department of Health and Biological Sciences (S Ahmad PhD), Abasyn University, Peshawar, Pakistan; Department of Epidemiology and Biostatistics (A Ahmadi PhD, A Mohammadian-Hafshejani PhD), Shahrekord University of Medical Sciences, Shahrekord, Iran; Department of Biosciences (H Ahmed PhD), COMSATS Institute of Information Technology, Islamabad, Pakistan; Institute of Public Health (L A Ahmed PhD, A S Bhagavathula PharmD), United Arab Emirates University, Al Ain, United Arab Emirates;

Department of Epidemiology (M B Ahmed MPH), Department of Health, Behavior and Society (A T Gizaw MPH), Jimma University, Jimma, Ethiopia; Australian Center for Precision Health (M B Ahmed MPH), University of South Australia, Adelaide, SA, Australia; Geriatric and Long Term Care Department (H Al Hamad MD, B Sathian PhD), Rumailah Hospital (H Al Hamad MD), Hamad Medical Corporation, Doha, Qatar; Mayo Evidence-based Practice Center (F Alahdab MSc), Mayo Clinic Foundation for Medical Education and Research, Rochester, MN, USA; Health Information Management and Technology Department (T M Alanzi PhD), Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia (F M Alanezi PhD); Faculty of Nursing (F A N Alhalaqi PhD), Philadelphia University, Amman, Jordan; Psychological Sciences Association, Amman, Jordan (F A N Alhalaqi PhD); Pars Hospital (Y Alimohamadi PhD), Health Management and Economics Research Center (V Alipour PhD, J Arabloo PhD), Department of Health Economics (V Alipour PhD), Department of Internal Medicine (A Anoushirvani MD), School of Medicine (M Dodangeh MD), Minimally Invasive Surgery Research Center (A Kabir MD), Student Research Committee (M Noori MD), The Five Senses Health Institute (S Rezaei MD), Colorectal Research Center (A Sarveazad PhD), Trauma and Injury Research Center (M Taheri PhD), Department of Epidemiology (A Tiyyuri MSc), Iran University of Medical Sciences, Tehran, Iran (F Pashazadeh Kan BSN); Department of Epidemiology and Biostatistics (Y Alimohamadi PhD), Liver and Pancreatobiliary Diseases Research Center (S Alvand MD), Digestive Diseases Research Institute (A Anoushirvani MD, S Hariri MD, Prof R Malekzadeh MD, S Masoudi MSc, Prof A Pourshams MD, H Salimzadeh PhD, S G Sepanlou MD), Non-communicable Diseases Research Center (A Aryannejad MD, S Azadnajafabad MD, M Azangou-Khyavy MD, Prof F Farzadfar DSc, N Fattahi MD, S Ghamari MD, M Keykhaei MD, J Khanali MD, M Malekpour MD, S Momtazmanesh MD, M Rashidi MD, N Rezaei PhD, S Rezaadeh-Khadem MD), Experimental Medicine Research Center (A Aryannejad MD), Iranian Research Center for HIV/AIDS (IRCHA) (O Dadras DrPH), School of Medicine (H Farrokhpour MD, N Hafezi-Nejad MD, A Karimi MD, S Momtazmanesh MD), Students Scientific Research Center (SSRC) (M Keykhaei MD), Tehran Heart Center (E Mehrabi Nasab MD), Faculty of Medicine (E Mohammadi MD), Department of Cardiology (S Rashedi MD), Endocrinology and Metabolism Research Center (N Rezaei PhD), Research Center for Immunodeficiencies (Prof N Rezaei PhD), Faculty of Medicine (E Shaker MD, P Shobeiri MD), Cancer Research Center (R Shirkoobi PhD), Cancer Biology Research Center (R Shirkoobi PhD), Radiology Department (E Tabibian MD), Pediatric Allergy, Immunology, and Immunodeficiency Department (F Yazdanpanah MD), Department of Pharmacology (M Zahir MD), Tehran University of Medical Sciences, Tehran, Iran; Department of Health Policy and Management (Prof S M Aljunid PhD), Kuwait University, Safat, Kuwait; International Centre for Casemix and Clinical Coding (Prof S M Aljunid PhD), National University of Malaysia, Bandar Tun Razak, Malaysia; Department of Internal Medicine (M Alkhayyat MD, A D Singh MD), Department of Cardiovascular Medicine (M M Gad MD), Cleveland Clinic, Cleveland, OH, USA; College of Medicine (S Almustanyir MD), Alfaisal University, Riyadh, Saudi Arabia; Ministry of Health, Riyadh, Saudi Arabia (S Almustanyir MD); Department of Community Medicine (R M Al-Raddadi PhD), Rabigh Faculty of Medicine (A A Malik PhD), Department of Dental Public Health (Z S Natto DrPH), King Abdulaziz University, Jeddah, Saudi Arabia; Research Group in Hospital Management and Health Policies (Prof N Alvis-Guzman PhD), Universidad de la Costa (University of the Coast), Barranquilla, Colombia; Research Group in Health Economics (Prof N Alvis-Guzman PhD), University of Cartagena, Cartagena, Colombia; Department of Health Services Management (S Amini PhD), Khomoin University of Medical Sciences, Khomoin, Iran; Pharmacy Department (Prof R Ancuceanu PhD), Internal Medicine Department (M Hostiu PhD), Department of General Surgery (I Negoi PhD, B Socca PhD), Carol Davila University of Medicine and Pharmacy, Bucharest, Romania; Department of Epidemiology and Biostatistics (Prof A Ansari-Moghaddam PhD), Health Promotion Research Center (H Okati-Aliabad PhD), Zahedan University of Medical Sciences,

Zahedan, Iran; Department of Biostatistics and Epidemiology (Prof M Asghari Jafarabadi PhD), Student Research Committee (M Hosseini MD), Anesthesiology and Critical Care Department (Prof A Mahmoodpoor MD), Department of Radiology (M Mirza-Aghazadeh-Attari MD), Department of Immunology (M Soltani-Zangbar MSc), Department of Pediatric Allergy and Immunology (F Yazdanpanah MD), Tabriz University of Medical Sciences, Tabriz, Iran; Department of Biostatistics and Epidemiology (Prof M Asghari Jafarabadi PhD), Department of Immunology (S Athari PhD), Department of Anatomical Sciences (S Mohammadi PhD), Zanzan University of Medical Sciences, Zanzan, Iran; Gastro-enterology Department (F Ausloos MD), University of Liège, Liège, Belgium; School of Business (Prof M Ausloos PhD), University of Leicester, Leicester, UK; Department of Statistics and Econometrics (Prof M Ausloos PhD, Prof C Herteliu PhD, A Pana MD), Bucharest University of Economic Studies, Bucharest, Romania; Department of Surgery (A F Awedew MD), Addis Ababa University, Addis Ababa, Ethiopia; Department of Epidemiology and Preventive Medicine (M A Awoke MPH), University of Melbourne, Melbourne, VIC, Australia; School of Nursing (T M Ayana MSc), Biomedical Sciences Department (B B A Bodicha MSc, T Getachew MSc), Department of Medical Laboratory Science (M Woldemariam MSc), Arba Minch University, Arba Minch, Ethiopia; Department of Medical-Surgical Nursing (H Azami MSc), Department of Medicine (Prof F Eghbalian MD), Department of Pharmacology and Toxicology (R Haddadi PhD), Health Management and Economics (V Ramezani-Doroh PhD), Research Center for Molecular Medicine (A Taherkhani PhD), Hamadan University of Medical Sciences, Hamadan, Iran; School of Medicine (A Azari Jafari MD, S Mirmoeni MD), Shahroud University of Medical Sciences, Shahroud, Iran; Department of Forensic Science (A D Badiye MSc, N Kapoor MSc), Government Institute of Forensic Science, Nagpur, India; School of Medicine (S Bagherieh BSc), Department of Environmental Health Engineering (A Fatehizadeh PhD), Research Institute for Primordial Prevention of Non-Communicable Disease (S Hariri MD), Food Security Research Center (Z Mokhtari PhD), Isfahan University of Medical Sciences, Isfahan, Iran; Department of Parasitology (S Bahadory PhD), Department of Clinical Biochemistry (R Ezzeddini PhD), Tarbiat Modares University, Tehran, Iran; Department of Parasitology (S Bahadory PhD), Alborz University of Medical Sciences, Karaj, Iran; Unit of Biochemistry (A A Baig PhD), Universiti Sultan Zainal Abidin (Sultan Zainal Abidin University), Kuala Terengganu, Malaysia; Center for Clinical Research and Prevention (J L Baker PhD), Bispebjerg University Hospital, Frederiksberg, Denmark; Department of Hypertension (Prof M Banach PhD), Medical University of Lodz, Lodz, Poland; Polish Mothers' Memorial Hospital Research Institute, Lodz, Poland (Prof M Banach PhD); Department of Public & Environmental Health (A Barrow MPH), University of The Gambia, Brikama, The Gambia; Epidemiology and Disease Control Unit (A Barrow MPH), Ministry of Health, Kotu, The Gambia; School of Nursing (A Y Berhie MSc), University of Gondar, Bahir Dar, Ethiopia; Golestan Research Center of Gastroenterology and Hepatology (S Besharat PhD, G Roshandel PhD), Medical Genetics Department (H Sadeghi PhD), Golestan University of Medical Sciences, Gorgan, Iran; Department of Forensic Chemistry (D S Bhagat PhD), Government Institute of Forensic Science, Aurangabad, India; Department of Social and Clinical Pharmacy (A S Bhagavathula PharmD), Charles University, Hradec Kralova, Czech Republic; Institutes of Applied Health Research and Translational Medicine (N Bhala PhD), Queen Elizabeth Hospital Birmingham, Birmingham, UK; Institute of Applied Health Research (N Bhala PhD), NIHR Global Health Research Unit on Global Surgery (J C Glasbey MSc), University of Birmingham, Birmingham, UK; Department of Statistical and Computational Genomics (K Bhattacharyya MSc), National Institute of Biomedical Genomics, Kalyani, India; Department of Statistics (K Bhattacharyya MSc), University of Calcutta, Kolkata, India; Department of Anatomy (V S Bhojaraja MD), Department of Biochemistry (J K Shetty MD), Manipal University College Melaka, Melaka, Malaysia; Institute of Soil and Environmental Sciences (S Bibi PhD, S Ullah PhD), University of Agriculture, Faisalabad, Faisalabad, Pakistan; Social Determinants of Health Research Center (A Bijani PhD, M A Jahani PhD),



Babol University of Medical Sciences, Babol, Iran; Department of General Surgery and Medical-Surgical Specialties (Prof A Biondi PhD, M Vacante PhD), Clinical and Experimental Medicine (C Ledda PhD), Department of Medical and Surgical Sciences and Advanced Technologies (Prof M Veroux PhD), University of Catania, Catania, Italy; Department of Global Public Health and Primary Care (Prof T Bjørge PhD), University of Bergen, Bergen, Norway; Cancer Registry of Norway, Oslo, Norway (Prof T Bjørge PhD); Department of Epidemiology (D Braithwaite PhD), University of Florida, Gainesville, FL, USA; Cancer Population Sciences Program (D Braithwaite PhD), University of Florida Health Cancer Center, Gainesville, FL, USA; Division of Clinical Epidemiology and Aging Research (Prof H Brenner MD), German Cancer Research Center, Heidelberg, Germany; Clinical Pharmacy Department (Prof D Calina PhD), University of Medicine and Pharmacy of Craiova, Craiova, Romania; Program in Physical Therapy (C Cao MPH), Department of Surgery (Y Cao DSc), Washington University in St. Louis, St. Louis, MO, USA; Oncological Network, Prevention and Research Institute (G Gorini MD), Institute for Cancer Research, Prevention and Clinical Network, Florence, Italy (G Carreras PhD); Research Unit on Applied Molecular Biosciences (UCIBIO) (Prof F Carvalho PhD, V M Costa PhD), Department of Medicine (Prof N Cruz-Martins PhD), Associated Laboratory for Green Chemistry (LAQV) (Prof E Fernandes PhD), Laboratory for Process Engineering, Environment, Biotechnology and Energy (LEPABE) (J Loureiro PhD), University Hospital Center of Porto (J Morgado-da-Costa MSc), Department of Chemistry (R B Pereira PhD), University of Porto, Porto, Portugal; Mary MacKillop Institute for Health Research (Prof E Cerin PhD), Australian Catholic University, Melbourne, VIC, Australia; School of Public Health (Prof E Cerin PhD), University of Hong Kong, Hong Kong, China; Hospitalist Department (R Chakinala MD), Geisinger Health System, Danville, PA, USA; Department of Clinical Oncology (W C S Cho PhD), Queen Elizabeth Hospital, Hong Kong, China; Center for Biomedicine and Community Health (D Chu PhD), VNU-International School, Hanoi, Vietnam; Nova Medical School (J Conde PhD), Nova University of Lisbon, Lisbon, Portugal; Department of Health Sciences (Prof N Cruz-Martins PhD), Institute of Research and Advanced Training in Health Sciences and Technologies (CESPU), Famalicão, Portugal; School of Public Health (O Dadrás DrPH), Walailak University, Nakhon Si Thammarat, Thailand; Institute for Health Metrics and Evaluation (X Dai PhD), Prof L Dandona MD, Prof R Dandona PhD, Prof S I Hay DSc, Prof S S Lim PhD, A H Mokdad PhD, Prof M Naghavi PhD, S M Ostroff PhD, Prof T Vos PhD), Department of Health Metrics Sciences, School of Medicine (X Dai PhD, Prof R Dandona PhD, Prof S I Hay DSc, Prof S S Lim PhD, A H Mokdad PhD, Prof M Naghavi PhD, Prof T Vos PhD), Henry M Jackson School of International Studies (S M Ostroff PhD), University of Washington, Seattle, WA, USA; Department of Research (A Pandey PhD), Public Health Foundation of India, Gurugram, India (Prof L Dandona MD, Prof R Dandona PhD, G Kumar PhD); Indian Council of Medical Research, New Delhi, India (Prof L Dandona MD); Department of Human Nutrition (A Danielewicz PhD), Uniwersytet Warmińsko-Mazurski w Olsztynie (University of Warmia and Mazury in Olsztyn), Olsztyn, Poland; Department of Medical Laboratory Sciences (F M Demeke MSc), Bahir Dar University, Bahir Dar, Ethiopia; Institute of Public Health (G D Demissie MPH), University of Gondar, Gondar, Ethiopia; Division of Cardiology (R Desai MBBS), Atlanta Veterans Affairs Medical Center, Decatur, GA, USA; Department of Community Medicine (D Dhamnetiya MD, R P Jha MSc), Dr. Baba Saheb Ambedkar Medical College & Hospital, Delhi, India; Department of Epidemiology (M Dianatinasab MSc), Maastricht University, Maastricht, Netherlands; Department of Epidemiology (M Dianatinasab MSc), Nutrition Research Center (A Kazemi PhD), Non-communicable Disease Research Center (Prof R Malekzadeh MD, S G Sepanlou MD), Department of Clinical Biochemistry and Autophagy Research Center (S Shafiee PhD), Health Policy Research Center (S Shahabi PhD), Department of Internal Medicine (G R Sivandzadeh MD), Gastroenterohepatology Research Center (G R Sivandzadeh MD), Shiraz University of Medical Sciences, Shiraz, Iran; Center of Complexity Sciences (Prof D Diaz PhD), National Autonomous University of Mexico, Mexico City, Mexico; Faculty of Veterinary Medicine and Zootechnics (Prof D Diaz PhD), Autonomous University of Sinaloa, Culiacán Rosales, Mexico; Department of Parasitology and Mycology (M Didehdar PhD), Department of Epidemiology (R Moradzadeh PhD), Department of Pediatrics (J Nazari MD), Arak University of Medical Sciences, Arak, Iran; School of Health (S Doaei PhD), Gastrointestinal and Liver Diseases Research Center (S Hassanipour PhD, F Joukar PhD), Prof F Mansour-Ghanaei MD), Caspian Digestive Disease Research Center (S Hassanipour PhD, F Joukar PhD), Prof F Mansour-Ghanaei MD), Guilan University of Medical Sciences, Rasht, Iran; Institute for Global Health Innovations (L P Doan MSc), Faculty of Medicine (L P Doan MSc), Duy Tan University, Da Nang, Vietnam; Department of Medicine (D D Ejeta MD), Ambo University, Ambo, Ethiopia; Department of Epidemiology and Medical Statistics (M Ekholuenetale MSc), Faculty of Public Health (M Ekholuenetale MSc), Department of Community Medicine (O S Ilesanmi PhD), University of Ibadan, Ibadan, Nigeria; Department of Biological Sciences (T C Ekundayo PhD), University of Medical Sciences, Ondo, Ondo, Nigeria; Biomedical Informatics and Medical Statistics Department (I El Sayed PhD), Pathology Department (Prof I M Talaat PhD), Alexandria University, Alexandria, Egypt; Faculty of Medicine (M Elhadi MD), University of Tripoli, Tripoli, Libya; Department of Health informatics (D B Enyew MSc), Department of Nursing (G F W Mijena MSc), Haramaya University, Harar, Ethiopia; Medical Laboratory Sciences (T Eyayu MSc), Debre Tabor University, Debre Tabor, Ethiopia; Head of the Laboratory of Experimental Medicine (I R Fakhradiyev PhD), Kazakh National Medical University, Almaty, Kazakhstan; Department of Internal Medicine (U Farooque MD), Dow University of Health Sciences, Karachi, Pakistan; Endocrinology and Metabolism Research Institute (H Farrokhpour MD), Department of Epidemiology (S Rashedi MD, E Shaker MD), Department of Medicine (A Salimi MD), Department of International Studies (P Shobeiri MD), Non-Communicable Diseases Research Center (NCDRC), Tehran, Iran (E Mohammadi MD); Centre for Primary Health Care Network Management (H Fattahi PhD), Ministry of Health and Medical Education, Tehran, Iran; Student Research Committee (N Fattahi MD), Social Determinants of Health Research Center (G Moradi PhD), Department of Epidemiology and Biostatistics (G Moradi PhD), Kurdistan University of Medical Sciences, Sanandaj, Iran; Department of Medicinal Chemistry (M Fereidoonhezad PhD), Ahvaz Jundishapur University of Medical Sciences, Ahvaz, Iran; Department of Nursing (G Fetensa MSc), Wollega University, Nekemte, Ethiopia; Psychiatry Department (I Filip MD), Kaiser Permanente, Fontana, CA, USA; School of Health Sciences (I Filip MD), AT Still University, Mesa, AZ, USA; Institute of Gerontological Health Services and Nursing Research (F Fischer PhD), Ravensburg-Weingarten University of Applied Sciences, Weingarten, Germany; Department of Medical Parasitology (M Foroutan PhD), Faculty of Medicine (M Foroutan PhD), Abadan University of Medical Sciences, Abadan, Iran; Health Services Management Training Centre (P A Gaal PhD, T Joo MSc, V Szerencsés MA), Faculty of Health and Public Administration (M Szócska PhD), Semmelweis University, Budapest, Hungary; Department of Applied Social Sciences (P A Gaal PhD), Sapientia Hungarian University of Transylvania, Târgu-Mureş, Romania; Gillings School of Global Public Health (M M Gad MD), University of North Carolina Chapel Hill, Chapel Hill, NC, USA; Department of Environmental Health Sciences (S Gallus DSc), Mario Negri Institute for Pharmacological Research, Milan, Italy; Department of Radiology (T Garg MBBS), King Edward Memorial Hospital, Mumbai, India; School of Public Health (A Ghashghaee BSc), Institute for Prevention of Non-communicable Diseases (R Kalhor PhD), Health Services Management Department (R Kalhor PhD), Social Determinants of Health Research Center (S Rafiei PhD), Qazvin University of Medical Sciences, Qazvin, Iran; Research Group for Genomic Epidemiology (N Ghith PhD), Technical University of Denmark, Copenhagen, Denmark; Non-communicable Diseases Research Center (J Gholizadeh Navashenaq PhD, N Pourtaheri PhD), Bam University of Medical Sciences, Bam, Iran; Health Systems and Policy Research (M Golechha PhD), Indian Institute of Public Health, Gandhinagar, India; Department of Genetics (P Goleij MSc), Sana Institute of Higher Education, Sari, Iran; Department of Surgery (K B Gonfa MD), Madda Walabu University,

Bale Robe, Ethiopia; Harrington Heart and Vascular Institute (A Guha MD), Department of Nutrition and Preventive Medicine (Prof J Sanabria MD), Case Western Reserve University, Cleveland, OH, USA; Division of Cardiovascular Medicine (A Guha MD), Ohio State University, Columbus, OH, USA; Toxicology Department (S Gupta MSc), Shriram Institute for Industrial Research, Delhi, India; School of Medicine (V Gupta PhD), Deakin University, Geelong, VIC, Australia; Department of Clinical Medicine (Prof V K Gupta PhD), Macquarie University, Sydney, NSW, Australia; Department of Radiology and Radiological Science (N Hafezi-Nejad MD, S Sheikhbahaei MD), Johns Hopkins University, Baltimore, MD, USA; Research & Scientific Studies Unit (S Haque PhD), Jazan University, Jazan, Saudi Arabia; Department of Zoology and Entomology (A I Hasaballah PhD), Al Azhar University, Cairo, Egypt; School of Business (Prof C Herteliu PhD), London South Bank University, London, UK; Kasturba Medical College, Mangalore (R Holla MD, A Kamath MD), Manipal Academy of Higher Education, Manipal, India (A Kamath MD); Pattern Recognition and Machine Learning Lab (M Hosseinzadeh PhD), Gachon University, Seongnam, South Korea; College of Science and Engineering (Prof M Househ PhD), Hamad Bin Khalifa University, Doha, Qatar; Jockey Club School of Public Health and Primary Care (J Huang MD), The Chinese University of Hong Kong, Hong Kong, China; Department of Public Health and Community Medicine (Prof A Humayun PhD), Shaikh Khalifa Bin Zayed Al-Nahyan Medical College, Lahore, Pakistan; Department of Public Health (Prof I Iavicoli PhD), University of Naples Federico II, Naples, Italy; Department of Community Medicine (O S Ilesanmi PhD), Department of Medicine (A S Oguntade MSc), University College Hospital, Ibadan, Ibadan, Nigeria; Faculty of Medicine (I M Ilic PhD), Prof M M Santric-Milicevic PhD), School of Public Health and Health Management (Prof M M Santric-Milicevic PhD), University of Belgrade, Belgrade, Serbia; Department of Epidemiology (Prof M D Ilic PhD), Department of Global Health, Economics and Policy (Prof M Jakovljevic PhD), University of Kragujevac, Kragujevac, Serbia; Surveillance and Health Services Research (F Islami PhD), American Cancer Society, Atlanta, GA, USA; Department of Health Services Research (M Iwagami PhD), University of Tsukuba, Tsukuba, Japan; Department of Non-Communicable Disease Epidemiology (M Iwagami PhD), London School of Hygiene & Tropical Medicine, London, UK; Institute of Comparative Economic Studies (Prof M Jakovljevic PhD), Hosei University, Tokyo, Japan; Health Informatic Lab (T Javaheri PhD), Department of Computer Science (R Rawassizadeh PhD), Boston University, Boston, MA, USA; Department of Physiology (R Jayawardena PhD), University of Colombo, Colombo, Sri Lanka; School of Exercise and Nutrition Sciences (R Jayawardena PhD), Queensland University of Technology, Brisbane, QLD, Australia; Department of Community Medicine (R P Jha MSc), Banaras Hindu University, Varanasi, India; Department of Community Medicine (N Joseph MD), Manipal Academy of Higher Education, Mangalore, India; Department of Family Medicine and Public Health (J J Jozwiak PhD), University of Opole, Opole, Poland; School of Health Professions and Human Services (I M Karaye MD), Hofstra University, Hempstead, NY, USA; Surgery Research Unit (J H Kauppila MD), University of Oulu, Oulu, Finland; Department of Molecular Medicine and Surgery (J H Kauppila MD), Karolinska Institute, Stockholm, Sweden; Department of Public Health (Prof Y S Khader PhD), Jordan University of Science and Technology, Irbid, Jordan; Amity Institute of Forensic Sciences (H Khajuria PhD, B P Nayak PhD), Amity University, Noida, India; Department of Biophysics and Biochemistry (Prof R Khalilov PhD), Baku State University, Baku, Azerbaijan; Russian Institute for Advanced Study (Prof R Khalilov PhD), Moscow State Pedagogical University, Moscow, Russia; The Iranian Academy of Medical Sciences, Tehran, Iran (M Khayamzadeh MD); Department of Epidemiology (M Khodadost PhD), School of Public Health (M Khodadost PhD), Larestan University of Medical Sciences, Larestan, Iran; College of Medicine (H Kim BN), Ewha Womans University, Seoul, South Korea; Department of Genomics and Digital Health (M Kim MD), Samsung Advanced Institute for Health Sciences & Technology (SAIHST), Seoul, South Korea; Public Health Center (M Kim MD), Ministry of Health and Welfare, Wando, South Korea; School of Health Sciences (Prof A Kisa PhD),

Kristiania University College, Oslo, Norway; Department of Global Community Health and Behavioral Sciences (Prof A Kisa PhD), Tulane University, New Orleans, LA, USA; Department of Nursing and Health Promotion (S Kisa PhD), Oslo Metropolitan University, Oslo, Norway; Social Determinants of Health Research Center (H Koohestani PhD), Saveh University of Medical Sciences, Saveh, Iran; School of Population and Public Health (J A Kopec PhD), University of British Columbia, Vancouver, BC, Canada; Arthritis Research Canada, Richmond, BC, Canada (J A Kopec PhD); Microbiology & Molecular Cell Biology Department (R Koteeswaran MD), Eastern Virginia Medical School, Norfolk, VA, USA; Biomedical Research Networking Center for Mental Health Network (CIBERSAM) (A Koyanagi MD), San Juan de Dios Sanitary Park, Sant Boi de Llobregat, Spain; Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain (A Koyanagi MD); Department of Community Medicine (Y Krishnamoorthy MD), Employees' State Insurance Model Hospital, Chennai, India; Chemical and Biochemical Processing Division (M Kumar PhD), Central Institute for Research on Cotton Technology, Mumbai, India; Brigham and Women's Hospital (V Kumar MD), Oral Health Policy and Epidemiology Department (Z S Natto DrPH), Department of Internal Medicine (N Rajai MD), Harvard University, Boston, MA, USA; Department of Clinical Sciences and Community Health (Prof C La Vecchia MD), University of Milan, Milan, Italy; Department of Community and Family Medicine (F H Lami PhD), University of Baghdad, Baghdad, Iraq; Unit of Genetics and Public Health (Prof I Landires MD), Unit of Microbiology and Public Health (V Nuñez-Samudio PhD), Institute of Medical Sciences, Las Tablas, Panama; Department of Public Health (V Nuñez-Samudio PhD), Ministry of Health, Herrera, Panama (Prof I Landires MD); Pattern Recognition and Machine Learning Lab (Prof S Lee PhD), Gachon University, Seongnam, South Korea; The Office of Health Policy & Legislative Affairs (W Lee PhD), University of Texas, Galveston, TX, USA; Department of Medicine (Prof Y Lee PhD), School of Medical Sciences (Prof Y Lee PhD), University of Science Malaysia, Kota Bharu, Malaysia; Faculty of Science (E Leong PhD), Universiti Brunei Darussalam (University of Brunei Darussalam), Bandar Seri Begawan, Brunei; Department of Sociology (B Li PhD), Shenzhen University, Shenzhen, China; Department of Professional and Medical Education (S W Lobo PhD), Meharry Medical College, Nashville, TN, USA; Department of Biomedical Sciences (S W Lobo PhD), Mercer University, Macon, GA, USA; School of Health (J Loureiro PhD), Polytechnic Institute of Porto, Portugal; Department of General Surgery (Prof R Lunevicius DSc), Liverpool University Hospitals NHS Foundation Trust, Liverpool, UK; Department of Surgery (Prof R Lunevicius DSc), University of Liverpool, Liverpool, UK; Department of Biostatistics and Epidemiology, School of Public Health (F Madadzadeh PhD), Yazd University of Medical Sciences, Yazd, Iran; Department of Primary Care and Public Health (Prof A Majeed MD, Prof S Rawaf MD), WHO Collaborating Centre for Public Health Education and Training (D L Rawaf MD), Imperial College London, London, UK; University Institute of Public Health (A A Malik PhD), The University of Lahore, Lahore, Pakistan; School of Medicine and Surgery (Prof L G Mantovani DSc), University of Milan Bicocca, Monza, Italy; Value-Based Healthcare Unit (Prof L G Mantovani DSc), IRCCS MultiMedica, Sesto San Giovanni, Italy; Department of Nutrition and Dietetics (M Martorell PhD), Centre for Healthy Living (M Martorell PhD) University of Concepcion, Concepción, Chile; National Centre for Disease Informatics and Research (P Mathur PhD), Indian Council of Medical Research, Bengaluru, India; Department of Preventive Oncology (J K Meena MD), All India Institute of Medical Sciences, New Delhi, India; Peru Country Office (W Mendoza MD), United Nations Population Fund (UNFPA), Lima, Peru; University Research Institute (A A Mentis MD), National and Kapodistrian University of Athens, Athens, Greece; Clinical Microbiology and Parasitology Unit (T Mestrovic PhD), Dr. Zora Profzic Polyclinic, Zagreb, Croatia; University Centre Varazdin (T Mestrovic PhD), University North, Varazdin, Croatia; School of Public Health and Community Medicine (J Miao Jonasson PhD), University of Gothenburg, Gothenburg, Sweden; Center for Innovation in Medical Education (B Miazgowski MD), Department of Propeudeutics of Internal Diseases & Arterial Hypertension (Prof T Miazgowski MD), Pomeranian Medical



University, Szczecin, Poland (B Miazgowski MD); Social Determinants of Health Center (M Mirza-Aghazadeh-Attari MD), Urmia University of Medical Science, Urmia, Iran; Research Center for Biochemistry and Nutrition in Metabolic Diseases (H Mirzaei PhD), Kashan University of Medical Sciences, Kashan, Iran; Department of Surgical Oncology (Prof S Misra MCh), All India Institute of Medical Sciences, Jodhpur, India; Department of Biology (K A Mohammad PhD), Salahaddin University-Erbil, Erbil, Iraq; Infectious Diseases Research Center (S Mohammadi PhD), Golestan University of Medical Sciences, Gorgan, Iran; Health Systems and Policy Research Unit (S Mohammed PhD), Ahmadu Bello University, Zaria, Nigeria; Department of Health Care Management (S Mohammed PhD), Technical University of Berlin, Berlin, Germany; Dental Basic Sciences Department (T A Mohammed MSc), University of Duhok, Duhok, Iraq; Oncology Department (N Moka MD), Appalachian Regional Healthcare, Hazard, KY, USA; Department of Internal Medicine (N Moka MD), University of Kentucky, Lexington, KY, USA; Faculty of Life Sciences and Medicine (M Molokhia PhD), King's College London, London, UK; Clinical Epidemiology and Public Health Research Unit (L Monasta DSc), Burlo Garofolo Institute for Maternal and Child Health, Trieste, Italy; Computer, Electrical, and Mathematical Sciences and Engineering Division (P Moraga PhD), King Abdullah University of Science and Technology, Thuwal, Saudi Arabia; Department of Epidemiology and Biostatistics (S Mubarik MS, Prof C Yu PhD), School of Medicine (Z Zhang PhD), Wuhan University, Wuhan, China; Department of Surgery (F Multa MD), General University Hospital of Patras, Patras, Greece; Medical School (F Multa MD), University of Thessaly, Larissa, Greece; Laboratory of Public Health Indicators Analysis and Health Digitalization (M Naimzada MD), Moscow Institute of Physics and Technology, Dolgoprudny, Russia; Experimental Surgery and Oncology Laboratory (M Naimzada MD), Kursk State Medical University, Kursk, Russia; Department of Preventive Medicine and Public Health (Prof H Nam PhD), Chungnam National University School of Medicine, Daejeon, South Korea; Daejeon Regional Cancer Center (Prof H Nam PhD), Chungnam National University Hospital, Daejeon, South Korea; Department of General Surgery (I Negoi PhD), Emergency Hospital of Bucharest, Bucharest, Romania; Institute for Global Health Innovations (C T Nguyen MPH), Duy Tan University, Hanoi, Vietnam; Center of Excellence in Behavioral Medicine (S H Nguyen BS), Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam; Medical Research Council Clinical Trials Unit (N M Noor MRCP), Institute of Cardiovascular Science (A S Oguntade MSc), University College London, London, UK; Department of Gastroenterology (N M Noor MRCP), Cambridge University Hospitals, Cambridge, UK; Department of Nutrition (S Noori PhD), Jundishapur University of Medical Sciences, Ahvaz, Iran; Center of Excellence in Reproductive Health Innovation (CERHI) (C I Nzoputam MPH), University of Benin, Benin City, Nigeria; Administrative and Economic Sciences Department (Prof B Oancea PhD), University of Bucharest, Bucharest, Romania; Department of Community Health and Primary Care (O O Odukoya MSc), University of Lagos, Idi Araba, Nigeria; Department of Family and Preventive Medicine (O O Odukoya MSc), University of Utah, Salt Lake City, UT, USA; Department of Psychiatry and Behavioural Neurosciences (A T Olagunju MD), Department of Pathology and Molecular Medicine (T O Olagunju MD), McMaster University, Hamilton, ON, Canada; Department of Psychiatry (A T Olagunju MD), University of Lagos, Lagos, Nigeria; Non-communicable Disease Prevention Unit (S Ong FAMS), Ministry of Health, Bandar Seri Begawan, Brunei; Early Detection & Cancer Prevention Services (S Ong FAMS), Pantai Jerudong Specialist Centre, Bandar Seri Begawan, Brunei; National School of Public Health (A Padron-Monedero PhD), Institute of Health Carlos III, Madrid, Spain; Department of Epidemiology (R Pakzad PhD), Ilam University of Medical Sciences, Ilam, Iran; Department of Health Metrics (A Pana MD), Center for Health Outcomes & Evaluation, Bucharest, Romania; Department of Neurology and Public Health (U K Patel MD), Icahn School of Medicine at Mount Sinai, New York, NY, USA; Research Section (U Paudel PhD), Nepal Health Research Council, Kathmandu, Nepal; Faculty of Humanities and Social Sciences (U Paudel PhD), Tribhuvan University, Kathmandu, Nepal; Department of Zoology (Prof N Perumalsamy PhD), Yadava College, Madurai, India; Zoology (Prof N Perumalsamy PhD), Annai Fathima College, Madurai, Madurai District, India; Pennsylvania Cancer and Regenerative Medicine Center (R G Pestell MD), Baruch S Blumberg Institute, Doylestown, PA, USA; Department of Medicine (R G Pestell MD), Xavier University School of Medicine, Woodbury, NY, USA; Department of Public Health (Z Z Piracha PhD), Health Services Academy, Islamabad, Pakistan; Institute of Infection and Immunity (R C G Pollok FRCP), St George's University of London, London, UK; Department of Biochemistry (Prof A Prashant PhD), Jagadguru Sri Shivarathreeswara University, Mysuru, India; Biomedical Engineering Department (Prof M Rabiee PhD), Amirkabir University of Technology, Tehran, Iran; Department of Physics (N Rabiee PhD), Sharif University of Technology, Tehran, Iran; College of Medicine (A Radfar MD), University of Central Florida, Orlando, FL, USA; Department of Population Science and Human Resource Development (M Rahman DrPH), University of Rajshahi, Rajshahi, Bangladesh; Future Technology Research Center (A Rahmani PhD), National Yunlin University of Science and Technology, Yunlin, Taiwan (province of China); Department of Epidemiology (V Rahmanian PhD), Jahrom University of Medical Sciences, Jahrom, Iran; Department of Surgery (A Rajesh MD), University of Texas Health Science Center at San Antonio, San Antonio, TX, USA; Health Emergency Operation Center (K Ranabhat MPH), Ministry of Health & Population, Kathmandu, Nepal; Central Department of Public Health (K Ranabhat MPH), Institute of Medicine, Kathmandu, Nepal; Epidemiology and Biostatistics (Prof M Rezaei PhD), Department of Medical Biochemistry (M Sahebazzamani MSc), Rafsanjan University of Medical Sciences, Rafsanjan, Iran (A Rashidi MD); Department of Clinical Science (M Rashidi DVM), Islamic Azad University, Garmsar, Iran; Cellular and Molecular Research Center (M Rastegar PhD), Department of Epidemiology (A Tiyuri MSc), Birjand University of Medical Sciences, Birjand, Iran; University College London Hospitals, London, UK (D L Rawaf MD); Academic Public Health England (Prof S Rawaf MD), Public Health England, London, UK; Department of Immunology and Laboratory Sciences (M Razeghinia MSc), Medical Laboratory Sciences (M Sahebazzamani MSc), Sirjan School of Medical Sciences, Sirjan, Iran; Department of Immunology (M Razeghinia MSc), Kerman University of Medical Sciences, Kerman, Iran; School of Medicine (Prof A M N Renzaho PhD), Translational Health Research Institute (Prof A M N Renzaho PhD), Western Sydney University, Campbelltown, NSW, Australia; Network of Immunity in Infection, Malignancy and Autoimmunity (NIIMA) (Prof N Rezaei PhD), Universal Scientific Education and Research Network (USERN), Tehran, Iran; Eye and Skull Base Research Centers (S Rezaei MD), Rassoul Akram Hospital, Tehran, Iran; Department of Medical Pharmacology (M M Saber-Ayad MD), Cairo University, Giza, Egypt; Department of Gastroenterology (B Saberzadeh-Ardestani MD), Tehran University of Medical Sciences, Tehran, Iran; Research and Development (Prof U Saeed PhD), Islamabad Diagnostic Center Pakistan, Islamabad, Pakistan; Biological Production Division (Prof U Saeed PhD), National Institute of Health, Islamabad, Pakistan; Applied Biomedical Research Center (A Sahebkar PhD), Biotechnology Research Center (A Sahebkar PhD), Mashhad University of Medical Sciences, Mashhad, Iran; Department of Immunology (A Salek Farrokhi PhD), Semnan University of Medical Sciences and Health Services, Semnan, Iran; Research Center for Molecular Medicine (P Samadi PhD), Hamadan University of Medical Sciences, Hamadan, Iran; Emergency Department (M Samaei MD), Brown University, Providence, RI, USA; Department of Entomology (A M Samy PhD), Ain Shams University, Cairo, Egypt; Department of Surgery (Prof J Sanabria MD), Marshall University, Huntington, WV, USA; Research Development Coordination Section (M N S Saqib PhD), Pakistan Health Research Council, Islamabad, Pakistan; Faculty of Health & Social Sciences (B Sathian PhD), Bournemouth University, Bournemouth, UK; UGC Centre of Advanced Study in Psychology (M Satpathy PhD), Utkal University, Bhubaneswar, India; Udyam-Global Association for Sustainable Development, Bhubaneswar, India (M Satpathy PhD); Department of Health Sciences (I J C Schneider PhD), Federal University of Santa Catarina, Araranguá, Brazil; Department of Medical Statistics (M Šekerija PhD), University of Zagreb, Zagreb, Croatia; Department of Epidemiology and Prevention of Chronic

Non-communicable Diseases (M Šekerija PhD), Croatian Institute of Public Health, Zagreb, Croatia; National Heart, Lung, and Blood Institute (A Seylani BS), National Institute of Health, Rockville, MD, USA; Center for Biomedical Information Technology (F Sha PhD), Shenzhen Institutes of Advanced Technology, Shenzhen, China; Clinical Research Development Unit of Farshchian Heart Center (Z Shaghghi PhD), Hamedan University of Medical Sciences, Hamadan, Iran; Rajaei Cardiovascular Medical and Research Center (M Sharifian MD), Zabol University of Medical Sciences, Tehran, Iran; Department of Hematology-Oncology (S K Siddappa Malleshappa MD), Baystate Medical Center, Springfield, MA, USA; Department of Physical Education (Prof D A S Silva PhD), Federal University of Santa Catarina, Florianópolis, Brazil; Real World Insights (G Silva Julian MSc), IQVIA, São Paulo, Brazil; School of Medicine (Prof J A Singh MD), University of Alabama at Birmingham, Birmingham, AL, USA; Medicine Service (Prof J A Singh MD), US Department of Veterans Affairs (VA), Birmingham, AL, USA; Maternal and Child Health Division (M Siraj MSc), International Centre for Diarrhoeal Disease Research, Bangladesh, Dhaka, Bangladesh; Department No.16 (V Y Skryabin MD), Moscow Research and Practical Centre on Addictions, Moscow, Russia; Therapeutic Department (A A Skryabina MD), Balashiha Central Hospital, Balashikha, Russia; Surgery (B Socea PhD), “Sf. Pantelimon” Emergency Clinical Hospital Bucharest, Bucharest, Romania; Department of Neurosciences (M Solmi MD), University of Ottawa, Ottawa, ON, Canada; Taub Institute for Research on Alzheimer’s Disease and the Aging Brain (S Song PhD), Columbia University Medical Center, New York, NY, USA; Department of Medicine (Prof R Tabarés-Seisdedos PhD), University of Valencia, Valencia, Spain; Carlos III Health Institute (Prof R Tabarés-Seisdedos PhD), Biomedical Research Networking Center for Mental Health Network (CiberSAM), Madrid, Spain; Department of Surgery (K Tan PhD), National University of Singapore, Singapore, Singapore; Department of Cell Therapy and Applied Genomics (A Tbakhi MD), King Hussein Cancer Center, Amman, Jordan; Department of Nursing (B Tesfaye MSc), Debre Markos University, Debre Markos, Ethiopia; Department of Public Health and Medicine (D N Tollasa PhD), University of Newcastle, Newcastle, NSW, Australia; Nutritional Epidemiology Research Team (EREN) (M Touvier PhD), National Institute for Health and Medical Research (INSERM), Paris, France; Health, Medicine and Human Biology (M Touvier PhD), Sorbonne Paris Nord University, Bobigny, France; Department of Health Economics (B X Tran PhD), Hanoi Medical University, Hanoi, Vietnam; Department of Epidemiology and Biostatistics (B S Tusa MPH), Haramaya University, Haramaya, Ethiopia; Department of Allied Health Sciences (I Ullah PhD), Iqra National University, Peshawar, Pakistan; Pakistan Council for Science and Technology (I Ullah PhD), Ministry of Science and Technology, Islamabad, Pakistan; Clinical Cancer Research Center (S Valadan Tahbaz PhD, S Yahyazadeh Jabbari MD), Milad General Hospital, Tehran, Iran; Department of Microbiology (S Valadan Tahbaz PhD), Faculty of Medicine (M Zahir MD), Islamic Azad University, Tehran, Iran; Faculty of Information Technology (B Vo PhD), HUTECH University, Ho Chi Minh City, Vietnam; Department of Medicine (C Wang MPH), Vanderbilt University, Nashville, TN, USA; Competence Center of Mortality-Follow-Up of the German National Cohort (R Westerman DSc), Federal Institute for Population Research, Wiesbaden, Germany; Cancer Epidemiology and Prevention Research (L Yang PhD), Alberta Health Services, Calgary, BC, Canada; Department of Oncology (L Yang PhD), University of Calgary, Calgary, AB, Canada; Cancer Institute (D Yuce MD), Hacettepe University, Ankara, Turkey; Department of Clinical Pharmacy and Outcomes Sciences (I Yunusa PhD), University of South Carolina, Columbia, SC, USA; Epidemiology and Cancer Registry Sector (Prof V Zadnik PhD), Institute of Oncology Ljubljana, Ljubljana, Slovenia; Independent Consultant, Tehran, Iran (I Zare BSc).

# Contributors

The estimates of colorectal cancer were produced as part of GBD 2019. All the estimates are available in the public domain. Members of the core Institute for Health Metrics and Evaluation (IHME) had full access to the underlying data used to generate estimates presented in this Article. All other authors had access to and reviewed estimates as part of

the research evaluation process, which includes additional stages of internal IHME and external formal collaborator review. The corresponding author had final responsibility for the decision to submit the manuscript for publication. Contributions for all authors can be found in the appendix (p 1).

# Declaration of interests

R Ancuceanu reports consulting fees from AbbVie; payment or honoraria for lectures, presentations, speaker’s bureaus, manuscript writing, or educational events from AbbVie, Sandoz, and B Braun; all outside the submitted work. M Ausloos reports grants from the Romanian National Authority for Scientific Research and Innovation, CNDS-UEFISCDI, project number PN-III-P4-ID-PCCF-2016-0084 “Understanding and modelling time-space patterns of psychology-related inequalities and polarization” (October, 2018, to September, 2022), outside the submitted work. J Conde reports grants from the European Research Council Starting Grant (ERC-StG-2019-848325); patents planned, issued or pending for functionalised nanoparticles and compositions for cancer treatment and methods (US Application No. 62/334538), and TRPV2 Antagonists WO Application No. PCT/PT2018/050035; all outside the submitted work. I Fillip reports consulting fees from Avicenna Medical and Clinical Research Institute, outside the submitted work. N Ghith reports grants from Novo Nordisk Foundation as salary payment (NNF16OC0021856), outside the submitted work. A Guha reports grants from the American Heart Association as the Strategically Focused Research Network Grant in Disparities in Cardio-Oncology (#847740 and #863620), outside the submitted work. C Herteliu and A Pana reports grants or contracts from Romanian National Authority for Scientific Research and Innovation, CNDS-UEFISCDI, project number PN-III-P4-ID-PCCF-2016-0084 (October, 2018, to September, 2022) “Understanding and modelling time-space patterns of psychology-related inequalities and polarization” and Project number PN-III-P2-2-1-SOL-2020-2-0351 (June to October, 2020) “Approaches within public health management in the context of COVID-19 pandemic”, all outside the submitted work. C Herteliu reports grants from the Ministry of Labour and Social Justice, Romania, project number 30/PSCD/2018, “Agenda for skills Romania 2020–2025;” outside the submitted work. J Jozwiak reports payment or honoraria for lectures, presentations, speaker’s bureaus, manuscript writing, or educational events from Teva, Amgen, Synexus, Boehringer Ingelheim, Alab Laboratories, and Zentiva as personal fees, all outside the submitted work. J H Kaupila reports grants from the Sigrid Juselius Foundation and Finnish Cancer Foundation as research grants paid to their institutions; all outside the submitted work. J A Loureiro reports support for the present manuscript from Fundação para a Ciência e Tecnologia (FCT) as a salary payment under the Scientific Employment Stimulus (CEECINST/00049/2018) and from FCT/MCTES (Ministério da Ciência, Tecnologia e Ensino Superior) (PIDDAC) as Base Funding (UIDB/00511/2020 of LEPABE). A-F A Mentis reports grants or contracts from ELIDEK (Hellenic Foundation for Research and Innovation, MIMS-860) and EPANEK - MilkSafe (T2EAK-02222), all outside the submitted manuscript. O O Odukoya reports support for the present manuscript from the Fogarty International Center of the National Institutes of Health under the award number K43TW010704. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. A Radfar reports consulting fees from Avicenna Medical and Clinical Research Institute; leadership or fiduciary role in other board, society, committee, or advocacy group, paid or unpaid with MEDICHEM as a board member; all outside the submitted work. M Šekerija reports payment or honoraria for lectures, presentations, speaker’s bureaus, manuscript writing, or educational events from Roche and Johnson & Johnson, outside the submitted work. D A S Silva reports support for the present manuscript in part from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brazil (CAPES)—Finance Code 001 and in part by Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil (CNPq - 302028/2018-8), as payments made to their institution. J A Singh reports consulting fees from Crealta/Horizon, Medisys, Fidia, Two labs, Adept Field Solutions, Clinical Care Options, ClearView Healthcare Partners, Putnam Associates, Focus Forward,

Navigant Consulting, Spherix, MedIQ, UBM, Trio Health, Medscape, WebMD, and Practice Point Communications; and the National Institutes of Health and the American College of Rheumatology; payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing, or educational events from Simply Speaking; support for attending meetings or travel, or both from OMERACT, an international organisation that develops measures for clinical trials and receives arm's length funding from 12 pharmaceutical companies, when traveling to OMERACT meetings; participation on a data safety monitoring board or advisory board as a member of the FDA Arthritis Advisory Committee; leadership or fiduciary role in other board, society, committee or advocacy group, paid or unpaid, with OMERACT as a member of the steering committee, with the Veterans Affairs Rheumatology Field Advisory Committee as a member, and with the UAB Cochrane Musculoskeletal Group Satellite Center on Network Meta-analysis as a director and editor; stock or stock options in TPT Global Tech, Vaxart Pharmaceuticals and Charlotte's Web Holdings, and previously owned stock options in Amarin, Viking, and Moderna Pharmaceuticals; all outside the submitted work. M Solmi reports payment or honoraria for lectures, presentations, speaker's bureaus, manuscript writing, or educational events from Lundbeck; and participation on a data safety monitoring board or advisory board with Angelini; all outside the submitted work. T Vos reports support for the present manuscript from the Bill & Melinda Gates Foundation as payment to their institution. All other authors declare no competing interests.

#### Data sharing

To download the data used in these analyses, please visit the Global Health Data Exchange GBD 2019 results website.

#### Acknowledgments

This research was supported by funding from the Bill & Melinda Gates Foundation. J M Acuna acknowledges support from the Khalifa University College of Medicine and Health Sciences, Research and Data Intelligence Support Center, Khalifa University, Abu Dhabi, United Arab Emirates. S Ahmad thanks the Department of Health and Biological Sciences, Abasyn University, Peshawar, Pakistan, for their support. S M Aljunid acknowledges the Department of Health Policy and Management, Faculty of Public Health, Kuwait University, Kuwait, for approval and support to participate in this research project. F Carvalho and E Fernandes acknowledge support from Fundação para a Ciência e a Tecnologia, I.P. (FCT), in the scope of the project UIDP/04378/2020 and UIDB/04378/2020 of the Research Unit on Applied Molecular Biosciences UCIBIO and the project LA/P/0140/2020 of the Associate Laboratory Institute for Health and Bioeconomy i4HB; FCT/MCTES through the project UIDB/50006/2020. J Conde acknowledges the European Research Council Starting Grant (ERC-StG-2019-848325). V M Costa acknowledges the grant SFRH/BHD/110001/2015, received by Portuguese national funds through Fundação para a Ciência e Tecnologia (FCT), IP, under the Norma Transitória DL57/2016/CP1334/CT0006. N Ghith acknowledges support by a grant from Novo Nordisk Foundation (NNF16OC0021856) in the form of salary payments. J C Glasbey is supported by a National Institute of Health Research (NIHR) Academy Doctoral Research Fellow (NIHR300175). V K Gupta and V B Gupta acknowledge funding support from the National Health and Medical Research Council (NHMRC), Australia. S Haque acknowledges support from the Jazan University, Jazan, Saudi Arabia, for providing the access of the Saudi Digital Library for this research study. A Pana, M Ausloos, and C Herteliu are partially supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNDS-UEFISCDI, project number PN-III-P4-ID-PCCF-2016-0084. M Jakovljevic acknowledges support from the Ministry of Education Science and Technological Development of the Republic of Serbia through the Grant OI175014 and from the Science Fund of The Republic of Serbia through Grant Em-CEAS. I Landires is a member of the Sistema Nacional de Investigación (SNI), which is supported by Panama's Secretaría Nacional de Ciencia, Tecnología e Innovación (SENACYT). B Li was supported by the Guangdong Philosophy and Social Science Fund (GD21YSH06), the Shenzhen Philosophy and Social Science Fund (grant number SZ2020C015), and the Shenzhen Science and Technology Fund (grant number 20200805164059001). J A Loureiro was supported by

national funds through the Fundação para a Ciência e Tecnologia, under the Scientific Employment Stimulus (CEECINST/00049/2018).

V Nuñez-Samudio is a member of the Sistema Nacional de Investigación (SNI), which is supported by Panama's Secretaría Nacional de Ciencia, Tecnología e Innovación (SENACYT). O O Odukoya acknowledges support from the Fogarty International Center of the National Institutes of Health under the award number K43TW010704. The content is solely the responsibility of the authors and does not necessarily represent the official views of the US National Institutes of Health. A M Samy acknowledges the support from the Egyptian Fulbright Mission Program. D A S Silva acknowledges support in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brazil (CAPES)—Finance Code 001 and in part by Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil (CNPq - 302028/2018-8).

Editorial note: the *Lancet* Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

#### References

- 1 GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020; **396**: 1204–22.
- 2 Lynch HT, de la Chapelle A. Hereditary colorectal cancer. *N Engl J Med* 2003; **348**: 919–32.
- 3 Graff RE, Möller S, Passarelli MN, et al. Familial risk and heritability of colorectal cancer in the nordic twin study of cancer. *Clin Gastroenterol Hepatol* 2017; **15**: 1256–64.
- 4 Anand P, Kunnumakara AB, Sundaram C, et al. Cancer is a preventable disease that requires major lifestyle changes. *Pharm Res* 2008; **25**: 2097–116.
- 5 Jaspersion KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. *Gastroenterology* 2010; **138**: 2044–58.
- 6 Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. *Gut* 2017; **66**: 683–91.
- 7 GBD 2017 Colorectal Cancer Collaborators. The global, regional, and national burden of colorectal cancer and its attributable risk factors in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* 2019; **4**: 913–33.
- 8 UN Department of Economic and Social Affairs. United Nations Sustainable Development Goals. Goal 3: ensure healthy lives and promote well-being for all at all ages. <https://www.un.org/sustainabledevelopment/health/> (accessed March 22, 2022).
- 9 Bhat SK, East JE. Colorectal cancer: prevention and early diagnosis. *Medicine* 2015; **43**: 295–98.
- 10 Wang X, O'Connell K, Jeon J, et al. Combined effect of modifiable and non-modifiable risk factors for colorectal cancer risk in a pooled analysis of 11 population-based studies. *BMJ Open Gastroenterol* 2019; **6**: e000339.
- 11 Frampton M, Houlston RS. Modeling the prevention of colorectal cancer from the combined impact of host and behavioral risk factors. *Genet Med* 2017; **19**: 314–21.
- 12 GBD 2019 Demographics Collaborators. Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950–2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020; **396**: 1160–203.
- 13 GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 2020; **396**: 1223–49.
- 14 Murray CJ, Abbafati C, Abbas KM, et al. Five insights from the Global Burden of Disease Study 2019. *Lancet* 2020; **396**: 1135–59.
- 15 Global Burden of Disease 2019 Cancer Collaboration. Cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life years for 29 cancer groups from 2010 to 2019: a systematic analysis for the Global Burden of Disease Study 2019. *JAMA Oncol* 2022; **8**: 420–44.
- 16 GBD 2016 Healthcare Access and Quality Collaborators. Measuring performance on the Healthcare Access and Quality Index for 195 countries and territories and selected subnational locations: a systematic analysis from the Global Burden of Disease Study 2016. *Lancet* 2018; **391**: 2236–71.

- 17 Murray CJ. Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bull World Health Organ* 1994; **72**: 429–45.
- 18 World Cancer Research Fund/American Institute for Cancer Research. Food, nutrition, physical activity, and cancer: a global perspective. A summary of the third expert report. Continuous Update Project Expert Report 2018. <https://www.wcrf.org/wp-content/uploads/2021/02/Summary-of-Third-Expert-Report-2018.pdf> (accessed Feb 28, 2022).
- 19 WHO. WHO report on the global tobacco epidemic 2017: monitoring tobacco use and prevention policies. 2017. <https://apps.who.int/iris/handle/10665/255874> (accessed Sept 29, 2020).
- 20 Stevens GA, Singh GM, Lu Y, et al. National, regional, and global trends in adult overweight and obesity prevalences. *Popul Health Metr* 2012; **10**: 22.
- 21 Schreuders EH, Ruco A, Rabeneck L, et al. Colorectal cancer screening: a global overview of existing programmes. *Gut* 2015; **64**: 1637–49.
- 22 Zauber AG. The impact of screening on colorectal cancer mortality and incidence: has it really made a difference? *Dig Dis Sci* 2015; **60**: 681–91.
- 23 DeCosse JJ, Ngoi SS, Jacobson JS, Cennerazzo WJ. Gender and colorectal cancer. *Eur J Cancer Prev* 1993; **2**: 105–15.
- 24 Karastergiou K, Smith SR, Greenberg AS, Fried SK. Sex differences in human adipose tissues - the biology of pear shape. *Biol Sex Differ* 2012; **3**: 13.
- 25 Wilsnack RW, Wilsnack SC, Kristjanson AF, Vogeltanz-Holm ND, Gmel G. Gender and alcohol consumption: patterns from the multinational GENACIS project. *Addiction* 2009; **104**: 1487–500.
- 26 Murphy N, Strickler HD, Stanczyk FZ, et al. A prospective evaluation of endogenous sex hormone levels and colorectal cancer risk in postmenopausal women. *J Natl Cancer Inst* 2015; **107**: 107.
- 27 Luan NN, Wu L, Gong TT, Wang YL, Lin B, Wu QJ. Nonlinear reduction in risk for colorectal cancer by oral contraceptive use: a meta-analysis of epidemiological studies. *Cancer Causes Control* 2015; **26**: 65–78.
- 28 Yang Y, Wang G, He J, et al. Gender differences in colorectal cancer survival: a meta-analysis. *Int J Cancer* 2017; **141**: 1942–49.
- 29 Hendifar A, Yang D, Lenz F, et al. Gender disparities in metastatic colorectal cancer survival. *Clin Cancer Res* 2009; **15**: 6391–97.
- 30 Giacchetti S, Dugué PA, Innominato PF, et al. Sex moderates circadian chemotherapy effects on survival of patients with metastatic colorectal cancer: a meta-analysis. *Ann Oncol* 2012; **23**: 3110–16.
- 31 Kim SE, Paik HY, Yoon H, Lee JE, Kim N, Sung MK. Sex- and gender-specific disparities in colorectal cancer risk. *World J Gastroenterol* 2015; **21**: 5167–75.
- 32 Siegel RL, Torre LA, Soerjomataram I, et al. Global patterns and trends in colorectal cancer incidence in young adults. *Gut* 2019; **68**: 2179–85.
- 33 Vuik FE, Nieuwenburg SA, Bardou M, et al. Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. *Gut* 2019; **68**: 1820–26.
- 34 Lui RN, Tsoi KKF, Ho JMW, et al. Global increasing incidence of young-onset colorectal cancer across 5 continents: a joinpoint regression analysis of 1,922,167 cases. *Cancer Epidemiol Biomarkers Prev* 2019; **28**: 1275–82.
- 35 Araghi M, Soerjomataram I, Bardot A, et al. Changes in colorectal cancer incidence in seven high-income countries: a population-based study. *Lancet Gastroenterol Hepatol* 2019; **4**: 511–18.
- 36 Wolf AMD, Fontham ETH, Church TR, et al. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. *CA Cancer J Clin* 2018; **68**: 250–81.
- 37 Bailey CE, Hu CY, You YN, et al. Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975–2010. *JAMA Surg* 2015; **150**: 17–22.
- 38 Cavestro GM, Mannucci A, Zuppardo RA, Di Leo M, Stoffel E, Tonon G. Early onset sporadic colorectal cancer: Worrisome trends and oncogenic features. *Dig Liver Dis* 2018; **50**: 521–32.
- 39 Keum N, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat Rev Gastroenterol Hepatol* 2019; **16**: 713–32.
- 40 Young JP, Win AK, Rosty C, et al. Rising incidence of early-onset colorectal cancer in Australia over two decades: report and review. *J Gastroenterol Hepatol* 2015; **30**: 6–13.
- 41 Siegel RL, Fedewa SA, Anderson WF, et al. Colorectal cancer incidence patterns in the United States, 1974–2013. *J Natl Cancer Inst* 2017; **109**: djw322.
- 42 Hur J, Smith-Warner SA, Rimm EB, et al. Alcohol intake in early adulthood and risk of colorectal cancer: three large prospective cohort studies of men and women in the United States. *Eur J Epidemiol* 2021; **36**: 325–33.
- 43 Nguyen LH, Liu PH, Zheng X, et al. Sedentary behaviors, TV viewing time, and risk of young-onset colorectal cancer. *JNCI Cancer Spectr* 2018; **2**: pky073.
- 44 Liu P-H, Wu K, Ng K, et al. Association of obesity with risk of early-onset colorectal cancer among women. *JAMA Oncol* 2019; **5**: 37–44.
- 45 Patel P, De P. Trends in colorectal cancer incidence and related lifestyle risk factors in 15–49-year-olds in Canada, 1969–2010. *Cancer Epidemiol* 2016; **42**: 90–100.
- 46 Nimptsch K, Wu K. Is timing important? the role of diet and lifestyle during early life on colorectal neoplasia. *Curr Colorectal Cancer Rep* 2018; **14**: 1–11.
- 47 Yue Y, Hur J, Cao Y, et al. Prospective evaluation of dietary and lifestyle pattern indices with risk of colorectal cancer in a cohort of younger women. *Ann Oncol* 2021; **32**: 778–86.
- 48 Crosbie AB, Roche LM, Johnson LM, Pawlish KS, Paddock LE, Stroup AM. Trends in colorectal cancer incidence among younger adults—disparities by age, sex, race, ethnicity, and subsite. *Cancer Med* 2018; **7**: 4077–86.
- 49 Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; **71**: 209–49.



**ACADEMIA**

Accelerating the world's research.

# Inverse Analog Filters: History, Progress and Unresolved Issues

Raj Senani


*Electronics*

## Cite this paper

Downloaded from [Academia.edu](#) 

[Get the citation in MLA, APA, or Chicago styles](#)

## Related papers

[Download a PDF Pack](#) of the best related papers 

# Inverse Analog Filters: History, Progress and Unresolved Issues

Raj Senani <sup>1,\*</sup> , Data Ram Bhaskar <sup>2</sup>  and Ajishek Raj <sup>2</sup> 

<sup>1</sup> Analog Signal Processing Research Lab., Department of Electronics and Communication Engineering, Netaji Subhas University of Technology, New Delhi 110078, India

<sup>2</sup> Department of Electrical Engineering, Delhi Technological University, Delhi 110042, India; drbhaskar@dtu.ac.in (D.R.B.); ajishek\_phd2k16@dtu.ac.in (A.R.)

\* Correspondence: senani@nsut.ac.in

**Abstract:** This paper traces the history of the evolution of inverse analog filters (IAF) and presents a review of the progress made in this area to date. The paper, thus, presents the current state-of-the-art of IAFs by providing an appraisal of a variety of realizations of IAFs using commercially available active building blocks (ABB), such as operational amplifiers (Op-amp), operational transconductance amplifiers (OTA), current conveyors (CC) and current feedback operational amplifiers (CFOA) as well as those realized with newer active building blocks of more recent origin, such as operational transresistance amplifiers (OTRA), current differencing buffered amplifiers (CDBA) and variants of current conveyors which, although not available as off-the-shelf ICs yet, can be implemented as complementary metal–oxide–semiconductors (CMOS) or be realized in discrete form using other commercially available integrated circuits (IC). In the end, some issues related to IAFs have been highlighted which need further investigation.

**Keywords:** inverse analog filters; inverse high pass filter; inverse low pass filter; inverse band pass filter; inverse notch filter; inverse all pass filter



**Citation:** Senani, R.; Bhaskar, D.R.; Raj, A. Inverse Analog Filters: History, Progress and Unresolved Issues. *Electronics* **2022**, *11*, 841. <https://doi.org/10.3390/electronics11060841>

Academic Editors: Paolo Colantonio and Alessandro Cidronali

Received: 29 January 2022

Accepted: 2 March 2022

Published: 8 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Over the last twenty-five years, there has been considerable interest and research effort in the technical literature on the circuit realizations of *inverse analog filters*. In many areas such as communication, control and instrumentation, inverse filters *may be* required to correct the distortions of the signals caused by the signal processors. Since inverse filters have a frequency response which is reciprocal of the frequency response of the system which caused the distortions, it is expected that this can be corrected by using an inverse filter. A survey of the work done on the realization of inverse filters indicates that before the publication of the 1997 paper of Adrian Leuciuc [1], although there had been mention of digital inverse filters from time-to-time [2–6], only the 1964 paper of Burch, Green and Grote [6] had discussed earlier about the restoration and correction of time functions by the synthesis of inverse filters on analog computers. Thus, with the sole exception of [6] and before the 1997 work of Leuciuc [1], no procedure or circuit appears to have been presented in the open technical literature to realize any kind of inverse *analog* filters using the now prevalent analog circuit building blocks.

However, subsequent to the publication of [1], a number of authors have come up with various kinds of circuits for realizing IAFs [7–43] using a variety of active building blocks such as the four-terminal floating nullors (FTFN) [44,45], operational transconductance amplifiers (OTA) [46], current conveyors [47,48] and their many variants such as differential difference current conveyors (DDCC), second generation voltage conveyors (VCII), current feedback operational amplifiers (CFOA) [49], operational transresistance amplifiers (OTRA) [50], current differencing buffered amplifiers (CDBA) [51], current differencing transconductance amplifiers (CDTA) [52], voltage differencing transconductance amplifiers (VDTA) [53], etc. This paper gives an account of the progress made on the evolution of IAFs so far and points out some unresolved problems and issues related to them.



## 2. The Transfer Functions of Inverse Filters, Their Frequency Responses, the Significance of the Various Parameters and the Stability Issues

In this section, we discuss the frequency response of the various types of inverse filters, and attempt to provide the significance of the various parameters  $H_0$ ,  $\omega_0$  and  $Q_0$  on the shape of the frequency responses. We also comment upon the stability issues related to the various types of inverse filters.

### 2.1. Significance of the Coefficients of the Inverse Transfer Functions in the Characterization of Their Frequency Responses

The physical meaning of the parameters  $H_0$ ,  $Q_0$  and  $\omega_0$  is clear in the context of normal low-pass (LP) and high-pass (HP) filters, and that of  $H_0$ ,  $\omega_0/Q_0$  (BW) and  $\omega_0$  is also clear in the case of normal band-pass (BP) and band reject (BR) filters. However, so far, the significance of the various coefficients of the inverse transfer functions and their influence on the shapes of the corresponding frequency responses does not appear to have been elaborated explicitly in the earlier literature. In the following, we attempt to provide some insight into this.

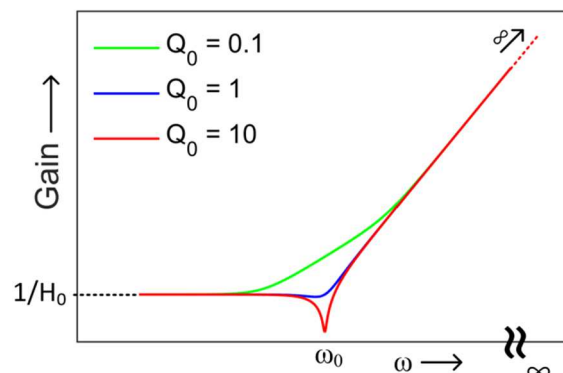
Let us consider an inverse low pass filter (ILPF) transfer function, which can be written as:

$$T(s)|_{\text{ILPF}} = \frac{s^2 + s\left(\frac{\omega_0}{Q_0}\right) + \omega_0^2}{H_0\omega_0^2} \quad (1)$$

From this transfer function, it can be verified that  $|T(j\omega)|$  at zero frequency (DC) turns out to be  $1/H_0$ ; at  $\omega = \infty$ ,  $|T(j\omega)|$  becomes  $\infty$ , and at a frequency  $\omega = \omega_0$ ,  $|T(j\omega)|$  is given by:

$$|T(s)|_{\omega=\omega_0} = \frac{1}{H_0Q_0} \quad (2)$$

Therefore, ' $\omega_0$ ' can be called the *corner frequency*; however, depending upon whether  $Q_0 = 1$ ,  $Q_0 > 1$  or  $Q_0 < 1$ , the resulting responses turn out to be as shown in Figure 1.



**Figure 1.** Frequency responses of ILPF for different values of  $Q_0$ .

Hence, the significance of the coefficients  $H_0$ ,  $Q_0$  and  $\omega_0$  is clear from this plot.

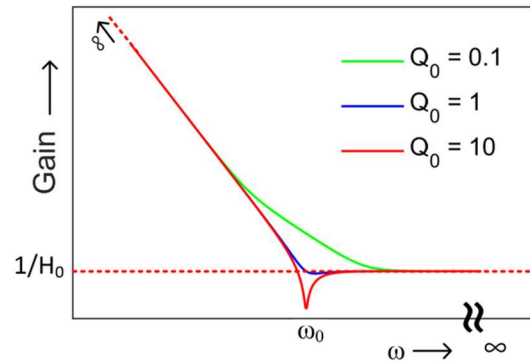
In a similar manner, the transfer function of inverse high pass filter (IHPF) may be expressed as:

$$T(s)|_{\text{IHPF}} = \frac{s^2 + s\left(\frac{\omega_0}{Q_0}\right) + \omega_0^2}{H_0s^2} \quad (3)$$

For the IHPF,  $|T(j\omega)|$  at  $\omega = 0$  turns out to be  $\infty$ ; at  $\omega = \infty$ ,  $|T(j\omega)|$  becomes 0 and at a frequency  $\omega = \omega_0$ ,  $|T(j\omega)|$  is given by:

$$|T(s)|_{\omega=\omega_0} = \frac{1}{H_0Q_0} \quad (4)$$

Therefore, in the case of IHPF also,  $\omega_0$  is the *corner frequency* and depending upon whether  $Q_0 = 1$ ,  $Q_0 > 1$  or  $Q_0$  less than 1, the resulting responses would be as displayed in Figure 2.



**Figure 2.** Frequency responses of IHPF for different values of  $Q_0$ .

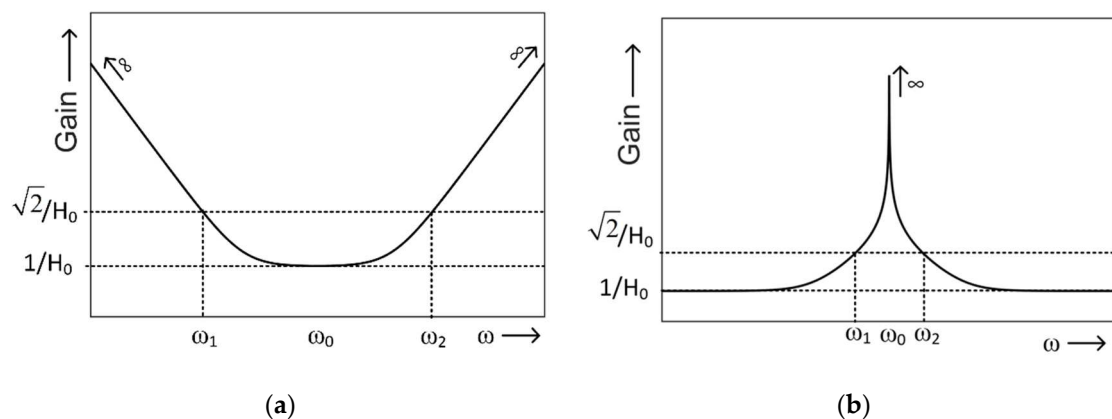
The transfer function of inverse band pass filter (IBPF) may be expressed as:

$$T(s)|_{\text{IBPF}} = \frac{s^2 + s\left(\frac{\omega_0}{Q_0}\right) + \omega_0^2}{H_0\left(\frac{\omega_0}{Q_0}\right)s} \quad (5)$$

The  $|T(j\omega)|_{\text{IBPF}}$  can be written as:

$$|T(j\omega)|_{\text{IBPF}} = \frac{\sqrt{(\omega_0^2 - \omega^2)^2 + (\omega\omega_0/Q_0)^2}}{H_0(\omega\omega_0/Q_0)} \quad (6)$$

From Equation (6), it may be seen that, for  $\omega = 0$  and  $\omega = \infty$  in both cases,  $|T(j\omega)|$  is infinite and at  $\omega = \omega_0$ ,  $|T(j\omega)|$  turns out to be  $1/H_0$ . The frequency response for IBPF has been shown in Figure 3a. Here,  $\omega_0$  is the *center frequency*.



**Figure 3.** Frequency response of (a) IBPF and (b) IBRF.

The transfer function of the IBRF can be written as:

$$T(s)|_{\text{IBRF}} = \frac{s^2 + s\left(\frac{\omega_0}{Q_0}\right) + \omega_0^2}{H_0(s^2 + \omega_0^2)} \quad (7)$$

The magnitude of the function in Equation (7) at  $s = j\omega$  can be written as:

$$|T(j\omega)|_{\text{IBRF}} = \frac{\sqrt{(\omega_0^2 - \omega^2)^2 + (\omega\omega_0/Q_0)^2}}{H_0(\omega_0^2 - \omega^2)} \quad (8)$$

From Equation (8), it can be seen that  $|T(j\omega)|$  of the notch filter at  $\omega = 0$  as well as  $\omega = \infty$ , is  $1/H_0$  whereas, at  $\omega = \omega_0$ ,  $|T(j\omega)|$  tends to infinity. Here also,  $\omega_0$  is the *center frequency* and the shape of response is as shown in Figure 3b.

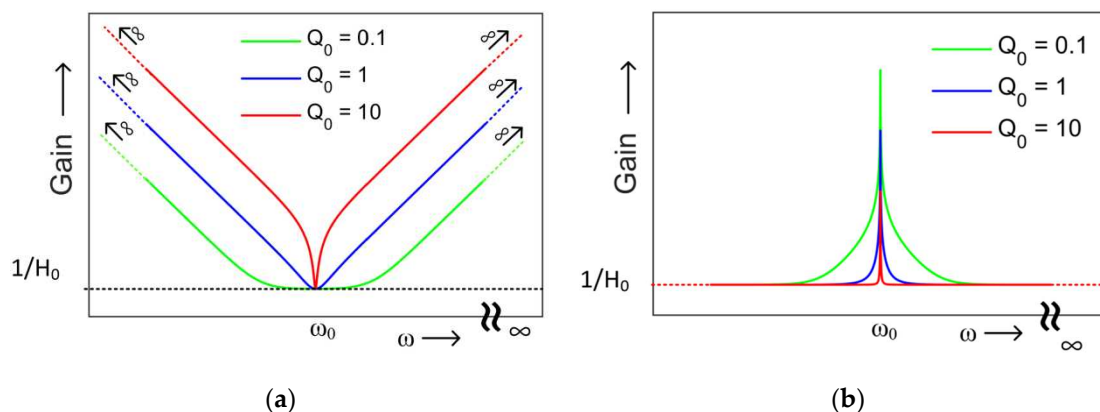
In case of IBPF as well as IBRF, a narrowing down (sharpness) of the shape of the frequency response is observed when  $Q_0$  is increased. For a quantitative interpretation of this, if we take  $\omega_0 = 1$  (normalized value) and determine the two frequencies ( $\omega_1$  and  $\omega_2$ ), on either side of the center frequency, at which the gain becomes  $\frac{\sqrt{2}}{H_0}$  (arbitrarily chosen to facilitate easy solution of the resulting quadratic equation) can be found from Equations (6) and (8) and turn out to be:

$$\omega_1 = \left[ \frac{1}{2} \left( 2 + \frac{1}{Q_0^2} \right) - \frac{1}{2Q_0} \sqrt{4 + \frac{1}{Q_0^2}} \right]^{\frac{1}{2}} \text{ and } \omega_2 = \left[ \frac{1}{2} \left( 2 + \frac{1}{Q_0^2} \right) + \frac{1}{2Q_0} \sqrt{4 + \frac{1}{Q_0^2}} \right]^{\frac{1}{2}}$$

from which it can be seen that the *width* of the IBP and IBR plots between these two frequencies is given by:

$$\omega_2 - \omega_1 = \frac{1}{Q_0} \sqrt{1 + \frac{1}{2Q_0^2}} \quad (9)$$

From Equation (9), it is clear that, the larger the value of  $Q_0$ , the narrower the *width* of the resulting frequency response at the gain value of  $\frac{\sqrt{2}}{H_0}$ , which is clearly visible from the plots of Figure 4a,b. On the contrary, for values of  $Q_0 < 1$ , the nature of the plot would be wider on either side of  $\omega_0$ . The frequency responses of IBPF and IBRF for different values of  $Q_0$  shown in Figure 4a,b confirm this interpretation.



**Figure 4.** Frequency responses of (a) IBPF and (b) IBRF for different values of  $Q_0$ .

Lastly, we do not deal with inverse all-pass filters (IAPF), since in our opinion, IAPF is a misnomer due to being an unstable system owing to having complex conjugate poles in the right half of the  $s$ -plane.

## 2.2. Stability Considerations

On the stability of inverse analog filters: Another potential problem which does not appear to have been noticed explicitly in the existing literature is concerned with the stability aspect of inverse filters. In case of the first order inverse filters, the ILP and IHP do not have stability issues; however, the inverse all pass filter transfer function clearly has its pole in the right half of the  $s$ -plane and, therefore, represents an

unstable system. Coming to second order inverse filters, ILP and IBP do not have stability problems, but in the case of IHP, there is pair of poles at its origin, the IBR has a pair of poles on the  $\pm j\omega$  axis and the inverse all pass (IAP) has clearly both poles in the right half of the s-plane. Thus, theoretically, all these cases represent unstable systems. Now, the fact that a number of authors have come up with first-order inverse all-pass filter [10,19,33], second order IHP [1,9,11–18,20,22,23,27,28,30,32,35,36,39–43], IBR [11,12,15,17,22,24,25,27,29,32,35,36] and IAP [8,9,24,27,29] filters, it is a matter of curiosity as to how the observance of instability was not reported by any of them. There could be two possible reasons for this:

- (i) As is well known, if the verification of the workability of the circuit is carried out only by frequency response analysis in SPICE without checking the transient response of the circuits, they may still exhibit the correct frequency responses and, thus, the circuit instability might go unnoticed in the absence of looking into its transient analysis.
- (ii) In those cases where experimental results of the frequency responses of such circuits are taken and the circuits are shown to work properly, the possible reason could be that the various parasitics of the active element used or their non-ideal parameters might have shifted the pole locations slightly to the left half of the s-plane (which appear feasible in case of IHP or IBR but unlikely in the case of inverse all-pass filters) and therefore, any instability could not have been observed.
- (iii) Of course, a stability issue of fractional order inverse filters (FOIFs) together with a check of their transient response is similarly necessary. In fact, Bhaskar et al. in [40] and Radwan et al. [41] have already taken due cognizance of this aspect.

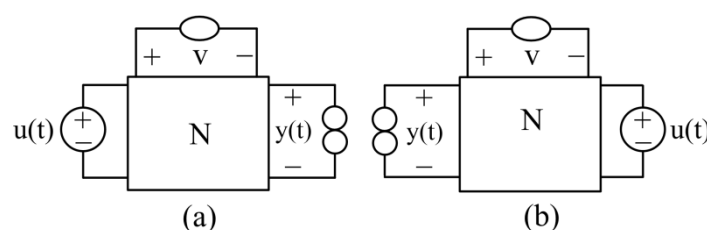
Thus, the stability issues related to the IHP, IBR and IAP call for further investigations in future.

### 3. Circuit Realizations of the Various Inverse Analog Filters

In this section, we outline some prominent inverse analog filter configurations using different kinds of analog circuit building blocks proposed during the last two decades or more.

#### 3.1. IAF Configurations Using Op-Amps and FTFNs

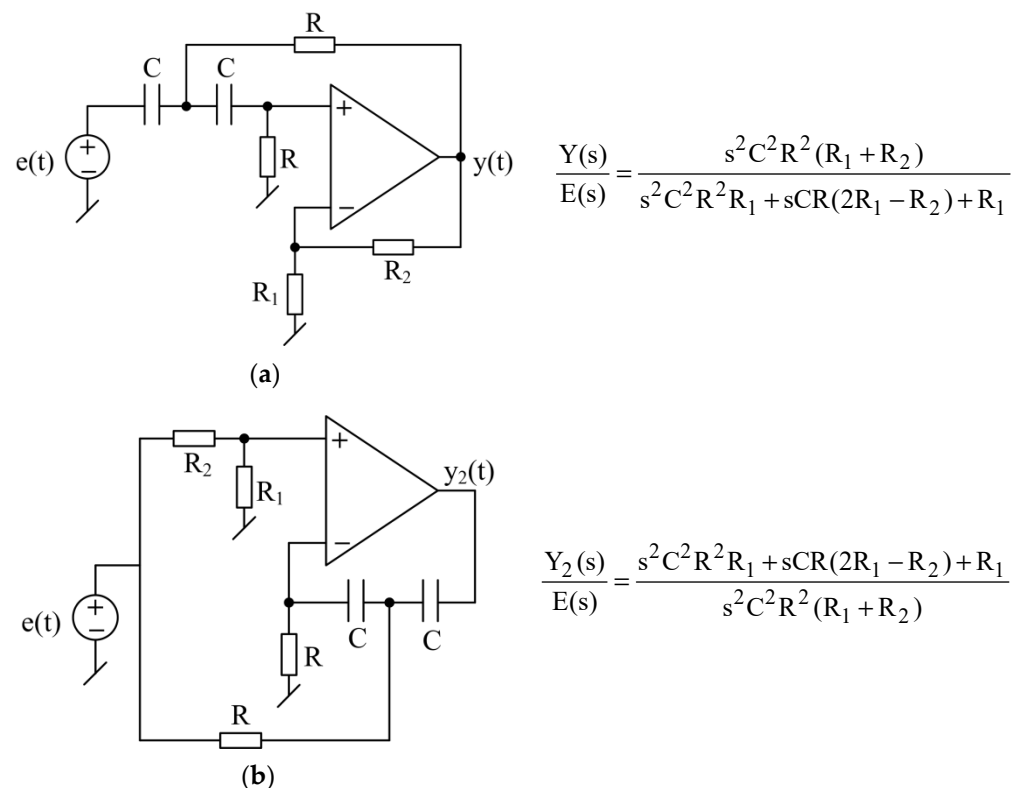
Leuciuc [1] appears to be the first to have presented a general methodology of finding the inverse of a given analog circuit based upon the use of *nullator* and *norator* which are two well-known pathological elements proposed by Carlin and Youla [54]. It may be recalled that the *nullator* is a one port (two-terminal element) characterized by the equations  $V = 0$ , and  $I = 0$ ; on the other hand, the *norator* is another one port (two-terminal element) characterized by  $V = \text{arbitrary}$  and  $I = \text{arbitrary}$ . The methodology proposed by Leuciuc [1] is depicted in Figure 5, according to which a network 'N' containing a nullator and norator and realizing a transfer function  $Y(s)/U(s) = T(s)$  will be converted into another network realizing the transfer function  $T(s)^{-1}$  if the driving source and norator are swapped.



**Figure 5.** Original circuit and its inverse obtained by swapping the driving source and the norator (a) original circuit (b) Inverse circuit.

Since a pair of a nullator and a norator (with one end of the norator connected to ground) represents an ideal op-amp [55], this idea is readily applicable in obtaining inverse of any given op-amp filter. Leuciuc [1], using this idea, demonstrated how the well-known

single op-amp-based Sallen-Key high-pass filter can be used to realize an inverse high-pass filter as shown here in Figure 6.



**Figure 6.** Sallen-Key high pass filter and its inverse by Leuciuc [1] (a) original filter (b) inverse filter.

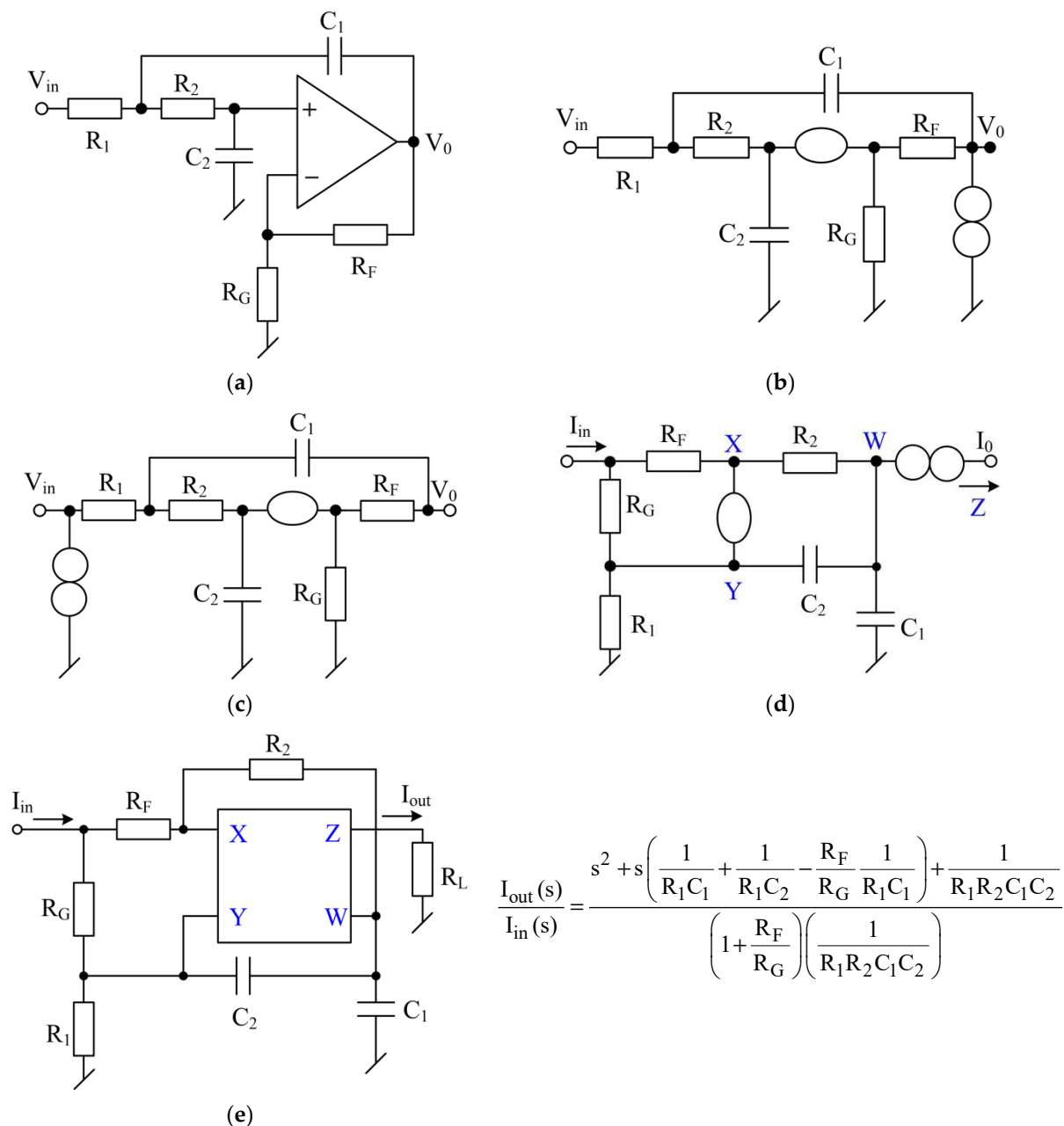
It is interesting to point out that the transformation of an op-amp circuit realizing a transfer function  $T(s)$  into another op-amp circuit realizing transfer function  $T(s)^{-1}$  can also be related to the *inverse active network theorem* (IANT) proposed by Rathore in [56] and Rathore and Singhi in [57]. As per IANT, if there is an op-amp RC network ‘N’ realizing a transfer function  $T(s)$ , then this can be transformed into a network realizing the reciprocal transfer function  $1/T(s)$  by bringing the connection of all elements going to the input signal to the output of the op-amp and connecting all elements connected to the output node to the input terminal, then assigning the correct polarities of the input terminals of the op-amp from the considerations of the stability of the resulting circuit. The circuit in Figure 6b can be considered to be obtainable from that of Figure 6a using IANT. Since Leuciuc did not demonstrate the practical results of his inverse filter of Figure 6b starting from zero frequency, it went unnoticed that this circuit cannot be functional because of the absence of a DC path in the negative feedback circuit.

In [37], a number of single-op-amp-based inverse filter circuits have been discussed.

### 3.2. IAF Configurations Using Four-Terminal-Floating-Nullors (FTFN)

This was followed by another procedure reported by Chipipop and Surakamponorn in [7] for the systematic realization of a current-mode four-terminal floating nullor (FTFN)-based inverse filter as shown in Figure 7 (the acronym ‘FTFN’ to represent a fully floating nullor was coined for the first time in [44,45]). In this procedure, starting from an op-amp circuit, a nullor-RC model is obtained by replacing the ideal op-amp with a pair of a nullator and norator (having one terminal connected to ground). Subsequently, the driving source and norator are swapped as per the transformations suggested by Leuciuc [1] to obtain an inverse filter in voltage mode. This is followed by an RC:CR duality transformation [58] from which an FTFN-based inverse filter in current-mode is obtained, which is converted into a physical realization by replacing the nullator and norator pair by an FTFN. In the

work of Chipipop and Surakampontrorn [7] the FTFN was realized using an operational-mirrored amplifier (OMA) formulation [59] using AD704 IC op-amp, Q2N2222A NPN transistors and Q2N2907A PNP transistors. Since the FTFN used for the simulation results does not have symmetrical/identical W and Z terminals, the simulation results given in [7] do not ensure that the derived inverse filter will still work satisfactorily in practice when the FTFN is realized by symmetrical structures such as those made from two CCII± or two CFOAs.



**Figure 7.** The procedure for realization of FTFN-based inverse filter in CM [7] (a) a single-op-amp filter (b) nullor equivalent (c) inverse function realization in voltage mode (VM) by swapping the voltage source by nullor (d) the RC:CR dual to realize the same function in CM as a current-ratio transfer function (e) the FTFN implementation of the inverse filter.

Wang and Lee [8] extended this approach and derived a current-mode IAF using FTFN corresponding to the Friend and Delyiannis [60] op-amp-RC biquad and also demonstrated how the same methodology can be applied to derive an FTFN-based current mode inverse



filter starting from a CCII-based VM filter, since a CCII- can also be represented by a three-terminal nullor (with the nullator and norator sharing a common terminal which represents the X-terminal of the CCII-, while the free end of the nullator represents the Y-terminal of the CCII- and the free end of the norator represents the terminal-Z). Wang and Lee in [8] have also shown the transformation of a high-input impedance filter into a current-mode inverse filter. In both cases, the verification of the workability of the circuits was demonstrated by only SPICE simulation results although it was not clarified as to which FTFN implementation was used.

Taking clues from the above mentioned FTFN-based approach, Abuelma'atti in [9] formulated a generalized single-FTFN network for realizing a CM transformation using as many as eleven admittances in the circuit and worked out a number of second-order inverse filters in CM as special cases therefrom. The simulation results of these inverse filters were obtained using two-CCII-based realization of the FTFN which was acknowledged in [9] to have been first outlined by Senani in [45], which is shown here in Figure 8. All the circuits employ a low number of passive components and their workability has also been demonstrated by a symmetrical FTFN structure; however, some of the circuits derived therein suffer from the drawback of being non-canonic.

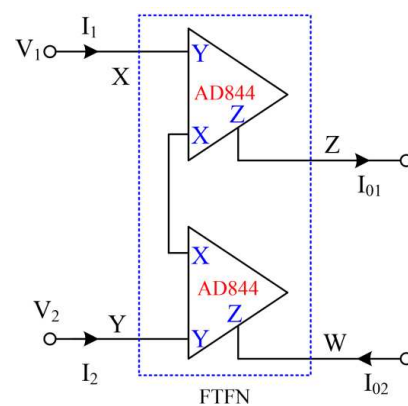


Figure 8. FTFN realization using two CCII [9,45].

Shah and Malik [10] proposed an all-pass inverse filter configuration with two inputs and a single output wherein the FTFN was again implemented using the two-AD844-based implementation, shown here in Figure 8. However, the authors have failed to notice that the first order IAP clearly has the pole located in the right half of the s-plane, and therefore, any circuit implementation of this will be unstable. Therefore, it would fail in the transient response test even though it might show correct frequency response in simulations.

### 3.3. IAF Configurations Using Current Feedback Operational Amplifiers (CFOA)

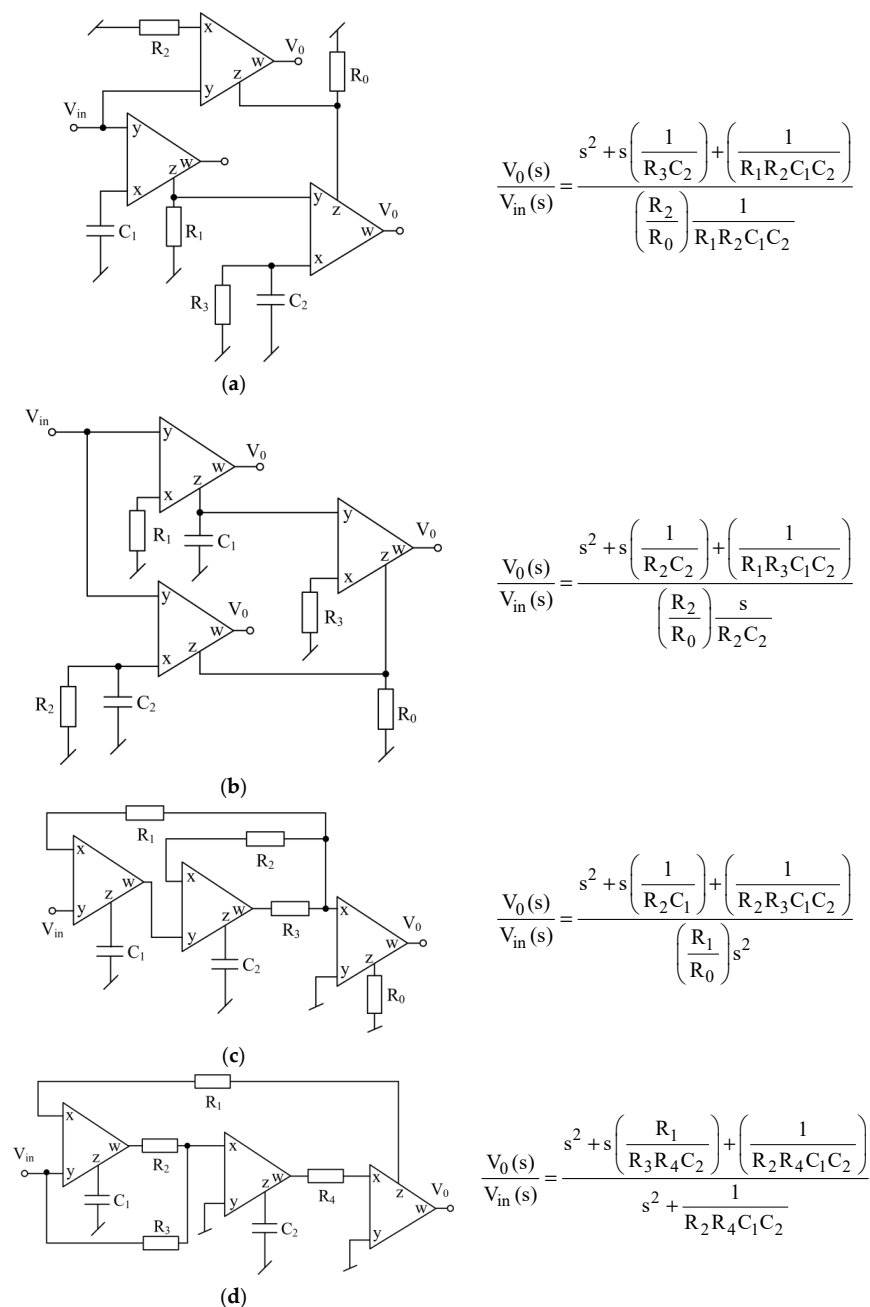
A CFOA is a four terminal ( $x$ ,  $y$ ,  $z$  and  $w$ ) active building blocks characterized by the following terminal equations:

$$i_y = 0, v_x = v_y, i_z = i_x \text{ and } v_w = v_z \quad (10)$$

Since there has been considerable interest in the use of CFOAs as active building blocks because of their several advantages over the traditional voltage mode op-amps (VOA), such as very high slew rates leading to a higher operational frequency range, elimination of the gain-bandwidth conflict and the feasibility of realizing any given function with the least possible number of passive components, most of the time, without requiring any component-matching conditions/equality constraints [49,61–65]. Due to these reasons, a number of researchers have formulated various kinds of IAF circuits too, using CFOAs.

In this context, Gupta, Bhaskar, Senani and Singh [11] were the first to have proposed four inverse filter configurations, each implementable from only three CFOAs for realizing inverse low-pass, inverse band-pass, inverse high-pass and inverse band

reject filters respectively. These realizations have been depicted here in Figure 9. The practically observed frequency responses of these circuits were also shown in [11]. All the circuits possessed the important and practically useful properties of infinite input impedance, zero output impedance and the use of two grounded capacitors as preferred for IC implementation [66–69]. The inverse low pass and band pass filters of Figure 9a,b have the novel feature of having all four resistors being grounded too, which is attractive for IC implementation as well as converting the circuits into voltage-controlled circuits by replacing the grounded resistors with FET/MOSFET-based linear voltage-controlled resistances.

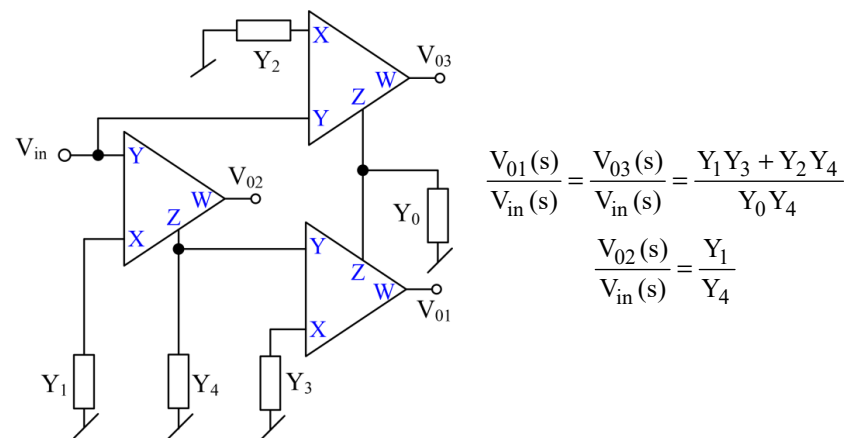


**Figure 9.** Proposed inverse filter configurations reported by Gupta, Bhaskar, Senani and Singh using CFOAs (a) ILP (b) IBP (c) IHP (d) IBR [11].

In [12], Gupta, Bhaskar and Senani further extended their work by presenting one new implementation of the IHP filter, four new circuits of the ILP filter and two new

configurations of the IBP filter, each employing three CFOAs, two grounded capacitors and three to five resistors along with the provision of independent controllability of the various coefficients of the realized transfer functions. These reported circuits have also been experimentally verified using AD844 type CFOA ICs in the frequency range of 100 Hz–1 MHz.

Subsequently, Wang, Chang, Yang and Tsai [13] proposed a general inverse filter scheme using CFOAs as shown in Figure 10. Curiously, however, the authors of [13] did cite the quoted work as a reference in their paper, but while projecting their general scheme of Figure 10, they did not spell it out *explicitly* that their configuration is, in fact, an adaptation of the two circuits of Figure 9a,b. It was demonstrated by them that, subject to appropriate (resistive/capacitive) choices of the various circuit admittances, this general circuit can realize twelve different types of inverse filter functions at each of the two outputs  $V_{01}$  and  $V_{02}$ . However, out of the twelve realized circuits shown by them, only three circuits represent *canonic* realizations; the other configurations employ three to four capacitors, and hence, are non-canonic.



**Figure 10.** Generalized inverse filter configuration using CFOAs reported by Wang, Chang, Yang and Tsai [13].

Garg, Bhagat and Jaint [14] presented a general inverse filter scheme which is also an adaptation of the quoted circuits from Figure 9a,b, the only difference being that instead of the three normal CFOAs, three modified CFOAs (MCFOA), characterized by  $i_z = \alpha_1 i_x$ ,  $i_y = \alpha_2 i_w$ ,  $v_x = \beta_1 v_y$ , and  $v_w = \beta_2 v_z$  [70,71] have been employed, each one implementable from three AD844-type CFOAs or a cascade of one CCII+ and one CCII−. It may, however, be noted [65] that the MCFOA is the same as the ‘Composite Current Conveyor’ formulated by Smith and Sedra in [72] which was comprised of one CCII+ (now realizable with one AD844-type CFOA) and one CCII− (now realizable with two AD844-type CFOAs) to realize various nonlinear circuit elements. For the verification of the workability of these inverse filters, a CMOS MCFOA has been used within the frequency range of 1 kHz–1 GHz. However, all the realized filters used three or more capacitors and hence, are not canonic. Moreover, the ABB used is not available as a commercial IC and needed to be implemented with three AD844s thereby, the proposed configuration requiring as many as nine CFOAs.

Patil and Sharma [15] presented four two-CFOA-based configurations for realizing ILP, IHP, IBP and IBR filters (the last one being inspired by [69]). All these reported circuits have been verified by simulation results in PSPICE using a macro model of AD844. Four of the five circuits have the advantage of being realizable with two only CFOAs. The configurations presented can also provide explicit current output, but at the expense of one additional CFOA.

### 3.4. IAF Configurations Using Operational Transconductance Amplifiers (OTA)

The OTA is characterized by  $I_0 = g_m (V_+ - V_-)$  and has been extensively used in the past to realize electronically controllable filters and oscillators. The bipolar OTA is also a commercially available device and many variants of this such as LM3080, LM13600/LM13700 are available as off-the-shelf ICs whose transconductance is a linear function of an external DC bias current ( $I_B$ ) and is *linearly* variable over a wide range of this current up to about four decades. On the other hand, a CMOS OTA as shown in Figure 11 can be analogously devised using a MOS differential pair along with a number of PMOS/NMOS current mirrors. However, for the CMOS OTA, the transconductance is proportional to  $(I_B)^{1/2}$ . CMOS OTAs with dual outputs, such as the one shown in Figure 11, have been employed by several authors [16,17].

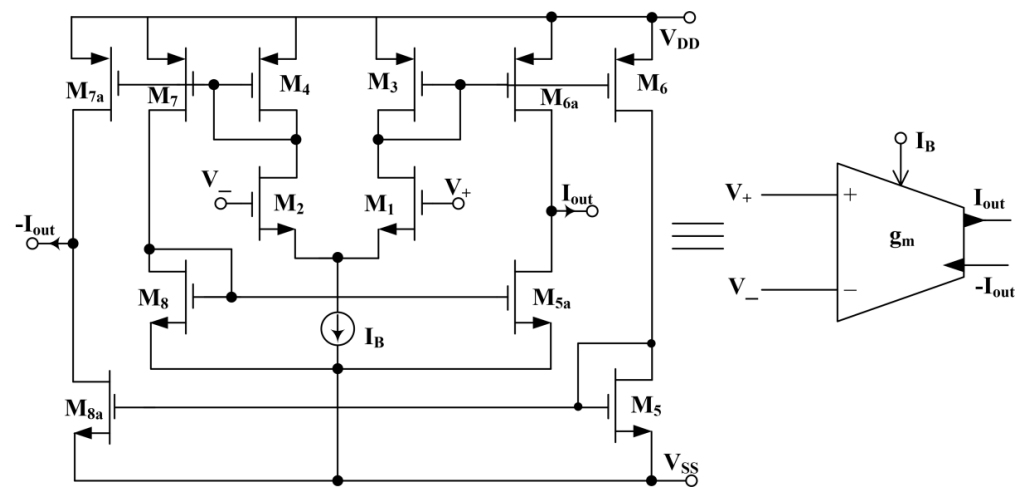


Figure 11. CMOS implementation and symbolic representation of the dual-output OTA.

In [16], Tsukutani, Sumi and Yabuki presented two six-OTA-based circuits, realizing ILP and IBP, and one five-OTA-based circuit realizing IHP as depicted here in Figure 12. All the circuits employ both GCs and have the advantage of offering orthogonal tunability of the parameters  $\omega_0$  and  $Q_0$  which has been successfully demonstrated by SPICE simulation results. The workability of inverse filter circuits shown in Figure 12 has been validated over a frequency range of 10 kHz–1 MHz using CMOS OTA displayed in Figure 11. However, if the circuits are implemented using commercially available IC OTAs, the actual number of IC OTAs required may be more in each case.

The transfer functions of the IAFs shown in Figure 12 have been presented here:

Figure 12a

ILPF

$$\frac{I_{out}(s)}{I_{in}(s)} = \frac{s^2 + s\left(\frac{g_{m3}}{C_1}\right) + \frac{g_{m2}g_{m4}}{g_{mb}C_1C_2}}{\left(\frac{g_{m2}g_{m4}g_{ma}}{g_{mb}C_1C_2}\right)}; I_{out}(s) = I_1(s) + I_3(s) + I_4(s)$$

Figure 12b

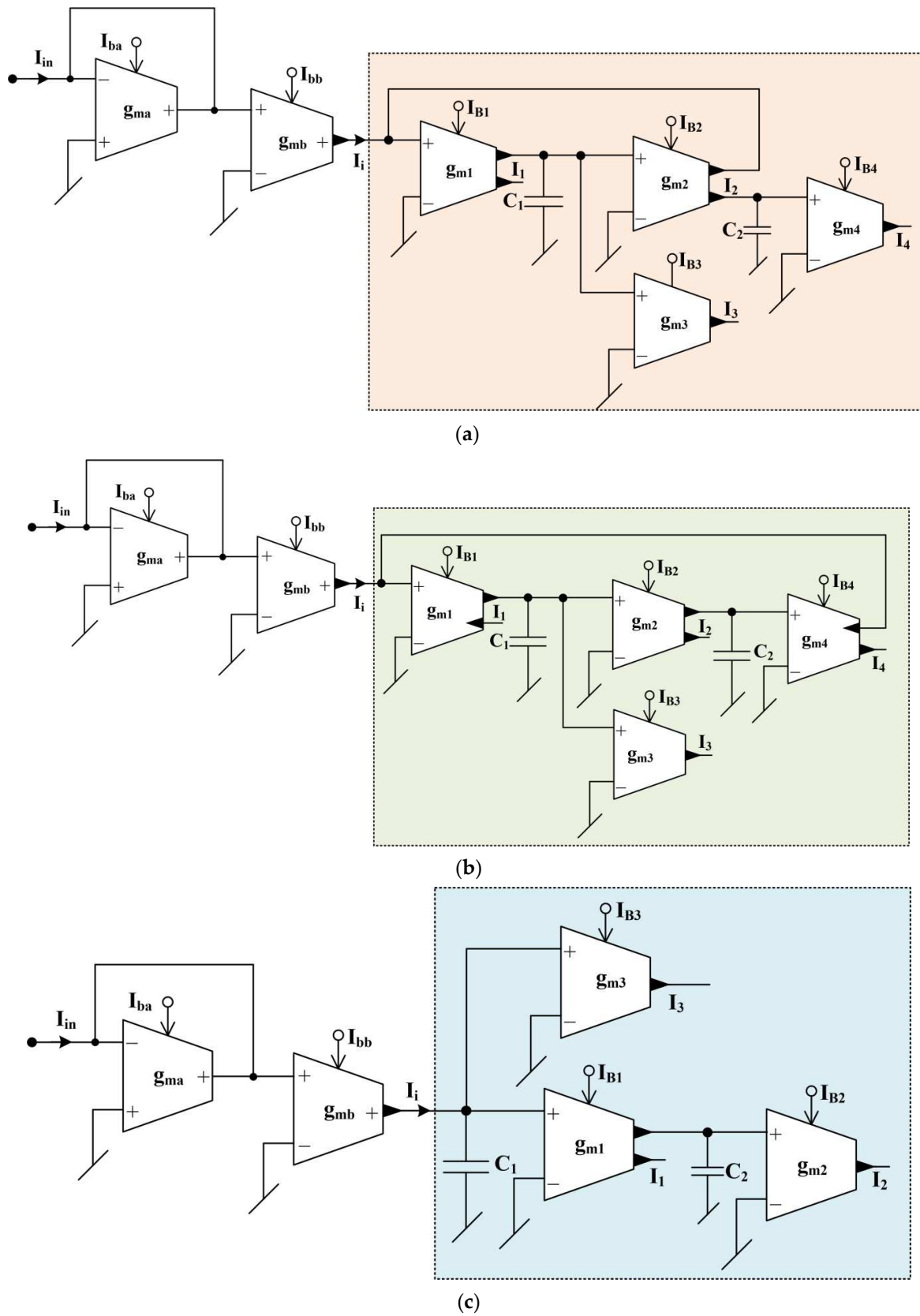
IBPF

$$\frac{I_{out}(s)}{I_{in}(s)} = \frac{s^2 + s\left(\frac{g_{m3}}{C_1}\right) + \frac{g_{m2}g_{m4}}{g_{mb}C_1C_2}}{s\left(\frac{g_{m2}g_{ma}}{g_{mb}C_1}\right)}; I_{out}(s) = I_1(s) + I_3(s) + I_4(s)$$

Figure 12c

IHPF

$$\frac{I_{out}(s)}{I_{in}(s)} = \frac{s^2 + s\left(\frac{g_{m3}}{C_1}\right) + \frac{g_{m1}g_{m2}}{C_1C_2}}{s^2\left(\frac{g_{ma}}{g_{mb}}\right)}; I_{out}(s) = I_1(s) + I_2(s) + I_3(s)$$

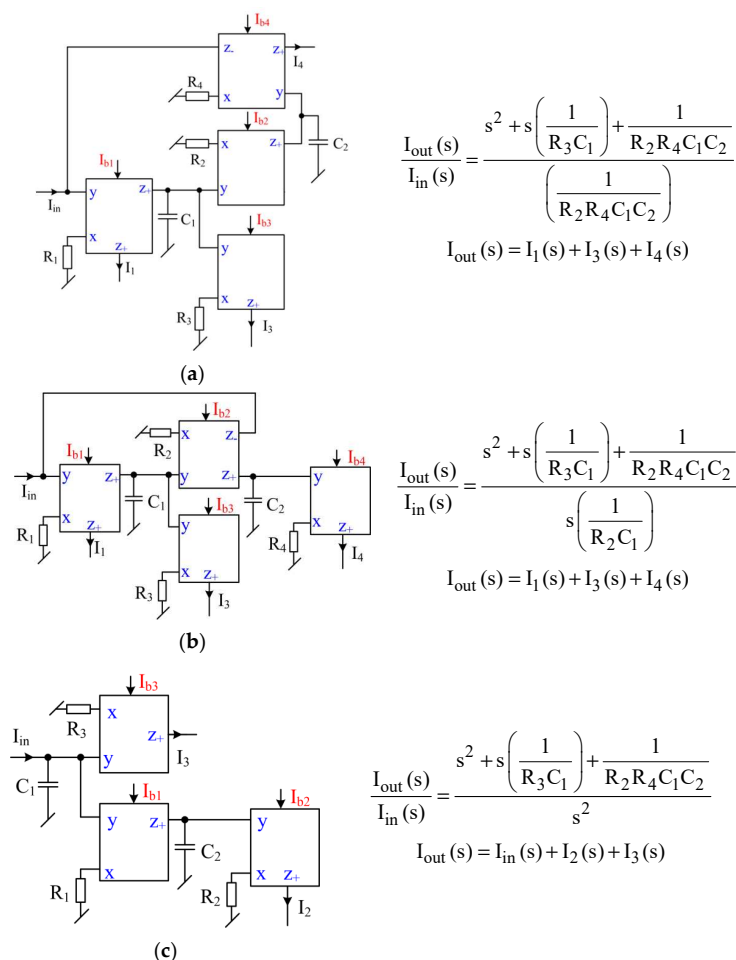


**Figure 12.** The basic current-mode inverse active filter configurations proposed by Tsukutani, Sumi and Yabuki in [16]. (a) ILPF; (b) IBPF; (c) IHPF.

In [17], Raj, Bhagat, Kumar and Bhaskar presented five new OTA-C inverse filters which include one IHP, two ILP, one IBP and one IBR, all possessing infinite input impedance and employment of both grounded capacitors, with  $\omega_0$  and  $Q_0$  being orthogonally adjustable. Their workability was demonstrated by a simple CMOS OTA with its  $g_m$  controllable by an external DC bias voltage.

### 3.5. IAF Configurations Using Second Generation Current Conveyors (CCII)

It may be recalled that a  $CCII_{\pm}$  is a three-port active building block [47,48] characterized by  $i_y = 0$ ,  $v_x = v_y$  and  $i_z = \pm i_x$ . Tsukutani and Kunugasa [18] have proposed three different configurations which simultaneously realize current mode ILP, IBP and IHP filters (Figure 13); two of them employ four current conveyors, while the third one employs only three CCII. However, whereas the first two circuits realize only ILP and IBP functions, the third one realizes only IHP response. Although it is not explicitly stated by the authors, these CCII-based circuits are derived directly from the OTA-C IAF circuits of [16], with each transconductor therein replaced by a CCII-based V-I convertor with a resistor connected from the X-terminal of the CCII to ground. All the circuits use all grounded passive components and the CCII involved was a standard translinear CMOS structure with two complementary Z-outputs (see Figure 14). To validate the functionality of these inverse filters, CMOS CCII, shown in Figure 14, has been used and the results showed the workability of these circuits in the frequency range of 10 kHz–10 MHz.



**Figure 13.** Current-mode inverse active configurations by Tsukutani and Kunugasa [18] (a) ILP (b) IBP (c) IHP.



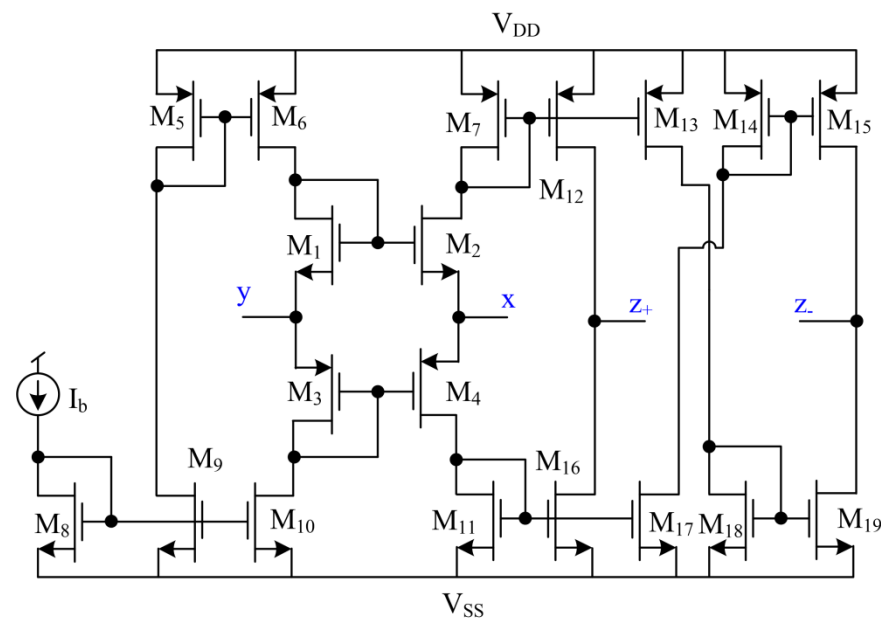


Figure 14. CMOS implementation of the CCII as employed in [18].

In spite of the advantage of all grounded passive elements, although the proposed circuits do have a high output impedance output available, none of them are able to provide the required (ideally zero) input resistance necessary for applying a CM input. The authors have also indicated that by adding a voltage to current converter (V-I) at the input and a resistive load along with a voltage follower at the output, all three CM IAFs can be converted into VM IAF circuits. These additional interfacing circuits too have been proposed to be realized by one CCII each.

A first order inverse all-pass filter using a single CCII- has been reported by Shah and Rather in [19], wherein the CCII- was implemented using two CCII+ s. However, the circuit clearly has a pole in the right half of the  $s$ -plane, and hence, is impractical due to being unstable.

In [20], Herencsar, Lahiri, Koton and Vrba have reported three new realizations of the inverse filters low-pass filter, inverse band-pass filter, and inverse high-pass filter employing differential difference current conveyors (DDCCs) [73], two grounded resistors and two grounded capacitors. It may be recalled that a DDCC is a five-terminal ( $x$ ,  $y_1$ ,  $y_2$ ,  $y_3$  and  $z$ ) active building block characterized by  $i_{yi} = 0$ ,  $i \rightarrow 1-3$ ,  $v_x = v_{y1} - v_{y2} + v_{y3}$  and  $i_z = \pm i_x$ . The reported inverse filter circuits have been validated using SPICE simulations and experimental measurements [20].

### 3.6. IAF Configurations Using Second Generation Voltage Conveyors (VCII)

An interesting variant of the CCII is the so-called *second-generation voltage conveyor* (VCII) [74] which is characterized by the equations:  $i_x = \pm \beta i_y$ ,  $V_z = \alpha V_x$  and  $V_y = 0$ , where nominally,  $\alpha$  and  $\beta$  are both unity. The use of VCII in realizing IAFs has been carried out by Al-Absi in [21] and Al-Shahrani and Al-Absi in [22]. The circuits proposed in both the papers are based around a CMOS implementation of the VCII as displayed in Figure 15 here, which has been claimed to offer the advantages of lower power consumption and less occupied chip area.

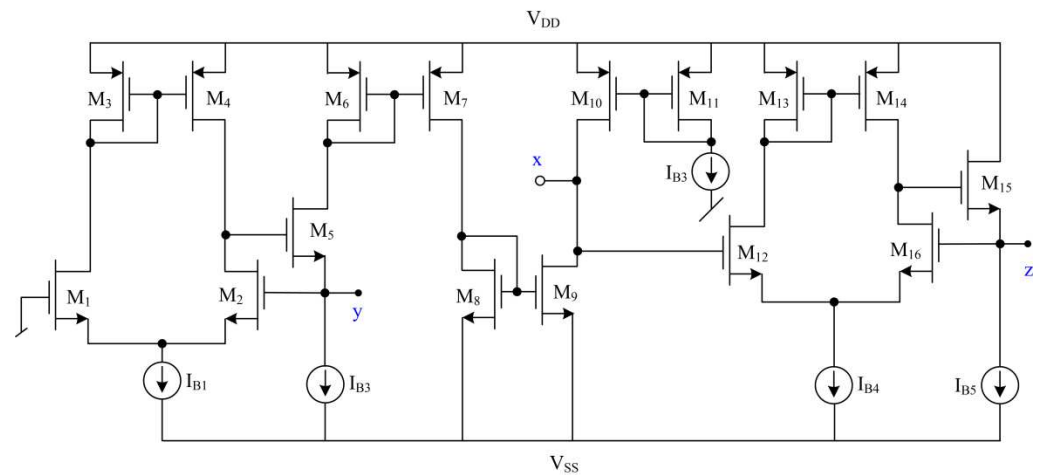


Figure 15. CMOS implementation of the VCII [75].

In [21], Al-Absi proposed two VCII-based circuits for realizing first order ILP and IHP and three-VCII-based circuit for realizing IBP; these are shown in Figure 16. Though all the circuits provide the desirable low-output impedance voltage output, none of them has ideally infinite input impedance.

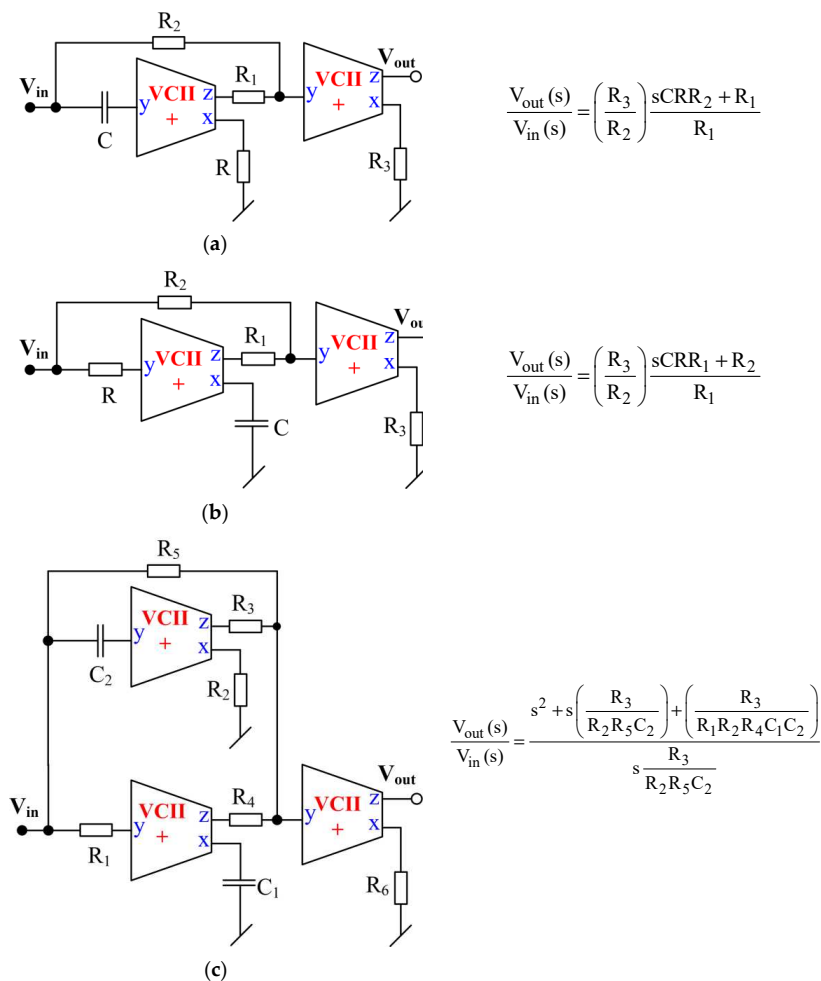


Figure 16. VCII-based inverse filters reported by Al-Absi (a) ILP (b) IHP and (c) IBP [21].

A claimed advantage of these circuits is that control of  $H_0$  and  $\omega_0$  is available in the cases of ILP, IHP and IBP. The workability of these reported circuits has been verified by

simulation results using CMOS VCII, shown in Figure 15, and their frequencies of operation were taken in the range of 1 Hz to 10 MHz.

This was followed by Al-Shahrani and Al-Absi in [22], wherein a modified-VCII (MVCII) proposed earlier in [75] was employed, which is characterized by the following equations:

$$I_{y1} = \pm \beta i_y, I_x = \beta i_y, V_z = \alpha V_x, V_y = 0 \text{ and } i_{z0} = i_z \quad (11)$$

Using MVCII, the authors of [22] presented four CM circuits realizing ILP, IHP, IBR and IBP respectively, using only two external resistors and two capacitors in each case. The IBR and IBP have the advantage of using both GCs, but due to being two-resistor and two-capacitor structures; they do not provide independent tunability of the various filter parameters. For the workability of these circuits, frequency and time responses using CMOS MVCII have been demonstrated.

### 3.7. IAF Configurations Using Operational Transresistance Amplifiers (OTRA) and Current Differencing Buffered Amplifiers (CDBA)

To overcome the limitations of the conventional op-amp, researchers had been continuously coming up with new active circuit building blocks which resulted in the evolution of several alternatives, of which OTAs, CCs, CFOAs are now available as off-the-shelf ICs. However, there are a number of other building blocks such as OTRA and CDBA, which have been investigated for their hardware implementation and applications in the past. Although both of these can be implemented by CMOS circuits, none of them have yet been made available as off-the-shelf ICs. Nevertheless, both of them can be realized using two commercially available AD844-type CFOAs.

An OTRA is characterized by:  $v_p = v_n = 0$  and  $v_o = R_m (i_p - i_n)$  and is shown symbolically in Figure 17.

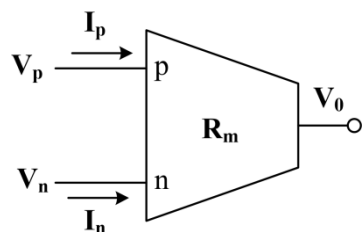


Figure 17. The symbolic representation of an OTRA.

On the other hand, CDBA is characterized by the terminal equations as:  $V_p = V_n = 0$ ,  $I_z = (I_p - I_n)$  and  $V_w = V_z$  and its symbolic representation is depicted in Figure 18.

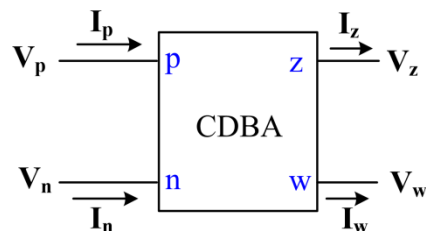
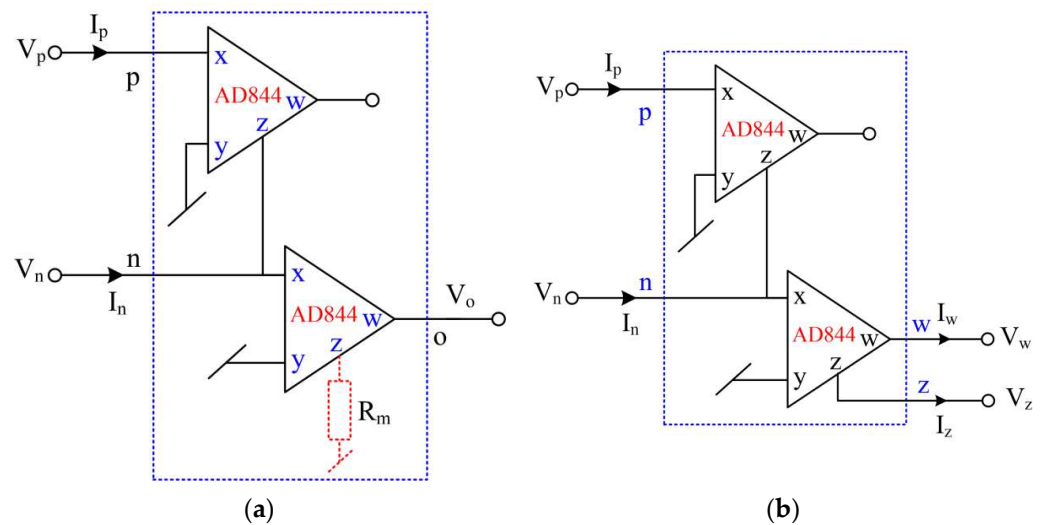


Figure 18. The symbolic representation of the CDBA.

Both of these are realizable from interconnection of two CFOAs as shown in Figure 19. In case of the CDBA implementation, if a load impedance is connected at terminal-Z, the voltage across this ( $V_z$ ) is then made available as  $V_w = V_z$ . On the other hand, for OTRA implementation, this load impedance connected at terminal-Z becomes the transresistance  $R_m$  and for making  $R_m = \infty$  as desired; this terminal, therefore, needs to be left open.

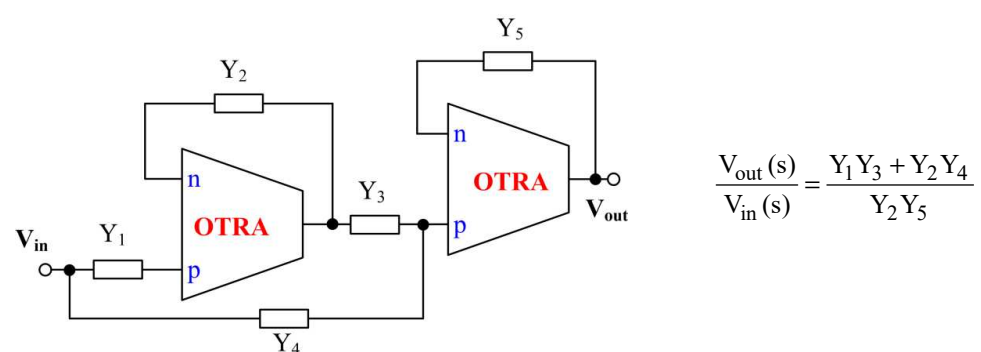


**Figure 19.** AD844 implementation of (a) OTRA ( $R_m = \infty$ ) (b) CDBA.

In the following, we outline some of the prominent works dealing with the realization of inverse filters using CDBA and OTRA.

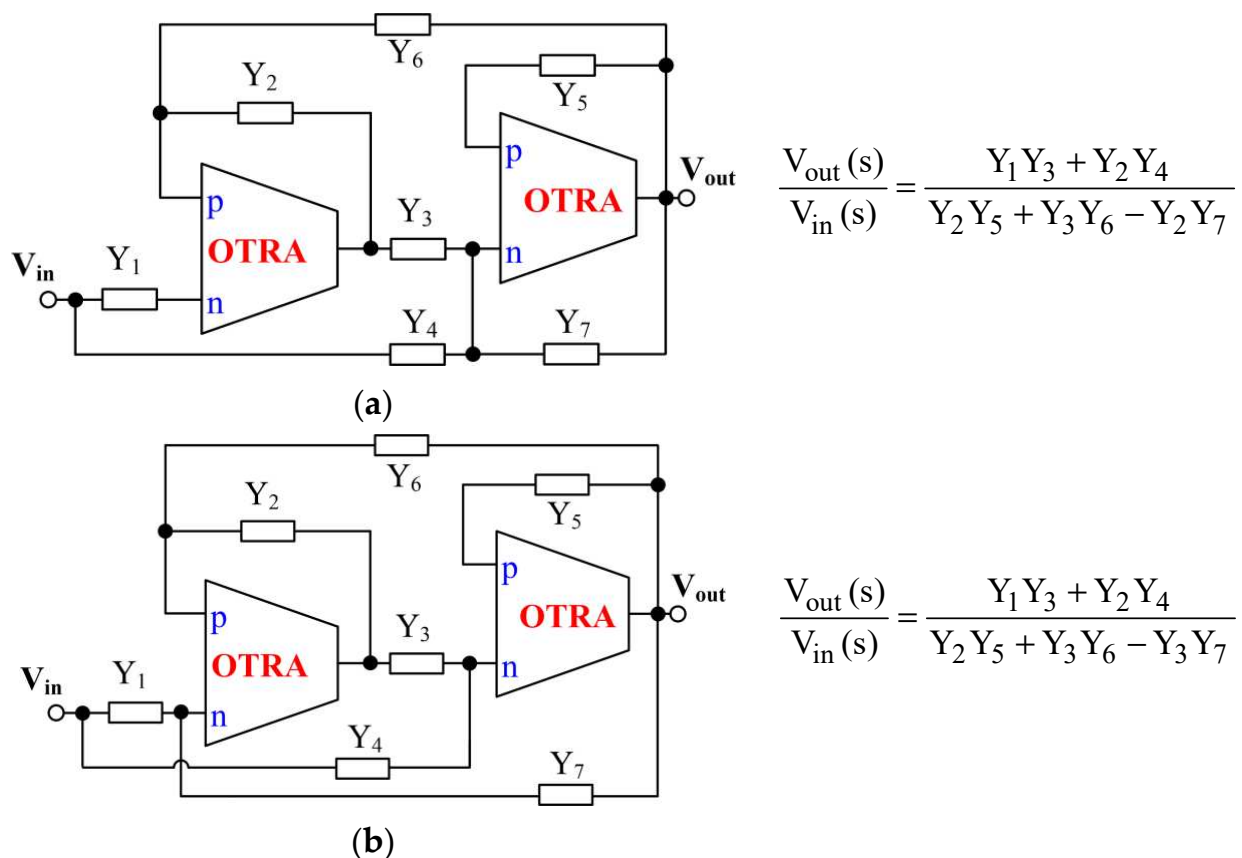
### 3.7.1. OTRA-Based Inverse Active Filters

In [23], Singh, Gupta and Senani proposed two OTRA and five admittance-based general inverse filter structures shown in Figure 20, from which an ILP, IBP and IHP, were shown to be realizable in VM via appropriate (resistive/capacitive) choice of the five circuit admittances. Although the IHP was non-canonic due to the employment of three capacitors along with equality constraints, the remaining two realizations are canonic in terms of capacitors and have the advantage of independent controllability of the various filter parameters. For the verification of the workability of the filter functions, PSPICE results using CMOS OTRA have been provided and experimental results have also been demonstrated using OTRA implemented with commercially available AD844-type CFOAs, as shown in Figure 19a.



**Figure 20.** A generalized OTRA-based structure for realizing inverse filters proposed by Singh, Gupta and Senani [23].

The configurations reported by Pradhan and Sharma in [24] were generalized by providing two different configurations adding two more admittances in the two-OTRA-based configurations of [23] as shown in Figure 21. The authors then came up with the generation of ten different second-order IBR filters and twelve inverse all-pass (IAP) filter circuits by making various resistive/capacitive choices of the seven admittances therein.



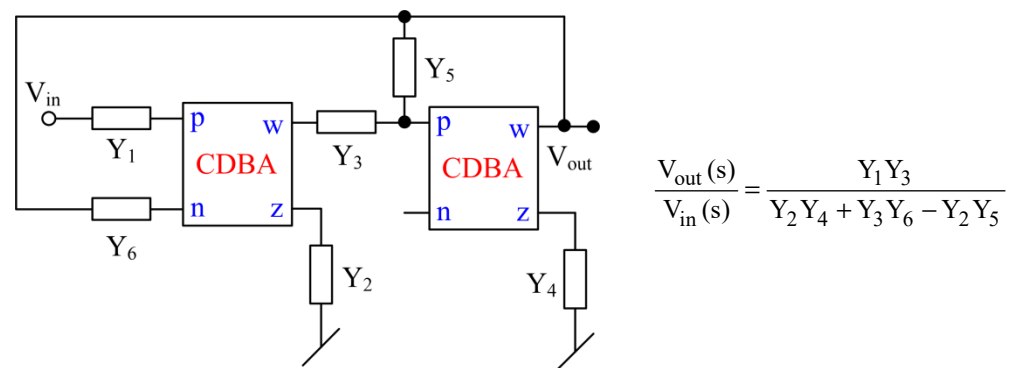
**Figure 21.** Generalized configuration of IBR and inverse all-pass filter employing OTRAs proposed by Pradhan and Sharma (a) IBRF (b) IAPF [24].

However, all the special cases required three to four capacitors and, hence, are non-canonic. Various simulation results using CMOS OTRA implemented with 0.18  $\mu\text{m}$  TSMC technology parameters have been provided to validate the workability of all the reported cases of IBRF and IAPF over the frequency range 1 kHz–10 MHz. However, since most of the results were verified through simulations, the inherent stability of IAPF appears to have gone unnoticed.

Banerjee, Borah, Ghosh and Mondal in [25] have come up with a configuration using a single OTRA, but the circuit requires as many as three resistors and three capacitors along with three switches to realize band reject filters; hence, these circuits too are non-canonic. Simulations as well as experimental results have been provided for the validation of the proposed circuits. The layout of these three configurations has also been demonstrated and their post-layout simulation results have been provided. The major drawbacks of these circuits are their non-canonicity, non-availability of ideally infinite input impedance and the inconvenience of obtaining different responses through three switches therein.

### 3.7.2. Inverse Active Filters Employing CDBA

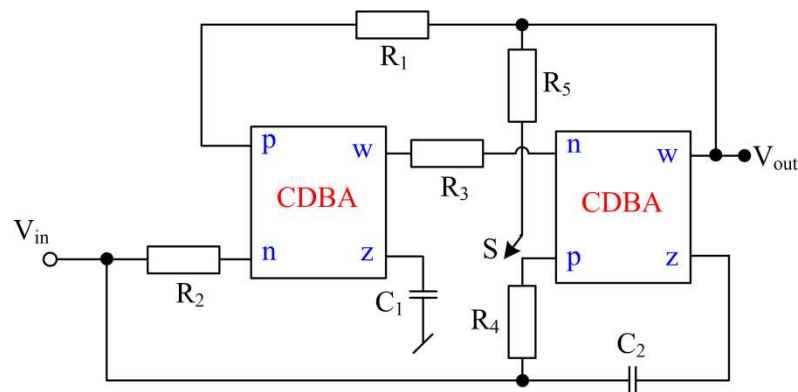
In [27], Pandey, Pandey, Negi and Garg proposed a CDBA-based universal inverse filter circuit consisting of six admittances whose circuit realization has been shown in Figure 22, from which various special cases have been obtained by appropriate RC impedances in place of six admittances of the circuit. However, excepting the circuit for realizing an ILP, which is canonic, all other choices lead to non-canonic structures. A PSPICE macro model of CFOA AD844 IC has been used to implement CDBA (shown in Figure 19b) and the same has been used to verify the workability of the various inverse filter functions.



**Figure 22.** CDBA-based multifunction inverse filter configuration reported by Pandey, Pandey, Negi and Garg [27].

Nasir and Ahmad in [28] presented a two-CDBA-based CM multifunction inverse filter, but the circuit suffers from the drawback of non-availability of explicit current output, besides one of the input terminals of the CDBA being left unutilized. The CDBA implemented using AD844 shown in Figure 19b has been used to validate the functionality of these inverse filter functions in the frequency range of 10 kHz–1 MHz.

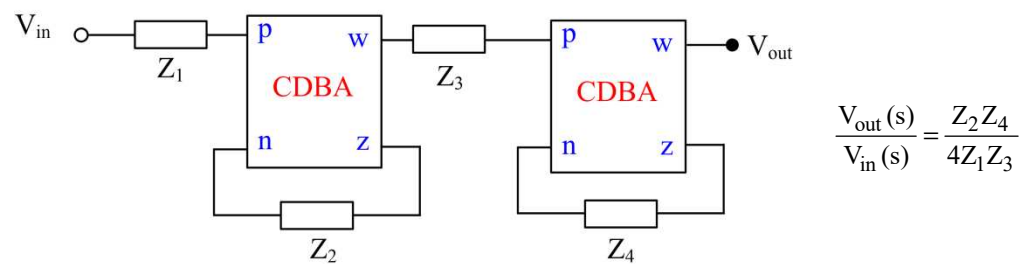
IBR and IAP filters employing two CDBAs were presented in [29] by Bhagat, Bhaskar and Kumar; these filters are based on the general configuration shown here in Figure 23, in which two different responses were obtainable by appropriate opening/closing of a switch from the same output. The circuit uses a GC and another virtually grounded capacitor. In spite of requiring an equality of two resistors, both  $\omega_0$  and BW can be tuned independently of each other. The workability of this proposition was demonstrated by PSPICE simulations using CMOS CDBA with 0.18  $\mu\text{m}$  CMOS technology parameters. However, there may be stability issues when the circuits are practically realized with CDBA implementation using two CFOA ICs.



**Figure 23.** CDBA-based general configuration for realizing IBR and inverse all-pass filters presented by Bhagat, Bhaskar and Kumar [29].

In [30], Bhagat, Bhaskar and Kumar proposed a two CDBA and four impedances-based circuit from which various inverse filters were shown to be realizable by appropriate resistive/capacitive choice of the various circuit elements. The notable feature of this circuit is that, from the same circuit, both normal filters and inverse filters are realizable with a canonic number of capacitors in all cases. The details of the twelve possible realizations resulting from the circuit of Figure 24 can be seen in [30]. The PSPICE simulation results of these normal filters and inverse filters using CMOS CDBA implemented with 0.18  $\mu\text{m}$  CMOS technology parameters have been provided over the frequency range 10 kHz–50 MHz.





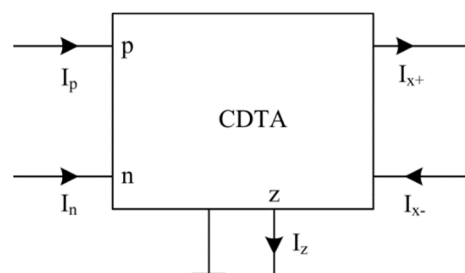
**Figure 24.** A multifunction filter/inverse filter configuration employing CDBAs reported by Bhagat, Bhaskar and Kumar [30].

Whereas most of the work has been focused on second-order inverse filters of various kinds, Borah, Singh and Ghosh in [31] have given a realization of a sixth-order inverse band-pass filter. The circuit consists of two CDBAs and eight impedances, which has been practically realized from four AD844-type CFOAs but employs nine capacitors and as many resistors. For the validation of the proposed IBP filter function, experimental results using CDBA implemented with AD844 have been taken and the layout of the proposed configuration has also been provided along with post-layout simulation results. The employment of more (nine) capacitors than really necessary, with most of them being floating, appears to be a clear disadvantage from the viewpoint of actual IC implementation.

While most of the works described above require two CDBAs, a single-CDBA-based inverse filter structure was proposed in [32] by Paul, Roy and Pal. However, out of the four types of filters realizable (ILP, IBP, IHP, IBR), only the ILP is realizable *canonically*; all others require three capacitors and, hence, are non-canonic. The various inverse filter functions have been validated using CMOS CDBA as well as CDBA implemented with AD844-type CFOAs.

### 3.8. Inverse Active Filters Employing Current-Differencing Transconductance Amplifier (CDTA)

The CDTA introduced by Biolek [52] is an active building block characterized by  $v_p = v_n = 0$ ,  $i_z = i_p - i_n$  and  $i_{x\pm} = \pm g_m v_z$ , symbolically represented as in Figure 25.

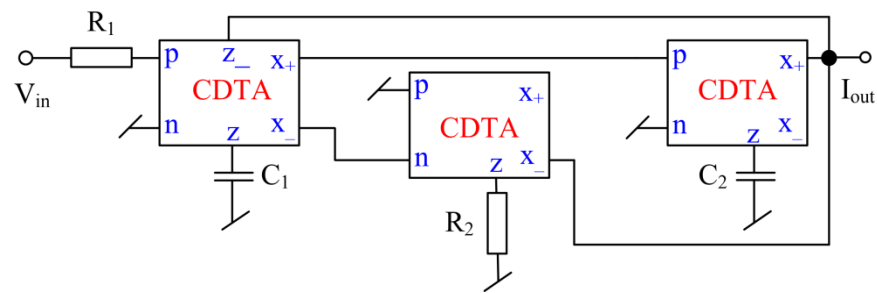


**Figure 25.** The symbolic representation of the CDTA [52].

A single-CDTA-based first order inverse all-pass filter was presented in [33] by Shah, Quadri and Iqbal. However, the authors have failed to notice the instability because of the verification through only SPICE simulations and that too, without checking the transient response.

In [34], Sharma, Kumar and Whig proposed a three-CDTA-based IHPF whose circuit configuration has been displayed in Figure 26.

The workability of the reported circuit has been verified in PSPICE using CMOS CDTA with 0.35  $\mu\text{m}$  CMOS technology parameters and the frequency responses have been demonstrated in the range of 1 Hz–100 MHz. The circuit suffers from the drawback of neither having infinite input impedance nor low output impedance.



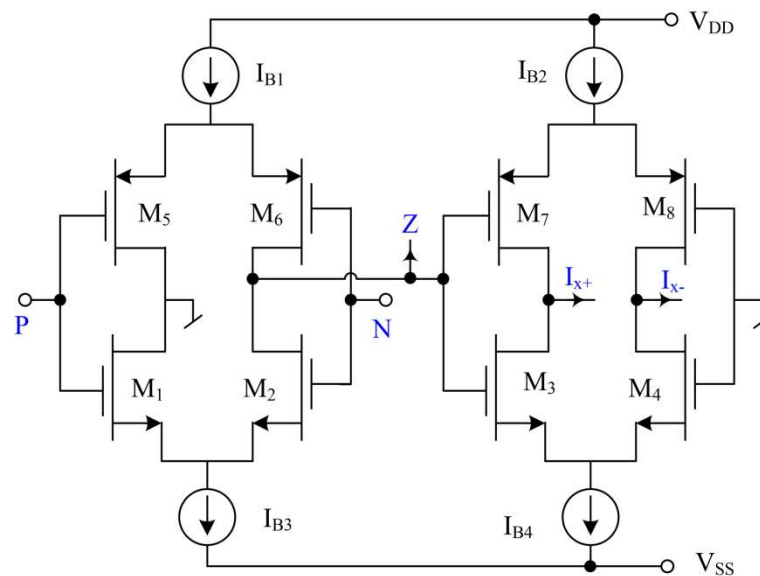
**Figure 26.** CDTA based inverse high-pass filter configuration reported by Sharma, Kumar and Whig [34].

### 3.9. Inverse Active Filters Employing Voltage Differencing Transconductance Amplifiers (VDTA)

A VDTA [35] is an active element which is characterized by:

$$I_z = g_{m1} (V_p - V_n) \text{ and } I_{x\pm} = \pm g_{m2} V_z \quad (12)$$

and is very closely related to an OTA. In fact, it is often made from a cascade of a differential-input single-output OTA and single-input complementary dual-output OTA or with three normal types of OTAs. However, most of the authors have used a specific CMOS implementation of the VDTA, as shown in Figure 27. The symbolic notation of the VDTA and its implementations in terms of normal OTAs take the forms as shown here in Figure 28.

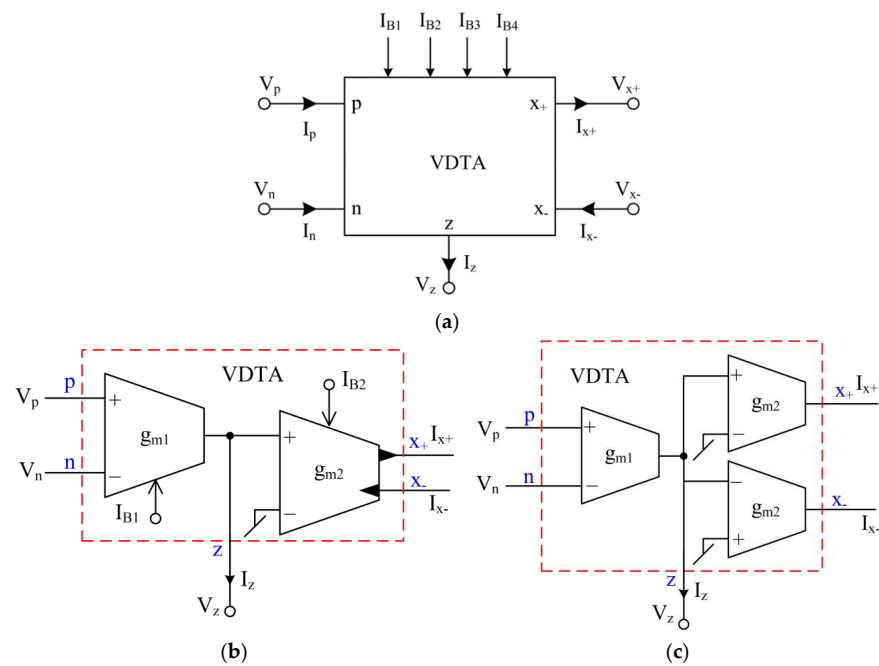


**Figure 27.** A popular CMOS implementation of the VDTA [35].

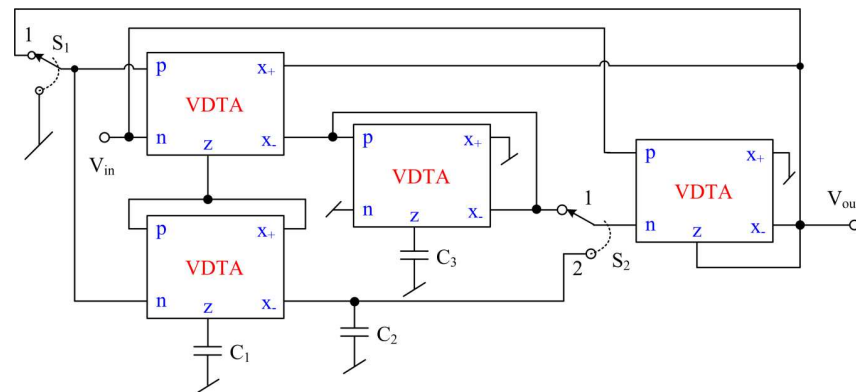
In [35], Kumar, Pandey and Paul presented a number of VDTA-based inverse filters using two to four VDTAs along with two grounded capacitors without requiring any resistors, as preferred for IC implementation. A unified inverse filter structure from [35] is shown in Figure 29.

The workability of all the realized inverse filters has been validated using CMOS VDTA with 0.18  $\mu\text{m}$  CMOS technology parameters from TSMC over a frequency range of 100 kHz–100 MHz.

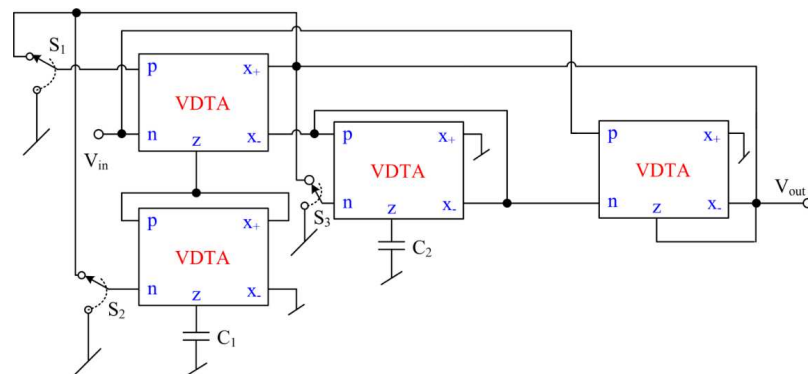
In another publication [36], the same authors presented another multifunction inverse filter similar to their earlier publication, which is shown in Figure 30 and contains three switches. By changing their positions, different filter functions can be achieved.



**Figure 28.** Symbolic implementation and OTA-based implementations of the VDTA; (a) the symbolic representation; (b,c) two different VDTA implementations using OTAs.



**Figure 29.** Unified inverse filter topology employing VDTAs proposed by Kumar, Pandey and Paul [35].



**Figure 30.** VDTA-based universal multifunction inverse filter reported by Kumar, Pandey and Paul [36].

The workability of this configuration in its various modes has been demonstrated by SPICE simulation results and hardware results with VDTA realized by three LM13700-type

IC OTAs, as per the schematic of Figure 28c. It is worth mentioning that the works [35,36] are the only ones where a study of total harmonic distortion and output noise variations has been investigated.

Lastly, it may be mentioned that, of late, there has been some interest in realizing fractional order inverse filters (FOIF) as well. It is obvious that the procedure applicable to fractional order normal filter design can also be applied on any inverse filter structure also to convert that into FOIF. Based upon this idea, a number of authors have recently come up with FOIF designs [76–82] using a variety of active elements. However, discussion of these is outside the scope of the present article.

In Table 1, we present a concise summary of the various inverse filter configurations discussed so far as a quick reference guide for the designers.

**Table 1.** A concise comparison of various IAFs reviewed in this article.

Reference, Year	Name and Number of ABB Used	Number of G + R + C	Whether All Grounded Capacitors Used	Whether USED ABB Is Available as Off-the-Shelf IC?	Inverse Filter Functions Realized
[1]	Op-amp (1)	$0 + 4 + 2$	No	Yes	IHP
[37]	Op-amp (1)	$0 + (1/2) + (1/2)$	No	Yes	IHP #, IBP
[16]	OTA (5/6)	$(5/6) + 0 + 2$	Yes	Yes	ILP, IHP, IBP
[17]	OTA (4/5)	$(4/5) + 0 + 2$	Yes	Yes	ILP, IHP, IBP, IBR
[18]	CCII (3/4)	$0 + (3/4) + 2$	Yes	Yes	ILP, IHP, IBP
[19]	CCII (1)	$0 + 2 + 1$	Yes	Yes	IAP #
[11]	CFOA (3)	$0 + 4 + 2$	Yes	Yes	ILP, IHP, IBP, IBR
[12]	CFOA (3)	$0 + (3/5) + 2$	Yes	Yes	ILP, IHP, IBP, IBR
[13]	CFOA (3)	$0 + (2/3) + (2/3)$	Yes	Yes	ILP, IHP, IBP
[15]	CFOA (2)	$0 + (4/6) + 2$	Yes	Yes	ILP, IHP, IBP, IBR
[14]	MCFOA (3)	$0 + (2/3) + (3/4)$	Yes	No	ILP, IHP, IBP
[7]	FTFN (1)	$0 + 5 + 2$	No	No	ILP
[8]	FTFN (1)	$0 + 8 + 2$	No	No	IAP
[9]	FTFN (1)	$0 + 4 + 2$	No	No	ILP, IHP, IBP, IBR, IAP
[10]	FTFN (1)	$0 + 3 + 1$	No	No	IAP #
[23]	OTRA (2)	$0 + (4/5) + 2$	No	No	ILP, IHP, IBP
[24]	OTRA (2)	$0 + (4/6) + (3/4)$	No	No	IBR, IAP
[25]	OTRA (1)	$0 + 3 + 3$	No	No	IBR
[27]	CDBA (2)	$0 + (2/4) + (2/4)$	No	No	ILP, IHP, IBP, IBR, IAP
[28]	CDBA (2)	$0 + (3/4) + (3/4)$	No	No	ILP, IHP, IBP
[29]	CDBA (2)	$0 + (4/5) + 2$	No	No	IBR, IAP
[30]	CDBA (2)	$0 + 3 + 2$	No	No	ILP, IHP, IBP, IBR
[31]	CDBA (2)	$0 + 9 + 9$	No	No	IBPF
[32]	CDBA (1)	$0 + (2/3) + (2/3)$	No	No	ILP, IHP, IBP, IBR
[33]	CDTA (1)	$0 + 1 + 1$	Yes	No	IAP #
[34]	CDTA (3)	$0 + 2 + 2$	Yes	No	ILP
[35]	VDTA (2/4)	$(4/8) + 0 + 2$	Yes	No	ILP, IHP, IBP, IBR
[36]	VDTA (4)	$8 + 0 + 2$	Yes	No	ILP, IHP, IBP, IBR
[21]	VCII (2/3)	$0 + (4/6) + (1/2)$	No	No	ILP #, IHP #, IBP
[22]	VCII (2)	$0 + 2 + 2$	Yes	Yes	ILP, IHP, IBP, IBR

G = Transconductor, R: Resistor, C: Capacitor, #: First order inverse filter.

#### 4. The Unresolved Issues

The following issues related to inverse analog filters have not been addressed adequately in the open literature to date, and are therefore still open for further research and investigations.

In spite of considerable work reported in the published literature on analog inverse filters using a variety of ABBs, consideration of applying these structures to actual real-life practical applications has been very rare.

If one looks at the IBPF characterized by:

$$\frac{V_0(s)}{V_{in}(s)} = \frac{K_1 s^2 + K_2 s + K_3}{K_4 s} \quad (13)$$

This can be re-written in the form of the following input–output equation:

$$V_0(s) = \left( K_P + \frac{1}{K_I s} + K_D s \right) V_{in}(s) \quad (14)$$

which, in time-domain can be expressed as:

$$V_0(t) = K_P V_{in}(t) + \frac{1}{K_I} \int V_{in} dt + K_D \frac{dV_{in}}{dt} \quad (15)$$

It therefore turns out that the IBP is the same thing as a PID controller. So, this is one practical application of IBP which can be considered to be known. However, to the best knowledge of the authors, except for the use of IBP filter as a PID controller, no explicit applications of any other types of inverse filters have so far been reported in the literature by anybody!

In this respect, it is worthwhile to mention that the synthesis of inverse filters on analog computers and their applications in the restoration and correction of time functions was carried out by Burch, Green and Grote in [6] as early as in 1964, i.e., before the easy commercial availability of the integrated circuit op-amps and any other IC circuit building blocks. Hence, it appears that the idea presented in [6] could now be attempted with the ready commercial availability of a large variety of analog circuit building blocks as the off-the-shelf ICs. Hence, putting all the ideas of the inverse filters to real-life practical applications is an unresolved and challenging problem which is still open to investigation.

It is worth mentioning that, with the exception of [35] and [36], the evaluation of the noise performance and dynamic range (DR) of IAFs have not been discussed in any of the earlier publications [1,7–43] on IAF so far, hence such data is not available for ant IAFs proposed in the quoted references. These aspects of most of the inverse filters proposed so far need further work.

Also, since the inverse filters feature (theoretically) infinite gain for some frequencies — for example, an inverse low-pass filter features infinite gain for high frequencies, and an inverse high-pass filter for low frequencies — this is a serious disadvantage, as very weak out-of-band signals may be amplified infinitely. Thus, the filter's own noise as well as the supply rail interference may cause 'clipping'. These aspects, too, need further investigation and resolution.

#### 5. Concluding Remarks

This paper has presented a comprehensive survey of a variety of inverse active filters reported in the literature so far. The salient features of a number of prominent circuits using a variety of active building blocks, ranging from the classical voltage mode op-amp, the OTA, to the popular current mode building blocks such as current conveyors and their variants, the CFOA and a number of building blocks of more recent origin such as CDTA, OTRA, CDBA, CDTA and VDTA have been detailed and their merits and limitations have been spelled out. A number of unresolved problems/issues related to inverse analog filters were highlighted. We have attempted to demonstrate the need for finding a correlation

between the coefficients of the various inverse transfer functions and the shape of the corresponding frequency responses. However, it is felt that additional research is needed to explore this issue further. The stability issues related to the first-order inverse all-pass filter and second-order IHP, IBR and IAP were pointed out. It was emphasized that further research is required to put most of the inverse filter types to practical use in real-life applications. In conclusion, the authors believe that a lot of research is still needed in order to bring the work done on inverse analog active filters [1–43] to a more practical footing.

**Author Contributions:** Conceptualization of this manuscript was by R.S.; the methodology by R.S., D.R.B. and A.R.; software verifications in SPICE by A.R.; validation by R.S., D.R.B. and A.R.; formal checking of the analysis by D.R.B. and A.R.; investigation by R.S.; resources by R.S., D.R.B. and A.R.; data curation by A.R.; writing—original draft preparation by R.S.; writing—review and editing by R.S., D.R.B. and A.R.; visualization by R.S.; supervision by R.S. and D.R.B.; project administration by R.S., D.R.B. and A.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** The first author (R.S.) wishes to thank Costas Psychalinos for extending the invitation to contribute a feature article for the special issue of this journal which was inspirational for the origin of this paper. The first author (R.S.) acknowledges the facilities provided by the ASP Research Lab. of the Department of ECE of NSUT and wishes to thank Tarun Rawat, S. P. Singh and Vice-Chancellor of NSUT J. P. Saini for the same. The other authors, namely, D.R.B. and A.R. acknowledge the facilities provided by ASP Lab. of the EE Department of Delhi Technological University (DTU) and wish to thank Pragati Kumar, Madhusudan Singh and Uma Nangia of the EE Department of DTU for their support. All the authors also wish to thank J. P. Saini, the Vice Chancellor of NSUT and DTU both, for his unflinching support. Thanks are due to all the anonymous reviewers for their constructive comments/suggestions which have been very useful in improving the presentation of this article. The authors particularly thank Reviewer # 4 for one of his very important comments, which we have taken the liberty of including nearly as-is, as the last paragraph of Section 4 of the revised version.

**Conflicts of Interest:** The authors declare that there is no conflict of interest.

## References

1. Leuciuc, A. Using nullors for realization of inverse transfer functions and characteristics. *Electron. Lett.* **1997**, *33*, 949–951. [\[CrossRef\]](#)
2. Tugnait, J. Identification and deconvolution of multichannel linear non-Gaussian processes using higher order statistics and inverse filter criteria. *IEEE Trans. Signal Process.* **1997**, *45*, 658–672. [\[CrossRef\]](#)
3. Kirkeby, O.; Nelson, P.A. Digital filter design for inversion problems in sound reproduction. *J. Audio Eng. Soc.* **1999**, *47*, 583–595.
4. Watanabe, A. Formant estimation method using inverse-filter control. *IEEE Trans. Speech Audio Process.* **2001**, *9*, 317–326. [\[CrossRef\]](#)
5. Zhang, Z.; Wang, D.; Wang, W.; Du, H.; Zu, J. A Group of Inverse Filters Based on Stabilized Solutions of Fredholm Integral Equations of the First Kind. In Proceedings of the 2008 IEEE Instrumentation and Measurement Technology Conference, Victoria, BC, Canada, 12–15 May 2008; pp. 668–671.
6. Burch, J.; Green, A.; Grote, H. Restoration and Correction of Time Functions by the Synthesis of Inverse Filters on Analog Computers. *IEEE Trans. Geosci. Electron.* **1964**, *2*, 19–24. [\[CrossRef\]](#)
7. Chipipop, B.; Surakamponorn, W. Realization of current-mode FTFN-based inverse filter. *Electron. Lett.* **1999**, *35*, 690–692. [\[CrossRef\]](#)
8. Wang, H.-Y.; Lee, C.-T. Using nullors for realisation of current-mode FTFN-based inverse filters. *Electron. Lett.* **1999**, *35*, 1889–1890. [\[CrossRef\]](#)
9. Abuelma'Atti, M.T. Identification of Cascadable Current-Mode Filters and Inverse-Filters Using Single FTFN. *Frequenz* **2000**, *54*, 284–289.
10. Shah, N.A.; Malik, M.A. FTFN based dual inputs current-mode all pass inverse filters. *Indian J. Radio Space Phys.* **2005**, *34*, 206–209.
11. Gupta, S.S.; Bhaskar, D.R.; Senani, R.; Singh, A.K. Inverse active filters employing CFOAs. *Electr. Eng.* **2009**, *91*, 23–26.
12. Gupta, S.S.; Bhaskar, D.R.; Senani, R. New analogue inverse filters realized with current feedback op-amp. *Int. J. Electron.* **2011**, *9*, 1103–1113. [\[CrossRef\]](#)
13. Wang, H.-Y.; Chang, S.-H.; Yang, T.-Y.; Tsai, P.-Y. A Novel Multifunction CFOA-Based Inverse Filter. *Circuits Syst.* **2011**, *2*, 14–17. [\[CrossRef\]](#)
14. Garg, K.; Bhagat, R.; Jain, B. A novel multifunction modified CFOA based inverse filter. In Proceedings of the 2012 IEEE 5th India International Conference on Power Electronics (IICPE), Delhi, India, 6–8 December 2012; pp. 1–5.



15. Patil, V.N.; Sharma, R.K. Novel inverse active filters employing CFOA. *Int. J. Sci. Res. Dev.* **2015**, *3*, 359–360.
16. Tsukutani, T.; Sumi, Y.; Yabuki, N. Electronically tunable inverse active filters employing OTAs and grounded capacitors. *Int. J. Electron. Lett.* **2014**, *4*, 166–176. [\[CrossRef\]](#)
17. Raj, A.; Bhagat, R.; Kumar, P.; Bhaskar, D.R. Grounded-Capacitor Analog Inverse Active Filters using CMOS OTAs. In Proceedings of the 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 26–27 August 2021; pp. 778–783.
18. Tsukutani, T.; Kunugasa, Y.; Yabuki, N. CCII-Based Inverse Active Filters with Grounded Passive Components. *J. Electr. Eng.* **2018**, *6*, 212–215. [\[CrossRef\]](#)
19. Shah, N.A.; Rather, M.F. Realization of voltage-mode CCII-based all pass filter and its inversion version. *Indian J. Pure Appl. Phys.* **2006**, *44*, 269–271.
20. Herencsar, N.; Lahiri, A.; Koton, J.; Vrba, K. Realizations of second-order inverse active filters using minimum passive components and DDCCs. In Proceedings of the 33rd International Conference on Telecommunications and Signal Processing-TSP, Vienna, Austria, 17–20 August 2010; pp. 38–41.
21. Al-Absi, M.A. Realization of inverse filters using second generation voltage conveyor (VCII). *Analog Integr. Circuits Signal Process.* **2021**, *109*, 29–32. [\[CrossRef\]](#)
22. Al-Shahrani, S.M.; Al-Absi, M.A. Efficient Inverse Filters based on Second-Generation Voltage Conveyor (VCII). *Arab. J. Sci. Eng.* **2021**, 1–6. [\[CrossRef\]](#)
23. Singh, A.K.; Gupta, A.; Senani, R. OTRA-Based Multi-Function Inverse Filter Configuration. *Adv. Electr. Electron. Eng.* **2018**, *15*, 846–856. [\[CrossRef\]](#)
24. Pradhan, A.; Sharma, R.K. Generation of OTRA-Based Inverse All Pass and Inverse Band Reject Filters. *Proc. Natl. Acad. Sci. India Sect. A Phys. Sci.* **2019**, *90*, 481–491. [\[CrossRef\]](#)
25. Banerjee, S.; Borah, S.S.; Ghosh, M.; Mondal, P. Three Novel Configurations of Second Order Inverse Band Reject Filter Using a Single Operational Transresistance Amplifier. In Proceedings of the TENCON 2019—2019 IEEE Region 10 Conference (TENCON), Kochi, India, 17–20 October 2019; pp. 2173–2178.
26. Prasad, D.; Tayal, D.; Yadav, A.; Singla, L.; Haseeb, Z. CNTFET-based OTRA and its Application as Inverse Low Pass Filter. *Int. J. Electron. Telecommun.* **2019**, *65*, 665–670.
27. Pandey, R.; Pandey, N.; Negi, T.; Garg, V. CDBA Based Universal Inverse Filter. *ISRN Electron.* **2013**, *2013*, 1–6. [\[CrossRef\]](#)
28. Nasir, A.R.; Ahmad, S.N. A new current-mode multifunction inverse filter using CDBAs. *Int. J. Comput. Sci. Inf. Secur.* **2013**, *11*, 50–53.
29. Bhagat, R.; Bhaskar, D.R.; Kumar, P. Inverse Band Reject and All Pass Filter Structure Employing CMOS CDBAs. *Int. J. Eng. Res. Technol.* **2019**, *8*, 39–44.
30. Bhagat, R.; Bhaskar, D.R.; Kumar, P. Multifunction Filter/Inverse Filter Configuration Employing CMOS CDBAs. *Int. J. Recent Technol. Eng.* **2019**, *8*, 8844–8853.
31. Borah, S.S.; Singh, A.; Ghosh, M. CMOS CDBA Based 6th Order Inverse Filter Realization for Low-Power Applications. In Proceedings of the 2020 IEEE Region 10 Conference (TENCON), Osaka, Japan, 16–19 November 2020; pp. 11–15.
32. Paul, T.K.; Roy, S.; Pal, R.R. Realization of Inverse Active Filters Using Single Current Differencing Buffered Amplifier. *J. Sci. Res.* **2021**, *13*, 85–99. [\[CrossRef\]](#)
33. Shah, N.A.; Quadri, M.; Iqbal, S.Z. High output impedance current-mode all pass inverse filter using CDTA. *Indian J. Pure Appl. Phys.* **2008**, *46*, 893–896.
34. Sharma, A.; Kumar, A.; Whig, P. On performance of CDTA-based novel analog inverse low pass filter using 0.35  $\mu\text{m}$  CMOS parameter. *Int. J. Sci. Technol. Manag.* **2015**, *4*, 594–601.
35. Kumar, P.; Pandey, N.; Paul, S.K. Realization of Resistor less and Electronically Tunable Inverse Filters Using VDTA. *J. Circ Syst. Comput.* **2018**, *28*, 1950143. [\[CrossRef\]](#)
36. Kumar, P.; Pandey, N.; Paul, S.K. Electronically Tunable VDTA-Based Multi-function Inverse Filter. *Iran. J. Sci. Technol. Trans. Electr. Eng.* **2021**, *45*, 247–257. [\[CrossRef\]](#)
37. Kamat, D.V. New Operational Amplifier based Inverse Filters. In Proceedings of the 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 12–14 June 2019; pp. 1177–1181.
38. Freeborn, T.J.; Elwakil, A.S.; Maundy, B. Approximated fractional-order inverse Chebyshev low pass filters. *Circuits Syst. Signal Process.* **2016**, *35*, 1973–1982. [\[CrossRef\]](#)
39. Bhaskar, D.R.; Kumar, M.; Kumar, P. Fractional order inverse filters using operational amplifier. *Analog Integr. Circuits Signal Process.* **2018**, *97*, 149–158. [\[CrossRef\]](#)
40. Bhaskar, D.R.; Kumar, M.; Kumar, P. Minimal Realization of Fractional-Order Inverse Filters. *IETE J. Res.* **2020**, 1–14. [\[CrossRef\]](#)
41. Hamed, E.M.; Said, L.A.; Madian, A.H.; Radwan, A. On the Approximations of CFOA-Based Fractional-Order Inverse Filters. *Circuits Syst. Signal Process.* **2019**, *39*, 2–29. [\[CrossRef\]](#)
42. Kumar, M.; Bhaskar, D.R.; Kumar, P. CFOA-Based New Structure of Fractional Order Inverse Filters. *Int. J. Recent Technol. Eng.* **2020**, *8*, 2277–3878.
43. Srivastava, J.; Bhagat, R.; Kumar, P. Analog Inverse Filters Using OTAs. In Proceedings of the 2020 6th International Conference on Control, Automation and Robotics (ICCAR), Singapore, 20–23 April 2020; pp. 627–631.
44. Senani, R. Generation of new two-amplifier synthetic floating inductors. *Electron. Lett.* **1987**, *23*, 1202–1203. [\[CrossRef\]](#)

45. Senani, R. A novel application of four-terminal floating nullors. *Proc. IEEE* **1987**, *75*, 1544–1546. [\[CrossRef\]](#)
46. Wheatley, C.F.; Wittlinger, H.A. OTA obsolete op-amp. *Proc. Nat. Electron. Conf.* **1969**, *4159*, 152–157.
47. Smith, K.; Sedra, A. The current conveyor—A new circuit building block. *Proc. IEEE* **1968**, *56*, 1368–1369. [\[CrossRef\]](#)
48. Sedra, A.; Smith, K. A second-generation current conveyor and its applications. *IEEE Trans. Circuit Theory* **1970**, *17*, 132–134. [\[CrossRef\]](#)
49. AD844: 60 MHz, 2000 V/ $\mu$ s, Monolithic Op Amp with Quad Low Noise Data Sheet (Rev. G). May 2017. Available online: [www.linear.com](http://www.linear.com) (accessed on 29 April 2019).
50. Chen, J.-J.; Tsao, H.-W.; Chen, C.-C. Operational transresistance amplifier using CMOS technology. *Electron. Lett.* **1992**, *28*, 2087–2088. [\[CrossRef\]](#)
51. Acar, C.; Ozoguz, S. A new versatile building block: Current differencing buffered amplifier suitable for analog signal-processing filters. *Microelectron. J.* **1999**, *30*, 157–160. [\[CrossRef\]](#)
52. Biolek, D. CDTA-building block for current-mode analog signal processing. In Proceedings of the 16th European Conference on Circuits Theory and Design, ECCTD'03, Krakow, Poland, 1–4 September 2003; Volume III, pp. 397–400.
53. Biolek, D.; Senani, R.; Biolková, V.; Kolka, Z. Active elements for analog signal processing: Classification, review, and new proposals. *Radioengineering* **2008**, *17*, 15–32.
54. Carlini, H.J.; Youla, D. Network Synthesis with Negative Resistors. *Proc. IRE* **1961**, *49*, 907–920. [\[CrossRef\]](#)
55. Mitra, S.K. A network transformation for active RC networks. *Proc. IEEE* **1967**, *55*, 2021–2022. [\[CrossRef\]](#)
56. Rathore, T.S. Inverse active networks. *Electron. Lett.* **1977**, *13*, 303–304. [\[CrossRef\]](#)
57. Rathore, T.S.; Singhi, B.M. Network transformations. *IEEE Trans. Circuits Syst.* **1980**, *27*, 57–59. [\[CrossRef\]](#)
58. Higashimura, M. Realisation of current-mode transfer function using four-terminal floating nullor. *Electron. Lett.* **1991**, *27*, 170–171. [\[CrossRef\]](#)
59. Normand, G. Floating-impedance realisation using a dual operational-mirrored amplifier. *Electron. Lett.* **1986**, *22*, 521–522. [\[CrossRef\]](#)
60. Anandamohan, P.V. New current-mode biquad on Friend-Deliyannis active RC biquad. *IEEE Trans. Circ. Syst. II Analog Digit. Signal Process.* **1995**, *42*, 225–228.
61. Toumazou, C.; Lidgey, F.J. Current-Feedback Op-Amps—A blessing in disguise. *IEEE Circ. Dev. Mag.* **1997**, *10*, 34–37.
62. Lidgey, F.J.; Hayatleh, K. Current-feedback operational amplifiers and applications. *Electron. Commun. Eng. J.* **1997**, *9*, 176–182. [\[CrossRef\]](#)
63. Soliman, A.M. Applications of the current feedback operational amplifiers. *Analog Integr. Circuits Signal Process.* **1996**, *11*, 265–302. [\[CrossRef\]](#)
64. Senani, R. Realisation of a Class of Analog Signal Processing/Signal Generation Circuits: Novel Configurations Using Current Feedback Op-Amps. *Frequenz* **1998**, *52*, 196–206. [\[CrossRef\]](#)
65. Senani, R.; Bhaskar, D.R.; Singh, A.K. *Current Feedback Operational Amplifiers and Their Applications*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2013.
66. Bhushan, M.; Newcomb, R.W. Grounding of capacitors in integrated circuits. *Electron. Lett.* **1967**, *3*, 148–149. [\[CrossRef\]](#)
67. Newcomb, R.W. *Active Integrated Circuit Synthesis*; Prentice-Hall: Hoboken, NJ, USA, 1967.
68. Gupta, S.S.; Senani, R. Realization of Current-mode SRCOs using All Grounded Passive Elements. *Freq. J. Telecommun.* **2003**, *57*, 26–37.
69. Sharma, R.K.; Senani, R. Multifunction CM/VM Biquads Realized with a Single CFOA and Grounded Capacitors. *AEU Int. J. Electron. Commun.* **2003**, *57*, 301–308. [\[CrossRef\]](#)
70. Yuce, E.; Minaei, S. A Modified CFOA and Its Applications to Simulated Inductors, Capacitance Multipliers, and Analog Filters. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2008**, *55*, 266–275. [\[CrossRef\]](#)
71. Swamy, M.N.S. Modified CFOA, its transpose, and applications. *Int. J. Circuit Theory Appl.* **2015**, *44*, 514–526. [\[CrossRef\]](#)
72. Smith, K.; Sedra, A. Realization of the Chua family of new nonlinear network elements using the current conveyor. *IEEE Trans. Circuit Theory* **1970**, *17*, 137–139. [\[CrossRef\]](#)
73. Chiu, W.; Liu, S.I.; Tsao, H.W.; Chen, J.J. CMOS differential difference current conveyors and their applications. *IEE Proc. Circuits Devices Syst.* **1996**, *143*, 91–96. [\[CrossRef\]](#)
74. Čajka, J.; Vrba, K. The voltage conveyor may have in fact found its way into circuit theory. *AEU—Int. J. Electron. Commun.* **2004**, *58*, 244–248. [\[CrossRef\]](#)
75. Safari, L.; Yuce, E.; Minaei, S.; Ferri, G.; Stornelli, V. A second-generation voltage conveyor (VCII)-based simulated grounded inductor. *Int. J. Circuit Theory Appl.* **2020**, *48*, 1180–1193. [\[CrossRef\]](#)
76. Radwan, A.; Soliman, A.M.; Elwakil, A.S.; Sedeek, A. On the stability of linear systems with fractional-order elements. *Chaos Solitons Fractals* **2009**, *40*, 2317–2328. [\[CrossRef\]](#)
77. Elwakil, A.S. Fractional-Order Circuits and Systems: An Emerging Interdisciplinary Research Area. *IEEE Circuits Syst. Mag.* **2010**, *10*, 40–50. [\[CrossRef\]](#)
78. Tsirimokou, G.; Psychalinos, C. Ultra-low voltage fractional-order circuits using current mirrors. *Int. J. Circuit Theory Appl.* **2016**, *44*, 109–126. [\[CrossRef\]](#)
79. Adhikary, A.; Sen, S.; Biswas, K. Practical Realization of Tunable Fractional Order Parallel Resonator and Fractional Order Filters. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2016**, *63*, 1142–1151. [\[CrossRef\]](#)

- 
80. Kartci, A.; Herencsar, N.; Koton, J.; Brancik, L.; Vrba, K.; Tsirimokou, G.; Psychalinos, C. Fractional-order oscillator design using unity-gain voltage buffers and OTAs. In Proceedings of the 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), Boston, MA, USA, 6–9 August 2017; pp. 555–558.
  81. Tsirimokou, G.; Psychalinos, C.; Elwakil, A.S. *Design of CMOS Analog Integrated Fractional-Order Circuits: Applications in Medicine and Biology*; Springer: Berlin/Heidelberg, Germany, 2017.
  82. Varshney, G.; Pandey, N.; Pandey, R. Electronically Tunable Multifunction Transadmittance-Mode Fractional-Order Filter. *Arab. J. Sci. Eng.* **2021**, *46*, 1067–1078. [[CrossRef](#)]

# K-means Embedded Deep Transform Learning for Hyperspectral Band Selection

Anurag Goel and Angshul Majumdar, *Senior Member, IEEE*

**Abstract**—In clustering based hyperspectral band selection techniques 2D images of each band are usually taken as input samples. Some form of feature extraction on these images is performed before they are input to the clustering algorithm. The clustering algorithm returns the cluster centroids; the bands closest to the centroids are selected as representative bands for each cluster. In this work, we propose a joint representation learning and clustering framework. We embed the popular K-means clustering loss into the newly developing framework of deep transform learning and solve the ensuing formulation via alternating direction method of multipliers. We combine clustering with feature extraction. Application of our proposed solution to the hyperspectral band selection problem shows that we improve over the state-of-the-art by a reasonable margin.

**Index Terms**—hyperspectral, band selection, unsupervised learning, deep learning, clustering.

## I. INTRODUCTION

THE problem of hyperspectral band selection is well known to the remote sensing community [1]. This is one of the strategies to reduce the volume of the hyperspectral datacube. The fundamental postulate of this approach is that since the different spectral bands are correlated, it is possible to choose one representative band from a correlated set.

Clustering is a prevalent strategy to achieve the aforesaid objective. The basic principle is straightforward. The images corresponding to different bands are fed as samples to a clustering algorithm. The algorithm returns a set of cluster centroids. It is unlikely that the centroid directly corresponds to an actual band. Therefore, the band closest to the centroid is chosen as the representative of the cluster. In this fashion, it is possible to choose a pre-defined number of bands.

The majority of recent studies concentrate on modifying clustering algorithms for hyperspectral band selection. Variants of subspace clustering based approaches are a mainstay in this respect [2, 3]. In these studies, a representation of the band images needs to be defined; for example in the aforesaid studies, dictionary learning was used for this purpose. Recent studies on deep learning based clustering combine the two stages – representation learning and clustering [4-7]. This results in a more optimal solution than the piece-meal techniques like [2, 3] where there is no

feedback between the representation learning and clustering operations.

Following the success of deep learning based clustering solutions for hyperspectral band selection, we propose a technique based on a similar vein. However, instead of embedding the clustering loss in a textbook deep learning framework like an autoencoder [4] or a convolutional autoencoder [5, 6], we embed the popular K-means clustering loss into the nascent framework of deep transform learning [8]. The framework has been used in the past for embedding subspace clustering [9].

The advantage of transform learning over autoencoder based approaches will be evident in the next section. In short, an autoencoder is composed of an encoder and a decoder; the clustering loss is embedded after the encoder. Since the network requires learning both encoder and decoder, it needs to learn twice the number of weights as that of a transform, which is practically just an encoder network. In the data constrained scenario of hyperspectral imaging, the requirement of learning twice the number of parameters (compared to transform) is likely to lead to overfitting. This is our motivation behind using deep transform learning.

Given that it is a letter, we can only briefly discuss existing deep learning based clustering studies on band selection in the next section. The generic shortcoming of existing approaches will also be described. The proposed clustering approach is discussed in section III. The engineering solution used for band selection is detailed in section IV. Experimental evaluation is carried out in section V. Conclusions of this work are discussed in section VI.

## II. BACKGROUND

### A. Deep learning based unsupervised band selection

Hyperspectral band selection strategies can be either supervised or unsupervised. Supervised strategies are those that select bands with some specific classification task in mind [10, 11]. Given the limited scope of this work, we are not delving into these cursorily related works. We will only discuss relevant studies on the topic of deep learning based band selection strategies.

An interesting strategy proposed in [4] uses an attention module to define masks that select hyperspectral bands such that the full hyperspectral datacube can be reconstructed from the masked samples. Spectral clustering is applied on top of the selected samples for band selection. This is a two-stage approach. The clustering solution is not embedded into the

A. Goel is associated with Delhi Technological University, Delhi, India and Indraprastha Institute of Information Technology Delhi, 110020, India (e-mail: [anuragg@iitd.ac.in](mailto:anuragg@iitd.ac.in)).

A. Majumdar is with Indraprastha Institute of Information Technology Delhi, 110020, India (e-mail: [angshul@iitd.ac.in](mailto:angshul@iitd.ac.in)).

attention based autoencoder network.

One of the recent studies [5] embeds a self-expression term akin to sparse subspace clustering in the bottleneck layer of a convolutional autoencoder. The idea herein is to learn representations that can select bands that lie on a subspace. The proposal in [6] builds upon [5], the former imposes a graph Laplacian penalty on the convolutional autoencoder and embeds spectral clustering in the bottleneck layer.

The idea proposed in [7] is similar to that of [6]; it uses attention for band selection and an autoencoder for reconstruction from the selected bands. The difference with [6] lies in the fact that [7] does not use any clustering. The selected bands from the attention masks are directly used for downstream processing.

In [13] a simple yet effective strategy is proposed for band selection. It uses the autoencoder framework for reconstruction. By looking at the activation function, the work decides which band to select. This too directly selects the bands without needing any form of clustering.

#### B. Advantage and disadvantage of the autoencoder framework

All the prior studies on deep learning based hyperspectral band selection uses the autoencoder based framework. This is because of the unsupervised setting. Autoencoder is the only framework that can learn unsupervised networks and is mathematically flexible enough to embed a variety of clustering losses.

In theory, one could have also used the restricted Boltzmann machine (RBM). However, the cost function of an RBM does not have the same mathematical flexibility as that of an autoencoder. Autoencoders can be trained via gradient descent/backpropagation, whereas the RBM cost function can only be solved approximately via contrastive divergence. This is the reason, RBMs were hardly ever used in clustering.

The problem with autoencoder is that it requires training an encoder and decoder network. The clustering loss is incorporated after the encoder network; the decoder portion is not relevant for clustering. However, one cannot remove the decoder network, as in that case there would be no layer to backpropagate from and would result in a trivial solution. The need of learning an encoder and a decoder layer leads to learning twice the number of parameters as that of a standard feedforward neural network with target labels. Arguably the encoder and the decoder weights can be tied; however, in practice, deep learning engineers have seen that autoencoders with tied weights do not perform very well. None of the aforesaid studies [4-7] has used a tied autoencoder. Owing to these limitations of autoencoder, it is likely to overfit in data constrained scenarios like hyperspectral image processing.

### III. DEEPLY TRANSFORMED K-MEANS CLUSTERING

The main issue with autoencoder based clustering approach has been discussed in the previous section. This is the reason we resort to deep transform learning and embed K-means in the said formulation.

The popular way to express K-means clustering is via the

following optimization:

$$\min_{h_{ij}, s} \sum_{i=1}^k \sum_{j=1}^n h_{ij} \|z_j - \mu_i\|_2^2 \text{ where } h_{ij} = \begin{cases} 1 & \text{if } z_j \in \text{Cluster } i \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where  $z_j$  denotes the  $j^{\text{th}}$  sample and  $\mu_i$  the  $i^{\text{th}}$  cluster centroid.

In [13], it was shown that (1) can be alternately represented in the form of matrix factorization.

$$\min_H \|Z - ZH^T (HH^T)^{-1} H\|_F^2 \quad (2)$$

where  $Z$  is the data matrix formed by stacking  $z_j$ 's as columns and  $H$  is the matrix of binary indicator variables  $h_{ij}$ .

The basic model for transform learning is  $TX=Z$  where the transform  $T$  analyses the data  $X$  to generate coefficients  $Z$  [14, 15]. The deeper version for the same was proposed in [8], wherein multiple layers of transforms were used to analyze the data. This was expressed as  $T_3 T_2 T_1 X = Z$ . The unsupervised formulation for deep transform learning is given by –

$$\min_{T_1, T_2, T_3, Z, H} \|T_3 T_2 T_1 X - Z\|_F^2 + \lambda \sum_{i=1}^3 (\|T_i\|_F^2 - \log \det T_i) \quad (3)$$

$s.t. T_2 T_1 X \geq 0, T_1 X \geq 0, Z \geq 0$

Here  $T_1$ ,  $T_2$  and  $T_3$  are three layers of transforms. In (3) the term  $\|T_i\|_F^2 - \log \det T_i$  is to prevent the trivial solutions and degenerate solutions to the transforms [14]. The non-negativity constraints in (3) translate to rectified linear unit (ReLU) type activations between the layers. The reason for using ReLU is two-fold. First, it is easy to accommodate in the optimization framework. Second, it is known for its function approximation capability [16].

In our proposed formulation, the input to the K-means ( $Z$ ) is the representation from deep transform learning. Incorporating the K-means cost into the said formulation leads to –

$$\min_{T_1, T_2, T_3, Z, H} \underbrace{\|T_3 T_2 T_1 X - Z\|_F^2 + \lambda \sum_{i=1}^3 (\|T_i\|_F^2 - \log \det T_i)}_{\text{Deep Transform Learning}} + \underbrace{\mu \|Z - ZH^T (HH^T)^{-1} H\|_F^2}_{\text{K-means}} \quad (4)$$

$$s.t. T_2 T_1 X \geq 0, T_1 X \geq 0, Z \geq 0; h_{ij} = \begin{cases} 1 & \text{if } z_j \in \text{Cluster } i \\ 0 & \text{otherwise} \end{cases}$$

To solve (4) we introduce two proxies  $T_1 X = X_2$  and  $T_2 T_1 X = X_3$ . The resulting augmented Lagrangian becomes

$$\min_{T_1, T_2, T_3, X_3, X_2, Z, H} \|T_3 X_3 - Z\|_F^2 + \gamma_1 \|T_2 X_2 - X_3\|_F^2 + \gamma_2 \|T_1 X - X_2\|_F^2 + \mu \|Z - ZH^T (HH^T)^{-1} H\|_F^2 + \lambda \sum_{i=1}^3 (\|T_i\|_F^2 - \log \det T_i) \quad (5)$$

$$s.t. X_3 \geq 0, X_2 \geq 0 \text{ and } Z \geq 0; h_{ij} = \begin{cases} 1 & \text{if } x_j \in \text{Cluster } i \\ 0 & \text{otherwise} \end{cases}$$

The hyper-parameters  $\gamma_1$  and  $\gamma_2$  are usually different, but in the given scenario they represent the relative significance of



different layers. We do not see any reason to give more significance to one layer over others; therefore we assign equal importance; hence we keep  $\gamma_1=\gamma_2=1$ ; this leads to

$$\begin{aligned} & \min_{T_1, T_2, T_3, X_3, Z, H} \|T_3 X_3 - Z\|_F^2 + \|T_2 X_2 - X_3\|_F^2 \\ & + \|T_1 X - X_2\|_F^2 + \mu \|Z - ZH^T (HH^T)^{-1} H\|_F^2 \\ & + \lambda \sum_{i=1}^3 (\|T_i\|_F^2 - \log \det T_i) \end{aligned} \quad (6)$$

$$s.t. X_3 \geq 0, X_2 \geq 0 \text{ and } Z \geq 0; h_{ij} = \begin{cases} 1 & \text{if } z_j \in \text{Cluster } i \\ 0 & \text{otherwise} \end{cases}$$

We solve (6) using the alternating direction method of multipliers (ADMM) [17]. Using ADMM (6) can be segregated into the following sub-problems –

$$P1: \min_{T_3} \|T_3 X_3 - Z\|_F^2 + \lambda (\|T_3\|_F^2 - \log \det T_3)$$

$$P2: \min_{T_2} \|T_2 X_2 - X_3\|_F^2 + \lambda (\|T_2\|_F^2 - \log \det T_2)$$

$$P3: \min_{T_1} \|T_1 X - X_2\|_F^2 + \lambda (\|T_1\|_F^2 - \log \det T_1)$$

$$P4: \min_{X_3} \|T_3 X_3 - Z\|_F^2 + \|T_2 X_2 - X_3\|_F^2 \text{ s.t. } X_3 \geq 0$$

$$P5: \min_{X_2} \|T_2 X_2 - X_3\|_F^2 + \|T_1 X - X_2\|_F^2 \text{ s.t. } X_2 \geq 0$$

$$P6: \min_Z \|T_3 X_3 - Z\|_F^2 + \mu \|Z - ZH^T (HH^T)^{-1} H\|_F^2 \text{ s.t. } Z \geq 0$$

$$P7: \min_H \|Z - ZH^T (HH^T)^{-1} H\|_F^2; h_{ij} = \begin{cases} 1 & \text{if } x_j \in \text{Cluster } i \\ 0 & \text{otherwise} \end{cases}$$

Solutions to P1-P3 is the standard transform update given in [15]. It has a closed form solution.

$$\min_T \|TX - Z\|_F^2 + \lambda (\|T\|_F^2 - \log \det T)$$

$$XX^T + \lambda I = LL^T \text{ (Cholesky Decomposition)}$$

$$L^{-1} XZ^T = USV^T \text{ (Singular Value Decomposition)}$$

$$T = 0.5U (S + (S^2 + 2\lambda I)^{1/2}) V^T L^{-1}$$

Solutions to sub-problems P4, P5 are akin to Tikhonov Regularization. The solution can be obtained as follows.

$$\min_X \|Z - TX\|_F^2 + \|X - C\|_F^2, T_2 X_2 \text{ or } T_1 X = C = \text{constant}$$

$$\Rightarrow \hat{X} = (T^T T + I)^{-1} (T^T Z + C)$$

Here we have not specifically accounted for the positivity constraint. This is because the exact solution would require forward-backward type splitting. We prefer the approximate solution where the aforesaid closed form solution is thresholded to keep only non-negative values. The same approach was followed in [9].

The solution to P6 is obtained by taking the derivative of it and equating it to zero.

$$\nabla_Z (\|T_3 X_3 - Z\|_F^2 + \mu \|ZK\|_F^2) = 0,$$

$$\text{where } I - H^T (HH^T)^{-1} H = K = \text{constant}$$

$$T_3 X_3 = Z(I + \mu K) \Rightarrow Z = T_3 X_3 (I + \mu K)^{-1}$$

In the solution of Z, we follow the same approach as before. We do not explicitly incorporate the non-negativity constraint but put the negative values in it to zero once it is solved in every iteration.

The solution to P7 is K-means clustering. We have used the vanilla solution in this work.

This concludes the derivation of the algorithm. The complexity of updating the transforms is dependent on the Cholesky and singular value decompositions; both of which have a complexity of  $O(n^3)$ . The update for the coefficients has a complexity of  $O(n^2)$ . K-means is ideally an NP-hard problem, but the algorithm used here has a complexity of  $O(t*k*n^2)$  where  $t$  is the number of loops and  $k$  is the number of clusters.

#### IV. APPLICATION TO HYPERSPECTRAL BAND SELECTION

Most of the clustering based band selection techniques work with individual band images, i.e. with  $x_i$  where  $i$  denotes the band. Some kind of a representative feature ( $z_i$ ) is learnt from  $x_i$ . This representation is passed into a clustering algorithm where the number of clusters is specified (this is the same as the number of bands to be chosen). The clustering algorithm outputs the cluster centroid. The band images that are closest to the cluster centroid in some sense (normed distance, cosine similarity, etc.) are finally selected.

We do not follow this approach for two reasons –

- 1) Trying to learn deep transforms from a few band images (~hundreds) will lead to over-fitting.
- 2) Clustering from a few samples (~hundreds) will not lead to robust cluster centroids.

We divide the entire hyperspectral datacube into spatial patches.  $P_{x_{ik}}$  then denotes the  $k^{th}$  patch of the  $i^{th}$  band image. These patches are input to our proposed deeply transformed K-means (DTKM) clustering algorithm. This requires solving the following problem –

$$\begin{aligned} & \min_{T_1, T_2, T_3, Z, H} \sum_k \left( \left( \sum_i \|T_3 T_2 T_1 P_{x_{ik}} - z_{ik}\|_F^2 \right) + \mu \|Z_i - Z_i H^T (HH^T)^{-1} H\|_F^2 \right) \\ & + \lambda (\|T_1\|_F^2 - \log \det T_1) + (\|T_2\|_F^2 - \log \det T_2) + (\|T_3\|_F^2 - \log \det T_3) \\ & s.t. T_2 T_1 X \geq 0, T_1 X \geq 0, Z \geq 0; h_{ij} = \begin{cases} 1 & \text{if } z_j \in \text{Cluster } i \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

The formulation states that K-means clustering is carried out on individual patches (instead of the images) and the transforms are globally learnt on all the patches.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Patch 1	15	29	54	11	79
Patch 2	32	29	42	4	77
Patch 3	15	45	56	7	84
Patch 4	15	7	46	4	77
Patch 5	87	16	56	16	77
Patch 6	6	71	15	5	69
Patch 7	15	33	56	4	51
Patch 8	54	29	55	5	26
Selected Band	15	29	56	4	77

Fig. 1. Table Look-up

By learning on a vast number of patches, the problem of over-fitting is resolved. Similarly, as the clustering is run



patch-wise robustness is ensured following a simple table look-up strategy. For each patch, we will get a band that is closest to the  $j^{th}$  cluster in Euclidean sense. We see which band is appearing to be the closest in the majority of patches and select the same. A toy example is given in Fig. 1. We consider an image with 8 patches and want to select 5 bands. The cluster centre corresponding to each patch is shown in the white portion of Fig. 1. The final selection, considering all the patches is obtained from the statistical mode of the column (shown in red).

One can argue that an identifiability problem may arise in such a scenario. This is avoided via deterministic initialisation of all cluster centres for each band.

## V. EXPERIMENTAL EVALUATION

We evaluate our proposed technique on two benchmark datasets – Indian Pines<sup>1</sup> and Pavia University<sup>2</sup>. The standard pre-processing steps are performed on these datasets.

- The Indian Pines dataset was collected by the Airborne Visible/Infrared Imaging Spectrometer in Northwestern Indiana, with a size of  $145 \times 145$  pixels with a spatial resolution of 20 m per pixel and 10-nm spectral resolution over the range of 400–2500 nm. As is the usual protocol, the work uses 200 bands, after removing 20 bands affected by atmospheric absorption. There are 16 classes.
- This Pavia university dataset is acquired by a reflective optics system imaging spectrometer (ROSIS). The image is of  $610 \times 340$  pixels covering the Engineering School at the University of Pavia, which was collected under the HySens project managed by the German Aerospace Agency (DLR). The ROSIS-03 sensor comprises 115 spectral channels ranging from 430 to 860 nm. In this dataset, 12 noisy channels have been removed and the remaining 103 spectral channels are investigated in this paper. The spatial resolution is 1.3 m per pixel. The available training samples of this data set cover 9 classes of interest.

There is no straightforward way to evaluate the efficacy of the selected bands. Therefore, it must be evaluated based on some other criterion. Here we have used classification as the evaluation criterion; i.e. after band selection, the selected bands are passed onto a classifier. In particular, we have used kernel sparse representation classifier [18] in this work since it is a highly cited work on hyperspectral classification.

We benchmark against several deep learning techniques – unsupervised deep subspace clustering (UDSC) [5], graph convolutional neural network (GCNN) [6] and BS-Nets [7]. We also compare with deeply transformed subspace clustering (DTSC) [9]; however, note that DTSC was never meant to be used for the said purpose and hence we use the table look-up approach for selecting the bands.

Our proposed technique requires the specification of two parameters –  $\lambda$  and  $\mu$ . We set the value of  $\lambda=1$  for all the experiments; this value has been found to work well on almost all transform learning based formulations. The parameter  $\mu$

controls the relative importance of the deep transform learning and the K-means costs. There is no reason to give more weightage to one than the other; therefore we set  $\mu=1$ . In the experiments, we used overlapping patches of size  $8 \times 8$ . The patches are obtained by shifting the patches 2 pixels in both horizontal and vertical direction, as and when needed. The best results were obtained for a three-layer architecture where the number of nodes in subsequent layers were halved, i.e. the architecture was 64-32-16.

The experiments were carried out on a 64 bit Intel Core i5-8265U CPU @ 1.60GHz, 16 GB RAM running Ubuntu. All the techniques were coded in Python. The run times of different algorithms are shown in Table I. One can see that our proposed method is the fastest. It is faster than DTSC because, unlike the latter, we do not need to solve a complex iterative optimization problem every time in the clustering step. Both our proposed method and DTSC are solved using ADMM and hence are considerably faster compared to the traditional backpropagation based approach used by UDSC, GCNN and BS-Nets.

TABLE I COMPARISON OF RUNTIMES IN SECONDS

Algorithm	Pavia University			Indian Pines		
# bands	10	20	30	10	20	30
UDSC	1002	1338	1609	553	819	1012
GCNN	1190	1581	2086	678	1006	1375
BS-Nets	984	1403	1760	425	664	808
DTSC	817	1119	1531	316	497	633
Proposed	<b>465</b>	<b>650</b>	<b>803</b>	<b>193</b>	<b>302</b>	<b>386</b>

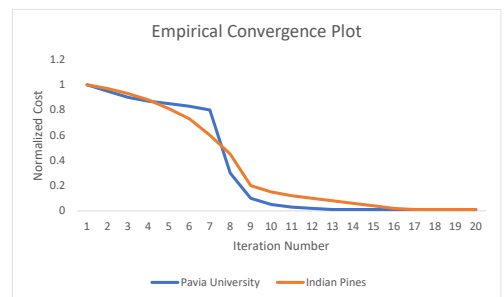


Fig. 2. Empirical convergence plots

Although we could not give any proof regarding the convergence of our proposed algorithm, we found that it does indeed converge to a local minimum given the choice of parameters. The empirical convergence plot is shown in Fig. 2 for 20 bands. Similar convergence results were obtained for the other number of bands.

For measuring classification accuracy we use the three standard measures in hyperspectral imaging – overall accuracy (OA), average accuracy (AA), and Kappa coefficient. In Fig. 3. the plots for two different datasets and three different metrics are shown.

<sup>1</sup> <https://paperswithcode.com/dataset/indian-pines>

<sup>2</sup> <https://paperswithcode.com/dataset/pavia-university>



Fig. 3. Detailed Results

From Fig. 3, one can note that our method yields the best results, especially for a fewer number of bands. Our proposed method is closely followed by GCNN and DTSC. GCNN sometimes performs better than ours when the number of chosen bands is higher. In general, for all the techniques, one can notice that the results improve as the number of bands increases from 5 to 15 or 20, after that the results deteriorate. The same has been observed by other studies like [5].

## VI. CONCLUSION

There are two major contributions of the work. First, it proposes a new deep learning based clustering approach by embedding the popular K-means algorithm into the deep transform learning framework. Second, it shows how the said clustering algorithm can be used for hyperspectral band selection. The proposed approach is not only more accurate with a fewer number of selected bands but also has the fastest run-times.

In this work, we have used a fully connected network for transforms. The unsupervised framework of convolutional transform learning has been in the works for a few years [19]. In the future, we will embed different clustering losses in the said framework and apply them to the problem of hyperspectral band selection.

## ACKNOWLEDGEMENT

The authors are thankfully to the Center for Artificial Intelligence at IIIT Delhi for its partial support.

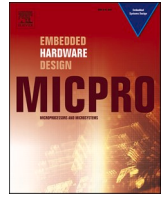
## VII. REFERENCES

- [1] W. Sun and Q. Du, "Hyperspectral Band Selection: A Review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, no. 2, pp. 118-139, 2019.
- [2] H. Zhai, H. Zhang, L. Zhang and P. Li, "Laplacian-Regularized Low-Rank Subspace Clustering for Hyperspectral Image Band Selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 3, pp. 1723-1740, 2019.
- [3] H. Zhai, H. Zhang, L. Zhang and P. Li, "Nonlocal Means Regularized Sketched Reweighted Sparse and Low-Rank Subspace Clustering for Large Hyperspectral Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 5, pp. 4164-4178, 2021.
- [4] Z. Dou, K. Gao, X. Zhang, H. Wang and L. Han, "Band Selection of Hyperspectral Images Using Attention-Based Autoencoders," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 1, pp. 147-151, 2021.
- [5] M. Zeng, Y. Cai, Z. Cai, X. Liu, P. Hu and J. Ku, "Unsupervised Hyperspectral Image Band Selection Based on Deep Subspace Clustering," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 12, pp. 1889-1893, 2019.
- [6] Y. Cai, Z. Zhang, Z. Cai, X. Liu, X. Jiang and Q. Yan, "Graph Convolutional Subspace Clustering: A Robust Subspace Clustering Framework for Hyperspectral Image," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 5, pp. 4191-4202, 2021.
- [7] Y. Cai, X. Liu and Z. Cai, "BS-Nets: An End-to-End Framework for Band Selection of Hyperspectral Image," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 3, pp. 1969-1984, 2020.
- [8] J. Maggu and A. Majumdar, "Unsupervised Deep Transform Learning," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6782-6786, 2018.
- [9] J. Maggu, A. Majumdar, E. Chouzenoux and G. Chierchia, "Deeply Transformed Subspace Clustering," *Signal Processing*, Vol. 174, 107628, 2020.
- [10] W. Kang, Y. Xiang, F. Wang and H. You, "DO-Net: Dual-Output Network for Land Cover Classification From Optical Remote Sensing Images," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1-5, 2022.
- [11] L. Mou, S. Saha, Y. Hua, F. Bovolo, L. Bruzzone and X. X. Zhu, "Deep Reinforcement Learning for Band Selection in Hyperspectral Image Classification," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-14, 2022.
- [12] H. Sun, J. Ren, H. Zhao, P. Yuen and J. Tschanerl, "Novel Gumbel-Softmax Trick Enabled Concrete Autoencoder With Entropy Constraints for Unsupervised Hyperspectral Band Selection," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-13, 2022.
- [13] C. Bauckhage, "K-means clustering is matrix factorization," *arXiv preprint arXiv:1512.07548*, 2015.
- [14] S. Ravishanker and Y. Bresler, "Learning Sparsifying Transforms," *IEEE Transactions on Signal Processing*, vol. 61, no. 5, pp. 1072-1086, 2013.
- [15] S. Ravishanker, B. Wen and Y. Bresler, "Online Sparsifying Transform Learning-Part I: Algorithms," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 625-636, 2015.
- [16] B. Hanin, "Universal function approximation by deep neural nets with bounded width and relu activations," *Mathematics*, vol. 7, no. 10, pp.992, 2019.
- [17] Y. Wang, W. Yin and J. Zeng, "Global convergence of ADMM in nonconvex nonsmooth optimization," *Journal of Scientific Computing*, vol. 78, no.1, pp. 29-63, 2019.
- [18] Y. Chen, N. M. Nasrabadi and T. D. Tran, "Hyperspectral Image Classification via Kernel Sparse Representation," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 1, pp. 217-231, 2013.
- [19] J. Maggu, E. Chouzenoux, G. Chierchia and A. Majumdar, "Deep Convolutional Transform Learning", *ICONIP*, pp. 300-307, 2020



Contents lists available at ScienceDirect

## Microprocessors and Microsystems

journal homepage: [www.elsevier.com/locate/micpro](http://www.elsevier.com/locate/micpro)

# On improving the performance of dynamic positive-feedback source-coupled logic (D-PFSCL) through inclusion of transmission gates

Ranjana Sivaram<sup>a</sup>, Kirti Gupta<sup>b</sup>, Neeta Pandey<sup>a,\*</sup>

<sup>a</sup> Department of Electronics and Communication Engineering, Delhi Technological University, Delhi- 110042, India

<sup>b</sup> Department of Electronics and Communication Engineering, Bharati Vidyapeeth's College of Engineering, Delhi-110063, India

## ARTICLE INFO

## Keywords:

D-PFSCL

Dynamic circuits

Positive feedback

Transmission gates

## ABSTRACT

Dynamic positive feedback source-coupled logic (D-PFSCL) gates are used as a low power alternative in mixed signal applications over static positive feedback source coupled logic (PFSCCL) gates. In this paper, transmission gates (TG) are included in the dynamic positive-feedback source coupled logic (D-PFSCL) architecture. The proposed scheme allows implementation of logic functions which in existing schemes is restricted to NOR/OR only or three input AND/OR only realization forms. The significant reduction in the number of stages, dynamic current sources and self-timed buffer with proposed scheme leads to efficient D-PFSCL design in comparison to the existing ones. Simulations using PTM 65 nm technology node at frequency of 1 GHz are performed. A maximum improvement of 40.8%, 73% and 97% in delay, dynamic power consumption and EDP respectively is been observed for D-PFSCL exclusive-OR(XOR2) gate implemented using proposed scheme with respect to existing D-PFSCL schemes. Further, a study on the trade-off between power dissipation and frequency of operation is included for the completeness.

## 1. Introduction

High performance circuits are imperative in high-resolution mixed-signal applications. The resolution of the analog circuits is restricted by the switching current in digital circuits. The conventional CMOS logic is not preferred in this realm due to generation of large switching noise [1–4]. Consequently, alternate logic styles are explored where transients during switching event get canceled and/or work on reduced voltage swing. Source coupled logic (SCL) [4–8] is one among these styles which works on current steering principle with reduced voltage swing. The conventional single-ended SCL gates employ a reference voltage source in pull down network which is further replaced by a positive feedback that results in considerable speed improvement. The gates are known as positive feedback source coupled logic (PFSCCL) gates [7]. However, the PFSCCL gates consume static power due to presence of constant current source which mitigates its usage in low power applications.

A low power alternative is the dynamic version of PFSCCL style introduced as dynamic PFSCCL style (D-PFSCL) in literature [9]. A D-PFSCL gate use pre-charge transistors and dynamic current source for high performance. There exist two schemes for circuit realization. The basic gate in conventional D<sub>4</sub>-PFSCL scheme realize a function in

NOR/OR form which infers cascading of D-PFSCL gates for realization of complex logic functions [9]. Since, the dynamic gates cannot be cascaded directly; therefore, intermittent self-timed buffers (STBs) are placed for proper operation [9]. Recently, a new scheme for function realization in D-PFSCL is presented that embeds AND/OR functionality in the gate itself by introducing a low threshold transistor in the pull down network [10]. This scheme [10] reduces overall gate count/stages/ intermittent buffers for complex logic function realization in comparison to [9]. The scheme [10] is referred to as D-PFSCL-N in the context of this paper. It is obvious that cascading of D-PFSCL gates deteriorates the overall performance. Thus, it is required that the logic function realization style for a D-PFSCL gate must be worked upon so that it is not restricted to any particular form. In this paper, an alternate scheme wherein transmission gates (TG) are used to realize logic function is proposed.

The paper is organized in four sections. The existing schemes in D-PFSCL style are briefly reviewed in Section 2. The proposed scheme is illustrated through various configurations in Section 3. The verification of the proposal is done through simulations and performance comparison with the existing ones. The impact of process variation in process, power supply on various performance parameter of a gate realized based

\* Corresponding author at: Electronics and Communication Engineering, Delhi Technological University, India.

E-mail address: [n66pandey@rediffmail.com](mailto:n66pandey@rediffmail.com) (N. Pandey).

<https://doi.org/10.1016/j.micpro.2022.104521>

Received 23 September 2021; Received in revised form 24 March 2022; Accepted 29 March 2022

Available online 31 March 2022

0141-9331/© 2022 Elsevier B.V. All rights reserved.

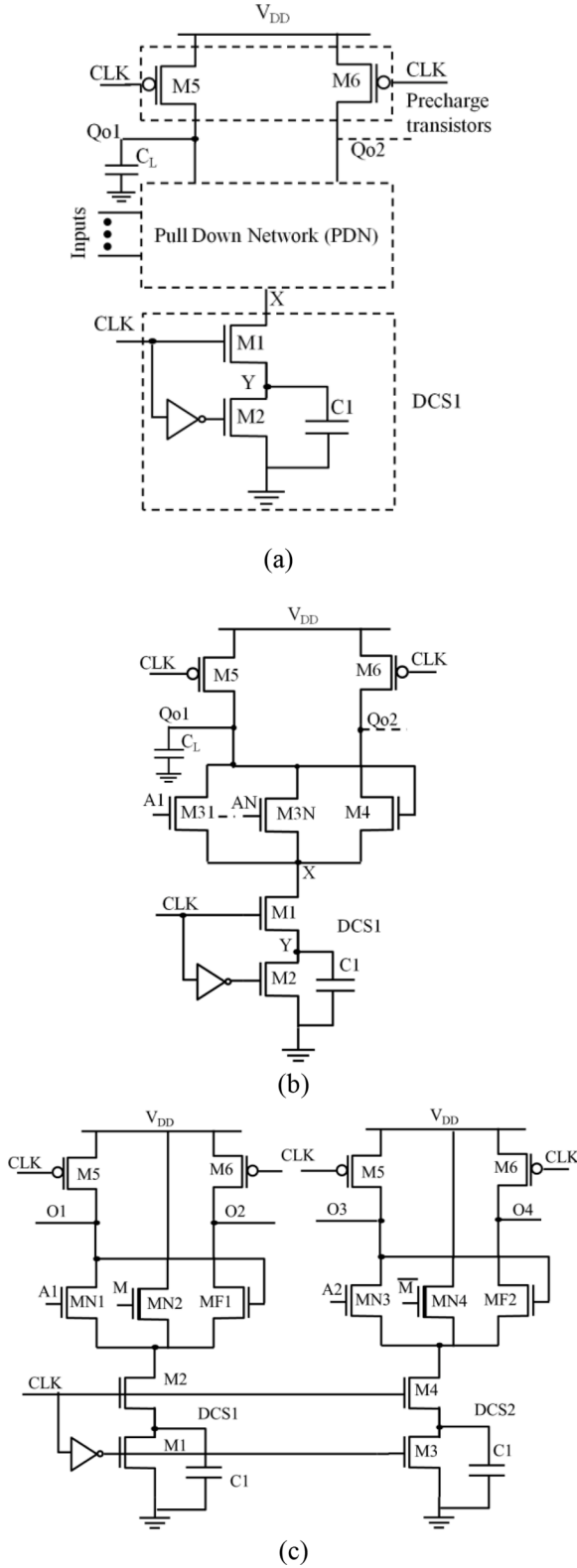


Fig. 1. (a) Generic architecture of a D-PFSCl gate (b) generic NOR based D-PFSCl gate [9] (c) generic D-PFSCl-N gate [10].

on the proposed scheme is studied and comparison with the existing is performed. The paper is concluded in Section 4.

## 2. Existing schemes

The basic structure of a D-PFSCl gate consist of precharge transistors, a pull down evaluation network (PDN) and a dynamic current source (DCS) [9] as shown in Fig. 1a. A clock signal (CLK) drives the precharge transistors and the DCS. Depending on the value of the CLK signal, the circuit operates in precharge phase (CLK = '0') or evaluation phase (CLK = 'V<sub>DD</sub>'). In the precharge phase, the output node is precharged to V<sub>DD</sub>, the capacitor C1 in DCS is discharged to ground potential and the changes in the inputs may take place. It is worth noting that the changes in the input have no effect at the output [9]. Conversely, in the evaluation phase, the circuit gets is evaluated. If a current path is established between the output node Qo1 and the ground, the output reduces to V<sub>DD</sub>-V<sub>SWING</sub> else it remains at V<sub>DD</sub>. The capacitor C1 gets charged in each evaluation phase either due to charge transfer either from output node or the power supply [9]. In literature, there exist two different schemes of realizing a logic function in the PDN evaluation network. The first scheme is elaborated in Fig. 1b where the input transistors are connected in parallel [9] and the output is further coupled through a feedback transistor. In other words, the NOR based realization of logic function is followed. Recently, an efficient realization scheme is proposed [10]. The triple tail cell concept is applied for the function realization while using the same implementation of DCS and precharge transistors. This modification introduced AND/OR functionality in D-PFSCl style. A generalized D-PFSCl-N gate to implement a three-input logic function is shown in Fig 1c. The scheme uses two triple tail cells driven by two complementary inputs which ensures that at particular instance only one of the two triple tail cell is activated and affects the output. The other cell is deactivated and does not influence the logic level at the output. The middle transistor in each of the triple tail cell has low threshold voltage transistor for proper activation and deactivation of the cell. Both of the D-PFSCl variants employ self-timed buffers in between the two stages since the direct connection of D-PFSCl gates leads to erroneous output. A self-timed buffer is cascade of two clocked inverters ensuring the sequential evaluation of the D-PFSCl style [4,10].

The NOR based realization scheme is efficient for realizing NOR functions while for AND/OR functions it results in high gate count which degrades the performance of the system in terms of both of power and delay values. The second realization based on triple tail cell implements AND/OR functionality but is restricted to three inputs per gate and require low threshold voltage transistors. Additionally, they require self-timed buffers which add to the cascading, delay and power. This paper proposes a new scheme based on including TG in PDN for logic function realization. The scheme is elaborated further in the next section.

## 3. Proposed scheme

This scheme is based on the inclusion of transmission gates in the PDN of a D-PFSCl gate. It consists of TG based logic network and a D-PFSCl gate. A generalized gate based on the proposed scheme is shown in Fig. 2. The logic function is implemented using the TGs and the corresponding output is fed to the D-PFSCl gate. During the precharge phase, the output node is charged to V<sub>DD</sub> while the capacitor C1 is discharged to ground potential. Additionally, the logic function is evaluated through the TG based network and the final output from the network is available at the input of the D-PFSCl gate. In the subsequent evaluation phase, the D-PFSCl gate evaluates the implemented logic. In Fig. 2, the proposed D-PFSCl for various functionalities is shown.

As inputs in dynamic circuits may make a single transition during evaluation, the output of the transmission gate has to be ready before the start of the evaluation phase, meaning that the propagation delay through the TG based network should be less than the precharge dura-



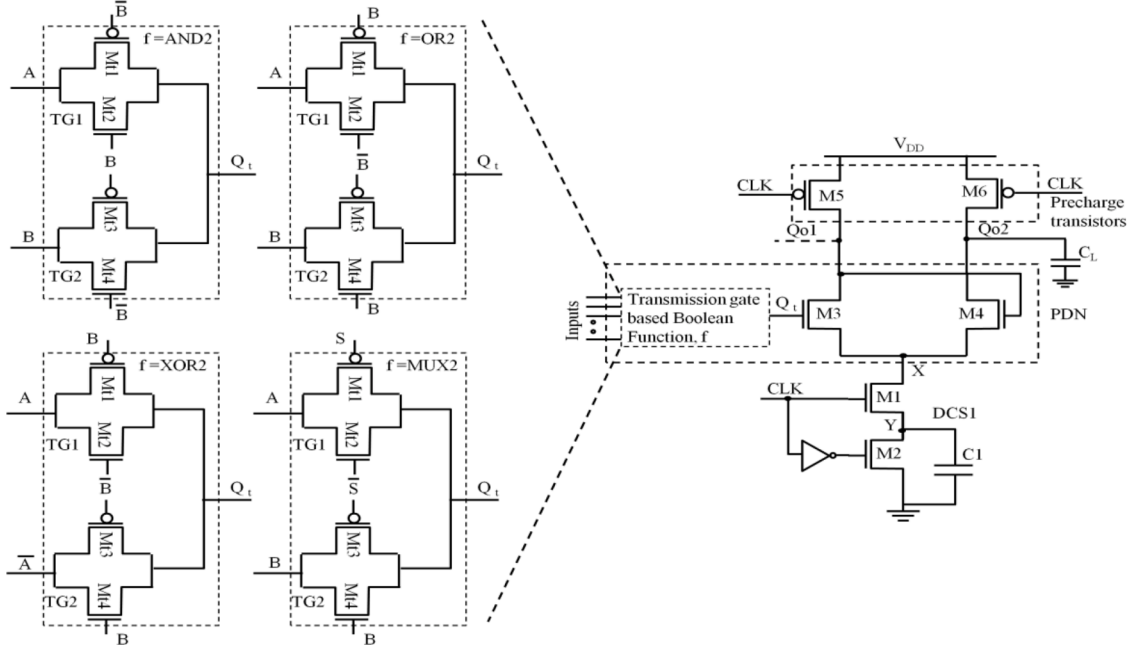


Fig. 2. Generic D-PFSCl gate based on the proposed scheme with output Qo1 or output Qo2.

tion of the clock. If a single TG stage has a combinational delay of  $T_{comb}$ , then for a TG based network with  $N$  stages, the total combinational delay is  $NT_{comb}$ . Assuming the duty cycle of the clock to be 50%, the correct operation of the proposed D-PFSCl gate requires the following relation between  $NT_{comb}$  and clock period ( $T_{CLK}$ ).

$$NT_{comb} < T_{CLK}/2 \quad (1)$$

The effectiveness of the proposed scheme in circuit design is evaluated through the implementation of various functions and simulation using PTM 65 nm CMOS technology parameters in Symica DE tool with a power supply, voltage swing, clock frequency and load capacitance of 1.1 V, 0.4 V, 1 GHz and 50 fF respectively. Firstly, a two input exclusive-OR (XOR2) gate is designed in the existing schemes. The corresponding schematics are drawn in Fig. 3. The NOR based realization in Fig. 3a has three two inputs D-PFSCl NOR/OR gates arranged in two stages with intermittent STBs. The second realization employs two triple-tail cells having individual and separate DCSs. Thus, both of them employ multiple DCSs. However, the XOR2 gate realized using the proposed scheme implements XOR2 functionality in TG network which further drives the D-PFSCl gate in order to produce the final output as illustrated in Fig. 3c. The simulation waveforms for the proposed D-PFSCl XOR2 gate is shown in Fig. 3d followed by output timing waveforms in Fig. 3e which clearly shows the fast transition in the proposed D-PFSCl XOR2 gate in comparison to the XOR2 gate based on existing schemes.

Further, process Monte Carlo simulations are carried out to study the impact of process variations on precharge delay, evaluation delay and dynamic power. The simulations are performed by varying the process parameter - threshold voltage with Gaussian distribution having  $\pm 3\sigma = 105$  mV for CMOS 65 nm technology [11]. The observations are summarized in Table I. The results signify that the performance of the circuit is overall improved by following the proposed scheme. Also, the proposed D-PFSCl XOR2 gate shows a 6.5%, 12.7% and 3.1% variation in precharge delay, evaluation delay and dynamic power respectively. The obtained values are least in comparison to the XOR2 gate design based

on existing schemes. The Monte Carlo results for the proposed D-PFSCl XOR2 gate are plotted in Fig. 4. The process corner analysis at FF, FS, SF and SS for all the XOR2 gate designs is summarized in Fig. 5. It is observed that the proposed D-PFSCl XOR2 gate operates correctly in all corners. The best case and the worst case evaluation delay is obtained at FF (29 ps) with maximum power consumption (28.5  $\mu$ W) and SS corner (45 ps) with minimum power consumption (22.4  $\mu$ W) respectively. For the precharge delay, it is observed that best case delay corresponds to SF process corner with 19.2 ps and worst case delay corresponds to FS process corner with 35.4 ps due to the Fast and Slow PMOS precharge transistors.

The effect of supply voltage reduction on these parameters is plotted in Fig. 6. An increasing trend in delay values (Fig. 6a, b) is observed with reducing supply voltage while a decreasing trend exists for dynamic power (Fig. 6c). The changes in energy delay product (EDP) were also noted and a minimum EDP point is identified for all the XOR2 gates as shown in Fig. 6d. At the minimum EDP point, the proposed D-PFSCl XOR2 gate shows a reduction of 98% and 72% in EDP value with respect to D-PFSCl [9] and D-PFSCl-N [10] based XOR2 gate respectively.

As the reduction in supply voltage affects the delay, therefore the maximum clock frequency at which the gate can operate correctly is analyzed through the results plotted in Fig. 7a. It is seen that the proposed D-PFSCl XOR2 gate can operate at higher frequency as compared to the other two designs. From the figure, it is also observed that the maximum operating frequency increases with increasing supply voltage, with a maximum operating frequency of 11 GHz for proposed D-PFSCl XOR2 at supply voltage of 1.1 V. Further, an insight to dynamic power consumption with clock frequency for a XOR2 gate is provided in Fig. 7b. It is seen that for the same power consumption, the proposed architecture offers an improvement in speed of 450% while for the same speed; it offers an 80% reduction in power consumption with respect to D-PFSCl.

In order to provide the complete view of the investigation, the other common functionalities such as 2: 1 multiplexer (MUX2), 8:1

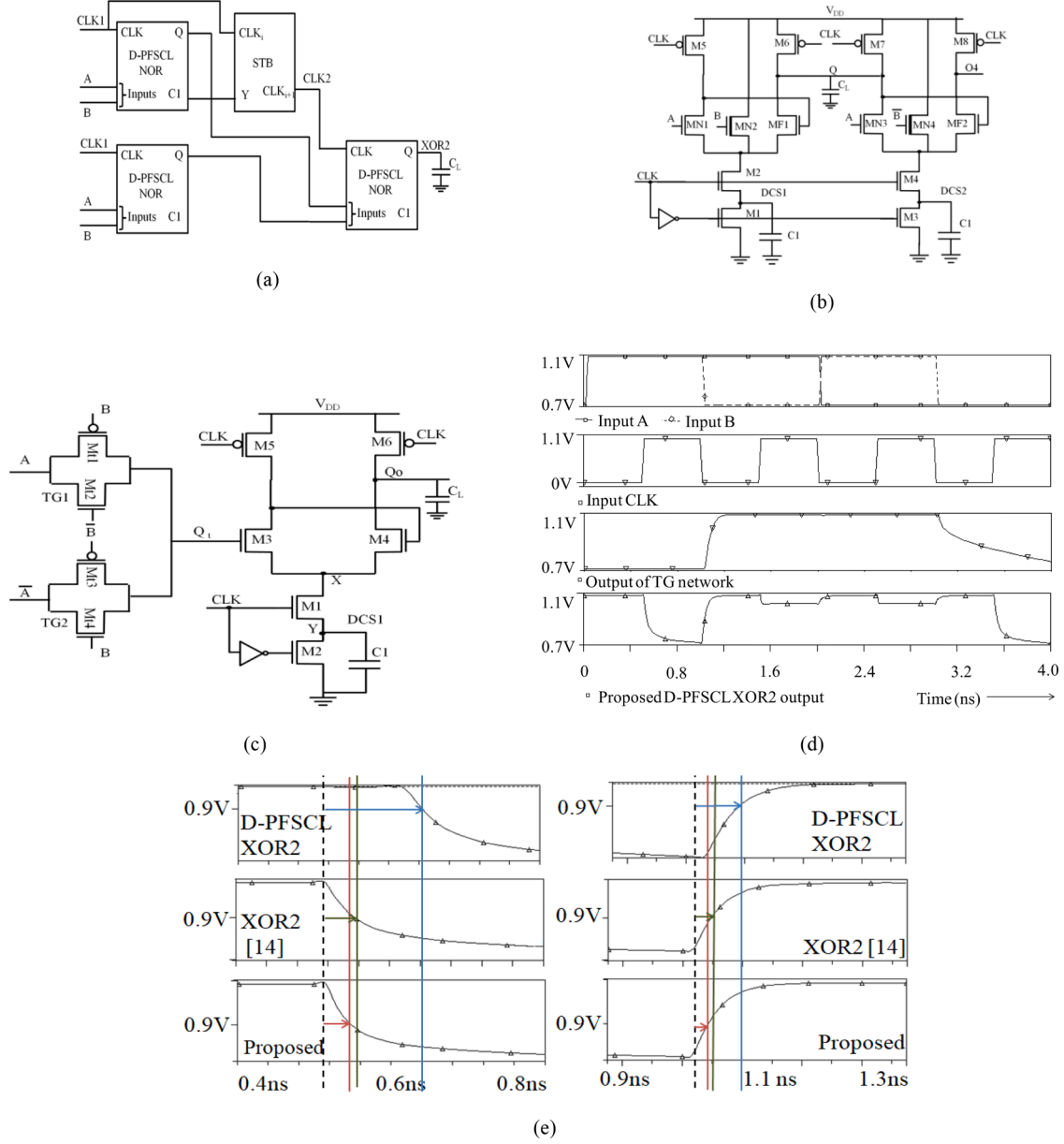


Fig. 3. XOR2 gate based on (a) DPFSCl [9] (b) D-PFSCl-N [10] (c) proposed D-PFSCl (d) Transitions at the XOR2 gate output.

Table I

D-PFSCl XOR2 gate - Monte Carlo simulation results.

Parameter	D-PFSCl [9]			D-PFSCl-N [10]			Proposed D-PFSCl		
	$\mu$	$\sigma$	%	$\mu$	$\sigma$	%	$\mu$	$\sigma$	%
$\tau_{pre}$ (ps)	37.6	2.67	7.1%	25.8	3.8	14.7%	27.3	1.77	6.5%
$\tau_{eval}$ (ps)	131.0	22.0	16.9%	53.3	12.0	22.5%	33.8	4.3	12.7%
$P_{dyn}$ ( $\mu$ W)	88.7	3.303	3.7%	59.3	2.54	4.2%	24.6	0.76	3.1%

$\mu$  = Mean,  $\sigma$  = Standard deviation, % = percentage variation.

multiplexer (MUX8), full adder are implemented and simulated in all the three schemes under the same simulation conditions. The findings are summarized in Table II. The advantage in terms of delay and power in the proposed scheme based function implementation is evident from the results.

It is pertinent to mention that each stage of the TG network has a delay of 136 ps (observed through simulations), therefore the proposed circuit will operate properly with 3 stage TG network with 1 GHz clock as the total delay of TG network ( $136 \text{ ps} \times 3 = 408 \text{ ps}$ ) < precharge

duration ( $1 \text{ GHz}/2 = 500 \text{ ps}$ ) for a 50% duty cycle clock. However, for larger logic circuits needing more number of stages of TG network, let us say MUX16, multiplier etc. cascading of multiple proposed D-PFSCl gates would be required.

#### 4. Conclusion

This paper proposes a modified scheme for logic function realization in D-PFSCl style. The embodiment of transmission gate in D-PFSCl is



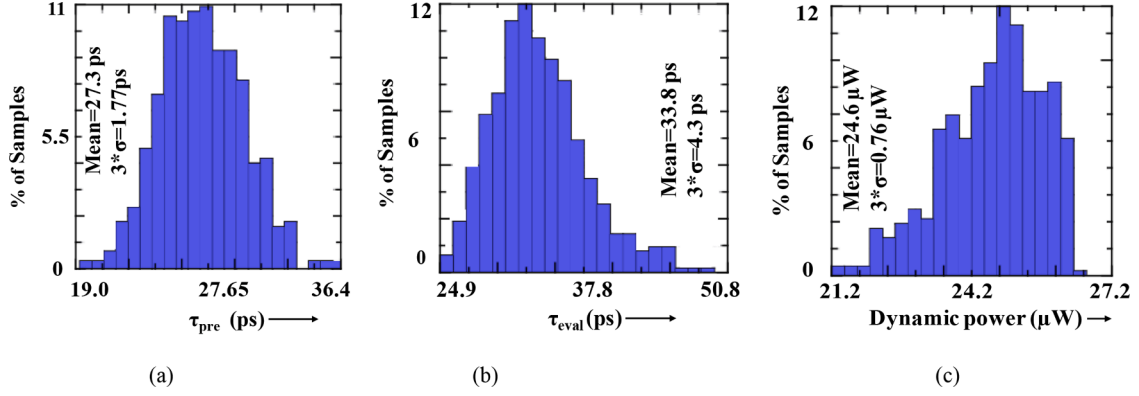


Fig. 4. Monte Carlo simulation results (500 runs) for the proposed D-PFSCS XOR2 gate (a) precharge delay (b) evaluation delay (d) dynamic power.

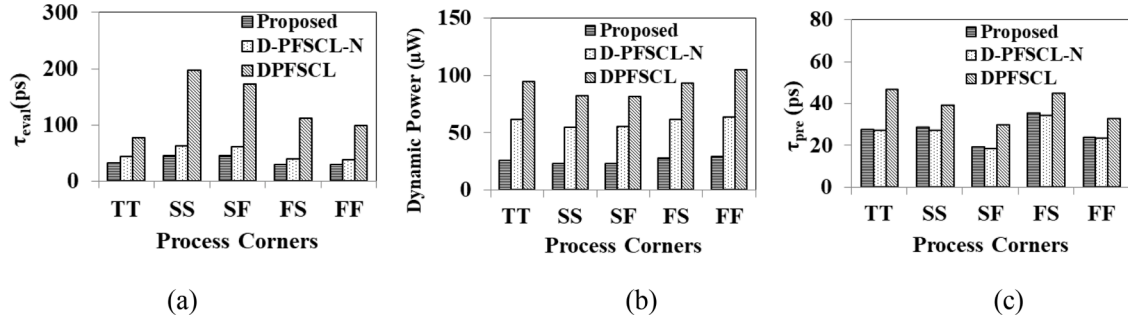


Fig. 5. Process corner results of the D-PFSCS XOR2 gate (a) precharge delay (b) evaluation delay (c) dynamic power.

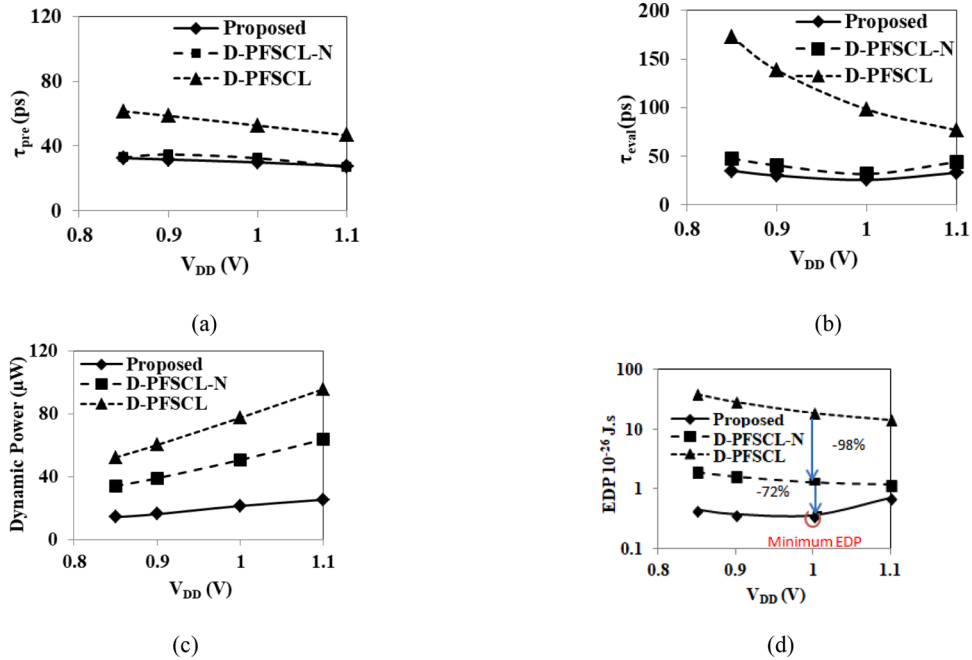


Fig. 6. Performance measurement with respect to power supply variations (a) precharge delay (b) evaluation delay (c) dynamic power (d) EDP.

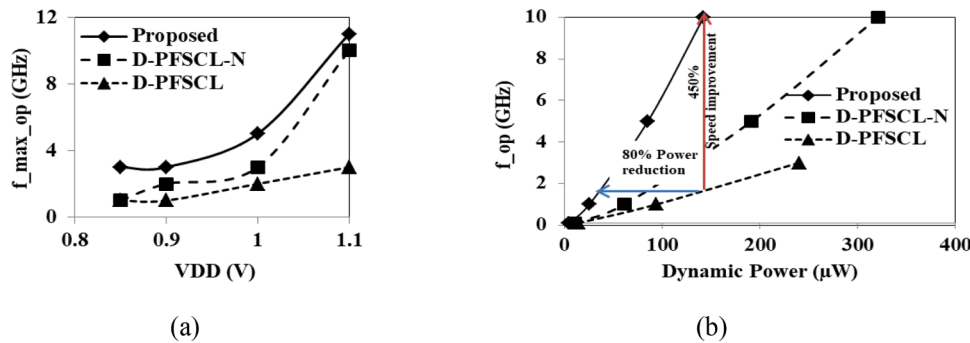


Fig. 7. (a) Maximum operating frequency with respect to different supply voltages (b) operating frequency with respect to dynamic power consumption.

Table II

Performance comparison of gates.

Function	Parameters	D-PFSCCL [9]	D-PFSCCL-N [10]	Proposed D-PFSCCL
MUX2	$\tau_{pre}$ (ps)	44	31	27.68
	$\tau_{eval}$ (ps)	109.7	32.6	33.09
	$P_{dyn}$ ( $\mu W$ )	90.81	64.05	31.06
MUX8	$\tau_{pre}$ (ps)	88.8	41.3	28.1
	$\tau_{eval}$ (ps)	454.6	37.1	33.1
	$P_{dyn}$ ( $\mu W$ )	1410	488.2	23.84
Full adder (SUM)	$\tau_{pre}$ (ps)	75.9	38.9	26.7
	$\tau_{eval}$ (ps)	361.1	126	34.02
	$P_{dyn}$ ( $\mu W$ )	524.7	185.5	26.74
Full adder (Carry)	$\tau_{pre}$ (ps)	45.7	46.5	28.04
	$\tau_{eval}$ (ps)	100.5	113.8	33.12
	$P_{dyn}$ ( $\mu W$ )	210.66	276.96	26.23

useful to reduce the number of required stages for realization of different logic gates and this is demonstrated in the paper. It is also being observed that adoption of the proposed scheme enhances the performance of the circuit in terms of both power and delay. The effect of parameter variation is also studied for D-PFSCCL XOR2 gate designed using proposed and available D-PFSCCL schemes.

## Declaration of Competing Interest

There is no conflict of interest.

## References

- [1] S. Donnay, G. Gielen (Eds.), *Substrate Noise Coupling in Mixed-Signal ASICs*, Springer US, Boston, MA, 2003, <https://doi.org/10.1007/b105382>.
- [2] S. Badel, MOS current-mode logic standard cells for high-speed low-noise applications, (2008). <https://doi.org/10.5075/EPFL-THESIS-4098>.
- [3] M. Rafiee, M.B. Ghaznavi-Ghouschi, Design of low-voltage shallow-depth differential source coupled logic using feedback and feedforward techniques, *Microelectronics J* 86 (2019) 140–149, <https://doi.org/10.1016/j.mejo.2019.03.002>.
- [4] K. Gupta, N. Pandey, M. Gupta, Model and Design of Improved Current Mode Logic Gates, Springer Singapore, Singapore, 2020, <https://doi.org/10.1007/978-981-15-0982-7>.
- [5] F. Besharati, A. Golmakani, S. Babayan-Mashhadi, Design of a LP current-mode comparator based on positive-feedback source coupled logic, in: *Electr. Eng. (ICEE)*, Iran. Conf., IEEE, 2018, pp. 224–227, <https://doi.org/10.1109/ICEE.2018.8472633>.
- [6] M. Damghanian, S.J. Azhari, A low-power 6-bit MOS CML flash ADC with a novel multi-segment encoder for UWB applications, *Integration* 57 (2017) 158–168, <https://doi.org/10.1016/J.VLSI.2017.01.006>.
- [7] M. Alioto, L. Pancioni, S. Rocchi, V. Vignoli, Modeling and evaluation of positive-feedback source-coupled logic, *IEEE Trans. Circuits Syst. I Regul. Pap.* 51 (2004) 2345–2355, <https://doi.org/10.1109/TCSII.2004.838149>.

- [8] M. Alioto, R. Mita, G. Palumbo, Design of high-speed power-efficient MOS current-mode logic frequency dividers, *IEEE Trans. Circuits Syst. II Express Briefs.* 53 (2006) 1165–1169, <https://doi.org/10.1109/TCSII.2006.882350>.
- [9] K. Gupta, N. Pandey, M. Gupta, Dynamic positive feedback source-coupled logic (D-PFSCCL), *Int. J. Electron.* 103 (2016) 1626–1638, <https://doi.org/10.1080/0020717.2016.1138519>.
- [10] R. Sivaram, K. Gupta, N. Pandey, A new realization scheme for dynamic PFSCCL style, *Integration* 75 (2020) 169–177, <https://doi.org/10.1016/j.vlsi.2020.05.004>.
- [11] K.J. Kuhn, Reducing variation in advanced logic technologies: approaches to process and design for manufacturability of nanoscale CMOS, *Tech. Dig. - Int. Electron Devices Meet. IEDM* (2007) 471–474, <https://doi.org/10.1109/IEDM.2007.4418976>.



**Ranjana Sivaram** received the B.Tech. degree in Electronics and Communication Engineering from Ambedkar Institute of Technology, IP University, New Delhi, India in 2010 and the M. Tech. degree in VLSI & Embedded Systems from Delhi Technological University, New Delhi in 2012. From 2012 onwards, she is working in Department of Telecommunications, Government of India.

She is currently pursuing part time PhD in VLSI from Delhi Technological University, New Delhi. Her current research interests include digital VLSI design.



**Kirti Gupta** received B.Tech. in Electronics and Communication Engineering from Indira Gandhi Institute of Technology, Delhi in 2002, M. Tech. in Information Technology from School of Information Technology in 2006. She received her Ph.D. in Electronics and Communication Engineering from Delhi Technological University, in 2016. Since 2002, she is with Bharati Vidyapeeth's College of Engineering, New Delhi and is presently serving as Professor in the same institute. A life member of ISTE, and senior member of IEEE, she has published more than 100 research papers in international, national journals and conferences. Her teaching and research interest is in digital VLSI design.



**Neeta Pandey** received her M.E. in Microelectronics from Birla Institute of Technology and Sciences, Pilani in 1991 and Ph.D. from Guru Gobind Singh Indraprastha University, Delhi in 2009. She has served in Central Electronics Engineering Research Institute, Pilani, Indian Institute of Technology, Delhi, Priyadarshini College of Computer Science, Noida and Bharati Vidyapeeth's College of Engineering, Delhi in various capacities. At present, she is a professor in the ECE department, Delhi Technological University. Her teaching and research interests include analog and digital VLSI design. A life member of ISTE, and senior member of IEEE, USA, she has coauthored over 100 papers in international, national journals of repute and conferences.

# Performance Evaluation using Online Machine Learning Packages for Streaming Data

Santosh Kumar Ray

Department of Information Technology  
Delhi Technological University  
Delhi, India  
raysantosh806@gmail.com

Seba Susan

Department of Information Technology  
Delhi Technological University  
Delhi, India  
seba\_406@yahoo.in

**Abstract**—Online machine learning concept concentrates on real-time and dynamic data. This dynamic data is in the form of data streams which is a continuous flow of data of infinite length. Recently many online machine learning tools were made available for evaluating performance metrics. Most of the approaches use online machine learning packages such as MLlib, spark streaming, scikit-multiflow, Creme and River. In this paper, performance evaluation metrics are evaluated for streaming data for three popular classification models: Logistic Regression (LR), K-Nearest Neighbors (KNN) and Gaussian Naive Bayes. For the evaluations, Creme and River online machine learning packages have been used. In the implementation phase, metrics are evaluated for four popular datasets: Elec2, Phishing, Bananas, and CreditCard. The accuracy, ROC-AUC, precision, recall and F1-score performance metrics are used for the comparison of the online classifiers. After analyzing the results of the four datasets, it is observed that the performance evaluation using the River online machine learning package is more accurate and reliable.

**Keywords**—Streaming data, Classification, Online machine learning, Performance evaluation.

## I. INTRODUCTION

In recent years, drastic advancements in technology have been observed that enlarges the scope of the application of technology for dealing with dynamically evolving data streams emanating from social media and the Internet-of-Things (IoT). Dynamic data streams [1] are processed one sample at a time. When a new sample is encountered, old samples are removed, and this process continues sequentially till the last sample in the data stream. The main issue with traditional machine learning is the assumption that the entire dataset is available [3], which is not feasible in real-world applications where the data is continuously evolving. Also, static data requires large memory and high processing speed. But online machine learning techniques resolve the problem of memory and processing time by considering the dynamic data known as a data stream that processes samples sequentially. Today, many online machine learning libraries are available to perform basic operations such as regression, classification, clustering, etc.; these online libraries boost up the accuracy of models [2, 6]. These libraries incorporate several well-known classifiers used for traditional machine learning. Logistic Regression is a classifier that deals with the linear model and shows a good accuracy by utilizing an efficient optimizing method (mini-batch gradients method, ordinary least square methods, ada bound, ada grad, etc.).

KNN is a lazy learning classifier that performs better for the numerical features but does not perform optimally for both categorical and numerical features taken collectively. In online machine learning, KNN also suffers from drifting of data, but this type of problem is resolved with the Self Adjusting Memory - KNN (SAM-KNN) [4] that handles the heterogeneous concepts of input data very efficiently and shows optimal performance. The Naive Bayes classifier [5] is suitable for large-size and high-dimensional datasets. It maintains the distribution of each class and respective features. In the online learning process, the classifier updates itself as new observations arrive. It is observed from the existing approaches on online machine learning that performance evaluation metrics play a significant role in deciding the optimality of a predictive model for streaming data.

In this paper, we present a detailed analysis using performance metrics for four streaming datasets using River [2] and Creme [6] online machine learning packages. The organization of this paper is as follows: Section II describes related work, Section III discusses performance metrics, Section IV presents the process flow, Section V analyses the results, and Section VI concludes the paper.

## II. RELATED WORK

In supervised machine learning, after deploying the model, performance evaluation is a crucial stage since it helps to highlight the advantages and limitations of the model. Many performance metrics such as accuracy, ROC-AUC, precision, recall and F1-score have been used in literature to evaluate the performance of online machine learning classifiers for streaming datasets [2, 5, 6, 7]. Some genuine issues due to data drift are described in [16, 17] when there is no prior knowledge of data generator and the generating function of new data is unknown for the learner hence the trained model may be unpredictable. Solutions to the drifting problem could be a) adaptive windowing [15] b) learners with modified training set [16]. In [8], the authors grow the network hidden layer incrementally during training as the input samples are fed into the model one by one. Class-imbalance is yet another issue that affects learning from data streams [17, 22].

Recently, various online machine learning packages based on Python have been developed for analyzing streaming data in different forms such as audio, video, text, images, etc. We discuss some of these packages here. MLlib [9] is an open-

source online machine library that emerged from spark's distribution machine learning library. Using this approach, implementation of a machine learning model takes place in rapid ways in case of regression, classification, clustering, etc. A most popular online machine learning package is scikit-multiflow [10] for streaming data that takes the data streams for training the model and facilitates multi-output learning. This framework is inspired by scikit learn [14], Massive Online Analysis (MOA) [12][13], and Multi-labels Environment for Knowledge Analysis (MEKA) [11]. Creme [6] and River [2] are Python based machine learning packages introduced recently for learning from streaming data. Creme was introduced by Halford et al. in 2019 and was claimed to be faster than Keras and Tensorflow for online machine learning. The package includes online classifiers, regressors and solutions for imbalanced learning. River was introduced by Montiel et al. in 2021 as a merger of scikit-multiflow and Creme. It contains additional functionalities and is faster than its parent packages.

### III. PERFORMANCE METRICS FOR STREAMING DATA

For an online machine learning model, performance evaluation is an important task. In this section, we focus on the performance metrics used in our experiments to evaluate the performance of the online classifiers on streaming datasets.

1. **ROC-AUC SCORE (Receiver Operating Characteristics – Area Under Curve):** The area under the curve (AUC) score is derived from the Receiver Operating Characteristics (ROC) curve which is the graph between True Positive Rate (TPR) and False Positive Rate (FPR). The value of ROC-AUC lies between [0,1]. A higher value indicates a more accurate result. A value close to 0.5 indicates a bad classifier.

2. **precision and recall:** precision is also known as positive predictive value. It is the ratio of the number of true positives to the total number of positives detected by the model. can be defined as

$$precision = \frac{t_p}{[t_p + f_p]} \quad (1)$$

where  $t_p$  means true positive and  $f_p$  means false positive. The recall is the true positive rate, also known as sensitivity.

$$recall = \frac{t_p}{[t_p + f_n]} \quad (2)$$

where  $t_p$  means true positive and  $f_n$  means false negative.

3. **F1-score:** F1-score is the harmonic mean of precision and recall values. A higher value indicates good accuracy. Mathematically, F1-score is defined as

$$F1 - score = \frac{2 \times precision \times recall}{[precision + recall]} \quad (3)$$

4. **accuracy:** accuracy is the percentage of correct predictions. It is defined as

$$accuracy = \frac{[t_p + t_n]}{[t_p + t_n + f_p + f_n]} \quad (4)$$

where  $t_p$  = true positive,  $t_n$  = true negative,  $f_p$  = false positive,  $f_n$  = false negative.

### IV. PROCESS FLOW FRAMEWORK

In the online machine learning approach, the main challenge is to find the model that yields a good accuracy for the given dataset. It is clear that patterns of the stream are unknown beforehand, which can affect the efficiency and accuracy of the system.

In this framework, the input data is processed and transferred to a data stream that is the concatenation of n samples  $D_1, D_2, D_3, \dots, D_n$ .

$$D = \text{Concatenate}(D_1, D_2, D_3, \dots, D_n) \quad (5)$$

In this paper, the streaming data is processed by three online classifier models - Logistic Regression (LR), K-Nearest Neighbor (KNN) and Gaussian Naïve Bayes (GNB). For complete online machine learning functionality and its evaluation, the Creme and River machine learning packages are used.

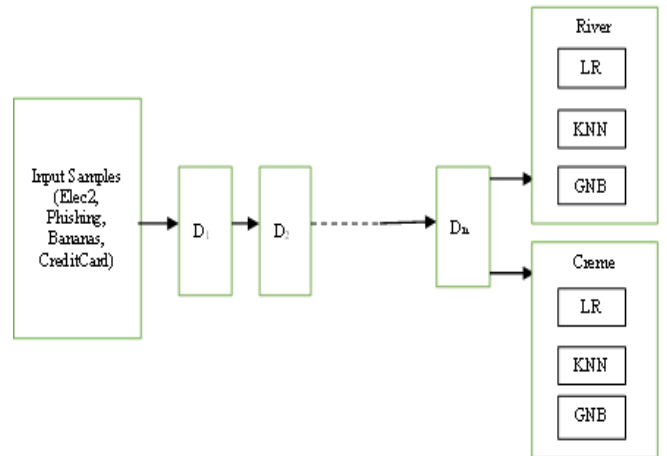


Figure 1. Process flow framework for the performance analysis using River and Creme online machine learning packages.

### V. EXPERIMENTS AND RESULTS

Online machine learning packages Creme and River are used to perform the evaluations. The following classifiers are used to perform online or incremental learning: Logistic Regression (LR), K-Nearest Neighbors (KNN) and Gaussian Naïve Bayes (GNB) classifiers. The experiments are performed using Python software version 3.8 on a 1.8 GHz Intel core PC. The code took a few seconds to execute. In this research, real-time streaming datasets are used for the experiments, namely, Elec2, Phishing, Bananas, and CreditCard, the details of which are given below. The accuracy, ROC-AUC, precision, recall and F1-score performance metrics are used to evaluate the predictions made by the online classifiers. The results are shown in Table I.

TABLE I. PERFORMANCE OF THE FOUR STREAMING DATASETS EVALUATED USING RIVER AND CREME ONLINE MACHINE LEARNING PACKAGES (SHOWN FOR LR, KNN AND GNB CLASSIFIERS)

S.No.	Dataset	Metrics	River			Creme		
			LR	KNN	GNB	LR	KNN	GNB
1	Elec2	accuracy	89.5%	85.8%	76.14%	83.73%	87.08%	76.13%
		ROC-AUC	0.8909	0.8519	0.8052	0.8275	0.8710	0.7357
		precision	0.885	0.845	0.8155	0.8398	0.8580	0.8155
		recall	0.865	0.814	0.5659	0.7623	0.8330	0.5658
		F1-score	0.875	0.829	0.6681	0.7992	0.8460	0.6681
2	Phishing	accuracy	89.0%	88.01%	88.0%	89.20%	75.92%	87.92%
		ROC-AUC	0.8919	0.8809	0.8814	0.8922	0.7590	0.8814
		precision	0.849	0.850	0.842	0.8642	0.7400	0.8417
		recall	0.911	0.883	0.8940	0.8942	0.6950	0.8923
		F1-score	0.879	0.866	0.867	0.8789	0.7170	0.8663
3	Bananas	accuracy	51.9%	88.5%	61.2%	53.74%	51.32%	61.15%
		ROC-AUC	0.4966	0.8905	0.5830	0.5007	0.5130	0.5830
		precision	0.443	0.877	0.640	0.4508	0.4580	0.6397
		recall	0.279	0.865	0.306	0.1465	0.4700	0.3064
		F1-score	0.342	0.871	0.414	0.2211	0.4640	0.4143
4	CreditCard	accuracy	99.9%	99.9%	98.1%	99.94%	99.86%	98.14%
		ROC-AUC	0.8707	0.9551	0.9129	0.8607	0.9990	0.9095
		precision	0.783	0.959	0.073	0.8853	0.9690	0.0732
		recall	0.742	0.427	0.837	0.7215	0.1890	0.8374
		F1-score	0.762	0.591	0.134	0.7951	0.3160	0.1346

1. *Electricity (Elec2)* [18]: The Elec2 electricity dataset is taken from the Australian New South Wales Electricity Market, where prices are varying. Prices of electricity are affected by the demand and supply of the market itself and set every five minutes. This data set contains 45,312 instances, where class labels identify the changes in the price (two possible classes: up or down) relative to a moving average of the last 24 hours. An important point of this dataset is that it exhibits temporal dependencies.

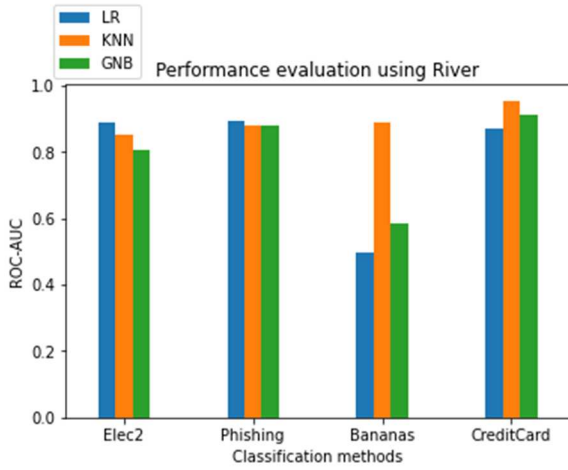
2. *Phishing* [19]: This dataset contains features from web pages that are classified as phishing or not. In phishing, dataset classifications are binary classifications types and it has 1250 samples with 9 features.

3. *Bananas* [20]: An artificial dataset where instances are related to clusters with a banana shape. There are two attributes At1 and At2 corresponding to the x-axis and y-axis, respectively. The class label (-1 or 1) represents one of the two banana shapes in the dataset.

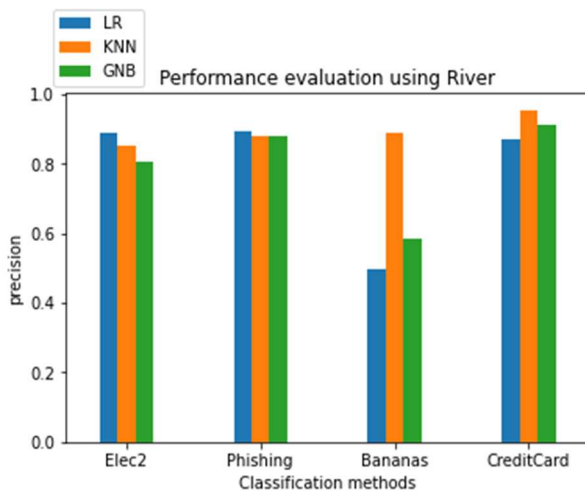
4. *CreditCard* [21]: This dataset contains transactions performed through credit cards in September 2013 by European cardholders. All transactions that occurred over two days are considered. Total samples in this dataset are 284807 and 492 fraud transactions are detected. This dataset is highly imbalanced and has 30 features. There are two classes. The class label 1 indicates fraud detection and 0 represents a normal transaction.

#### A. Performance visualizations using River

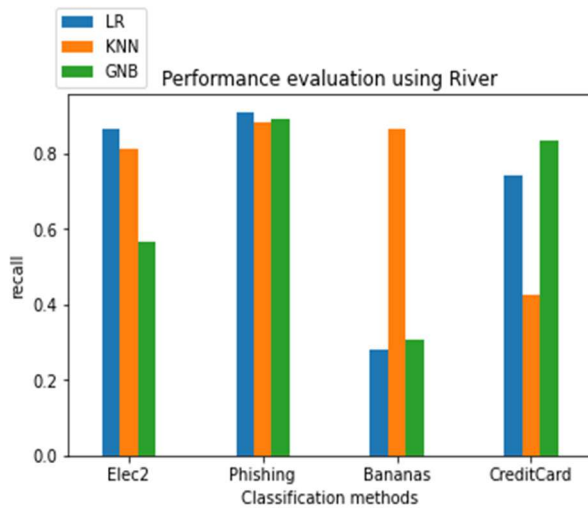
Performance evaluation of online machine learning models tests the consistency of models. In Fig. 2(a), Fig. 2(b), Fig. 2(c) and Fig. 2(d), the bar plots visualize the ROC-AUC, precision, recall and F1-score performance metrics using River online machine learning package. The metrics are computed for four streaming datasets (Elec2, Phishing, Bananas, CreditCard). For each dataset, the results of logistic regression (LR), K-nearest neighbor (KNN) and Gaussian Naive Bayes (GNB) online classifiers are shown. For the Elec2 dataset, LR and KNN show a better ROC-AUC value while the GNB classifier yields a low value. For the other metrics such as precision, recall and F1-score, the performance of KNN classifier is better while LR shows average performance, and overall performance of GNB is better in the case of Elec2 dataset. In the case of the Phishing dataset, performances of all three classifiers are similar with respect to all four-performance metrics. In the Bananas dataset, the performance of the KNN classifier is significantly better for all four metrics but the other classifiers show very low performance. CreditCard is a large-sized imbalanced streaming dataset. The LR classifier performs best for this dataset. Performance scores using River online machine learning package are summarized in Table I for all datasets.



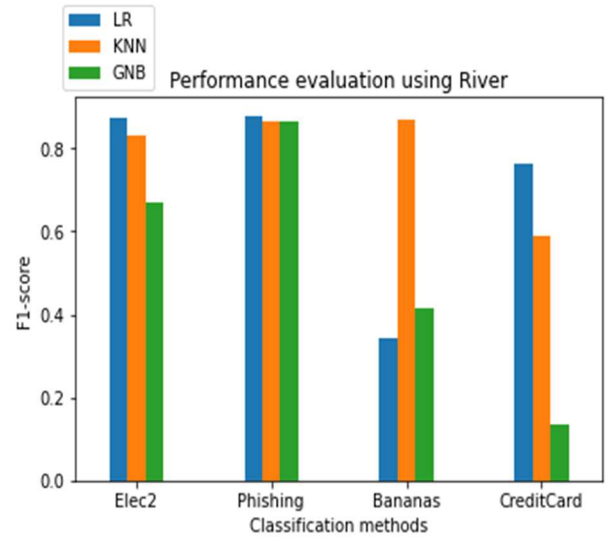
(a)



(b)



(c)



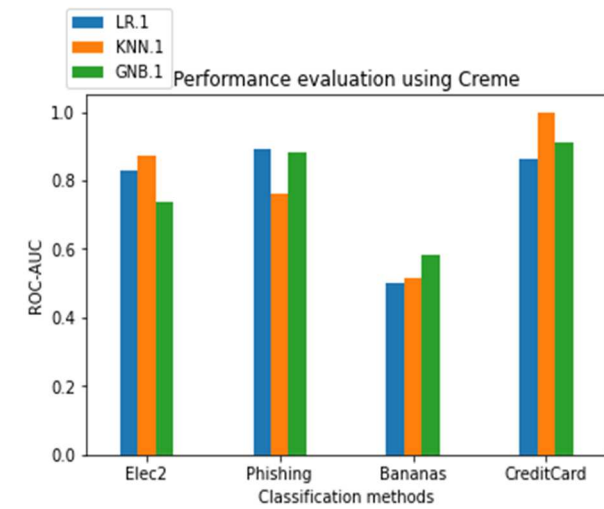
(d)

Figure 2. Bar plots of classifiers (LR, KNN, GNB) for the four datasets (Elec2, Phishing, Bananas, CreditCard) using River for the performance metrics (a) ROC-AUC (b) precision (c) recall (d) F1-score.

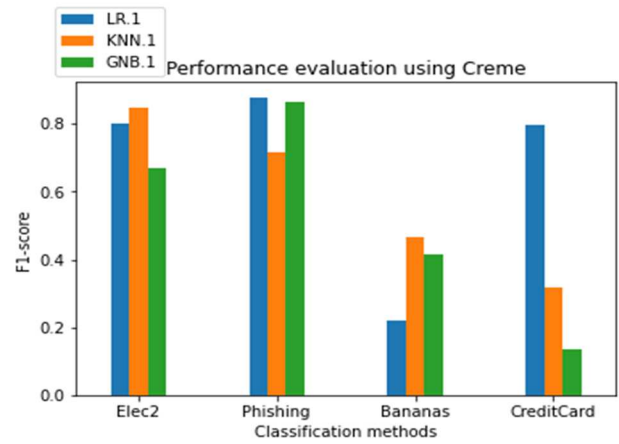
#### B. Performance visualizations using Creme

Similar experimental conditions are applied for the Creme online machine learning package, for the same datasets, and same online classifiers. The ROC-AUC, F1-score, precision and recall performance metrics are plotted in Fig. 3(a), Fig. 3(b), Fig. 3(c) and Fig. 3(d) respectively, and summarized in Table I. The average performance scores of Creme are lower in comparison to River. For the Creme online machine learning package, the performance of KNN classifier is better but the overall performances of LR and GNB classifiers are not as high as that of the River package. The performance of the Bananas dataset, however, is low even with the KNN classifier. Online machine learning shows an average performance for small-size streaming datasets and when the size of the dataset is large then it shows better performance. Creme takes more processing time than River. Creme supports the mini-batch concept, but the outcome is not more efficient. It can extend the functionality of machine learning techniques at an industrial scale but it is time-consuming and requires more memory.



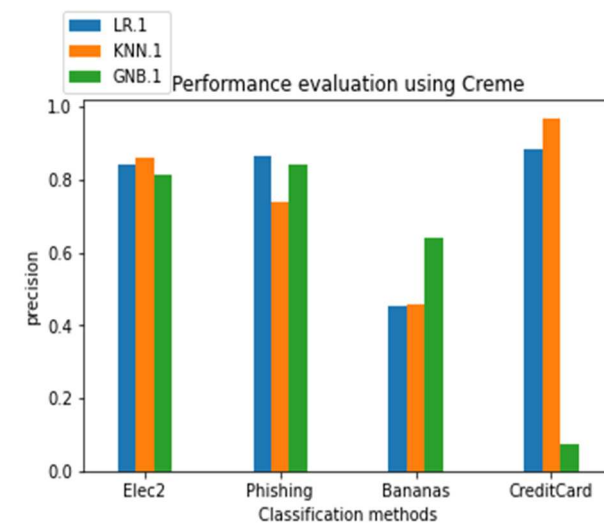


(a)

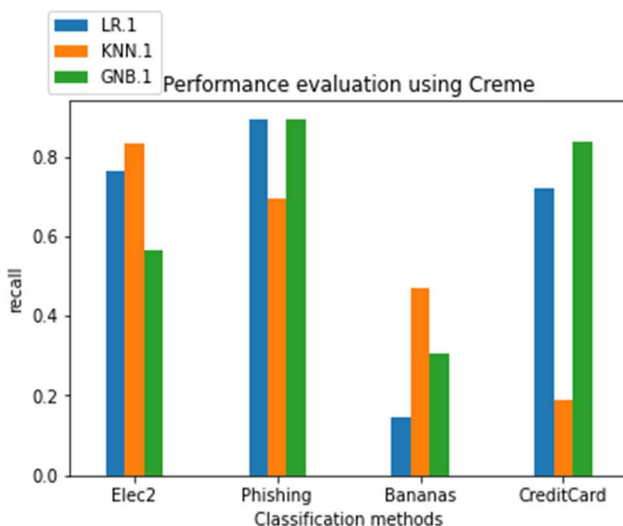


(d)

Figure 3. Bar plots of classifiers (LR, KNN, GNB) for the four datasets (Elec2, Phishing, Bananas, CreditCard) using Creme for the performance metrics (a) ROC-AUC (b) precision (c) recall (d) F1-score.



(b)



(c)

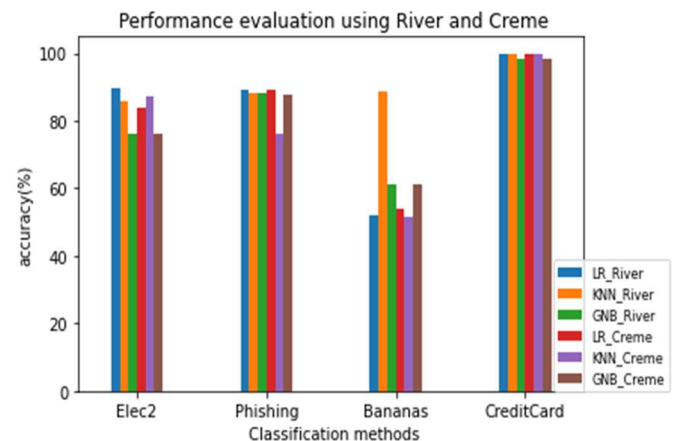


Figure 4. Bar plot of classifiers (LR, KNN, GNB) accuracy (%) for the four datasets (Elec2, Phishing, Bananas, CreditCard) using River and Creme.

## VI. CONCLUSION

There are certain challenges involved in using traditional machine learning for real-time streaming data. This concept concentrates on static data and predicts the result of the entire data collectively. It may cause a problem in the case of real-time and dynamic data in which case, the samples are processed one at a time and discarded after use. Online or incremental learning methodologies are found apt for learning from streaming data. In this paper, we conduct a comparative performance analysis of the River and Creme online machine learning packages for four streaming datasets. The accuracy, ROC-AUC, precision, recall, F1-score performance metrics are used to evaluate the predictions of three online classifiers (Logistic Regression, KNN, Gaussian Naive Bayes). It is observed that River outperformed Creme in all the classification tasks. The KNN classifier performed consistently well for all datasets. Out of the four streaming datasets used in our experiments, Bananas yielded lower scores as compared to Elec2, Phishing and CreditCard datasets. It is also noted that accuracy is not a reliable metric by itself when the dataset is imbalanced, as proved in the case of CreditCard dataset. The other performance metrics (ROC-AUC, precision, recall and F1-score) are better measures of the classification performance in case of imbalanced datasets. Online machine learning packages are very efficient for processing data streams. The most important concept is to develop a system that evaluates the streaming data efficiently provided the fact that this data is large size, continuous and dynamic in nature. Today many streaming data processing systems are facing challenges [7] such as ready availability of data without latency, lack of high throughputs, requirements of trade-off between memory and length of streaming data, etc. These factors may cause a drop in the performance of the system. New stream data evaluating techniques are required that will handle easily the main challenges: latency, throughputs, memory, and fixed-rate performance.

## REFERENCES

- [1] Žliobaitė, Indrė, Albert Bifet, Bernhard Pfahringer, and Geoffrey Holmes. "Active learning with drifting streaming data." *IEEE transactions on neural networks and learning systems* 25, no. 1 (2013): 27-39.
- [2] Montiel, Jacob, Max Halford, Saulo Martiello Mastelini, Geoffrey Bolmier, Raphael Sourty, Robin Vaysse, Adil Zouitine et al. "River: machine learning for streaming data in Python." (2021): 1-8.
- [3] Putatunda, Sayan. "Supervised Learning for Streaming Data." In *Practical Machine Learning for Streaming Data with Python*, pp. 57-96. Apress, Berkeley, CA, 2021.
- [4] Losing, Viktor, Barbara Hammer, and Heiko Wersing. "KNN classifier with self-adjusting memory for heterogeneous concept drift." In *2016 IEEE 16th international conference on data mining (ICDM)*, pp. 291-300. IEEE, 2016.
- [5] Nguyen, Tien Thanh, Thi Thu Thuy Nguyen, Xuan Cuong Pham, Alan Wee-Chung Liew, and James C. Bezdek. "An ensemble-based online learning algorithm for streaming data." *arXiv preprint arXiv:1704.07938* (2017).
- [6] Max Halford, Geoffrey Bolmier, Raphael Sourty, Robin Vaysse, and Adil Zouitine. *creme*, a Python library for online machine learning, 2019. URL <https://github.com/MaxHalford/creme>.
- [7] Karimov, Jeyhun, Tilmann Rabl, Asterios Katsifodimos, Roman Samarev, Henri Heiskanen, and Volker Markl. "Benchmarking distributed stream data processing systems." In *2018 IEEE 34th International Conference on Data Engineering (ICDE)*, pp. 1507-1518. IEEE, 2018.
- [8] Susan, Seba, and Mayank Dwivedi. "Dynamic growth of hidden-layer neurons using the non-extensive entropy." In *2014 Fourth International Conference on Communication Systems and Network Technologies*, pp. 491-495. IEEE, 2014.
- [9] Meng, Xiangrui, Joseph Bradley, Burak Yavuz, Evan Sparks, Shivaram Venkataraman, Davies Liu, Jeremy Freeman et al. "Millib: Machine learning in apache spark." *The Journal of Machine Learning Research* 17, no. 1 (2016): 1235-1241.
- [10] Montiel, Jacob, Jesse Read, Albert Bifet, and Talel Abdesslem. "Scikit-multiflow: A multi-output streaming framework." *The Journal of Machine Learning Research* 19, no. 1 (2018): 2915-2914.
- [11] Read, Jesse, Peter Reutemann, Bernhard Pfahringer, and Geoffrey Holmes. "Meka: a multi-label/multi-target extension to weka." (2016): 1-5.
- [12] Bifet, Albert, Ricard Gavaldà, Geoff Holmes, and Bernhard Pfahringer. *Machine learning for data streams: with practical examples in MOA*. MIT press, 2018.
- [13] Bifet, Albert, Geoff Holmes, Bernhard Pfahringer, Philipp Kranen, Hardy Kremer, Timm Jansen, and Thomas Seidl. "Moa: Massive online analysis, a framework for stream classification and clustering." In *Proceedings of the First Workshop on Applications of Pattern Analysis*, pp. 44-50. PMLR, 2010.
- [14] Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel et al. "Scikit-learn: Machine learning in Python." *The Journal of machine Learning research* 12 (2011): 2825-2830.
- [15] Bifet, Albert, and Ricard Gavaldà. "Learning from time-changing data with adaptive windowing." In *Proceedings of the 2007 SIAM international conference on data mining*, pp. 443-448. Society for Industrial and Applied Mathematics, 2007.
- [16] Gama, Joao, Pedro Medas, Gladys Castillo, and Pedro Rodrigues. "Learning with drift detection." In *Brazilian symposium on artificial intelligence*, pp. 286-295. Springer, Berlin, Heidelberg, 2004.
- [17] Hoens, T. Ryan, Robi Polikar, and Nitesh V. Chawla. "Learning from streaming data with concept drift and imbalance: an overview." *Progress in Artificial Intelligence* 1, no. 1 (2012): 89-101.
- [18] Harries, Michael, and New South Wales. "Splice-2 comparative evaluation: Electricity pricing." (1999).
- [19] <http://archive.ics.uci.edu/ml/datasets/Website+Phishing>.
- [20] <https://www.openml.org/d/1460>.
- [21] <https://www.kaggle.com/mlg-ulb/creditcardfraud>.
- [22] Susan, Seba, and Amitesh Kumar. "The balancing trick: Optimized sampling of imbalanced datasets—A brief survey of the recent State of the Art." *Engineering Reports* 3, no. 4 (2021): e12298.

# Real Time Analysis of Material Removal Rate and Surface Roughness for Turning of Al-6061 using ANN and GA

Abhishek Jha<sup>1\*</sup>, Baibhav Kumar<sup>2</sup>, Ashok Kumar Madan<sup>3</sup>

<sup>1,2</sup>Under Graduate Student, Department of Mechanical Engineering, Delhi Technological University, New Delhi, India

<sup>3</sup>Professor, Department of Mechanical Engineering, Delhi Technological University, New Delhi, India

**Abstract:** The paper shows and includes targeted supervision of optimizing machining parameters. Material Removal Rate and surface roughness being integral to machining efficacy of any workpiece, simulation-based modeling helps in failure mitigation. The potential of ANN-GA mathematical approach for prediction and optimization of MRR and surface roughness of AL 6061, an analysis based statistical study has been discussed. The computational model between the desired output and the inputs have been configured using Multiple Regression- Genetic Algorithm and Artificial Neural Network (ANN) methods. The closeness in predicted and optimized data sets were mapped using integrational ANN with GA to interpolate efficacy in optimality.

**Keywords:** Artificial Neural Network (ANN), Material Removal Rate (MRR), Genetic Algorithm (GA), Surface roughness, Depth of cut, Feed rate, Regression analysis, ANOVA.

## 1. Introduction

The industrial machining requirements calls for the need of precision and manufactured product efficacy. The diverse machining processes and its role in shaping final product has crucial role making products industrially ready. Turning operation on horizontal lathe machine is one of those integral machining operation, where by cylindrical surfaces of required diameter consideration and surface profile is achieved. Turning being work piece driven and single point cutting tool driven machining, economics of industrial manufacturing necessities for optimal machining parameter selection to simulate desired output without any catastrophic manufacturing. ANN being a widely looked up tool where by multiple machining parameter can be architected into input, hidden and output layers, ability to use learning outcomes in training sets. Aluminum 6061 alloy is one of integral industrial friendly alloy which makes its presence in several applications like aviation components, automobile parts, weapon casings, and high vacuum chambers. These industries look for low allowance operational values and utmost precision is desirable in these respective manufacturing. The inclusion of Artificial Neural Networks mitigates with missing data and machining parameters learn by examples,

making real-time simulated modelling possible. With exclusively integrating with GA, optimization of Discrete MRR outputs in turning operation can be predicted even with due considerations of minute errors in the inputs, with quick inheritance and good accuracy.

## 2. Literature Review

Ramu I. et.al [7] explained and demonstrated a systematic procedure of Taguchi parameter design and applying it to the data on turning. The second was to find out the optimal combination of process parameters based on S/N ratio and to know the significance of each parameter by performing ANOVA analysis.

T.M. Chenthil Jegana et.al [1] This paper covers an important aspect and brings how ANN should be equipped with such functionalities as storing information, reasoning, decision making, learning, and integration of these into the process. In particular, the learning characteristic is a unique feature of the ANN Turning operation optimization.

Tiagrajah V. jannharimal et.al [8] discusses effective, Multi Objective Genetic Algorithm (MOGA) will act as an optimizer of the developed model. Turning input parameters such as feed rate, cutting speed and depth of cut were considered as input variables and surface roughness, specific power consumption and cutting force were used as output variables.

S. Nes, Eli et.al [2] highlights tool geometry on the surface finish obtained in turning of AISI 1040 steel. In order to find out the effect of tool geometry parameters on the surface roughness during turning, response surface methodology (RSM) was used and a prediction model was developed related to average surface roughness (Ra) using experimental data. The results explained that the tool nose radius was the dominant factor on the surface roughness.

Ndaruhadi, P.Y.M.W et.al [4] helps in assessing surface roughness value it was found that feed did not significantly influence surface roughness. Among the influencing factor, the rank is tool type, cutting speed, and cutting direction.

\*Corresponding author: abhishekjha\_me20a8\_31@dtu.ac.in

Dhabale, R. et.al [6] presents a new idea to generate process plan from feature-based modeling, based on an inclusive of geometric modeling method that proposes both feature-based modeling and feature information storage.

#### A. Experimental Studies and Datasets

Table 1  
Datasets

Run order	Speed (RPM)	Feed Rate (mm/min)	Depth of cut (mm)	MRR (mm <sup>3</sup> /min)	Ra (μs)
1	180	0.2	0.2	113.927	1.04
2	180	0.2	0.4	212.018	0.98
3	180	0.2	0.6	261.698	2.2
4	180	0.315	0.2	672.129	2.44
5	180	0.315	0.4	333.569	3.84
6	180	0.315	0.6	897.738	2.4
7	180	0.4	0.2	22.443	2.06
8	180	0.4	0.4	953.491	2.3
9	450	0.4	0.6	901.732	3.66
10	450	0.2	0.2	780.859	0.9
11	450	0.2	0.4	528.915	0.94
12	450	0.2	0.6	145.761	2.9
13	450	0.315	0.2	176.76	1.42
14	450	0.315	0.4	303.637	3.38
15	450	0.315	0.6	2263.05	1.34
16	450	0.4	0.2	40.271	1.74

The above table of datasets depicts input parameters considered for turning operation Like the feed rate, depth of cut, cutting speeds and corresponding Material Removal Rate obtained. The machining parameters considered was kept in accordance with industrial requirements of turning operations on AL 6061. AL 6061 being industrially proclaimed as a medium to high strength heat-treatable alloy with desirable strength. It also has very good corrosion resistance and very good weldability. The chemical composition consideration 6061 aluminium is 97.9% Al, 0.6% Si, 1.0%Mg, 0.2%Cr, and 0.28% Cu. The density of 6061 aluminium alloy is 2.7 g/cm<sup>3</sup> with great formability and workpiece can be machined using turning operation to meet industrial demands.

The objective of carrying our research on given dataset is to model conventional turning operation to infer optimal value of material removal rate and surface roughness to enhance precision and accuracy of machining.

### 3. Methodology

The experimental results were assessed on design and modeling level by taking considerations of Taguchi Method and MRR surface roughness was assessed based on response of predicted data statistically without remotely performing set experiments. The experimental results derived in original experimental research are in synchronization to the ANN predicted results.

The MATLAB NN toolbox is integrated for training and testing of neural network model. The results inferred using Artificial Neural Network (ANN) indicate good agreement between the predicted output values and that of experimental values. Thereafter, Genetic Algorithm (GA) is made inclusive with neural network model to determine the optimal machining parameters feed rate, depth of cut and cutting speed to get

desired industrial demand in accordance to needs of optimal Material removal rate and surface roughness in turning operation.

The due considerations of exploration in data analysis were looked into for establishing relationship between input and output variables, The mentioned integration assessed the non-linearity between parameters. Due to unavailability of solutions for non- linearity, ANN was integrated into our project for its ability to learn and model complex solutions.

### 4. Development of Artificial Neural Networks

Artificial Neural networks are the framework that are intended to reproduce the working of a human mind. With the assistance of involvement and information Artificial Neural Network works on itself utilizing artificial neurons. Artificial Neural Network comprise of rudimentary units called as neurons which takes at least one sources of info and produces a result.

Computations that are carried out on each neuron if ANN are as follows:

$$z[i] = w[i] * A[i - 1] + b[i]$$

$$A[i] = f(Z[i])$$

Here the notations are as follows:

$W^{[i]}$  represents the weight of the neuron

$A^{[i-1]}$  represents the output derived from last layer

$b^{[i]}$  represents the bias of the connection

$f(x)$  represents the nonlinear activation function

A neural network consists of several layers which are grouped as input layer, hidden layer and output layer. Here layer means the set of parallel neurons without having connections between them. Neurons from input layer are connected to hidden layer which in turn is connected to the output layer in the whole Artificial Neural Networks System

In our model we have used Gradient Descent back propagation method for the best tuning of the neurons presents in each layer.

For our model we have used MATLAB NN tool for training and obtaining of Artificial Neural Network. and results associated with the Neural networks.

For ANN in MATLAB, we first imported classified our dataset into input set which include input parameters viz. Cutting Speed, Depth of Cut, Feed Rate and after classification of Input data we had made another classification which was target set that include data of output parameters viz. Material Removal Rate and Surface Roughness.

A neural network comprises of various configurations such as type of network used, training function, learning function, Transfer function, number of neurons in input, hidden and output layers and number of epochs. To get the best result it is very important to have best value for these configurations. The table below has listed down the functions and specifications that we used for our neural network.

Levenberg-Marquardt algorithm is taken as the training function due to its wide applications in solving non-linear problems by curve fitting and Mean Square error function is chosen as the performance function for accurate model.

To get the accurate neural network model by tuning weights and bias efficiently we have used Gradient descent back propagation technique.

Table 2  
Neural Network parameters

Network Type	Feed Forward Back Propagation
Training function	TRAINLM
Adaptive Learning Function	LEARNGDM
Performance Function	MSE
Number of Layers	5
Number of Neurons in hidden layers	[5, 3, 4]
Transfer Function	TANSIG
Number of epochs	1000

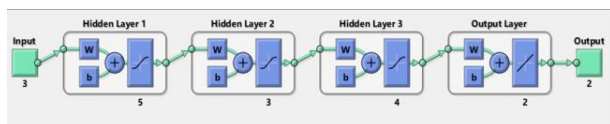


Fig. 1. Neural Network model in Matlab

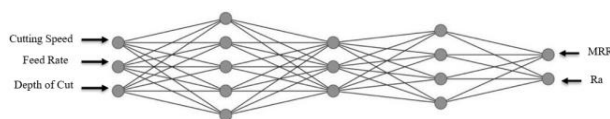


Fig. 2. Neural Network model architecture

Stochastic Gradient Descent is the optimization algorithm used for adaptive learning as it minimizes the gradient and adjusts weights accordingly. Weights and biases are tuned in an iterative manner to obtain optimum values for least error.

## 5. Results of Artificial Neural Networks

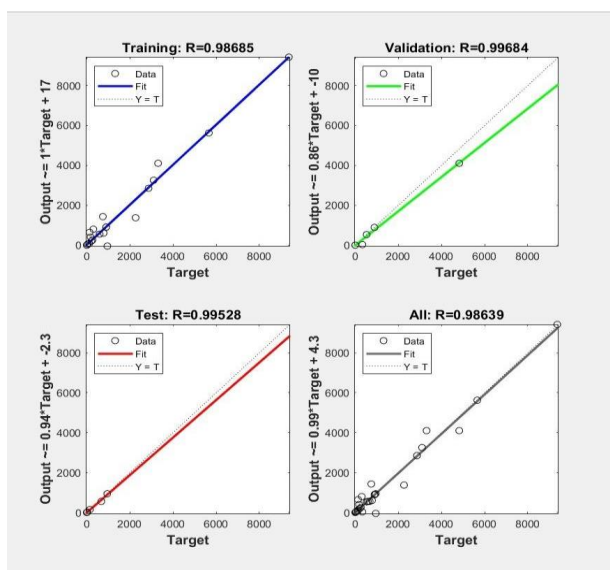


Fig. 3. Regression plots

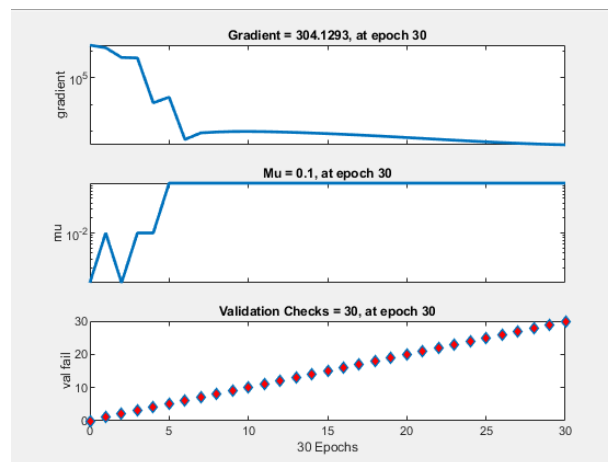


Fig. 4. Validation performance

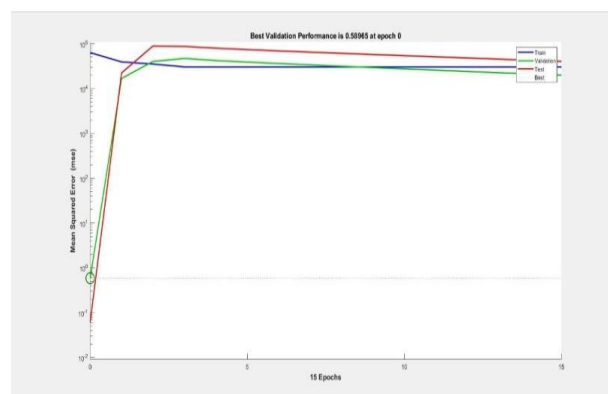


Fig. 5. Gradient descent plots

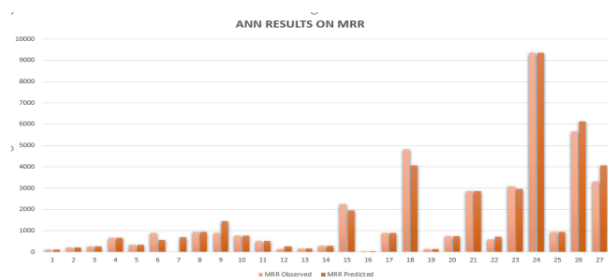


Fig. 6. Observed vs. Predicted MRR

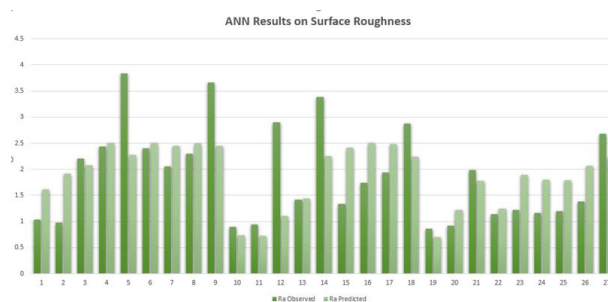


Fig. 7. Observed vs. Predicted Ra

Table 3  
ANN results

S.No.	Experimental		Predicted	
	MRR (mm <sup>3</sup> /min)	Ra (μs)	MRR (mm <sup>3</sup> /min)	Ra (μs)
1	113.927	1.04	114.165	1.61
2	212.018	0.98	212.155	1.91
3	261.698	2.20	261.80	2.08
4	672.129	2.44	672.159	2.51
5	333.569	3.84	333.641	2.27
6	897.738	2.40	572.567	2.50
7	22.443	2.06	691.35	2.44
8	953.491	2.30	953.427	2.49
9	901.732	3.66	1461.130	2.44
10	780.859	0.9	781.045	0.73
11	528.915	0.94	529.102	0.72
12	145.761	2.9	260.810	1.11
13	176.76	1.42	177.068	1.44
14	303.637	3.38	303.654	2.24
15	2263.05	1.34	1942.553	2.41
16	40.271	1.74	40.401	2.50
17	890.513	1.94	890.598	2.48
18	4822.2	2.88	4064.131	2.24
19	130.586	0.86	130.751	0.70
20	744.814	0.92	744.989	1.21
21	2862.91	1.98	2863.012	1.78
22	592.809	1.14	720.44	1.24
23	3098.48	1.22	2958.771	1.89
24	9365.24	1.16	9365.190	1.79
25	948.858	1.20	948.953	1.78
26	5658.57	1.38	6147.365	2.07
27	3305.82	2.68	4064.131	2.24

Table 3 above shows the experimental values and predicted values of Material Removal Rate (MRR) and Surface Roughness (Ra) achieved from Artificial Neural Networks. Regression results are also plotted in Fig.4 for training, testing and validation set where our model achieved the accuracy of 98.68% for training and 99.52% for testing set. The plots represent the fitting of model where R represents the square of observed and predicted relations. More the value close to 1 better are the results.

Apart from that Gradient descent visualization in Fig. 5 and Validation performance in Fig.6 are also plotted. A comparison between the observed and predicted values of Material Removal Rate and Surface Roughness are also plotted in Fig. 7 and Fig. 8 which shows how much predicted values are deviated from observed response.

## 6. Optimization Using Genetic Algorithm

Genetic algorithm considerations include evolutionary revolution optimization which centers around natural selection, inheritance, mutations and crossovers.

The genetic algorithm considered to be trial and tested classical evolutionary algorithm. With inception of random means that in order to find a solution using the GA, random changes are integrated. Note that GA may be called Simple GA due to its ease input compared to other EAs.

GA is underlined on Darwin's theory of evolution. It is a slow gradual process that works by making changes to the making slight and slow changes. Also, GA makes slight changes to its solutions slowly until getting the best solution.

In a genetic algorithm, efficacy ensured population of candidate solutions (called individuals, creatures, or phenotypes) to an optimization problem is evolved toward better solutions. The choice of opting Genetic Algorithm in optimizing turning operations points out to its easy compatibility and integrability compared to other parallel optimization algorithms like Ant Colony optimization, swarm optimization and others.

The GA is known to imitate the natural selection as proposed by Darwin and decodes algorithm best suited to give best possible predictions, optimality is configured for both MRR and surface roughness. The major flow practice includes following:

### 1) Selection (Reproduction)

It is initialized first operator is based on selection criteria. It chooses the chromosomes from the population of parents, cross over and produce offspring. It is evident from evolution theory of "Survival of the fittest" given by Darwin.

There are many techniques involved for reproduction or selection operator such as,

- Tournament selection
- Ranked position selection
- Steady state selection

### 2) Cross Over

Population is enriched with better individuals after reproduction phase. Then crossover operator is integrated to the mating pool and create better strings. Crossover operator improves clones of good strings but does not create new ones. By recombining good individuals, the process is likely to create even better individuals.

### 3) Mutation

Mutation is a background operator. Mutation of a bit includes flipping it by changing 0 to 1 and vice-versa. After crossover, the mutation operator subjects to the strings to mutation. It facilitates a sudden change in a gene within a chromosome.

Thus, it allows the algorithm to see for the solution far away from the current ones. It ensures that the search algorithm is not trapped on a local optimum. The main consideration of this background operator is to prevent early convergence and maintain diversity within the population. The added advantage of performing modal state with behavioral integration is appreciable.

## A. Abbreviations and Acronyms

Response surface methodology is an integral mathematical tool which associates with modelling and assessment of problems in which a response of interest is influenced by several variables, and the objective is to reach out to optimal value. The inclusion of RSM,

Relationship between the preferred response and independent input variables could be produced as:

$$Y = f(x_1, x_2, x_3, \dots, x_n) \pm e_r$$

Where,

y -preferred response,

f - response function (or response surface),

x<sub>1</sub>, x<sub>2</sub>, x<sub>3</sub>, ..., x<sub>n</sub> - independent input variables,



er - fitting error

The predicted surface response in accordance to response function as obtained from graph plotting of function  $f$ , the closeness of  $f$  will help in reaching to true results. The second order polynomial regression model was framed which results out to be:

$$Y = c_0 + \sum_{i=1}^n c_i X_i + \sum_{i=1}^n b_i X_i^2 + \sum_{i,j=1}^n c_{ij} X_i X_j + e_r$$

The following considerations of assumptions of surface roughness is often shown with linear, crossed and squared product terms of  $X_i$ 's design finds the second-order response surface very precisely.

The second-order response surface representing the surface roughness ( $R_a$ ,  $\mu m$ ) can be expressed as a function of cutting speed, depth of cut and feed rate.

$$MRR = a_0 + a_1(A) + a_2(B) + a_3(C) + a_4(A^2) + a_5(B^2) + a_6(C^2) + a_7(AB) + a_8(BC) + a_9(CA)$$

$$Ra = a_0 + a_1(A) + a_2(B) + a_3(C) + a_4(A^2) + a_5(B^2) + a_6(C^2) + a_7(AB) + a_8(BC) + a_9(CA)$$

The following phase of documentation and experimentation saw the inclusion of machining parameters like depth of cut, feed rate and cutting speed on surface roughness and trials on included data sets were carried and machining on turning operation in dry condition was studied.

Here:

A = Cutting Speed

B = Feed Rate

C = Depth of Cut

### B. Regression Analysis

The efficacy of desired optimality condition is obtained with the integration of regression analysis. The regression analysis makes the use of experimental data to calculate the equation coefficients which shows the study results system at any point in the experimentally range. The inclusion of regression analysis in output-oriented machining parameters like feed rate, depth of cut and cutting speed was considered.

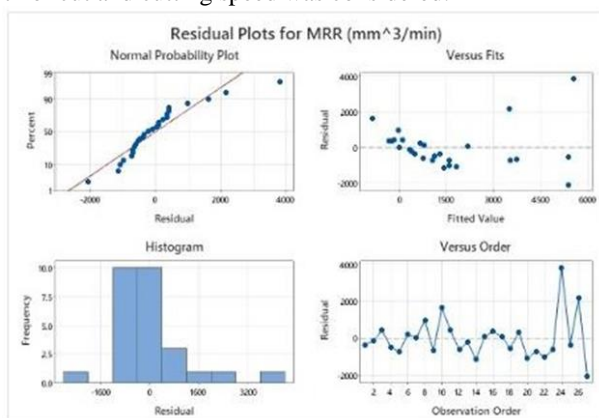


Fig. 8. Residual plots for MRR

By taking reference from Table 1 we have found out the coefficients for Material Removal Rate and Surface Roughness

and hence find the regression equation for the above given parameters and if we put the input parameters viz. Cutting Velocity, Feed Rate, Depth of cut into the equation we will get the desired experimental Material Removal Rate and Surface Roughness and then compare it with the predicted values.

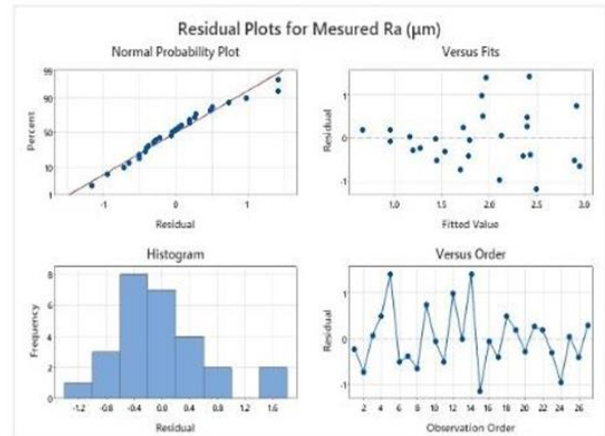


Fig. 9. Residual plots for Ra

### Regression Equation for Material Removal Rate (MRR):

$$MRR \text{ (mm}^3/\text{min)} = -3250 - 21.72A + 50978B - 2958C + 0.01396A^2 - 96052B^2 - 3983C^2 + 19.0AB + 3847BC + 20.84AC$$

### Regression Equation for Measured Surface roughness ( $R_a$ ):

$$\text{Measured } R_a \text{ (}\mu\text{m)} = -1.16 + 0.00175A + 12.8 \text{ feed rate} \cdot B + 0.17C - 0.000002A^2 - 10.8B^2 + 1.67C^2 - 0.00578AB + 1.67BC - 0.00578AC + 1.9B^2C + 0.00120AC$$

The computation and variance analysis were performed using ANOVA analysis where sums of squares and separated and then variance among these are checked due to different parameters and accordingly best fitted regression equation is formed. The ANOVA analysis terms were computed in accordance to MINITAB 21 calculation process flow. The design tree interface of ANOVA helped in data integration and regressive predicted values closure to statistically proven data sets.

In Fig. 8 and Fig. 9 residuals plot are shown for Material Removal Rate and Surface Roughness respectively. Normal Probability plot shows the fit of the distribution to the data and shows how the data is fitted along the distribution line. Versus Fits helps in verifying the random distribution and constant variance. Versus order plot helps in verifying that the residuals are independent from one another and showing no particular pattern among the residuals.

### C. Fitness Function

After finding coefficients for the regression equation from the Regression Analysis we got our fitness function for Material Removal Rate (MRR) and Surface Roughness ( $R_a$ ). Fitness Function is very important in applying Genetic Algorithm Optimization. Fitness Function is the equation which is dependent on the inputs given for the desired outputs for e.g. Here are inputs are Cutting Speed, Feed Rate, Depth of Cut and

are output are Material Removal Rate and Surface Roughness. Fitness Function tells how good or fit the data is according to the aims set for the outputs. Our output will be optimized on the basis of fitness function equation

*Fitness Function for Material Removal Rate (MRR):*

$$y(1) = -3250 - 21.72 * x(1) + 50978 * x(2) - 2958 * x(3) + 0.01396 * (x(1)^2) - 96052 * (x(2)^2) - 3983 * (x(3)^2) + 19.0 * (x(1) * x(2)) + 3847 * (x(2) * x(3)) + 20.84 * (x(1) * x(3))$$

*Fitness Function for Surface Roughness:*

$$y(2) = -1.16 + 0.00175 * x(1) + 12.8 * x(2) + 0.17 * x(3) - 0.000002 * (x(1)^2) - 10.8 * (x(2)^2) + 1.67 * (x(3)^2) - 0.00578 * (x(1) * x(2)) + 1.9 * (x(2) * x(3)) + 0.00120 * (x(1) * x(3))$$

## 7. Results of Optimization using GA

After developing fitness function, we have defined the bounds for the inputs and population size were determined which was taken as 50. Tournament Selection was used and for the reproduction the crossover function was set to 0.6. Adaptive feasible function was used as the crossover function and Single point mutation function was used.

The upper and lower bounds for the input parameters are given below:

$$\begin{aligned} 180 &\leq \text{Cutting Speed} \leq 710 \\ 0.2 &\leq \text{Feed Rate} \leq 0.4 \\ 0.2 &\leq \text{Depth of Cut} \leq 0.6 \end{aligned}$$

Result for material removal rate and surface roughness obtained from genetic algorithm is,

Table 4  
Optimized results from GA

S.No.	Material Removal Rate (mm <sup>3</sup> /min)	Surface Roughness(μm)
1.	698.66	0.91

The results obtained from genetic algorithm for Material Removal rate and Surface Roughness are summarized in the Table 4. We obtained total of 18 optimized results out of which are criteria for selection was that Material Removal Rate should be maximum and for Surface Roughness was chosen to be less than equal to 1μm. The surface Roughness was chosen as per the industry requirements.

## 8. Conclusion

In our study the optimization of conventional turning operation was carried out with considerations of machining parameters like feed rate, cutting speed and depth of cut and optimal MRR was predicted, industrially viable using ANN and GA. For studying relationship between linearity of plots between actual and predicted, ANN model was developed and accuracy up to – were achieved on training and datasets. The model showed the closeness to input values and values pertaining to optimality was selected. The accounted error fall well within industrial limits were achieved. The integration of GA optimization led to calibrate targeted machining parameter

suitable index and achieve minimal loss of material finish compromise.

## 9. Future Scope

We are hopeful with successful integration of ANN-GA tools in being proactive prediction methodology, industrial machining parameters like nose radius, rake angle can be included for enhancement of current working process. The added advantage of computational contingent ability of process flowing complex machining ability and minimum machining time while considering technological and material constraints. The promising solutions to existing gap in behavioral prediction in intelligent machining and suitable solution for automatic selection of the machining parameters may open new avenues to real time-based manufacturing process optimization.

The design mode failure incorporation can be minimized up to greater extent and accuracy in desired manufactured product can be achieved leading to overall increase in productivity. The cycle time of failure and material enhancement can be viewed in real perspective.

## References

- [1] T. M. Chenthil Jegana R. Chitrab and R. Ezhilarasuc, "Analysis for optimization of machining parameters of CNC turning steel", IOP Conf. Ser. Mater. Sci. Eng. 2018,
- [2] S. Neseli, S. Yaldiz, and E. T'urkes, "Optimization of tool geometry parameters for turning operations based on the response surface methodology", Measurement, vol. 44, no. 3, pp. 580–587, 2011.
- [3] Gupta M, Kumar S, "Multi-objective optimization of cutting parameters in turning using grey relational analysis", Int. J. Ind. Eng. Comput. 2013, 4, 547–558.
- [4] Ndaruhadi, P.Y.M.W.; Sharif, S.; Noordin, M.Y.; Kurniawan, D., "Effect of Cutting Parameters on Surface Roughness in Turning of Bone", Adv. Mater. Res. 2014, 845, 708–712.
- [5] Asilt'urk and H. Akkus, "Determining the effect of cutting parameters on surface roughness in hard turning using the taguchi method", Measurement vol. 44, no. 9, pp. 1697–1704, 2011.
- [6] Dhabale, R.; Jatti, V.S.; Singh, T.P., "Feature Based Modeling and automated process plan generation for turning Components", Procedia Mater. Sci. 2014, 6, 961–966.
- [7] Ramu I, Srinivas, P, Venkatesh, "Analysis for optimization of machining parameters of CNC turning steel", IOP Conf. Ser. Mater. Sci. Eng. 2018.
- [8] Tiagrajah V. Jannharimal, Noozia Ahmad, "Multi Objective Optimization for Turning Operation using Hybrid Extreme Learning Machine and Multi Objective Genetic Algorithm", vol. 7, no. 4, 2018.
- [9] V. S. Sharma, Rakesh Sehgal, "Estimation of cutting forces and surface roughness for hard turning using neural networks" Journal of Intelligent Manufacturing 19(4):473-483, August 2008.
- [10] Azlan Mohd. Jain, "Integrated ANN-GA for estimating the minimum value for machining performance", International Journal of Production Research. January 2012 50(1):191-213.
- [11] M. Solimanpur and F. Ranjdoostfard, "Optimisation of cutting parameters using a multi-objective genetic algorithm", vol. 47, no. 21, pp. 6019-6036, 2009.
- [12] S. K. Nayak, J. K. Patro, S. Dewangan, and S. Gangopadhyay, "Multi-objective optimization of machining parameters during dry turning of AISI 304 austenitic stainless steel using grey relational analysis", Procedia Materials Science, vol. 6, pp. 701-708, 2014.
- [13] S. Xie and Y. Guo, "A review on machine learning in 3D printing: applications, potential, and challenges", Journal of Computational Information Systems, vol. 7, no. 5, pp. 1714-1721, 2011.
- [14] S. Yang and U. Natarajan, "Multi-objective optimization of cutting parameters in turning process using differential evolution and non-dominated sorting genetic algorithm- approaches", The International Journal of Advanced Manufacturing Technology, vol. 49, no. 5-8, pp. 773-784, 2010.



# S-DCNN: stacked deep convolutional neural networks for malware classification

Anil Singh Parihar<sup>1</sup> · Shashank Kumar<sup>1</sup> · Savya Khosla<sup>1</sup>

Received: 24 June 2021 / Revised: 5 November 2021 / Accepted: 9 February 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

Malware classification continues to be exceedingly difficult due to the exponential growth in the number and variants of malicious files. It is crucial to classify malicious files based on their intent, activity, and threat to have a robust malware protection and post-attack recovery system in place. This paper proposes a novel deep learning-based model, S-DCNN, to classify malware binary files into their respective malware families efficiently. S-DCNN uses the image-based representation of the malware binaries and leverages the concepts of transfer learning and ensemble learning. The model incorporates three deep convolutional neural networks, namely ResNet50, Xception, and EfficientNet-B4. The ensemble technique is used to combine these component models' predictions and a multilayered perceptron is used as a meta classifier. The ensemble technique fuses the diverse knowledge of the component models, resulting in high generalizability and low variance of the S-DCNN. Further, it eliminates the use of feature engineering, reverse engineering, disassembly, and other domain-specific techniques earlier used for malware classification. To establish S-DCNN's robustness and generalizability, the performance of proposed model is evaluated on the Maling dataset, a dataset collected from VirusShare, and packed malware dataset counterparts of both Maling and VirusShare datasets. The proposed method achieves a state-of-the-art 10-fold accuracy of 99.43% on the Maling dataset and an accuracy of 99.65% on the VirusShare dataset.

**Keywords** Deep convolutional neural networks · Ensemble model · Malware classification · Pattern recognition · Security · Transfer learning

## 1 Introduction

The software that aims to harm a system without the user's knowledge or consent is malicious software (abbreviated as malware). Malware typically performs harmful actions that violate the system's confidentiality, integrity, and availability. Such harmful actions make

---

✉ Anil Singh Parihar  
parihar.anil@gmail.com

<sup>1</sup> Machine Learning Research Lab, Department of Computer Science & Engineering,  
Delhi Technological University, Delhi, India

the system prone to unauthorized access and modification of sensitive information. The ongoing research in this field revealed the staggering growth rate of malware (i.e., malicious) attacks and the pressing need to alleviate them. For instance, McAfee Labs reported [2] that in the first quarter of 2019, Ransomware attacks increased by 118%, and new classes of the same are also identified. Similarly, Kaspersky Security Network (KSN) published a report [16] showing that malware attacked more than 70% of the users, which is devised to collect users' data without their knowledge. With the increasing dependency of individuals, institutions, and organizations on computers and databases for their several services, such as the ability to store important and, perhaps, sensitive information, there is an evident need to restrain the ever-increasing rate of malware attacks.

In literature, Signature-based methods and Heuristic-based methods are two conventional techniques for malware classification. The signature-based malware classification method compares the suspected file with the database (a blocklist) containing the signatures of already known malicious software. It detects and classifies the dubious file if a minimum of 8-bit code pattern matches the signatures database. However, this method poses a critical drawback. Since the signature database is extracted from already known malicious software, all the signatures are well-known to the malware authors. Thus, malware authors can easily elude this technique by using basic obfuscation methods. Another drawback lies in its inability to classify malware files that are variants of already known malware. Hence, this anti-malware approach proves futile and limiting when dealing with new and updated malware programs. Heuristic-based methods are dependent on comparing suspected malware's code with already known malware present in the heuristic database. The system marks the suspected file as a potential threat if the comparison is similar to the database. However, the size of the required database is increasing exponentially with the increasing number of malicious software.

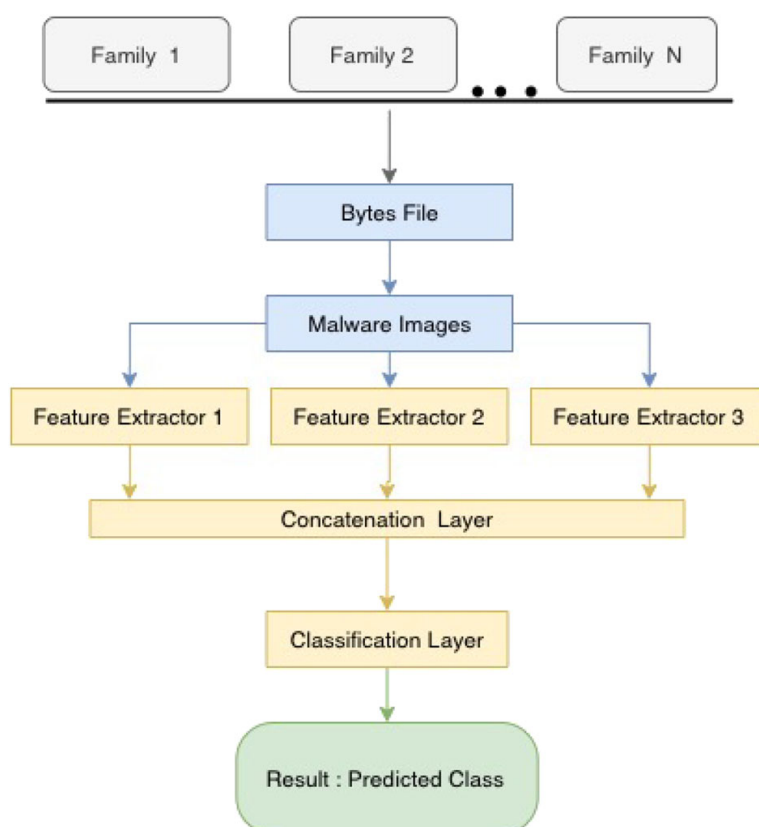
As the dependency on technology is increasing, thus malware analysis becomes essential. Malware analysis involves learning the source, operation, intent, and likely outcomes of malicious software. Static analysis and dynamic analysis are two techniques for malware analysis. The static analysis is a signature-based method that analyses the suspicious file by dismantling it without any execution. The static analysis method performs reverse-engineering on an executable version of the malware to gather the human interpretable code. The static analysis method carefully investigates the reverse-engineered code to analyze the program's intent. The investigation in static analysis is carried out through several techniques involving fingerprints, debugging, memory dumping, and many others. On the other hand, dynamic analysis is a behavior-based method that executes the malware in a safe, isolated and controlled setting. The dynamic analysis executes the suspicious file and observes the execution for any suspicious activity. Since this risks harming the system, execution occurs in an isolated virtual environment, such as a sandbox. A debugger is used to ensure a proper understanding of the purpose and function of the suspicious file. However, malware authors can elude these techniques with simple obfuscation techniques.

In the literature, researchers proposed various algorithms for the detection and classification of malware. A set of algorithms [27, 29] depend on extensive feature engineering to build a database of malware features. However, as the adversaries generate new malware every day, updating the database containing the features with the samples of new malware files becomes tough. Further, traditional machine learning methods perform well in a small domain [21, 22] where the dataset is small, and the target classes are relatively smaller. However, as the complexity increases and the number of target classes grows, machine learning (ML) methods show their limited performance. The traditional ML methods require more time to be trained on large datasets and lack the ability to learn the highly

multidimensional features required for accurate classification. Hence, the traditional methods become cost-ineffective. Therefore, it is not practical to use traditional approaches to malware classification in the real-world setup. Recently, researchers [8, 15, 32] proposed deep learning-based techniques for malware classification. These techniques have shown considerable improvement in malware detection and classification tasks.

In this paper, a deep learning-based method (S-DCNN) is proposed to perform the task of malware classification. The proposed S-DCNN architecture uses the image-based representation of malicious files to train a classifier for malware binaries as shown in Fig. 1. It can be noted from Fig. 1 that the proposed S-DCNN extracts image based features from the malware files. The proposed S-DCNN model uses the paradigms of transfer learning to leverage the capabilities of state-of-the-art computer vision models. By stacking sub-models having significantly different learning mechanisms, the ensemble yields a robust classifier. Further, due to the texture-based analysis of malware images, the proposed model is resilient to obfuscation techniques like section encryption and packing [30]. The main contributions of this work are as follows:

- A novel ensemble of deep convolutional neural networks is proposed for performing the task of malware classification. The proposed model is simple, robust, computationally efficient, and thus, ideal for real-world applications. Further, it does not require feature engineering and domain expertise.
- The performance of models trained on non-obfuscated malware is evaluated on obfuscated malware exhibiting various levels of obfuscation. To our knowledge CNN models are being evaluated in this setting for the first time.



**Fig. 1** Flow-graph of proposed S-DCNN for malware classification



- A comprehensive evaluation of the proposed model is performed using several evaluation metrics and carry out extensive experimentation for model design to ensure the adequacy of the model's architecture. Based on a comparison with recent malware classification techniques, it is concluded that the proposed model exhibits superior performance.

The organization of the rest of this article is as follows. Section 2 presents a review of the recent techniques for malware classification. A detailed discussion of the proposed S-DCNN model is shown in Section 3. Section 4 demonstrates the results and experiments to establish the competence of the proposed model. Finally, the conclusion of the proposed model is shown in Section 5.

## 2 Related work

Malware detection is the task of identifying whether a given executable file is malicious or not. Once a file is identified as malicious, malware classification is performed to identify its malware family. Nowadays, machine learning approaches solve various problems like event forecasting [20], collision avoidance [37], patient outcome prediction [6] and others. Researchers [25, 29, 38, 45] proposed malware detection and classification based on machine learning techniques, namely SVM, AdaBoost, Decision Trees, and others. Schultz et al. [38] proposed an algorithm for static malware analysis based on features extracted from program executables, strings, and byte n-grams. The algorithm uses the Multinomial Naïve Bayes to classify malware files and achieves an accuracy of 97.11%. Schultz et al. used the number of unique calls to the system used within DLLs, a list of printable characters, byte sequence, and others as features for classification. Zhang et al. [45] proposed an algorithm for ransomware classification. The algorithm uses n-grams of opcodes as input for static malware analysis. Moskovitch et al. [25] proposed the use of opcode-based n-grams as features and experimented with various machine learning methods like Decision Trees, Naïve Bayes, Artificial Neural Networks, and others. Narayanan et al. [29] used PCA to extract global level features, which are used as input to the nearest neighbor classifier to classify the input files into one of 8 malware families. Angelo et al. [9] proposed association rules based malware classification. It uses subsequences of API calls and markov chain to perform malware classification. These techniques uses handcraft features for malware classification which restricts the performance of the algorithms. The proposed S-DCNN uses automatic feature extractor which helps in improving the performance.

Deep Learning (DL) has garnered significant interest in various fields [4, 17, 18, 24] over the past few years. DL methods provide solutions for most of the challenges better than the traditional methods. Jain et al. [14] analysed the performance of malware classification based on CNN and extreme learning machines. Saxe et al. [36] proposed a four-layered Neural Network for Binary Malware Classification. The method trains four fully connected layers on features extracted from the processed malware binaries. Pascanu et al. [34] developed an algorithm that uses Recurrent Neural Networks to capture the relationships across time in the event streams of API calls encoded into 114 high-level events. Alsulami et al. [1] proposed to train Convolutional Recurrent Neural Network on the Microsoft Windows prefetch files. It classifies samples that belong to rare malware families. Further, the model is capable of tuning on new malware samples. Kolosnjaji et al. [19] proposed an architecture consisting of CNN and LSTM as components. The CNN component captured the local relationships, while the LSTM layers captured the long-term dependencies. The input data



consisted of sequences of API calls made to the system. Yuan et al. [44] proposed a byte-level malware classifier using the markov images. The algorithm generates markov images from the binary files and then performs the classification of the sample. These algorithms uses the processed binary files as the features that limits the capabilities of classifiers and sometimes require larger amount of time.

Nataraj et al. [30] proposed a new type of feature extraction method. Nataraj et al. proposed transforming binary samples into grayscale images, which helps in classification using any general image classification pipeline. In another work, Nataraj et al. [31] demonstrated the superiority of binary texture analysis over other feature extraction methods for malware classification. GIST feature [33] is used for extracting a texture-based feature from the generated malware images. Nataraj et al. performed k-Nearest Neighbors-based classification on the extracted features. Naeem et al. [27] proposed an efficient framework named Malware Image Classification System for IoT infrastructure. The system first converts the incoming malware files into grayscale images and then extracts the global and local features for classification from the grayscale image. Gibert et al. [11, 12] developed a file-agnostic method for malware classification using a CNN-SVM architecture. The output features of the CNN are given as input to an SVM to classify the malware files into their respective families. Cui et al. [8] proposed the use of bat algorithm for data balancing and employed CNN-based architecture to achieve an accuracy of 94.50% on Maling Dataset[30]. Nisa et al. [32] proposed a distinguished feature extraction method from an input image. The feature extractor [32] combines the Scale Feature Texture Analyzer (SFTA) and features from deep learning architectures (like InceptionV3 and AlexNet, pre-trained on the ImageNet dataset). The combined features are used for training of different variants of Decision Trees, SVM, KNN, and other classifiers. These methods uses machine learning based classifiers which limits the performance of the algorithm by sometimes avoiding the relationship among the features.

Vasan et al. [40] proposed IMCFN, a CNN-based model for malware classification. The method converts malware binaries into colored images and uses data augmentation techniques to fine-tune a deep convolutional neural network previously trained on the ImageNet dataset. In another work, Vasan et al. [41] proposed an ensemble of deep learning models for classifying malware files. The architecture comprises Resnet50 and VGG16 fine-tuned on the malware images. The three one-vs-all SVMs trained on the features extracted from the models. The k-fold cross-validation for the model is not given to support the generalizability. Saadat et al. [35] proposed a CNN-XGBoost model that uses CNN for feature extraction and XGBoost for classification. Saadat et al. demonstrated that CNN-XGBoost architecture performs better than CNN-SVM, CNN-Softmax, and other configurations. Venkatraman et al. [42] proposed a unified hybrid deep learning and visualization technique for malware detection and classification. For classification purposes, the application of hybrid image-based approaches with deep learning architectures is investigated. Verma et al. [43] proposed a first and second-order texture statistics for multiclass malware classification. The algorithm analyses the texture of the file to classify the malware. Transfer learning enables the methods to solve the problems of other domains [3]. Gao et al. [10] use transfer learning based semi-supervised method for malware classification. Gao et al. proposed an RNN-based byte classifier and an asm classifier. Çayır et al. [5] introduced an ensemble of Capsule Networks (CapsNet) for malware classification. Treating the CapsNet as weak classifiers, an ensemble using the bootstrap aggregation is developed. The CNNs extract a 128-dimensional feature vector from the input malware images, then fed the features to the capsule network.

Most of the existing works concentrate on extracting features from the image-based representations of the malware files and then classifying them using different machine learning techniques. Many researchers used CNN-based architectures for feature extraction from malware images. The techniques like Capsule Networks, CNN-SVM, and CNN-XGBoost give modest results only. As per observation, the methods that perform very well use a combination of multiple feature extractors to generate feature representations for malware files and then use convoluted machine learning models to perform classification from these features. However, the lack of explanation behind the choice of models used for the feature extraction and the rationale behind the particular combination leaves a knowledge gap. A diverse and comprehensive set of feature representations from the input images with even a simple classifier can achieve very high accuracy. To this end, This paper presents S-DCNN, an ensemble of carefully chosen CNN-based models that generates a robust feature representation of the input so that even a simple ANN classifier can achieve state-of-the-art results using these features.

### 3 Methodology

This section contains a detailed discussion of the proposed method and architecture. Three deep convolutional networks, namely ResNet50, Xception, and EfficientNet-B4, are trained for the concerned task. Their concatenated outputs are used as features for a multilayered perceptron, which gives the final output. The proposed method is capable of tackling standard obfuscation techniques such as section encryption and packing due to the use of texture-based analysis of the malware. Moreover, since the proposed model is an ensemble of models of considerably different learning mechanisms, it forms a good representation of the malware images. Consequently, it is robust and generalizable (the ablation experiments in Section 4 further substantiate this).

The upcoming subsection explains the dataset, the sub-models, the fine-tuning process and the training of the final model.

#### 3.1 Dataset

The proposed method in this paper uses the Maling Dataset [30]. The dataset contains 9339 data points corresponding to 25 classes. These data points are image-based representations of malware binaries, and the classes are their respective malware families.

Further, the 343 samples corresponding to 5 malware families from VirusShare.com are scrapped to test the generalizability of the proposed model. Table 1 gives the collected data's statistics. To test the robustness of the proposed model against obfuscation, a packed

**Table 1** Dataset downloaded from VirusShare

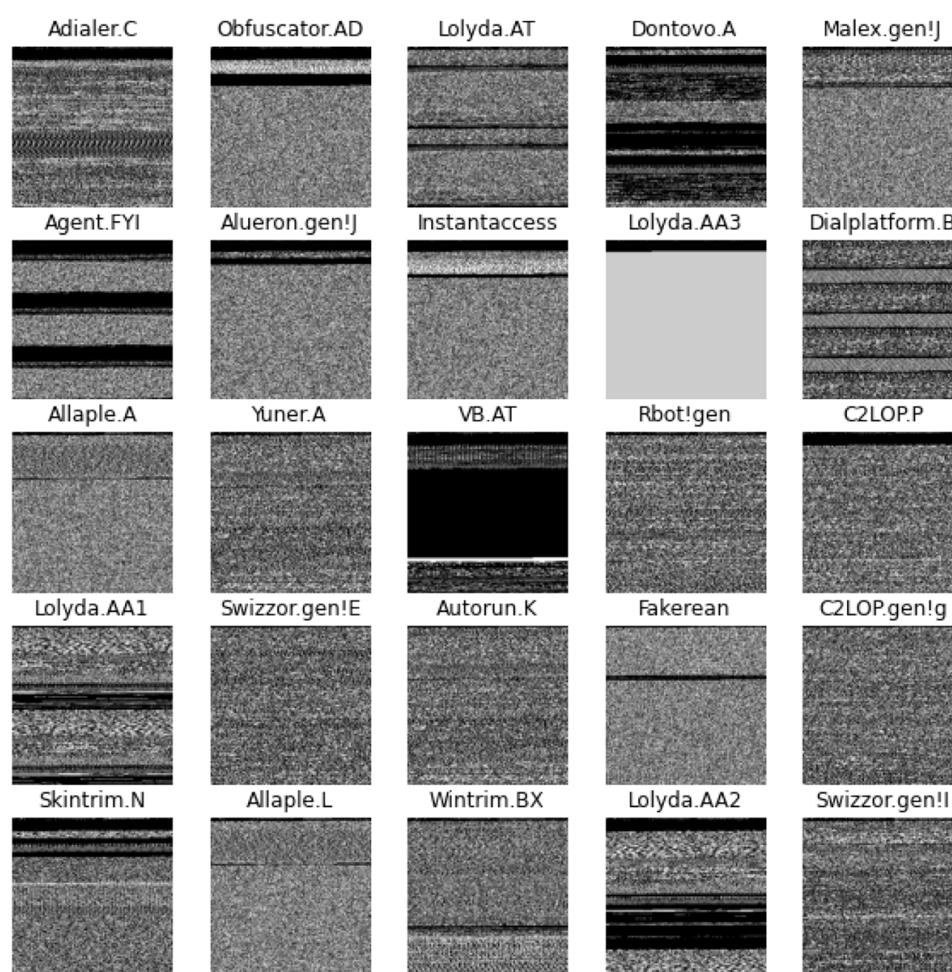
Family	Number of samples
Lolyda.BF	100
VB	116
Prepsram	39
Eksor.A	40
Wintrim.BF	48

malware dataset is generated from the VirusShare dataset for testing purposes. The images are generated for the packed and unpacked dataset by combining the bits of malware binaries into 8-bit vectors and organizing these vectors into a two-dimensional array. The 2D array is then rendered as a grayscale image.

Figure 2 shows a sample of image-based representation for each malware family, and Fig. 3 shows the data distribution. The resized images of uniform shape  $224 \times 224$  are considered for the use case.

### 3.2 Transfer learning

Transfer Learning is a learning paradigm in which a model trained for one task is used as an initialization point for a different task. Transfer learning is extremely useful as it saves training time, leads to better-performing models, and solves the issue of data deficiency. The pre-trained models on the ImageNet dataset as the starting point are used for leveraging the concept of transfer learning. Then, the fine-tuning of these models for the task of malware classification is performed. Once fine-tuned, the models act as feature extractors for the malware images. The sub-models that form a part of the ensemble architecture are ResNet50, Xception, and EfficientNet-B4.



**Fig. 2** Image-based representation of malware binaries

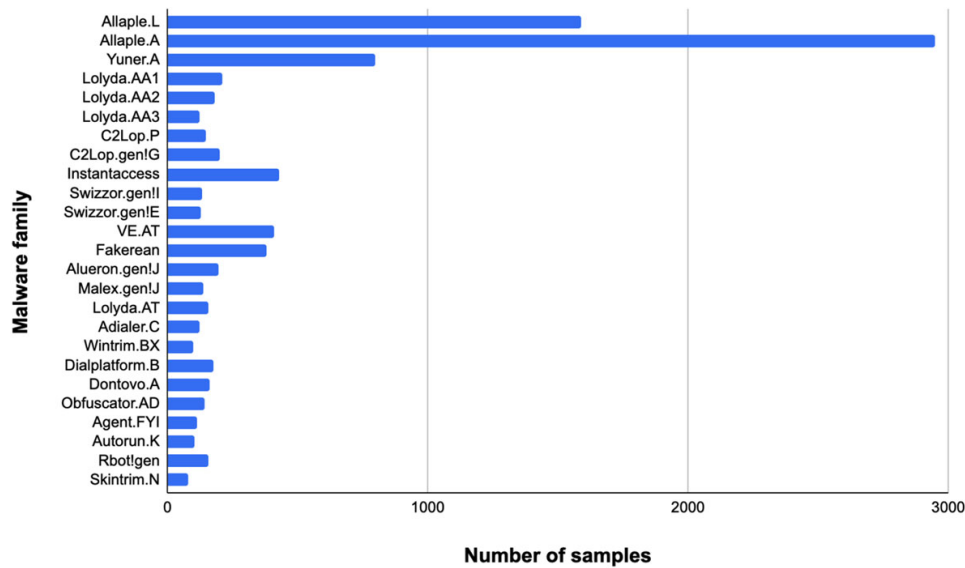


Fig. 3 Data distribution

### 3.2.1 ResNet50

Residual Network [13], or ResNet for short, is a deep convolutional neural network that deploys a series of residual blocks containing skip connections. These skip connections address the issue of diminishing accuracy with an increasing number of layers in the model. Thus, it facilitates the formulation of profound neural networks. ResNet50 is a Residual Network having 50 layers. This model has led to a series of breakthroughs in the domain of image classification. Consequently, it is a prevalent choice for building transfer-learning-based models.

### 3.2.2 Xception

Xception [7] is a type of convolutional neural network that uses the concept of modified depthwise separable convolutions. In general, depthwise separable convolution uses depthwise convolution (channel-wise spatial convolution) followed by pointwise convolution ( $1 \times 1$  convolution). The modified depthwise separable convolutions used in Xception perform these operations in reverse order; i.e., pointwise convolution is followed by depthwise convolution. These modified depthwise separable convolutions are used throughout the architecture as inception modules. Further, Xception deploys residual connections like ResNet to give an accuracy boost to the deep learning architecture.

### 3.2.3 EfficientNet-B4

Tan and Le proposed EfficientNet [39] that puts forward a systematic way of scaling up convolutional neural networks. It explores scaling across three dimensions - width (the number of channels), depth (the number of layers), and resolution (resolution of the input image). It proposes compound scaling wherein these three dimensions are systematically scaled to ensure optimal performance (accuracy) and efficiency (FLOPS). EfficientNet establishes its efficacy in the domain of transfer learning as it achieves state-of-the-art accuracy in 5 out of the eight standard transfer learning datasets. EfficientNet-B4 is used in the proposed ensemble model.

### 3.2.4 Fine-tuning

Each of the models mentioned earlier is fine-tuned to perform malware classification on the Maling dataset. These models are initialized with the pre-trained ImageNet weights, and the top-most layer of 1000 neurons is removed. While fine-tuning, the batch normalization layers are kept non-trainable, and a dropout is applied on the final feature map of the sub-models. Finally, a fully connected layer containing 25 classes is appended to generate the final output. The customized models are then trained using the categorical cross-entropy loss given by (1).

$$L_{cce} = \sum_i y_i \cdot \log(p_i) \quad (1)$$

Here,  $L_{cce}$  represents the categorical cross-entropy loss,  $y_i$  represents the true class label of  $i^{th}$  data sample, and  $p_i$  denotes the model's prediction for  $i^{th}$  sample.

### 3.3 Ensemble learning

Figure 4 shows the complete model architecture of S-DCNN. Three models, viz. ResNet50, Xception, and EfficientNet-B4 are stacked together to form the ensemble. The feature vectors derived from these models are concatenated to form a single input feature vector. The concatenated feature vector is fed as input to a multilayered perceptron (MLP). While training the final ensemble model, the three fine-tuned sub-models act as feature extractors and are not further trained. In contrast, the MLP is trained using the categorical cross-entropy loss as given by (1).

## 4 Results and experiments

This section shows a discussion about the implementation details, results and ablation experiments.

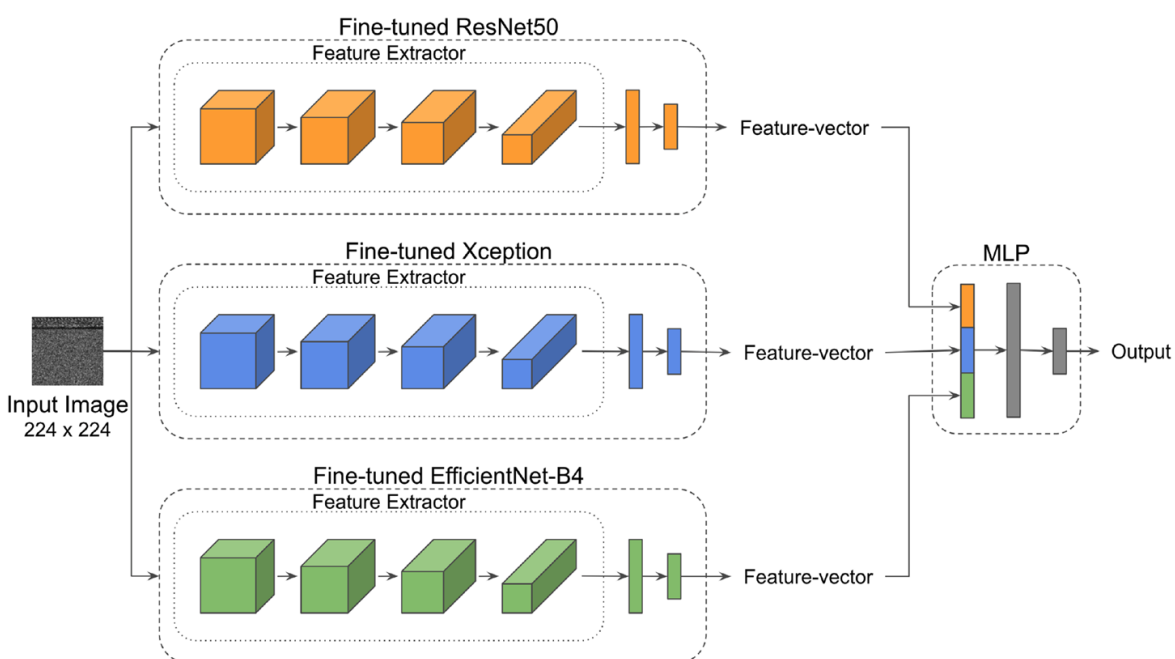


Fig. 4 S-DCNN model architecture



## 4.1 Implementation details

### 4.1.1 Model design and parameters

After trying out several different deep convolutional models, It is observed that the ensemble of ResNet50, Xception, and EfficientNet-B4 performed the best for our use case. The input size of the grayscale image is fixed to  $224 \times 224$  for all models.

For the models mentioned above, the model uses the ImageNet weights as the starting point. The topmost fully connected layer containing 1000 neurons is removed and the rest of the model is used as a feature extractor (as highlighted in Fig. 4). After removing the top layer, the dropout is applied on the extracted output having a dropout rate of 0.1. The method flattens the feature map and appends a fully connected layer containing 25 neurons to generate the output. These models are trained using Adam optimizer. Further, the training uses a learning rate scheduler with a polynomial decay factor of 0.01 and an early stopping mechanism. The initial learning rate is  $5e-5$ , and the batch size is 32.

The 25-dimensional outputs from the three models are concatenated for the ensemble model to form a 75-dimensional feature vector. This feature vector is used as the input to a multilayered perceptron (MLP). The hidden layer of MLP is a fully connected layer of 75 neurons and a dropout with a rate of 0.5. The output layer is a fully connected layer of 25 neurons, which generates the final output. The model uses ReLU activation on the hidden layer and SoftMax classifier on the output layer. The ensemble is trained using Adam optimizer and a batch size of 32. Again, the learning rate scheduler with a polynomial decay factor as 0.01 and an early stopping mechanism to train the MLP is used. The initial learning rate is  $1e-3$ . A single Nvidia Tesla T4 GPU is used for all the experiments.

### 4.1.2 Evaluation metrics

Since the size of the dataset is small and there is no pre-defined train-test split, it would be inappropriate to randomly divide the dataset into a fixed training and testing sets and then evaluate the model's performance. Thus, the stratified 10-fold cross-validation is performed to evaluate the performance of S-DCNN. The 10-fold cross-validation ensures that the proposed model is robust and has a good generalization ability.

The statistical measures used to evaluate the performance of proposed model are accuracy, precision, recall, and F1-score. Taking  $TP$ ,  $TN$ ,  $FP$ , and  $FN$  as the number of true-positive, true-negative, false-positive, and false-negative predictions respectively. Equations (2), (3), (4), and (5) are the formulae of the concerned evaluation metrics:

$$Accuracy (A) = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Precision (P) = \frac{TP}{TP + FP} \quad (3)$$

$$Recall (R) = \frac{TP}{TP + FN} \quad (4)$$

$$F1 - score (F) = \frac{2 \times P \times R}{P + R} \quad (5)$$



For 10-fold cross-validation scores, the dataset is randomly shuffled and divided into ten groups. Then the model was trained and evaluated ten times such that each group got to be the testing set once while the other nine combined to form the training set. Equation (6) gives the generic form of 10-fold evaluation metrics used to evaluate the model's performance.

$$E = \frac{1}{10} \times \sum_{i=1}^{10} E_i \quad (6)$$

Here,  $E$  represents the evaluation metrics i.e.,  $E \in \{A, P, R, F\}$ , and  $E_i$  is the value of the evaluation metric for  $i^{th}$  fold.

## 4.2 Results

Table 2 shows the 10-fold cross-validation results of the proposed model and the sub-models. One may observe that the ensemble of ResNet50, Xception, and EfficientNet-B4 performs better than the three models individually. For example, the ensemble performs 0.52 percentage points better in 10-fold accuracy and 0.53 percentage points better in 10-fold F1-score from the best performing sub-model. Highlighted values show the results of the proposed method.

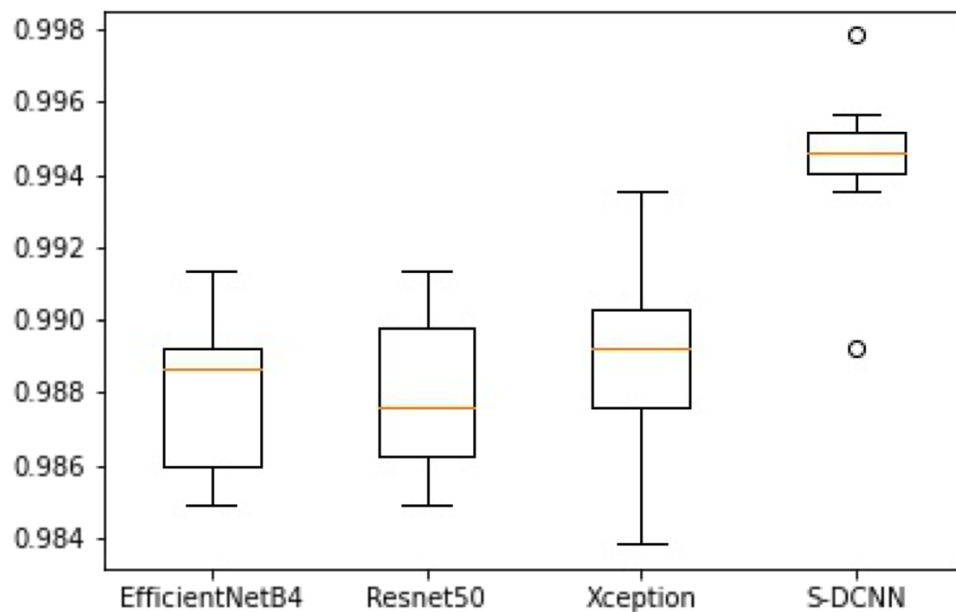
Further, the box plot in Fig. 5 compares the distribution of accuracies achieved by S-DCNN and the constituent models in different folds. It can be observed that the ensemble model achieves very stable and high accuracy across folds while achieving a fold-wise maximum accuracy of 99.78%. The mean accuracy achieved by the S-DCNN (99.43%) is far above the maximum accuracy achieved by any base model in any fold. This observation establishes a clear advantage of the S-DCNN over Resnet-50, EfficientNet-B4, and Xception.

In literature, 10-fold accuracy is widely used as the evaluation metric for the malware classification task on the Maling dataset. Table 3 compares the performance of S-DCNN with other recent models. One may note from the table that S-DCNN gives the best accuracy values. Further, S-DCNN demonstrates excellent efficiency with the prediction time per sample of 0.062 seconds on the NVIDIA Tesla T4 GPU. Highlighted values indicate the results of the proposed method.

Figure 6 shows the performance of S-DCNN using a confusion matrix. Due to the 10-fold cross-validation, ten confusion matrices are generated (one for every fold) and then added to give the final matrix. As can be seen, the model got a little confused between swizzor.gen!i and swizzor.gen!i. The confusion is because of a high level of similarity between samples of these classes. Further, it misclassified some samples from swizzor.gen!i class as c2lop.p and some samples of c2lop.gen!g class as c2lop.p. Other than these, the model performed well for all other classes. These results also demonstrate that S-DCNN is resilient to obfuscation. For example, despite the multilayered encryption of the code section in for allaple.l and allaple.a, the model is able to classify all samples of these

**Table 2** Classification performance of S-DCNN and the sub-models

Model	Accuracy	Precision	Recall	F1-score
ResNet50	98.80%	98.87%	98.75%	98.81%
Xception	98.91%	98.95%	98.85%	98.90%
EfficientNet-B4	98.80%	98.90%	98.72%	98.81%
<b>S-DCNN</b>	<b>99.43%</b>	<b>99.44%</b>	<b>99.43%</b>	<b>99.43%</b>



**Fig. 5** Box Plot of the sub-models and the ensemble model

classes correctly, thereby demonstrating that it can handle sectional encryption in malware files. Similarly, S-DCNN is able to classify all samples of the obfuscator.ad class correctly.

S-DCNN achieves a 3-fold cross-validation accuracy of 99.65% on the malware executables collected from VirusShare.com. Further, to demonstrate the robustness of the proposed ensemble model against obfuscation, The testing is performed on packed malware files after training them on unpacked malware files. UPX packer is used for packing the executable files collected from VirusShare.com. Table 4 shows the accuracies achieved by S-DCNN and its component models. It can be observed that S-DCNN outperforms the component models, which establishes a clear advantage. The results of the proposed method are shown in bold.

**Table 3** Comparative study of different malware classification models

Model	Accuracy
Nataraj et al. [30]	97.18%
Liu et al. [23]	98.90%
Kalash et al. [15]	98.50%
Naeem et al. [28]	98.00%
Cui et al. [8]	94.50%
Gibert et al. [11]	98.48%
Naeem et al. [26]	98.18%
Nisa et al. [32]	99.30%
Yuan et al. [44]	97.30%
Verma et al. [43]	98.97%
<b>S-DCNN</b>	<b>99.43%</b>

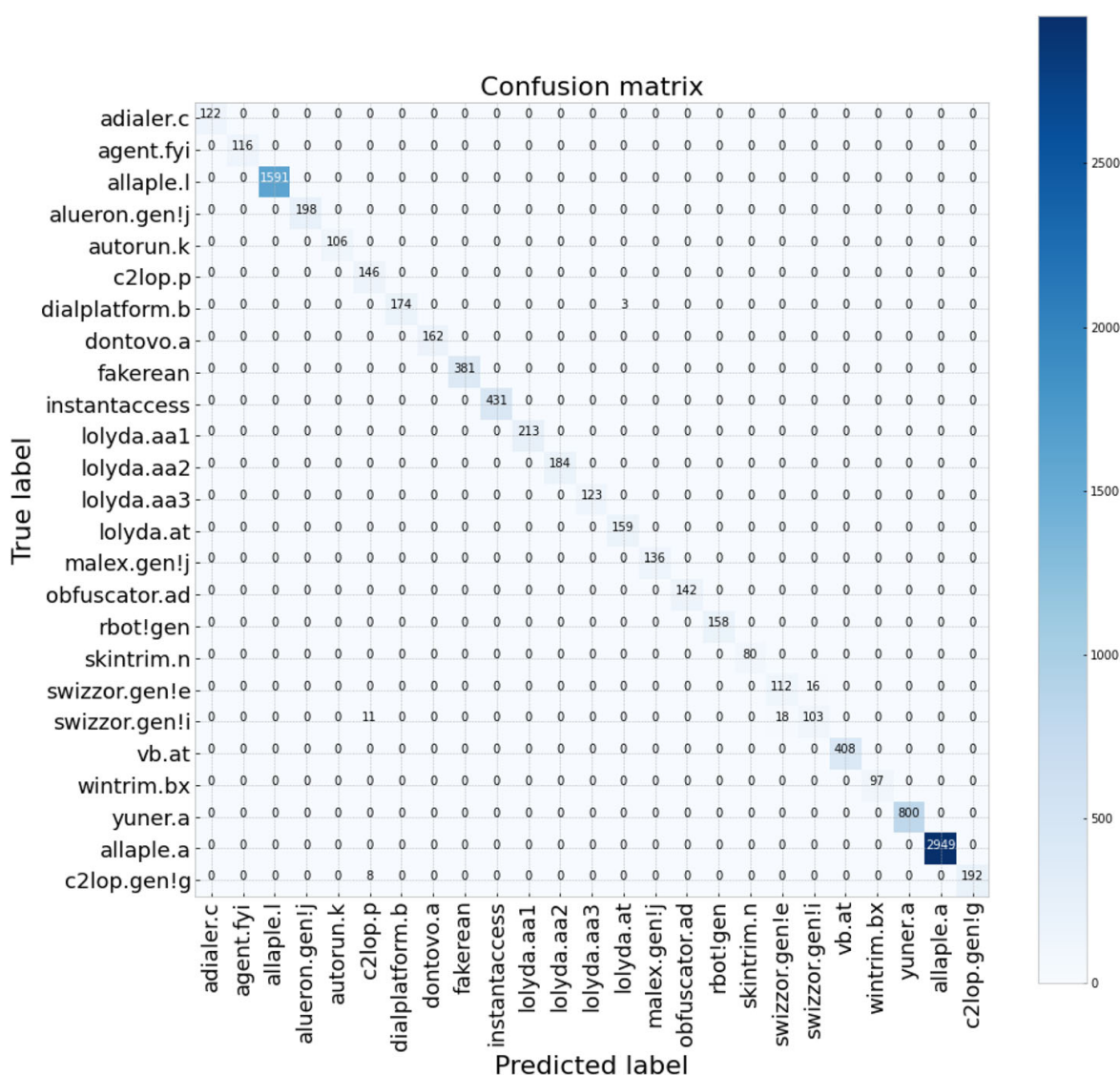


Fig. 6 Confusion Matrix for S-DCNN

## 4.3 Ablation experiments

### 4.3.1 Model selection

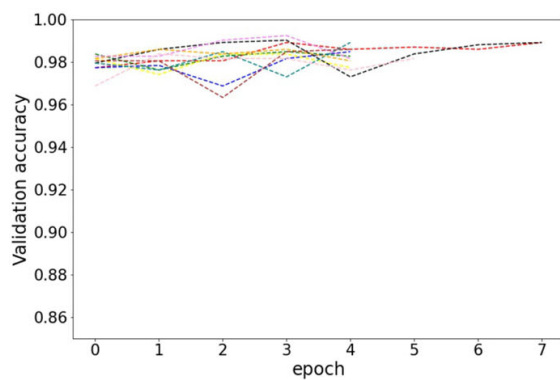
In computer vision, some of the most recent advances are made by leveraging the ideas of residual connections, inception modules, and developing efficiently scaled model architectures. It is hypothesized that an ensemble comprising of all the techniques mentioned above is capable of learning robust and diverse features from any given input. Thus, the

**Table 4** Classification performance of S-DCNN on packed malware executables

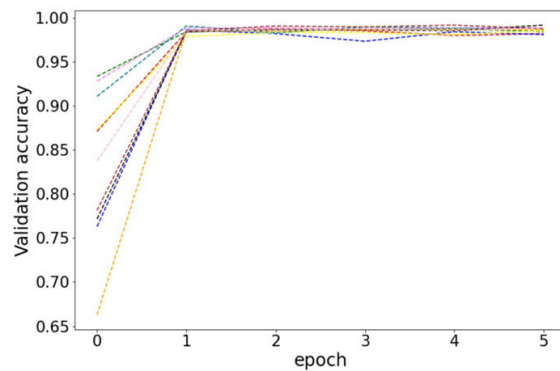
Model	Accuracy
ResNet50	85.40
Xception	86.30
EfficientNet-B4	86.40
<b>S-DCNN</b>	<b>89.70%</b>

**Table 5** Performance comparison of different deep convolution networks

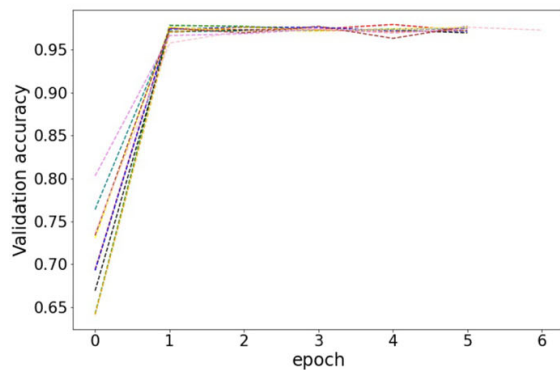
Model	Accuracy	Precision	Recall	F1-score
VGG16	98.17%	98.17%	98.06%	98.11%
ResNet50	98.80%	98.87%	98.75%	98.81%
InceptionV3	97.63%	97.69%	97.56%	97.63%
Xception	98.91%	98.95%	98.85%	98.90%
EfficientNet-B3	98.74%	98.80%	98.69%	98.74%
EfficientNet-B4	98.80%	98.90%	98.72%	98.81%



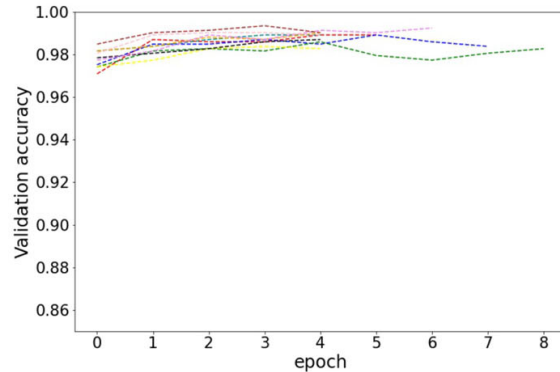
(a) VGG16



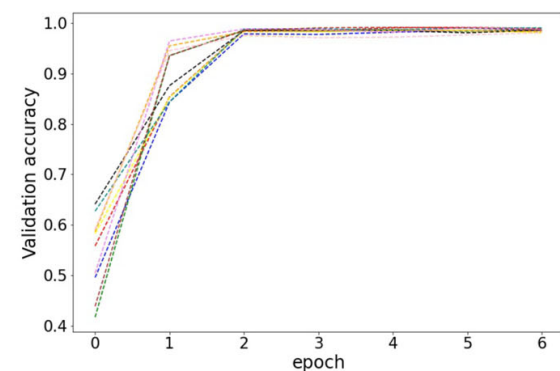
(b) ResNet50



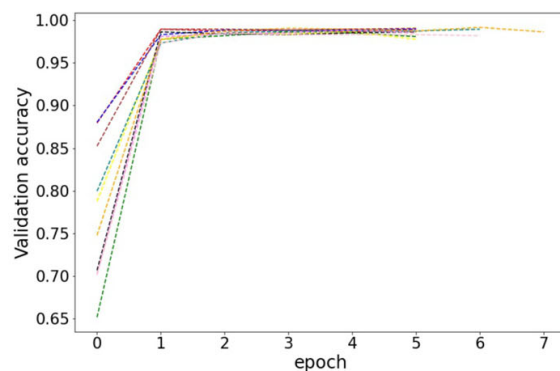
(c) inception



(d) Xception



(e) EfficientNet-B3



(f) EfficientNet-B4

**Fig. 7** Validation accuracy vs Epochs for the 10 folds of the models

**Table 6** Performance comparison of different ensemble models

Ensemble	Accuracy	Precision	Recall	F1-score
V + R + X	99.11%	99.38%	98.38%	98.87%
V + R + E4	99.19%	99.33%	97.70%	98.48%
V + X + E4	98.99%	99.01%	98.91%	98.95%
<b>R + X + E4</b>	<b>99.43%</b>	<b>99.44%</b>	<b>99.43%</b>	<b>99.43%</b>
V + X + R + E4	99.26%	99.42%	98.69%	99.05%

experimentation with six different deep convolutional networks viz., VGG16, ResNet50, InceptionV3, Xception, EfficientNet-B3, and EfficientNet-B4 is performed to ensure that the ensemble model is built using the appropriate combination of sub-models. These particular models are considered from their respective families to ensure an appropriate trade-off between the model's performance and its inference time. Further, these models are well-understood and well-established in the field of transfer learning. Table 5 encapsulates the results for these models. As can be seen from the table, on average, Xception, ResNet50, and EfficientNet-B4 give the best results (10-fold cross-validation scores). Figure 7 shows the graph for validation accuracy vs. epochs for the 10-folds of models. The results of the proposed method are shown in bold.

Further experimentations reveal that even though the individual models achieve great results, the models individually are not capable of state-of-the-art performance. Thus, The different combinations of ensemble architectures are formulated for evaluating the performance of the possible sub-models. Since InceptionV3 and Xception are quite similar and the former's results are considerably lower than the latter model, InceptionV3 is excluded from the choice of sub-model. Similarly, EfficientNet-B3 and EfficientNet-B4 essentially belong to the same family. Thus, based on their stand-alone performance, the model selection for the ensemble is restricted to only EfficientNet-B4. Different possible combinations of ensemble architectures from VGG16, ResNet50, Xception, and EfficientNet-B4 are evaluated. These models employ very different learning techniques. Consequently, the models learn different representations of malware images, which help the final ensemble model make sound decisions. Table 6 summarizes the results of different combinations of ensemble architectures. This paper uses V, R, X, and E4 as placeholders for VGG16, ResNet50, Xception, and EfficientNet-B4, respectively. As shown in the table, the combination of ResNet50, Xception, and EfficientNet-B4 performs significantly better than the other combinations. The model configuration used in the proposed method is highlighted in bold

### 4.3.2 Ensemble technique

Using ResNet50, Xception, and EfficientNet-B4 as sub-models, The experimentation is performed with three ensemble techniques - hard-voting ensemble, soft-voting ensemble, and stacking generalization ensemble. Table 7 encapsulates the results obtained with each of these techniques. The results of the proposed method are indicated in bold.

In a hard-voting ensemble, predictions are made using each sub-model. For each data point, each prediction is considered as a vote for the predicted class, and the class that received the maximum sum of votes is reported as the final prediction. In a soft-voting ensemble, each sub-model gives the class probability. These probabilities are added, and the class that received the maximum sum of probabilities is reported as the final prediction. A drawback of this technique is that each model contributes equally to the prediction

**Table 7** Comparative study of different ensemble techniques

Ensemble technique	Accuracy	Precision	Recall	F1-score
Hard-voting	98.92%	97.39%	97.13%	97.26%
Soft-voting	99.03%	97.53%	97.38%	97.45%
<b>Stacked</b>	<b>99.43%</b>	<b>99.44%</b>	<b>99.43%</b>	<b>99.43%</b>

irrespective of its performance. Thus, the results obtained with this technique are a little sub-optimal.

Further, a stacked generalization ensemble is used for the experiments. In this technique, a learning algorithm is used to combine the predictions of the sub-models. In particular, a multilayered perceptron (MLP) is used as a meta classifier that took the concatenated output of the sub-models as input and learned the best possible way to combine the input predictions to form the output predictions. One may notice from Table 7 that the stacked generalization ensemble yields significantly better results as compared to the hard-voting ensemble and soft-voting ensemble.

The major findings of the experimentation are as follows:

1. ResNet50, Xception, and EfficientNet-B4 as sub-models performs better feature extraction than the other deep-learning based models.
2. The multilayered perceptron based classifier shows that the light classifiers can perform the classification effectively with a robust feature extractor.
3. The performance of the sub-models without ensemble is limited. Further, the ensemble of sub-models helps in achieving the best performance.

#### 4.4 Computational complexity

The analysis of computational cost for an algorithm is important. This section presents the average running time analysis to show the relative computational time required per algorithm. The analysis is performed on a system with Tesla T4 GPU. The malware file from the Malimg and VirusShare datasets are used to analyse the average running time of the proposed algorithms. The parameters in the code of various algorithms are used as given in

**Table 8** Average running time per sample (in seconds)

Model	Run time
Nataraj et al. [30]	0.118
Liu et al. [23]	0.097
Kalash et al. [15]	0.091
Naeem et al. [28]	0.081
Cui et al. [8]	0.078
Gibert et al. [11]	0.102
Naeem et al. [26]	0.195
Nisa et al. [32]	0.099
Yuan et al. [44]	0.103
Verma et al. [43]	0.071
<b>S-DCNN</b>	<b>0.062</b>



the papers. Table 8 shows the average computation time to process a sample malware for the algorithms. Most of the algorithms uses either CNN-based deep classifiers or complex machine learning classifiers which requires comparably larger time. The proposed S-DCNN uses a light Multilayered perceptron based classifier which improves the speed of the proposed method. It can be noticed from Table 8 that the proposed S-DCNN outperforms the other contemporary algorithms. The results of the proposed method are shown in bold.

## 5 Conclusion

This paper proposes a stacked deep convolutional model to address malware classification problem. The proposed method uses the concept of transfer learning and ensemble learning. It formulates a robust and efficient architecture for malware classification by ensemble of the sub-models. The ensemble model extracts the features from the malware file to perform the multilayered perceptron based classification. The effective ensemble enables a simple classifier to classify the samples effectively. It is demonstrated that the proposed model is able to tackle the challenges posed by the unbalanced nature of the real-world datasets and is resilient to standard obfuscation techniques. Further, it is established that the proposed model is effective with the help of a comprehensive evaluation using suitable statistical metrics, extensive result comparison with existing models, and diverse ablation experiments.

## Declarations

**Conflict of Interests** 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.

## References

1. Alsulami B, Mancoridis S (2018) Behavioral malware classification using convolutional recurrent neural networks. In: 2018 13th international conference on malicious and unwanted software (MALWARE), pp 103–111. <https://doi.org/10.1109/MALWARE.2018.8659358>
2. Beek C, Dunton T, Fokker J, Grobman S, Hux T, Polzer T, Lopez MR, Roccia T, Saavedra-Morales J, Samani R, Sherstobitof R (2019) Accessed on December 27, 2020 McAfee labs threats report. <https://www.mcafee.com/enterprise/en-us/assets/reports/rp-quarterly-threats-aug-2019.pdf>
3. Bhowmik A, Kumar S, Bhat N (2019) Eye disease prediction from optical coherence tomography images with transfer learning. In: International conference on engineering applications of neural networks. Springer, pp 104–114
4. Bhowmik A, Kumar S, Bhat N (2021) Evolution of automatic visual description techniques-a methodological survey. *Multimed Tools Appl*, 1–45
5. Çayır A, Ünal U, Dağ H (2021) Random capsnet forest model for imbalanced malware type classification task. *Comput Secur* 102:102133. <https://doi.org/10.1016/j.cose.2020.102133>
6. Chaudhary P, Gupta DK, Singh S (2021) Outcome prediction of patients for different stages of sepsis using machine learning models. In: *Advances in communication and computational technology*. Springer, Singapore, pp 1085–1098
7. Chollet F (2017) Xception: deep learning with depthwise separable convolutions. In: 2017 IEEE Conference on computer vision and pattern recognition (CVPR), pp 1800–1807. <https://doi.org/10.1109/CVPR.2017.195>
8. Cui Z, Xue F, Cai X, Cao Y, ge Wang G, Chen J (2018) Detection of malicious code variants based on deep learning. *IEEE Trans Industr Inform* 14(7):3187–3196. <https://doi.org/10.1109/tii.2018.2822680>
9. D'Angelo G, Ficco M, Palmieri F (2021) Association rule-based malware classification using common subsequences of api calls. *Appl Soft Comput* 105:107234. <https://doi.org/10.1016/j.asoc.2021.107234>

10. Gao X, Hu C, Shan C, Liu B, Niu Z, Xie H (2020) Malware classification for the cloud via semi-supervised transfer learning. *J Inform Secur Applic* 55:102661. <https://doi.org/10.1016/j.jisa.2020.102661>
11. Gibert D, Mateu C, Planes J, Vicens R (2018) Using convolutional neural networks for classification of malware represented as images. *J Comput Virol Hack Techniques* 15(1):15–28. <https://doi.org/10.1007/s11416-018-0323-0>
12. Gibert D, Mateu C, Planes J (2020) Hydra: a multimodal deep learning framework for malware classification. *Comput Secur* 95:101873. <https://doi.org/10.1016/j.cose.2020.101873>
13. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: 2016 IEEE Conference on computer vision and pattern recognition (CVPR), pp 770–778. <https://doi.org/10.1109/CVPR.2016.90>
14. Jain M, Andreopoulos W, Stamp M (2020) Convolutional neural networks and extreme learning machines for malware classification. *J Comput Virol Hack Techniques* 16(3):229–244
15. Kalash M, Rochan M, Mohammed N, Bruce NDB, Wang Y, Iqbal F (2018) Malware classification with deep convolutional neural networks. In: 2018 9th IFIP international conference on new technologies, mobility and security (NTMS), pp 1–5. <https://doi.org/10.1109/NTMS.2018.8328749>
16. Kaspersky (2020) Accessed on December 27, 2020 Protecting your personal data online at every point. <https://media.kasperskydaily.com/wp-content/uploads/sites/92/2020/01/27103216/International-Privacy-Day-2020-Kaspersky-report.pdf>
17. Katyal S, Kumar S, Sakhuja R, Gupta S (2018) Object detection in foggy conditions by fusion of saliency map and yolo. In: In 2018 12th international conference on sensing technology (ICST), pp 154–159. <https://doi.org/10.1109/ICSensT.2018.8603632>
18. Kaur G, Singh S, Rani R, Kumar R, Malik A (2021) High-quality reversible data hiding scheme using sorting and enhanced pairwise pcc. *IET Image Processing*. <https://doi.org/10.1049/ipr2.12212>
19. Kolosnjaji B, Zarras A, Webster G, Eckert C (2016) Deep learning for classification of malware system call sequences. In: *AI 2016: advances in artificial intelligence*. Springer International Publishing, pp 137–149. [https://doi.org/10.1007/978-3-319-50127-7\\_11](https://doi.org/10.1007/978-3-319-50127-7_11)
20. Kumar M, Gupta DK, Singh S (2021) Extreme event forecasting using machine learning models. In: *Advances in communication and computational technology*. Springer, Singapore, pp 1503–1514
21. Kumar N, Kumar R, Malik A, Singh S (2021) Low bandwidth data hiding for multimedia systems based on bit redundancy. *Multimed Tools Appl*, 1–19
22. Kumar R, Chand S, Singh S (2019) An optimal high capacity reversible data hiding scheme using move to front coding for lzw codes. *Multimed Tools Appl* 78(16):22977–23001
23. Liu L, Wang B, Yu B, Zhong Q (2017) Automatic malware classification and new malware detection using machine learning. *Front Inform Technol Electron Eng* 18:1336–1347
24. Malik A, Kumar R, Singh S (2018) A new image steganography technique based on pixel intensity and similarity in secret message. In: *International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*. IEEE, pp 828–831
25. Moskovitch R, Feher C, Tzachar N, Berger E, Gitelman M, Dolev S, Elovici Y (2008) Unknown malcode detection using OPCODE representation. In: *Intelligence and security informatics*. Springer, Berlin, pp 204–215. [https://doi.org/10.1007/978-3-540-89900-6\\_21](https://doi.org/10.1007/978-3-540-89900-6_21)
26. Naeem H (2019) Detection of malicious activities in internet of things environment based on binary visualization and machine intelligence. *Wirel Pers Commun* 108(4):2609–2629. <https://doi.org/10.1007/s11277-019-06540-6>
27. Naeem H, Guo B, Naeem MR (2018) A light-weight malware static visual analysis for iot infrastructure. In: 2018 International conference on artificial intelligence and big data (ICAIBD), pp 240–244. <https://doi.org/10.1109/ICAIBD.2018.8396202>
28. Naeem H, Guo B, Naeem MR, Ullah F, Aldabbas H, Javed MS (2019) Identification of malicious code variants based on image visualization. *Comput Electr Eng* 76:225–237. <https://doi.org/10.1016/j.compeleceng.2019.03.015>
29. Narayanan BN, Djaneye-Boundjou O, Kebede TM (2016) Performance analysis of machine learning and pattern recognition algorithms for malware classification. In: 2016 IEEE national aerospace and electronics conference (NAECON) and ohio innovation summit (OIS), pp 338–342. <https://doi.org/10.1109/NAECON.2016.7856826>
30. Nataraj L, Karthikeyan S, Jacob G, Manjunath BS (2011) Malware images: visualization and automatic classification. In: the 8th International symposium on visualization for cyber security. Association for Computing Machinery, New York, NY, USA, VizSec '11. <https://doi.org/10.1145/2016904.2016908>
31. Nataraj L, Yegneswaran V, Porras P, Zhang J (2011) A comparative assessment of malware classification using binary texture analysis and dynamic analysis. In: *Proceedings of the 4th ACM workshop on*

- security and artificial intelligence. Association for Computing Machinery, New York, NY, USA, AISec '11, pp 21–30. <https://doi.org/10.1145/2046684.2046689>
32. Nisa M, Shah JH, Kanwal S, Raza M, Khan MA, Damaševičius R, Blažauskas T (2020) Hybrid malware classification method using segmentation-based fractal texture analysis and deep convolution neural network features. *Appl Sci* 10(14):4966. <https://doi.org/10.3390/app10144966>
  33. Oliva A, Torralba A (2001) Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int J Comput Vis* 42(3):145–175
  34. Pascanu R, Stokes JW, Sanossian H, Marinescu M, Thomas A (2015) Malware classification with recurrent networks. In: IEEE International conference on acoustics, speech and signal processing (ICASSP), pp 1916–1920. <https://doi.org/10.1109/ICASSP.2015.7178304>
  35. Saadat S, Raymond VJ (2020) Malware classification using CNN-XGBoost model. In: Artificial intelligence techniques for advanced computing applications. Springer, Singapore, pp 191–202. [https://doi.org/10.1007/978-981-15-5329-5\\_19](https://doi.org/10.1007/978-981-15-5329-5_19)
  36. Saxe J, Berlin K (2015) Deep neural network based malware detection using two dimensional binary program features. In: 2015 10th international conference on malicious and unwanted software (MALWARE), pp 11–20. <https://doi.org/10.1109/MALWARE.2015.7413680>
  37. Saxena A, Gupta DK, Singh S (2021) An animal detection and collision avoidance system using deep learning. In: Advances in communication and computational technology. Springer, Singapore, pp 1069–1084
  38. Schultz MG, Eskin E, Zadok F, Stolfo SJ (2001) Data mining methods for detection of new malicious executables. In: Proceedings 200 IEEE symposium on security and privacy S P 2001, pp 38–49. <https://doi.org/10.1109/SECPRI.2001.924286>
  39. Tan M, Le Q (2019) EfficientNet: rethinking model scaling for convolutional neural networks. In: Proceedings of the 36th international conference on machine learning, PMLR, proceedings of machine learning research, vol 97, pp 6105–6114
  40. Vasani D, Alazab M, Wassan S, Naeem H, Safaei B, Zheng Q (2020) Imcfn: image-based malware classification using fine-tuned convolutional neural network architecture. *Comput Netw* 171:107138. <https://doi.org/10.1016/j.comnet.2020.107138>
  41. Vasani D, Alazab M, Wassan S, Safaei B, Zheng Q (2020) Image-based malware classification using ensemble of cnn architectures (imcec). *Comput Secur* 92:101748. <https://doi.org/10.1016/j.cose.2020.101748>
  42. Venkatraman S, Alazab M, Vinayakumar R (2019) A hybrid deep learning image-based analysis for effective malware detection. *J Inform Secur Applic* 47:377–389. <https://doi.org/10.1016/j.jisa.2019.06.006>
  43. Verma V, Mutttoo SK, Singh VB (2020) Multiclass malware classification via first- and second-order texture statistics. *Comput Secur* 97:101895. <https://doi.org/10.1016/j.cose.2020.101895>
  44. Yuan B, Wang J, Liu D, Guo W, Wu P, Bao X (2020) Byte-level malware classification based on markov images and deep learning. *Comput Secur* 92:101740. <https://doi.org/10.1016/j.cose.2020.101740>
  45. Zhang H, Xiao X, Mercaldo F, Ni S, Martinelli F, Sangaiah AK (2019) Classification of ransomware families with machine learning based on n-gram of opcodes. *Futur Gener Comput Syst* 90:211–221. <https://doi.org/10.1016/j.future.2018.07.052>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

# Temperature Analysis of a Novel 5nm Stacked Oxide Top Bottom Gated Junctionless FinFET for improved switching and efficiency

Lakshya Gangwani  
Department of Applied Physics  
Delhi Technological University  
Delhi, India  
lakshyagangwani1@gmail.com

Sumanyu Hajela  
Department of Applied Physics  
Delhi Technological University  
Delhi, India  
sumanyu4@gmail.com

**Abstract**— Temperature Analysis of a Novel Stacked Oxide Top Bottom Gated Junctionless Fin-shaped Field Effect Transistor (TBG JL FinFET) is described in this research paper. This examination reveals the alteration of performance parameters with temperature for the Novel 5nm Stacked Oxide TBG JL FinFET. The Temperature Analysis is done on Silvaco TCAD tools for the Novel Stacked Oxide Top-Bottom Gated Junctionless FinFET and compared with the Conventional Junctionless FinFET structure. The device shows enhanced output performance at 200K while having less electric field near drain end concurrently, thus reducing the Short Channel Effects (SCEs) by a greater extent. The results from device simulation reveal that the Novel Stacked Oxide Top-Bottom Gated Junctionless FinFET has better analog performance over varied temperature range and lowers the Short Channel Effects (SCEs) a cut above the Conventional Junctionless FinFET.

**Keywords**— FinFET, Top-Bottom Gated, Junctionless, Stacked Oxide, Temperature, Switching Ratio

## I. INTRODUCTION

As we progress further there are new advancements and contraptions springing up consistently. These advancements are intended to serve the majority and make ordinary assignments more helpful and speedy. In the midst of such improvements it turns out to be considerably more basic to chip away at the productivity and asset utilization of our devices [1]. Numerous advancements are being created to do exactly that. At the point when we talk about the ambit of correspondence innovation and all the more explicitly strong state and semiconductor industry we notice a pattern [2]. This pattern beginning right from MOSFET gadgets till SOI and FinFET innovation can be seen moving towards better execution with respect than speed and force utilization [3]. This is being accomplished by downsizing the gadgets and making them more compact and advantageous. Yet, this includes some significant downfalls. As we downsize specific unfortunate ascribes become more predominant. These incorporate short channel impacts like Drain Induced Barrier Lowering, Velocity Saturation and Subthreshold Swing. These can create setbacks and inefficacy [4]. Accordingly to diminish these impacts there is a requirement for more advancement past the 20nm scale. The new door designs for FinFET innovation is a positive development. These progressions are being embraced by

different monsters in the Semiconductor and Chip fabricating industry [5]. Therefore in this paper, the performance of the Novel Device has been monitored over different temperatures [6].

## II. DEVICE STRUCTURE

A 3D structure of the Novel Stacked Oxide Top-Bottom Gated Junctionless FinFET structure is shown in the Fig. 1. Stacking of the oxide layer has been done wherein a layer of 7nm Hafnium Oxide is implanted on top of 3nm Silicon Dioxide. This stacking of oxide layers is performed for achieving better output efficiency and switching performance.

Gates are present both above and below the substrate in the case of Novel Stacked Oxide TBG JL FinFET [7]. This is done to further develop the gate controllability of the device in comparison to Conventional Junctionless FinFET.  $\text{HfO}_2$  has a dielectric constant in the range of 20 to 25 whereas  $\text{SiO}_2$  has dielectric constant of 3.7 [8]. For the simulation purposes in TCAD, the dielectric constant of Hafnium Oxide ( $\text{HfO}_2$ ) is taken as 25 [9]. The oxide layer stacking and top and bottom fins prompts improvement in the analog exhibition and Ion/Ioff ratio of the novel structure when contrasted with the Conventional JL FinFET device [10].

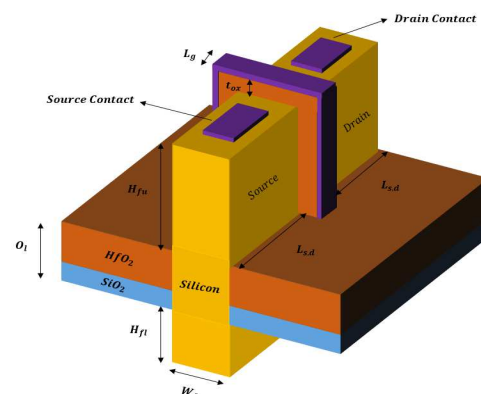


Fig. 1 - Systematic 3D structure of the Novel Stacked Oxide Top Bottom Gated Junctionless FinFET Structure

The device that was simulated is of n-type and has constant doping concentration of as  $5.0 \times 10^{18} \text{ (cm}^{-3}\text{)}$  for source, drain and channel.

Table I. STRUCTURAL PARAMETERS SPECIFICATIONS

Parameters	Dimensions of Simulated Devices	
	Novel Stacked Oxide TBG JL FinFET	Conventional JL FinFET
Gate Length ( $L_g$ )	5nm	5nm
Upper Fin Height ( $H_{fu}$ )	12nm	12nm
Lower Fin Height ( $H_{fl}$ )	5nm	0nm
Fin Width ( $W_s$ )	5nm	5nm
Oxide Layer Thickness ( $O_t$ )	7nm $HfO_2$ + 3nm $SiO_2$	10nm $SiO_2$
Substrate Width ( $W_s$ )	5nm	5nm
Source, Drain Length ( $L_{s,d}$ )	20nm	20nm

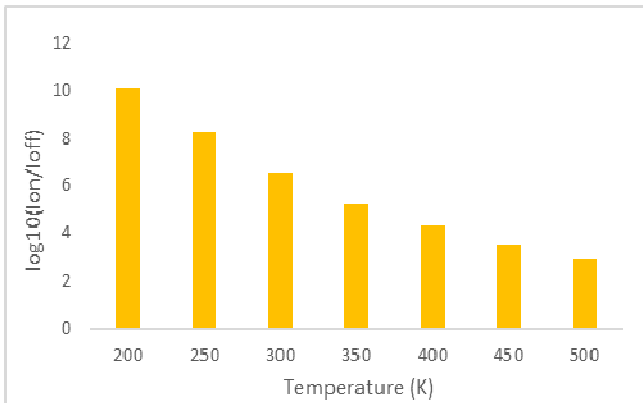
Table I gives the device dimensions for the Novel Stacked Oxide Top Bottom Gated Junctionless FinFET device and Conventional Junctionless FinFET device.

### III. SIMULATION METHODOLOGY

Silvaco Atlas TCAD Tool has been utilized for re-enacting the gadgets [11]. The improvement of device parameters on Silvaco Atlas follows specific methodology. Most importantly the Detail of construction is done for example network lattice definition by fixing the dispersing and the framework focuses for the device, doping fixation definition and explaining the locales. Besides Material Definition is done and actual models for the device are chosen [11]. This is trailed by Mathematical Strategy Determination which is utilized to decide the Numerical Models for the device simulation. At last the Determination of Arrangement and Result Investigation is accomplished for deciphering the outcomes and extraction of yield boundaries [12].

Quantum effects become an integral factor when the device is downsized past specific measurements. Around 5nm measurements, there is an expanded inconstancy and potential for opens and short-circuits in the device because of quantum effects [3]. The device temperature is varied from 200K to 500K for the Novel Stacked Oxide TBG JL FinFET device.

### IV. RESULTS AND DISCUSSIONS

Fig. 2 -  $\log_{10}(I_{on}/I_{off})$  variation with Temperature for Novel Stacked Oxide TBG JL FinFET

**Fig. 2** depicts the variation of  $\log_{10}(I_{on}/I_{off})$  with Temperature ranging from 200K to 500K for the Novel Stacked Oxide TBG JL FinFET. It plainly portrays that the  $I_{on}/I_{off}$  ratio is best at 200K which is in the order of  $10^{10}$  and thus the device shows least losses at 200K [13].

Scaling down of devices results in a large increase in  $I_{off}$  values. This can be attributed to leakage current and Short Channel Effects (SCEs) [14]. We have to maximize the switching ratio ( $I_{on}/I_{off}$ ) to reduce losses [15]. On increasing the temperature, the scattering of charge carrier magnifies, thus leading to decreased  $I_{on}$ . Moreover, due to increase in temperature, the drain current increases in the sub threshold region which in turn decreases the threshold voltage ( $V_{th}$ ) and thus leads to performance degradation as far as  $I_{on}$ ,  $I_{off}$  and switching ratio are concerned.

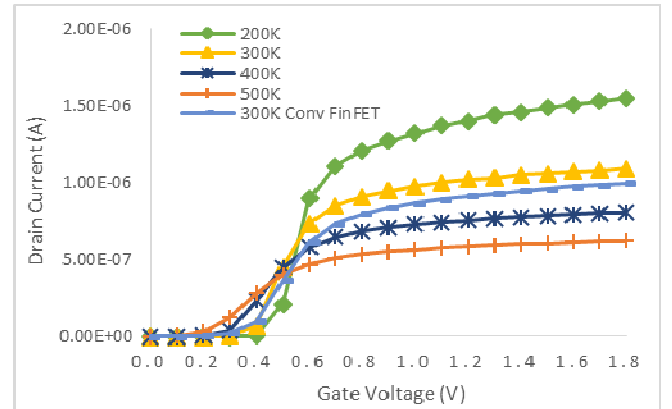


Fig. 3 - Novel Stacked Oxide TBG JL FinFET Transfer Characteristics at various temperatures

**Fig. 3** depicts the transfer characteristics of the Novel Stacked Oxide TBG JL FinFET at different temperatures from 200K to 500K. We notice that at 300K the Novel Stacked Oxide structure has increased output current when contrasted with the Conventional JL FinFET structure for similar gate voltage values.

Thus we have improved transfer characteristics [16]. Further we observe that as the temperature decreases from 500K to 200K the output current increases almost 40 times. Thus indicating enhanced performance at 200K.

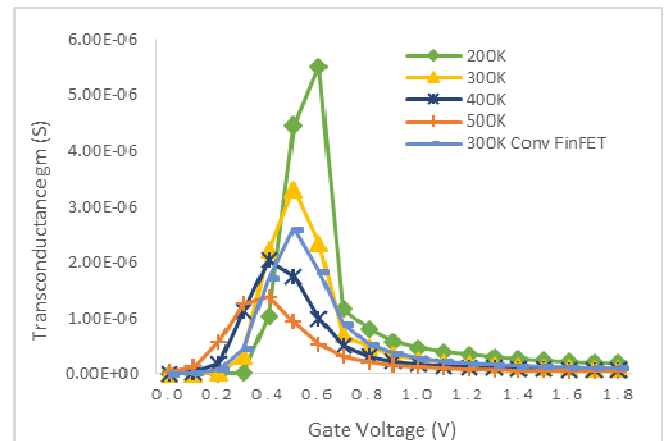


Fig. 4 - Transconductance for Novel Stacked Oxide TBG JL FinFET at various temperatures



So we basically obtain  $g_m$  by taking the slope of  $I_{ds}$ - $V_{gs}$  curve [17]. **Fig. 4** gives the transconductance ( $g_m$ ) of Novel Stacked Oxide TBG JL FinFET at different temperatures from 200K to 500K. Furthermore, the equivalent is given for the Conventional JL FinFET device structure at 300K for comparison. We observe for the same input gate voltage values, the Novel Stacked Oxide structure gives elevated values of  $g_m$  when contrasted with the Conventional FinFET structure and therefore shows improved behavior [14]. And further at 200K we observe a sharp increase in transconductance at 0.6V.

The transconductance declines with rise in temperature owing to the reduction in carrier mobility due to scattering [16]. Thus the proposed device has superior characteristics with respect to transconductance [18]. It also shows improved control of gate voltage over drain current at 200K.

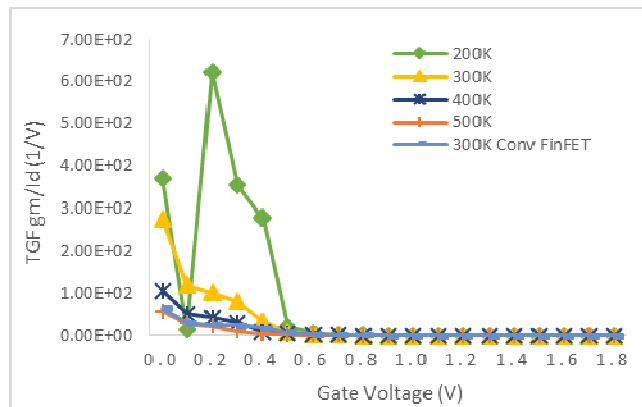


Fig. 5 - Transconductance Generation Factor variation with temperature for Novel Stacked Oxide TBG JL FinFET

**Fig. 5** shows the progressions in TGF as for changes in  $V_{gs}$  at various temperatures from 200K to 500K for the Novel structure and the Conventional JL FinFET structure for 300K [13]. The maximum TGF appears around  $V_{gs}$  of 0.0 volt but at 200K we observe a spike at 0.25 volts thus showing highly efficient behaviour [12].

On increasing the values of Temperature for lower values of Gate voltages, the Drain current keeps on escalating and thus leading to reduction in TGF as it is inversely proportional to drain current. It very well may be seen unmistakably that the Novel device has elevated values of TGF when contrasted with the Conventional FinFET structure at 300K as shown and thus has more efficiency and performance [19].

**Fig. 6** portrays the Novel device output characteristics for temperature ranging from 200K to 500K. Essentially this diagram shows the variety in drain current (output current) with the adjustment of drain voltage [13]. So to further develop the gate controllability of the device, the output current ought to fluctuate with changes in gate voltage (input voltage) rather than the drain voltage. According to the graph, at 500K the device shows least dependence of Drain Voltage on the drain current as the current variation is lowest of all. Thus the Novel Stacked Oxide TBG JL FinFET shows improved output characteristics for the temperature range of 400K to 500K [20].

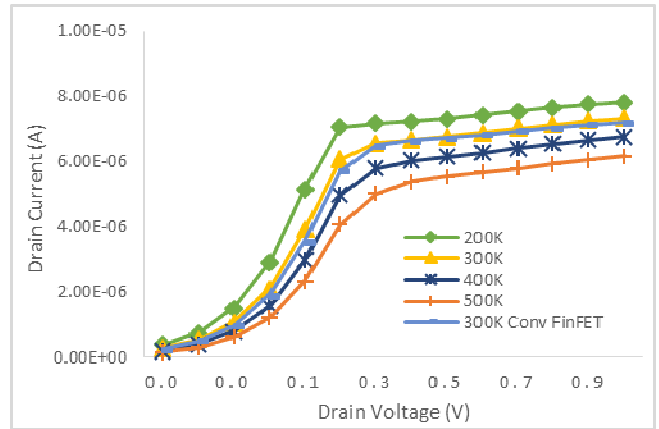


Fig. 6 - Output Characteristics for Novel Device at different temperatures

Output Conductance ( $g_d$ ) is a performance parameter that basically connotes the progressions in output current with the adjustment of drain predisposition [21]. Equation (1) gives the formula for Output Conductance.

$$g_d = \frac{\partial I_d}{\partial V_d} \quad (1)$$

To accomplish more power over Drain Induced Barrier Lowering (DIBL) and channel length modulation, lower upsides of  $g_d$  are favored [22]. Along these lines lower output conductance helps in lessening the short channel impacts (SCEs) in the designed Novel structure [23].

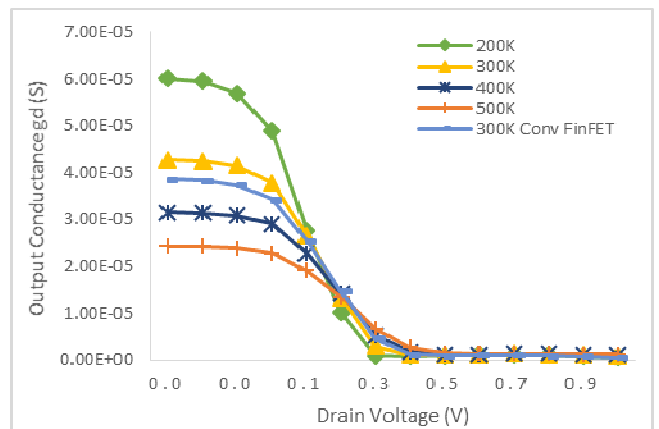


Fig. 7 - Output Conductance variation with Temperature for Novel device

**Fig. 7** portrays the variance of output conductance ( $g_d$ ) for the Novel device structure over temperature range of 200K to 500K. The design shows less reliance of output current  $I_d$  on drain bias for higher temperatures (400K-500K) [15]. Change in output current ought to be just about as low as conceivable because of drain bias. Likewise device properties like Intrinsic Gain and Early Voltage ( $V_A$ ) are improved with lower upsides of  $g_d$  as given by [24]:



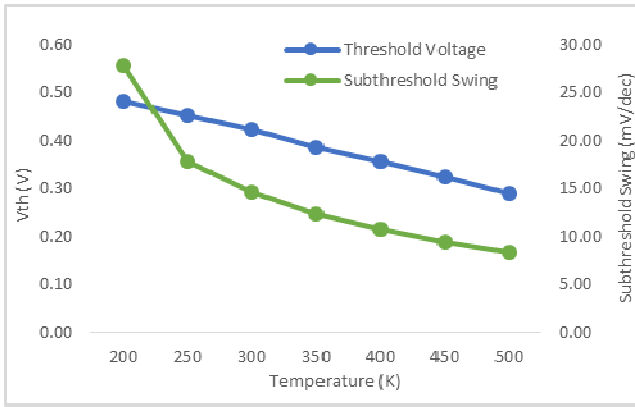


Fig. 8 - Subthreshold Swing and Threshold Voltage variation for Novel Stacked Oxide TBG JL FinFET with temperature

**Fig. 8** portrays the variance of Subthreshold Swing and Threshold Voltage with Temperature for the Novel device. We can obviously construe that with expansion in temperature, both the threshold voltage ( $V_{th}$ ) and Subthreshold Swing decreases [25]. The Subthreshold Swing shows minimum value of 8.25mV/decade at 500K, thus implying better channel control and improved  $I_{on}/I_{off}$  ratio when contrasted with the Conventional FinFET structure [17]. Additionally limiting the Subthreshold Swing helps in lessening the leakage current consequently prompting more energy proficient devices. The Novel device has Subthreshold Swing esteems under 60mV/decade for every one of the temperatures going from 200K to 500K. The diagram likewise shows that Threshold Voltage diminishes with expansion in temperature. The Threshold Voltage ( $V_{th}$ ) is lowest at 0.29V at temperature of 500K. Lower Threshold Voltage suggests less input voltage is needed to turn ON the device and less power utilization [13].

The **Fig. 9** depicts the Electric field contours for the temperature ranging from 200K to 500K for the Novel device [26]. The device shows lower upsides of electric field close to the channel end even at higher temperatures as seen in figure 9(c) and 9(d), thus reducing the carrier scattering and thereby improving the device performance [13]. Contours exhibits the enhancement of Novel Stacked Oxide TBG JL FinFET in terms of *hot carrier effects*. High electric field close to the channel end can prompt carrier warming, thus leading to increased energies and device degradation [5].

The Figure 9(a) shows the device at 200K and it has lowest Electric Field and improved switching ratio of the order of  $10^{10}$ , thus making it an ideal choice for analog performance devices [14]. Further on increasing the temperature, the field in the device increases close to the channel end as seen in figure 9(b) and figure 9(c). But the electric field close to the channel end for higher temperatures is still less than that for Conventional FinFET structure.

Velocity saturation is caused because of high electric field close to the drain end [26]. Carriers reach their maximum velocities due to high electric field values thereby prompting degradation of transconductance and delays in switching times [27].

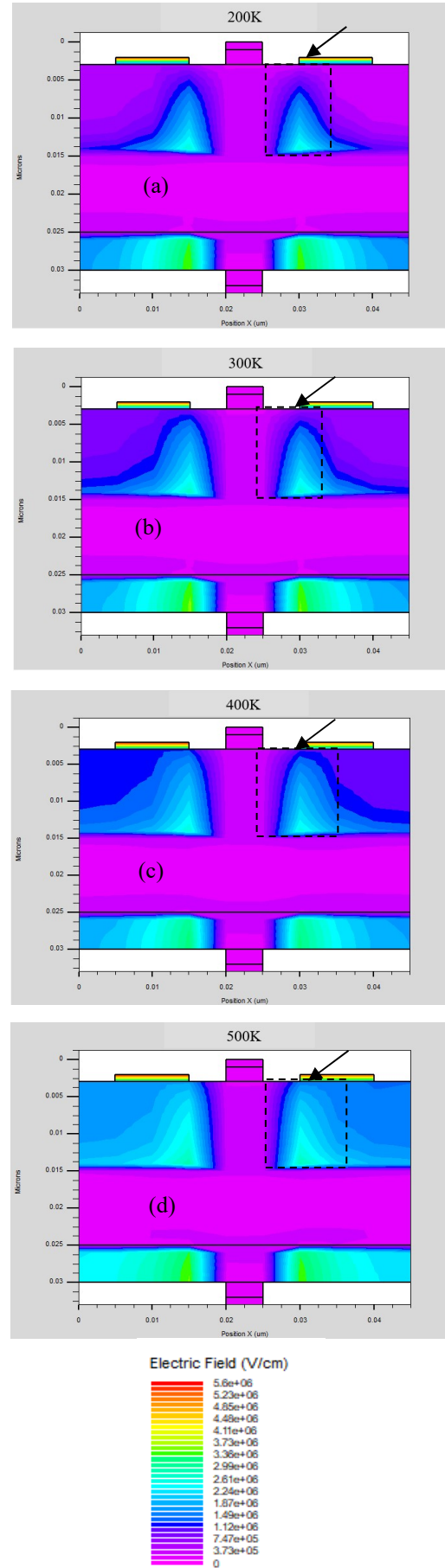


Fig. 9 - Novel device Electric Field Contours for 200K, 300K, 400K and 500K

Thus the Novel Stacked Oxide TBG JL FinFET device brings down the Short Channel Effects because of lower electric field close to the drain end as portrayed in figure 9.

## V. CONCLUSION

In this research investigation, we have carried out the temperature analysis for the Novel Stacked Oxide Top-Bottom Gated Junctionless FinFET with gate length 5nm. The novel device shows multiple times improvement in switching ratio when contrasted with the Conventional FinFET structure at 300K. Switching Ratio of the order of  $10^{10}$  is observed at 200K ( $10^5$  times improvement over Conventional FinFET structure) for the Novel device [28]. Further we observe that as the temperature decreases from 500K to 200K, the output current increases almost 40 times. It has also shown considerably enhanced output at 200K and reduced electric field close to the drain end at this temperature along these lines decreasing the hot carrier impacts. The device also showed improved efficiency with spikes in the values of Transconductance and Transconductance Generation Factor at 200K [6, 7]. Thus we have good device function for temperature range of 200K to 300K for the Novel Stacked Oxide TBG JL FinFET structure. We have also observed lowered threshold voltage (39.5% reduction from 200K) at 500K and thus can be operated at lower voltages even at higher temperatures [29].

## ACKNOWLEDGMENT

The authors are thankful to Dr. Rishu Chaujar and the Department of Applied Physics, Delhi Technological University, for aiding the research work.

## REFERENCES

- [1] G.E. Moore (1998) Cramming more components onto integrated circuits. *Proc IEEE* 86(1):82–85. <https://doi.org/10.1109/JPROC.1998.658762>
- [2] TSMC. Leading Edge Technology. <http://www.tsmc.com/english/dedicate-Foundry/technology/16nm.htm>. Accessed 4th February 2021
- [3] Z. Ramezani, M. Anvarifard, A.I. Sadegh, N. Mahdavi, Alireza. A Nanoscale - Modified band energy junctionless transistor with considerable progress on the electrical and frequency issue. *Materials Science in Semiconductor Processing* 2020. 107. 104849. [10.1016/j.mssp.2019.104849](https://doi.org/10.1016/j.mssp.2019.104849).
- [4] M. Anvarifard, Creation of a new high voltage device with capable of enhancing driving current and breakdown voltage. *Materials Science in Semiconductor Processing* 2017. 60. 60-65. [10.1016/j.mssp.2016.12.030](https://doi.org/10.1016/j.mssp.2016.12.030).
- [5] Semiconductor Engineering. IBM, Intel and TSMC Roll out FinFETs. <http://semiengineering.com/ibm-intel-and-tsmc-roll-outfinfets>. Accessed 4th February 2021
- [6] L. Gangwani et al 2021 *J. Phys.: Conf. Ser.* 1921 012100 <https://doi.org/10.1088/1742-6596/1921/1/012100>
- [7] L. Gangwani, R. Chakravarti and R. Chaujar, "Lower Fin Modulation Analysis for a Novel 5nm Top Bottom Gated Junctionless FinFET for improved performance," 2021 *Devices for Integrated Circuit (DevIC)*, 2021, pp. 1-5, doi: 10.1109/DevIC50843.2021.9455834.
- [8] M. Kailasam, M. Govindasamy. Impact of High-K Gate Dielectrics on Short Channel Effects of DG N-FinFET 2020.
- [9] P.W. Peacock, K. Xiong, K. Tse, J. Robertson, Bonding and interface states of Si: HfO<sub>2</sub> and Si: ZrO<sub>2</sub> interfaces. *Phys. Rev. B* 2006, 73, 075328.
- [10] B. Kumar, R. Chaujar Analog and RF Performance Evaluation of Junctionless Accumulation Mode (JAM) Gate Stack Gate All Around (GS-GAA) FinFET. *Silicon* 13, 919–927 (2021). <https://doi.org/10.1007/s12633-020-00910-7>
- [11] I. SILVACO, "ATLAS User's Manual," Santa Clara, CA, Ver. vol. 5, 2011.
- [12] M. P. Kashyap and R. Chaujar, "Temperature investigation of a Novel 3nm TF-Bulk FinFET for Improved Performance," 2020 *IEEE 20th International Conference on Nanotechnology (IEEE-NANO)*, Montreal, QC, Canada, 2020, pp. 382-387, doi: 10.1109/NANO47656.2020.9183594.
- [13] B. Kumar, R. Chaujar, TCAD Temperature Analysis of Gate Stack Gate All Around (GS-GAA) FinFET for Improved RF and Wireless Performance. *Silicon* 13, 3741–3753 (2021). <https://doi.org/10.1007/s12633-021-01040-4>
- [14] D.Y. Jeon, S.J. Park, M. Mouis, S. Barraud, G.T. Kim, G. Ghibaudo Low-temperature electrical characterization of junctionless transistors. *Solid State Electron* 2013 80:135–141. <https://doi.org/10.1016/j.sse.2012.10.018>
- [15] K.R. Barman, S. Baishya Study of temperature effect on analog/RF and linearity performance of dual material gate (DMG) vertical super-thin body (VSTB) FET. *Silicon* 2020. <https://doi.org/10.1007/s12633-020-00561-8>
- [16] R. Saha, B. Bhowmick, S. Baishya Temperature effect on RF/ analog and linearity parameters in DMG FinFET. *Appl Phys A Mater Sci Process* 124(9):1–10 (2018). <https://doi.org/10.1007/s00339-018-2068-5>
- [17] V. Subramanian, B. Parvais, J. Borremans et al. Planar bulk MOSFETs versus FinFETs: an analog/RF perspective. *Electron Devices, IEEE Transactions* 2007 on. 53. 3071 - 3079. [10.1109/TED.2006.885649](https://doi.org/10.1109/TED.2006.885649).
- [18] M. Tan, R. Ismail, Modeling of Nanoscale MOSFET Performance in the Velocity Saturation Region. *Elektrika: Journal of Electrical Engineering* 2007.
- [19] E. Yu, K. Heo, and S. Cho, "Characterization and Optimization of Inverted-T FinFET under Nanoscale Dimensions," *IEEE Trans. Electron Devices*, vol. 65, no. 8, pp. 3521–3527, 2018, DOI:10.1109/TED.2018.2846478
- [20] M. R. K. Akanda and Q. D. M. Khosru, "Analysis of output transconductance of FinFETs incorporating quantum mechanical and temperature effects with 3D temperature distribution," 2011 *International Semiconductor Device Research Symposium (ISDRS)*, 2011, pp. 1-2, doi: 10.1109/ISDRS.2011.6135292.
- [21] C. Hyungki, S. Changhwan *Semicond. Sci. Technol* 2019. 34 025004
- [22] X. Yang et al., "Analysis of short channel effects for 14nm and beyond Si-bulk FinFET," 2016 *China Semiconductor Technology International Conference (CSTIC)*, 2016, pp. 1-3, doi: 10.1109/CSTIC.2016.7464044.
- [23] M. H. Han, C. Y. Chang, H. Bin Chen, J. J. Wu, Y. C. Cheng, and Y. C. Wu, "Performance comparison between bulk and SOI junctionless transistors," *IEEE Electron Device Lett.*, vol. 34, no. 2, pp. 169–171, 2013, DOI:10.1109/LED.2012.2231395.
- [24] V. Kilchytska, N. Collaert, R. Rooyackers et al. Perspective of FinFETs for analog applications. *ESSCIRC 2004 - Proceedings of the 34th European Solid-State Device Research Conference*. 65 - 68. [10.1109/ESSDER.2004.1356489](https://doi.org/10.1109/ESSDER.2004.1356489).
- [25] T. Rudenko, V. Kilchytska, N. Collaert et al. Carrier Mobility in Undoped Triple-Gate FinFET Structures and Limitations of Its Description in Terms of Top and Sidewall Channel Mobilities. *Electron Devices* 2009, *IEEE Transactions on*. 55. 3532 - 3541. [10.1109/TED.2008.2006776](https://doi.org/10.1109/TED.2008.2006776).
- [26] M.C. Cheng, J. Smith, W. Jia, R. Coleman. An Effective Thermal Model for FinFET Structure. *Electron Devices, IEEE Transactions* 2014 on. 61. 202-206. [10.1109/TED.2013.2291512](https://doi.org/10.1109/TED.2013.2291512)
- [27] S. Banchhor, K. D. Kumar, A. Dwivedi and B. Anand, "A New Aspect of Saturation Phenomenon in FinFETs and Its Implication on Analog Circuits," in *IEEE Transactions on Electron Devices*, vol. 66, no. 7, pp. 2863-2868, July 2019, doi: 10.1109/TED.2019.2914867.
- [28] C. Xu et al. "Analytical Thermal Model for Self-Heating in Advanced FinFET Devices With Implications for Design and Reliability." *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 32 (2013): 1045-1058.
- [29] A. B. Sachid, M. C. Chen, and C. Hu, "Bulk FinFET With Low-κ Spacers for Continued Scaling," *IEEE Trans. Electron Devices*, vol. 64, no. 4, pp. 1861–1864, 2017, DOI:10.1109/TED.2017.2664798