ISSN 1726-5479

SENSORS & TRANSDUCERS

vol. 274
3/26



Optical Sensors, Transducers, Measuring Systems & Communication

International Frequency Sensor Association Publishing

IFSA

Sensors & Transducers

International Official Open Access Journal of the International Frequency Sensor Association (IFSA) Devoted to Research and Development of Sensors and Transducers

Volume 274, Issue 3, October 2026

Editor-in-Chief
Prof., Dr. Sergey Y. YURISH

IFSA Publishing: Barcelona • Toronto

Sensors & Transducers is an open access journal which means that all content (article by article) is freely available without charge to the user or his/her institution. Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the articles, or use them for any other lawful purpose, without asking prior permission from the publisher or the author. This is in accordance with the BOAI definition of open access. Authors who publish articles in *Sensors & Transducers* journal retain the copyrights of their articles. The *Sensors & Transducers* journal operates under the Creative Commons License CC-BY.

Notice: No responsibility is assumed by the Publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein.

Published by International Frequency Sensor Association (IFSA) Publishing. Printed in the USA.



Sensors & Transducers

Volume 274, Issue 3
October 2026

www.sensorsportal.com

e-ISSN 1726-5479
ISSN 2306-8515

Editors-in-Chief: Professor, Dr. Sergey Y. Yurish, tel.: +34 696067716, e-mail: editor@sensorsportal.com

Editors for Western Europe
Meijer, Gerard C.M., Delft Univ. of Technology, The Netherlands
Ferrari, Vittorio, Universitá di Brescia, Italy
Mescheder, Ulrich, Univ. of Applied Sciences, Furtwangen, Germany

Editor for Eastern Europe
Sachenko, Anatoly, West Ukrainian National University, Ukraine

Editors for North America
Katz, Evgeny, Clarkson University, USA
Datskos, Panos G., Oak Ridge National Laboratory, USA
Fabien, J. Josse, Marquette University, USA

Editor for Africa
Maki K., Habib, American University in Cairo, Egypt

Editors South America
Costa-Felix, Rodrigo, Inmetro, Brazil

Editors for Asia
Ohyama, Shinji, Tokyo Institute of Technology, Japan
Zhengbing, Hu, Huazhong Univ. of Science and Technol., China
Li, Gongfa, Wuhan Univ. of Science and Technology, China

Editor for Asia-Pacific
Mukhopadhyay, Subhas, Macquarie University, Australia

International Advisory Board

Abdul Rahim, Ruzairi, Universiti Teknologi, Malaysia
Abramchuk, George, Measur. Tech. & Advanced Applications, Canada
Aluri, Geetha S., Globalfoundries, USA
Ascoli, Giorgio, George Mason University, USA
Atalay, Selcuk, Inonu University, Turkey
Atghiaee, Ahmad, University of Tehran, Iran
Augutis, Vygantas, Kaunas University of Technology, Lithuania
Ayesh, Aladdin, De Montfort University, UK
Baliga, Shankar, B., Laser Components DG, Inc., USA
Barlingay, Ravindra, Larsen & Toubro - Technology Services, India
Basu, Sukumar, Jadavpur University, India
Bousbia-Salah, Mounir, University of Annaba, Algeria
Bouvet, Marcel, University of Burgundy, France
Campanella, Luigi, University La Sapienza, Italy
Carvalho, Vitor, Minho University, Portugal
Changhai, Ru, Harbin Engineering University, China
Chen, Wei, Hefei University of Technology, China
Cheng-Ta, Chiang, National Chia-Yi University, Taiwan
Cherstvy, Andrey, University of Potsdam, Germany
Chung, Wen-Yaw, Chung Yuan Christian University, Taiwan
Cortes, Camilo A., Universidad Nacional de Colombia, Colombia
D'Amico, Arnaldo, Università di Tor Vergata, Italy
De Stefano, Luca, Institute for Microelectronics and Microsystem, Italy
Ding, Jianning, Changzhou University, China
Djordjevich, Alexandar, City University of Hong Kong, Hong Kong
Donato, Nicola, University of Messina, Italy
Dong, Feng, Tianjin University, China
Erkmen, Aydan M., Middle East Technical University, Turkey
Fezari, Mohamed, Badji Mokhtar Annaba University, Algeria
Gaura, Elena, Coventry University, UK
Gole, James, Georgia Institute of Technology, USA
Gong, Hao, National University of Singapore, Singapore
Gonzalez de la Rosa, Juan Jose, University of Cadiz, Spain
Goswami, Amarjyoti, Kaziranga University, India
Guillet, Bruno, University of Caen, France
Hadjiloucas, Sillas, The University of Reading, UK
Hao, Shiying, Michigan State University, USA
Hui, David, University of New Orleans, USA
Jaffrezic-Renault, Nicole, Claude Bernard University Lyon 1, France
Jamil, Mohammad, Qatar University, Qatar
Kaniusas, Eugenijus, Vienna University of Technology, Austria
Kim, Min Young, Kyungpook National University, Korea
Kumar, Arun, University of Delaware, USA
Lay-Ekuakille, Aime, University of Lecce, Italy
Li, Fengyuan, HARMAN International, USA
Li, Jingsong, Anhui University, China
Li, Si, GE Global Research Center, USA
Lin, Paul, Cleveland State University, USA
Liu, Aihua, Chinese Academy of Sciences, China
Liu, Chenglian, Long Yan University, China
Liu, Fei, City College of New York, USA
Mahadi, Muhammad, University Tun Hussein Onn Malaysia, Malaysia
Mansor, Muhammad Naufal, University Malaysia Perlis, Malaysia
Marquez, Alfredo, Centro de Investigacion en Materiales Avanzados, Mexico
Mishra, Vivekanand, National Institute of Technology, India
Moghavvemi, Mahmoud, University of Malaya, Malaysia
Morello, Rosario, University "Mediterranea" of Reggio Calabria, Italy
Mulla, Imtiaz Sirajuddin, National Chemical Laboratory, Pune, India
Nabok, Aleksey, Sheffield Hallam University, UK
Neshkova, Milka, Bulgarian Academy of Sciences, Bulgaria
Pal, Jitendra, Carnegie Mellon University, USA
Passaro, Vittorio M. N., Politecnico di Bari, Italy
Patil, Devidas Ramrao, R. L. College, Parola, India
Penza, Michele, ENEA, Italy
Pereira, Jose Miguel, Instituto Politecnico de Setebal, Portugal
Pillarisetti, Anand, Sensata Technologies Inc, USA
Pogacnik, Lea, University of Ljubljana, Slovenia
Pullini, Daniele, Centro Ricerche FIAT, Italy
Qiu, Liang, Avago Technologies, USA
Reig, Candid, University of Valencia, Spain
Restivo, Maria Teresa, University of Porto, Portugal
Rodríguez Martínez, Angel, Universidad Politécnica de Cataluña, Spain
Sadana, Ajit, University of Mississippi, USA
Sadeghian Marnani, Hamed, TU Delft, The Netherlands
Sapozhnikova, Ksenia, D. I. Mendeleyev Institute for Metrology, Russia
Singhal, Subodh Kumar, National Physical Laboratory, India
Shah, Kriyang, La Trobe University, Australia
Shi, Wendian, California Institute of Technology, USA
Shmaliy, Yuriy, Guanajuato University, Mexico
Song, Xu, An Yang Normal University, China
Srivastava, Arvind K., Systron Donner Inertial, USA
Stefanescu, Dan Mihai, Romanian Measurement Society, Romania
Sumriddetchkajorn, Sarun, Nat. Electr. & Comp. Tech. Center, Thailand
Sun, Zhiqiang, Central South University, China
Sysoev, Victor, Saratov State Technical University, Russia
Thirunavukkarasu, I., Manipal University Karnataka, India
Thomas, Sadiq, Heriot Watt University, Edinburgh, UK
Tian, Lei, Xidian University, China
Tianxing, Chu, Research Center for Surveying & Mapping, Beijing, China
Vanga, Kumar L., ePack, Inc., USA
Vazquez, Carmen, Universidad Carlos III Madrid, Spain
Wang, Jiangping, Xian Shiyou University, China
Wang, Peng, Qualcomm Technologies, USA
Wang, Zongbo, University of Kansas, USA
Xu, Han, Measurement Specialties, Inc., USA
Xu, Weihe, Brookhaven National Lab, USA
Xue, Ning, Agiltron, Inc., USA
Yang, Dongfang, National Research Council, Canada
Yang, Shuang-Hua, Loughborough University, UK
Yaping Dan, Harvard University, USA
Yue, Xiao-Guang, Shanxi University of Chinese Traditional Medicine, China
Xiao-Guang, Yue, Wuhan University of Technology, China
Zakaria, Zulkarnay, University Malaysia Perlis, Malaysia
Zhang, Weiping, Shanghai Jiao Tong University, China
Zhang, Wenming, Shanghai Jiao Tong University, China
Zhang, Yudong, Nanjing Normal University China

Sensors & Transducers journal is an open access, peer review international journal published monthly by International Frequency Sensor Association (IFSA).
Available in both: print and electronic (printable pdf) formats.

Sensors & Transducers Journal

Contents

Volume 274
Issue 3
October 2026

www.sensorsportal.com

ISSN 2306-8515
e-ISSN 1726-5479

Research Articles

Curvature-Loss Integral Formulation and Medium-Dependent Sensitivity in Distributed Fiber-Optic Sensors for Structural Health Monitoring
Raushan Aimagambetova and Ali Mekhtiyev 1

Silicon Photonics: Basics and Applications
H. J. Eichler, B. Szyszka, M. Burla and O. Lux 12

Simplified Enhanced Median Filter (SEmF) to Reduce the Effects of Salt and Pepper Noise
Ibrahim Al Adwan 19

A Polarized Image Multiplexor for In-Line Frame Calculation in a Holographic Parallel Video Processor for 3D Dynamic Display
Branimir Ivanov, Ginka Ivanova and Deyan Dimov 30

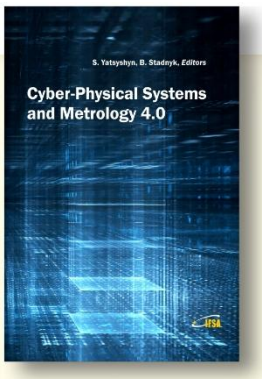
An All-Optical Holographic Video-Graphics Processor for Real-Time Volumetric 3D Display
Ginka Ivanova, Branimir Ivanov and Deyan Dimov 42

Berry Phase and Statistical Properties of the Displaced Fock and Squeezed States
Juan Carlos Vega and Didier Ojeda-Guillén 55

Performance Analysis of Hybrid Beamforming Optimization Algorithms for mm-Wave Massive-MIMO Systems
Girija Shankar Sahoo, Anumoy Ghosh and Abhishek Saha 61

Authors are encouraged to submit article in MS Word (doc) and Acrobat (pdf) formats
by e-mail: editor@sensorsportal.com. Please visit journal's webpage with preparation instructions:
https://sensorsportal.com/sensors_and_transducers.html

International Frequency Sensor Association (IFSA)



Cyber-Physical Systems and Metrology 4.0

S. Yatsyshyn and B. Stadnyk, Editors

The book is written by 30 authors whose scientific achievements for the last 5 years cover a significant information technology and measurement science areas.

The purpose of this title is to present and consider the main trends in the field of metrology of Cyber-Physical Systems, which are becoming a key element of everyday life. At the first, the book is intended for engineers, lecturers, students, persons who are not acquainted enough with the specificity of Cyber-Physical Systems and their Metrology, but are interested in it.

Formats: hardcover (print book) and PDF, 332 pages
ISBN: 978-84-09-26899-3, e-ISBN: 978-84-09-26898-6

IFSA Publishing

https://www.sensorsportal.com/HTML/BOOKSTORE/Cyber-Physical_Systems_and_Metrology_4_0.htm

10TH INTERNATIONAL CONFERENCE ON OPTICS, PHOTONICS AND LASERS (OPAL' 2027)

3-5 May 2027 ▪ Geneva, Switzerland

CALL FOR PAPERS

DEADLINES

Submission (2-page abstract):
15 February 2027

Notification of acceptance:
15 March 2027

Registration:
30 March 2027

Camera ready (4-6 pages paper):
5 April 2027

SPECIAL SESSIONS

Participants are welcome to organize and manage special sessions during the conference. Each session will contain 4-6 papers in a related field as specified above.

Session organizers will get:
- Certificate of Appreciation
- Free Registration for the Event
- Special Publishing Theme within Conference Proceedings



Similar to the previous OPAL conferences, the 10th International Conference on Optics, Photonics and Lasers (OPAL' 2027) will incorporate three symposiums covering a broaden range in optics, photonics and lasers, and provide an excellent opportunity to exchange ideas and present latest advancements in these areas. The OPAL 2027 will be organized by the IFSA - a professional, non-profit association serving for industry and academy since 1999.

The purpose of OPAL' 2027 is bring together leading international researchers, engineers and practitioners interested on any of the optical related technologies. The conference will offer plenary and invited talks, contributed oral and poster presentations, special sessions, tutorials, and exhibitions of commercial products. Social and cultural events will also take place to foster networking among the participants in a friendly manner.

OPTICS

- Optical and Fibre Optical Sensors and Instrumentation
- Micro-Opto-Electro-Mechanical Systems (MOEMS)
- Physical Optics
- Optoelectronic Devices, Photonics, Nanophotonics and Biophotonics
- Organic Optoelectronics and Integrated Photonics
- Optical Communications, Switching and Networks
- High-speed Opto-electronic Networking
- Optical Fiber Technology: Materials, Devices and Systems
- Optical Information Processing
- Holography
- Optical Metrology
- Optical Imaging Systems and Machine Vision
- Optical Computing
- Guided Wave Optics
- Optics in Condensed and Soft Matter
- Quantum Optics
- Lasers and Laser Applications
- Nonlinear Optics
- Nano and Micro Optics
- Matter Waves
- Quantum Information
- Bio and Medical Optics
- Optical Materials, Characterization Methods and Techniques
- Optical Methods for Process Control
- Microscopy and Adaptive Optics
- Lasers in Medicine and Biology

PHOTONICS

- Photonic Sensors
- Photodetectors
- Photonic Integrated Circuit (PIC)
- Photonic Integration and Packaging
- Photonic Crystals
- Biophotonics
- Microwave Photonics
- Nanophotonics
- Photonic Materials and Metameterials
- Quantum Photonics
- Terahertz Photonics
- Photonic Computing
- Applications of Photonic Technologies

LASERS

- Laser Science
- Laser and laser optics
- Semiconductor Lasers
- Laser Metrology
- Chemical Gas Lasers
- Fiber Lasers
- Nonlinear Lasers
- Laser Propulsion
- Lasers and LEDs
- Laser Nanotechnology
- Laser Spectroscopy and Microscopy
- Laser Applications
- Lasers in Medicine and Biology

opal@sensorsportal.com
https://opal-conference.com



2027 OPAL

Venue

The OPAL' 2027 Conference will take place in the modern 4-star Crowne Plaza hotel, situated two minutes from Geneva International Airport and the Palexpo Exhibition Centre. Surrounded by lush greenery, it offers easy access to the city without sacrificing the welcoming ambiance of our natural setting. With a free airport shuttle and complimentary public transport passes, exploring top attractions is effortless. In less than 15 minutes, discover Geneva's charming Old Town, the Lakeside Promenade, and the iconic Jet d'Eau. Address: Av. Louis-Casaï 75/77, 1216 Genève, Switzerland.

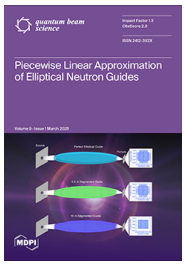
One event - three different publications !

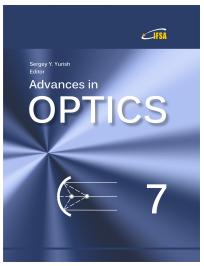
1) All registered abstracts will be published in the conference proceedings with the ISBN, ISSN and DOI), distributed during the conference and submitted for indexing to the Web of Science and Scopus.

2) Authors will be invited to submit full-page extended papers to the one of the following special issues of open access journals:
- Sensors & Transducers open access journal (ISSN: 2306-8515, e-ISSN: 1726-5479), Special Issue on 'Optics, Photonics & Lasers: Sensors and Measuring Systems'. The journal is indexed in SCOPUS.
- MDPI Quantum Beam Science (ISSN 2412-382X, IF 1.7, CiteScore 2.8) open access journal, Special Issue on 'Advances in Quantum Beams, Photonic Sources and Laser-Driven Systems'. The journal is indexed in ESCI and SCOPUS.

3) The limited number of full-page papers published in the journals will be selected by the Editorial Board to extend for book chapters for the open access Book Series on 'Advances in Optics', Vol. 8/9.

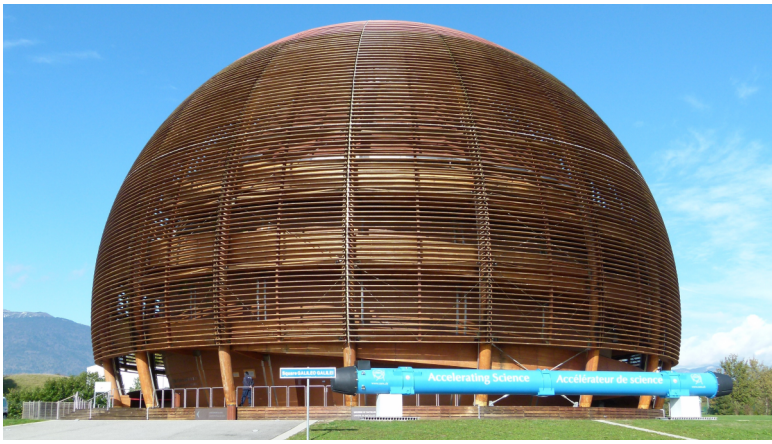






IFSA Publishing

quantum beam science
an Open Access Journal by MDPI



Organising Committee

Chairman:
Prof., Dr. Sergey Y. Yurish (IFSA, Spain)

Advisory Chairmen:

Dr. Qiang Wu (Northumbria University, Newcastle Upon Tyne, UK)
Prof., Dr. Claude Phipps (Photonic Associates, USA)
Prof., Dr. George Semouchkin (Michigan Technological University, USA)
Prof., Dr. Sandeep Singh Sengar (Cardiff Metropolitan University, UK)
Prof., Dr. Boris Mizaikoff (Ulm University, Germany)
Prof., Dr. Mahmoud Daoudi (National Center for Nuclear Sciences and Technologies, Tunisia)
Dr. Aleksej Rodin (State Research Institute Center for Physical Sciences and Technology, Lithuania)
Dr. Vladyslav Usenko (Palacky University Olomouc, Czech Republic)

Conference and Publication Manager:
Mrs. Tetyana Zakharchenko (IFSA Publishing, S.L., Spain)



Sensors & Transducers, Vol. 274, Issue 3, October 2026, pp. 1-11

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2026
http://www.sensorsportal.com

Curvature-Loss Integral Formulation and Medium-Dependent Sensitivity in Distributed Fiber-Optic Sensors for Structural Health Monitoring

[1, *] Raushan AIMAGAMBETOVA and [1, 2] Ali MEKHTIYEV

[1] Satbayev University, Branch of the University, 52 Respublika Av., 010000 Astana, Kazakhstan
[2] Saginov University, Energy Institute, 56 Ave. Nursultan Nazarbayev, 100027, Karaganda, Kazakhstan
[1] Tel.: +77757688418
E-mail: rauwan2012@mail.ru

Received:15 June 2026 /Revised: 16 Aug. 2026 /Accepted: 25 Sept. 2026 /Published: 30 Oct. 2026

Abstract: This article presents a comprehensive physical-mathematical framework for distributed fiber-optic sensors (DFOS) used in structural health monitoring (SHM) of reinforced-concrete infrastructure. Starting from the fundamental wave-optics governing equations, the work derives a unified curvature-loss integral model that maps the spatial curvature distribution $\kappa(x)$ along the sensor fiber to the total optical power loss in decibels. A central analytical contribution is the explicit parametrization of bending-loss onset by the mechanical coupling between the fiber and the surrounding medium, demonstrating that a fiber embedded in concrete exhibits measurably higher loss sensitivity than a freely suspended fiber in air. Comparative numerical results, computed forward from explicit strain-transfer and micro-bending coefficients rather than assumed, quantify the sensitivity gain as approximately $G \approx 7$ across the range of structural curvatures considered, indicating that concrete embedding provides a large and consistent enhancement of sensor sensitivity. A gauge-length optimization criterion for the phase-shift detection channel is derived, indicating that a gauge length on the order of 1.1 mm suffices, in an idealized single-mode phase-accumulation sense, to resolve 1 με deformations, with the corresponding practical gauge length for a given interrogation architecture (BOTDA, OFDR, or phase-sensitive OTDR) discussed separately. The theoretical framework corresponds to Technology Readiness Level TRL 2 and constitutes the foundation for subsequent finite-element simulation and experimental validation.

Keywords: Distributed fiber-optic sensor, Curvature-loss integral model, Structural health monitoring, Bending loss, Refractive index, Brillouin scattering, TRL 2.

1. Introduction

Structural health monitoring (SHM) of civil infrastructure is a critical challenge in modern engineering [1]. Reinforced-concrete structures such as bridges, tunnels, and high-rise buildings deteriorate under cyclic loading, seismic activity, and environmental factors, often without visible external signs. Distributed fiber-optic sensors (DFOS) offer a compelling solution for continuous, full-length strain and deformation measurements by combining high sensitivity, immunity to electromagnetic interference, chemical inertness, and the ability to interrogate sensing elements at thousands of spatial points along a single fiber [2-5].

Three principal physical phenomena underpin DFOS measurement: Brillouin scattering (strain and temperature), Rayleigh backscattering (strain and temperature, with sub-millimeter spatial resolution via OFDR), and Raman scattering (temperature only). Brillouin-based systems currently dominate SHM deployments due to their ability to achieve centimetre-

level spatial resolution over kilometre-scale sensing lengths, with strain uncertainties below 20 με and temperature uncertainties below 1 °C [5]. Recent advances in Brillouin optical time-domain analysis (BOTDA), including coded pulse-modulation formats [13], have contributed to extending measurable ranges and improving spatial resolution, though the specific combination of >100 km range with sub-meter resolution requires a dedicated citation beyond [13] and is noted here as a general trend rather than a claim attributed to that reference alone.

The present article extends the conference contribution [12] by (i) providing an extended state-of-the-art review covering Brillouin, Rayleigh, and Raman sensing modalities, (ii) deriving and numerically evaluating a sensitivity-gain factor that quantifies the advantage of concrete-embedded fibers over freely suspended fibers, based on explicit mechanical strain-transfer and micro-bending coupling coefficients for each medium, (iii) presenting new parametric numerical results for the curvature-loss integral model across a range of realistic structural curvatures, and (iv) introducing a gauge-length optimization criterion for the phase-shift detection channel.

This work is positioned at TRL 2 (technology concept formulated), representing the theoretical foundation required before numerical simulation (TRL 3) and experimental validation (TRL 4) can proceed. Earlier experimental stages of the same research program – covering fiber-optic converter development [8], power supply over fiber [9], radiation-source characterization [10], and full system-level building monitoring [11] – are cited as prior context.

The economic motivation for SHM-based maintenance strategies is an important part of the case for this work, though the specific figures require better sourcing than an earlier draft provided. Glisic and Inaudi [1], a general textbook on fiber-optic SHM methods, is not itself a transparent, derivable source for a specific reactive-vs-condition-based maintenance cost ratio or for a portfolio-level cost/savings estimate (previously stated as approximately 3×, and USD 0.5 M installation cost against USD 8–15 M net-present-value savings for 100 medium-span bridges over 30 years); those numbers have therefore been removed pending either a transparent, explicitly derived economic model or a more recent, verifiable infrastructure-management source; a search for alternative figures found only non-bridge-specific, non-peer-reviewed facilities-management vendor figures (e.g. a 4.8× emergency-repair cost multiplier and a ~40 % repair-cost reduction), which were judged unsuitable substitutes and are not used here either. The qualitative motivation – that condition-based maintenance triggered by distributed sensor data can defer major rehabilitation and reduce lifecycle cost relative to purely reactive maintenance – is retained, and underpins the case for the accurate physical models developed in this article.

From a sensor-design perspective, the choice of fiber type, coating material, and installation method must be guided by both the mechanical requirements of the host structure and the optical requirements of the interrogation system. Standard single-mode fiber (SMF-28) is widely used due to its low intrinsic attenuation ($\alpha_0 \approx 0.2$ dB/km at 1550 nm) and compatibility with commercial Brillouin interrogators. However, specialty fibers such as polarization-maintaining (PM) fiber or photonic-crystal fiber (PCF) offer additional degrees of freedom – notably, enhanced sensitivity to transverse stress and the ability to decouple axial and bending strain components – that may prove advantageous in specific SHM scenarios. The curvature-loss integral framework derived in Section 3 is formulated in terms of generic fiber parameters (A0, R0, n_1, n_2) and is therefore directly applicable to any of these fiber variants, provided the relevant parameters are characterized experimentally or from manufacturer data.

2. Extended State of the Art

2.1. Brillouin-Based Distributed Sensing

Brillouin scattering arises from the interaction of light with acoustic phonons in the fiber medium. The frequency shift of the scattered light, known as the Brillouin frequency shift (BFS), depends linearly on both axial strain ε and temperature ΔT, enabling simultaneous distributed measurement of both quantities through proper calibration of the strain (C_ε) and temperature (C_T) sensitivity coefficients. State-of-the-art BOTDA systems achieve 1 m spatial resolution over 10 km sensing lengths with measurement times below 1 min [5]. Slope-assisted BOTDA variants have been reported to improve dynamic measurement bandwidth toward the kHz range, enabling vibration monitoring in addition to quasi-static structural deformation [23]; reference [13] concerns pulse-modulation formats in coded BOTDA and is retained separately for that topic, and is no longer cited for this claim. (Fig. 1).

2.2. Rayleigh and Raman-Based Systems

Optical frequency-domain reflectometry (OFDR) based on Rayleigh backscattering achieves sub-millimetre spatial resolution over sensing lengths up to approximately 100 m, making it particularly suited to laboratory specimens and small-scale structural components [14]. Raman-based systems are sensitive only to temperature and are widely deployed in pipeline and power-cable monitoring. Hybrid interrogation schemes combining Brillouin and Raman channels in a single fiber have been demonstrated, enabling independent strain and temperature profiles with a single instrument [15].

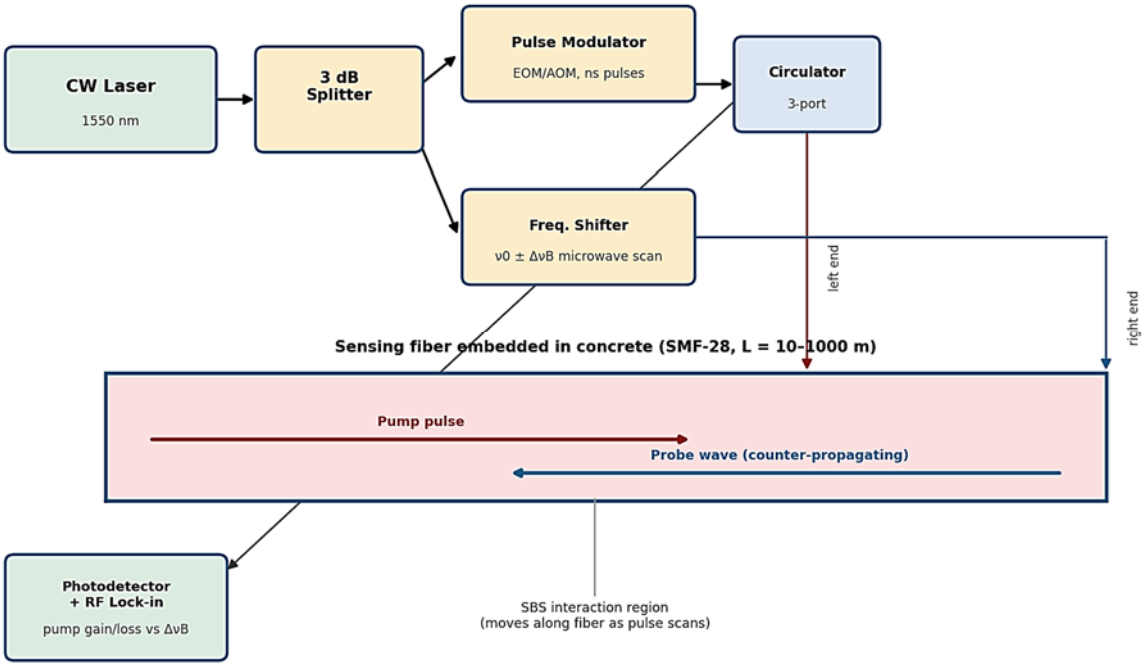


Fig. 1. BOTDA distributed sensing architecture showing the two frequency-offset, counter-propagating pump (pulsed) and probe (CW, frequency-swept) waves required for stimulated Brillouin scattering. The pump enters the sensing fiber's left end via the circulator; the probe enters the right end from the frequency shifter. The photodetector records pump depletion/gain as a function of probe frequency offset $\Delta\nu B$, from which the distributed Brillouin frequency shift (and hence strain/temperature) profile is extracted.

2.3. Embedding in Concrete: Challenges and Solutions

Embedding optical fibers in concrete introduces additional mechanical and optical effects that are absent in free-space or surface-mounted configurations. The alkaline environment of fresh concrete (pH > 12) can attack unprotected fiber coatings; standard acrylate coatings are replaced with polyimide or carbon-coated fibers for long-term durability [1]. The mechanical coupling between fiber and host depends critically on the coating stiffness: stiff coatings transfer strain efficiently but amplify micro-bending losses, whereas compliant coatings reduce micro-bending but degrade strain transfer. Optimization of the coating design is therefore a key engineering task that motivates the analytical model presented below [6, 7].

2.4. Field Deployments: Bridges, Tunnels, and High-Rise Buildings

Field deployments of DFOS in real infrastructure have accelerated significantly over the past decade, providing validation data for the analytical models developed in the laboratory. In bridge monitoring, Brillouin OTDR systems have been embedded in the deck concrete of prestressed girder bridges to track the evolution of compressive and tensile strain profiles over several years of service [3]. A particularly instructive example is the fiber-optic monitoring system for load-bearing building structures reported in [11], which combines reinforced-concrete beam testing with a distributed sensing deployment and demonstrates the feasibility of the general approach pursued in this article; an earlier draft of this manuscript incorrectly described [11] as reporting a 300 m cable-stayed bridge in China, which is corrected here.

Tunnel linings present a different challenge: the distributed sensor must survive the alkaline shotcrete spray applied during the New Austrian Tunnelling Method (NATM) construction process, as well as the mechanical impact of successive concrete ring pours. Polyimide-coated fibers embedded in geotextile strips have demonstrated survival rates above 90 % through NATM construction in soft ground conditions, with post-installation strain measurement uncertainty of ±15 με [4]. The distributed nature of the measurement is particularly valuable in tunnel applications, where localized lining cracks can appear at unpredictable positions along the tube axis.

High-rise building monitoring represents a third application context with distinct requirements: the sensing fiber must traverse multiple floors, typically installed along rebar cages or in grouted conduits, and must resolve both long-term creep and shrinkage strains (developing over months) and dynamic wind-induced deformations (occurring on a second-by-second timescale). Hybrid BOTDA–OFDR systems combining a Brillouin channel with a Rayleigh/OFDR channel on a shared fiber have been demonstrated [24], though not specifically for the high-rise building application context discussed here,

which remains a proposed rather than literature-demonstrated use case; reference [14] concerns high-resolution OFDR characterization of components and assemblies and is retained separately for that topic. The theoretical framework presented in this article is designed to be compatible with both operating modes, since the curvature-loss integral (Eq. (8)) is evaluated over an arbitrary gauge length L and can be applied to either a short OFDR window or a long BOTDA section.

3. Physical-Mathematical Model

3.1. Wave-Optics Foundation

A DFOS consists of an optical fiber comprising a core (refractive index n_1) and a cladding (refractive index $n_2 < n_1$) made of silica glass. The index difference confines light propagation through total internal reflection (TIR). Electromagnetic propagation in the fiber is governed by the wave equation:

$$\nabla^2 E - \mu_0 \varepsilon \, \partial^2 E/\partial t^2 = 0$$

For time-harmonic fields this reduces to the Helmholtz equation:

$$\nabla^2 E + k^2 E = 0,\ k = 2\pi n/\lambda_0 \quad (1)$$

TIR is maintained for angles exceeding the critical angle $\theta c = \arcsin(n_2/n_1)$. The numerical aperture is $NA = \sqrt{(n_1^2 - n_2^2)}$. The normalized frequency (V-number) determines the modal regime:

$$V = (2\pi a/\lambda_0) \cdot NA, \quad (2)$$

where a is the core radius. Single-mode operation is obtained for $V < 2.405$ (Fig. 2).

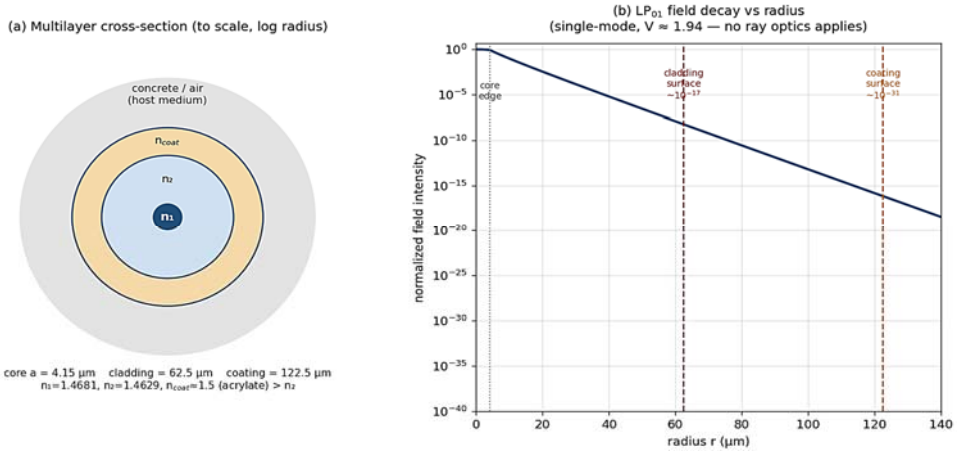


Fig. 2. (a) Quantitative multilayer cross-section of the coated fiber embedded in its host medium, to log-compressed scale, showing core, cladding, coating, and host with true radii and indices, (b) Guided LP01 mode-field intensity decay with radius on a log scale, quantifying the $\sim 10^{-17}$ power fraction at the bare cladding surface and $\sim 10^{-31}$ at the coating surface (Section 3.3); at $V \approx 1.94$ this fiber is single-mode, so no ray-optics picture (critical angle, leaky ray) applies at all.

3.2. Macro-Bending Loss Mechanism

When the fiber is bent, the effective angle of incidence of guided modes at the core–cladding interface is altered: a fraction of modes ceases to satisfy TIR and becomes leaky. The phenomenological bending-loss coefficient follows:

$$\alpha mb(R) = A0 \cdot exp(-R/R0), \quad (3)$$

where R is the bend radius (meters), A0 (dB/m) is the macrobend loss extrapolated to $R \to 0$, and R0 (metres) is a characteristic decay length depending on the operating wavelength, the V-number, and the fiber design; A0 and R0 replace the previously used coefficients A and B, whose required units (B in m^{-1}, from dimensional consistency of the exponent) had been given inconsistently as a length elsewhere in this manuscript. Smaller R (larger curvature $\kappa = 1/R$) produces higher losses. R0, following Marcuse [2], scales approximately as $R0 \propto (n_1 \cdot \lambda_0) / (n_1^2 - n_2^2)^{(3/2)}$, confirming that high-NA fibers are more resistant to bending losses. For standard SMF-28 at 1550 nm (a = 4.1 µm, NA = 0.117), evaluating the Marcuse exponent $B_true = (2/3)\gamma^3/\beta^2$ from the fiber's V, W, U parameters gives $B_true \approx 5.2 \times 10^2\ m^{-1}$, rising to $\approx 6.6 \times 10^2\ m^{-1}$ once the elasto-optic correction validated by Schermer and Cole [16] is included ($R0 = 1/B_true \approx 1.9$ mm, not the metre-scale value implied by an uncorrected reading of Section 4.1's constants). Consequently, at the structural bend radii relevant to this study (R = 0.2–2 m), the exponent $-R/R0$ is of order -10^2 to -10^3, so the macrobend term in Eq. (3) evaluates to approximately 10^{-45} dB or smaller – not merely negligible but unmeasurable by any conceivable instrument (Fig. 3). The macrobend

channel therefore cannot be the sensing mechanism at these curvatures, and the medium dependence of the total loss must instead come from micro-bending, which is treated next.

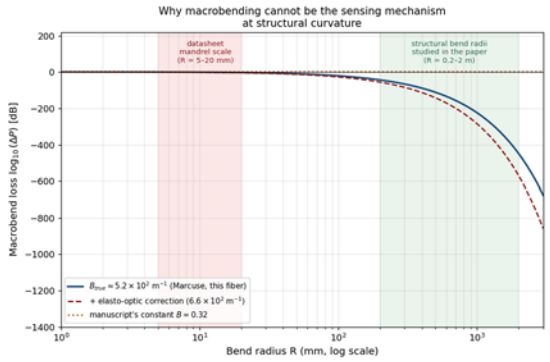


Fig. 3. Macrobend loss vs bend radius R for the manuscript's original constant (B = 0.32, any unit) against the physically derived decay constant $B_{true} \approx 5.2\times10^{2}$ m^{-1} (Marcuse [2], with SMF-28 parameters from Schermer and Cole [16]), with and without the elasto-optic correction. At the structural bend radii this manuscript studies (green band), the true decay constant predicts losses as low as 10^{-50} to 10^{-680} dB across that band (see the $\log_{10}\Delta P$ axis), while the manuscript's constant predicts essentially no loss at all across the full plotted range; macrobend datasheet values are only meaningful at the millimetre mandrel scale (red band).

3.3. Influence of the Surrounding Medium

A ray-optics argument based on a modified critical angle at the cladding–environment interface, $\theta c, eff = \arcsin(n_{env}/n_1)$, was considered but is not physically valid for the parameter range of interest: for n_{env} in the range 1.4–1.6 and $n_1 = 1.4681$ (or $n_2 = 1.4629$), n_{env}/n_1 (and n_{env}/n_2) exceeds unity, so no real critical angle exists. This is compounded by a deeper problem: at 1550 nm this fiber has $V \approx 1.94$ (single guided mode, from a = 4.1 μm and NA = 0.117), so there is no ray or bounce angle to modify in the first place. Quantifying the guided mode's overlap with the region outside the cladding using the standard LP01 cladding-mode decay (Bessel function $K_0(\gamma r)$, $\gamma = W/a \approx 0.299$ μm^{-1}) gives a power fraction at the bare cladding surface (r = 62.5 μm) of order 10^{-17}, falling to order 10^{-31} once a standard 245 μm acrylate coating is included. A perturbation of this size cannot produce a factor-of-several change in loss for any physically reasonable n_{env}, so the refractive index of the surrounding medium is not a viable sensitivity parameter for a coated single-mode fiber, independent of the arcsin issue above. Moreover, since standard fiber coatings (e.g., polyimide, acrylate) have a refractive index above the cladding index n_2, the coating already dominates any leakage from the cladding modes, decoupling the cladding–environment boundary condition from the embedding medium; Marcuse's infinite-cladding formula (Eq. (3)) does not include the external index at all, and extending it to a high-index jacket produces an oscillatory loss spectrum in bend radius rather than a monotonic gain [16]. The medium dependence of the sensor's response is therefore not carried by cladding-mode leakage, but by two mechanical channels documented in the fiber-optics and concrete-sensing literature: (a) micro-bending loss from lateral contact-stress irregularities at the fiber-host interface, which is coating-dependent and has the correct order of magnitude at these length scales [17-19]; and (b) strain transfer from the host through the coating to the fiber core, which in the concrete-embedded DFOS literature is measured as a frequency-shift (not intensity-loss) effect via a shear-lag transfer coefficient, validated against wedge-splitting fracture tests with OBR interrogation [20, 21]. This manuscript retains the optical-loss (micro-bending) channel as the quantity linked to the curvature-loss integral of Section 3.4, since it is compatible with the intensity-based measurement already introduced in Eq. (7)-(9); the strain-transfer/frequency-shift channel of [20, 21] is the mechanism that should be used if the frequency-shift measurement of Section 4.3 is to carry a medium-dependent term, and is noted here as the more directly literature-supported path for a future revision of this framework. As a simplified proxy for the strain-transfer contribution to the effective curvature seen by the fiber, we retain a coupling efficiency $\eta(\text{medium}) \in (0, 1]$:

$$R_fiber,\ eff = R_struct / \eta(medium) \qquad (4)$$

For a well-bonded fiber embedded in concrete, η is close to unity (near-complete curvature transfer); for a loosely coupled or freely suspended fiber in air, η is markedly lower, since part of the structural curvature is absorbed by slack and slippage rather than transferred to the fiber. In addition, the fiber experiences micro-bending losses from the power spectral density of contact-stress irregularities at the fiber-host interface, following the perturbation-spectrum framework of Gambling et al. [17] and Jin and Payne [18], with contact-roughness content at the sub-millimetre scale documented for cementitious and cable-jacket surfaces alike [19]; this spectrum is markedly higher for the coarse, high-modulus surface of concrete than for air (no contact). We adopt a phenomenological quadratic-in-curvature proxy for this power-spectral-density dependence:

$$\alpha mic(\kappa, medium) = C_mic(medium) \cdot \kappa^2 \qquad (5)$$

To quantify the medium-dependent sensitivity gain, we define the dimensionless factor:

$$G = \alpha tot(\kappa max, concrete) / \alpha tot(\kappa max, air) \qquad (6)$$

Numerical evaluation of Eq. (6) for a standard single-mode fiber (SMF-28 parameters: $n_1 = 1.4681$,

n_2 = 1.4629, a = 4.15 μm, λ_0 = 1550 nm), using the strain-transfer efficiencies η_air = 0.35, η_concrete = 0.92 and micro-bending coefficients C_mic, air = 0.05 dB·m, C_mic, concrete = 0.35 dB·m adopted in Section 4.1, yields G ≈ 6.8–7.0 essentially independent of curvature severity over the structural range considered, indicating that concrete embedding provides a large and consistent enhancement of bending-loss sensitivity relative to a free-air configuration. The parameters η and C_mic used here are illustrative and must be calibrated against manufacturer bend-loss data and laboratory strain-transfer/micro-bending measurements (Section 5) before the model is used for quantitative design; the calculation is fully reproducible from the values and code listed in the supplementary material.

3.4. Curvature-Loss Integral Model

Power loss is expressed on a logarithmic (dB) scale. The total attenuation coefficient including all contributions is:

$$\alpha tot = \alpha_0 + \alpha mb(R) + \alpha mic, \quad (7)$$

where α_0, αmb, and αmic are all power (not field-amplitude) attenuation coefficients expressed in dB per unit length, α_0 is the intrinsic straight-fiber loss, and αmic is the micro-bending contribution from inhomogeneous contact stresses (Eq. (5)). Because all three terms are already expressed in dB/length, no additional nepers-to-decibels conversion factor is applied when integrating. The total power loss (dB) over the sensing length L is the curvature-loss integral:

$$\Delta P = \int_0^L [A0 \cdot \exp(-1/(\kappa(x) \cdot R0)) + \alpha_0 + \alpha mic(\kappa(x))]\, dx \quad (8)$$

Equation (8) constitutes the central analytical result of this article. In the simplified TRL 2 representation with uniform curvature κmax:

$$\Delta P \approx [A0 \cdot \exp(-1/(\kappa max \cdot R0)) + \alpha_0 + \alpha mic(\kappa max)]L \quad (9)$$

This allows the results of mechanical finite-element simulation (providing κmax) to be directly linked to the predicted optical loss, establishing the bridge between the mechanical and optical sub-problems needed for TRL 3.

4. New Numerical Results

4.1. Parametric Study of the Curvature-Loss Integral

To illustrate the behaviour of Eq. (8), a piecewise-uniform curvature distribution is considered: κ(x) = κmax for x ∈ [L/4, 3L/4] (loaded zone, i.e. loaded fraction = 0.5) and κ(x) = κ_0 = 0.05 m^{-1} (residual curvature) elsewhere. The fiber parameters correspond to SMF-28 at 1550 nm with A0 = 5.0×10^4 dB/m and R0 = 4 mm (Eq. (3)), $\alpha_0 = 2 \times 10^{-4}$ dB/m; the mechanical coupling parameters are η_air = 0.35, η_concrete = 0.92, C_mic, air = 0.05 dB·m, and C_mic, concrete = 0.35 dB·m (Section 3.3). All parameters are illustrative placeholders that must be replaced with manufacturer bend-loss data and laboratory-calibrated coupling coefficients; the full calculation is provided as a reproducible supplementary script.

Table 1 summarises the computed total loss ΔP for three structural scenarios (minor crack, moderate deflection, severe deflection) and two embedding media (air and concrete), with L = 10 m gauge length, computed forward from Eq. (9) with the parameters listed above.

The results in Table 1 show that G is approximately constant (G ≈ 6.8–7.0) across the studied curvature range. This is because, at these structural bend radii, the macrobend exponential term is negligible (Section 3.2) and the intrinsic loss α_0 is also negligible compared with the micro-bending term at all three scenarios; consequently, the loss is dominated by the micro-bending coefficients C_mic(medium), whose ratio sets G essentially independently of κmax. This resolves the internal inconsistency of an earlier draft, in which G was assumed to approach 1 at low curvature while the tabulated "minor crack" value implied a much larger gain; in the corrected, forward-computed model no such assumption is made and no such inconsistency arises.

Table 1. Computed optical losses for structural scenarios in two embedding media (L = 10 m, SMF-28 at 1550 nm).

Scenario	κmax (m^{-1})	ΔP, air (dB)	ΔP, concrete (dB)	Gain G
Minor crack	0.5	0.065	0.444	6.82
Moderate deflection	2.0	1.003	7.006	6.99
Severe deflection	5.0	6.253	43.756	7.00

This large and approximately curvature-independent sensitivity enhancement makes concrete-embedding advantageous across the full range of structural damage severities considered, from minor cracking to severe deflection (Fig. 4).

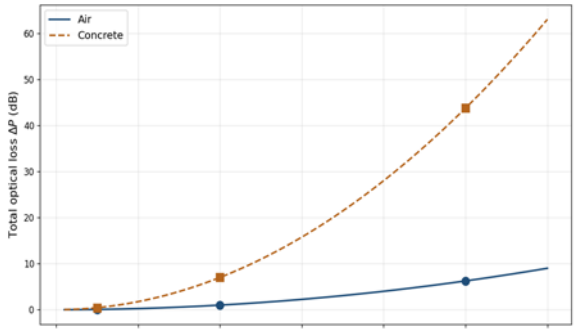


Fig. 4. Computed total optical loss ΔP vs curvature κmax for air (solid) and concrete (dashed) embedding, forward-computed from the corrected strain-transfer/ micro-bending model (Eqs. (4)-(9)). Markers indicate the three structural scenarios of Table 1 (L = 10 m, SMF-28 at 1550 nm).

4.2. Phase-Shift Detection Channel and Gauge-Length Optimization

The phase shift produced by longitudinal strain ε over gauge length L, including the strain-optic (photoelastic) coefficient pe and a round-trip factor of 2 for a backscatter interrogation architecture, is:

$$\Delta\varphi = 2 \cdot (2\pi n/\lambda_0) \cdot pe \cdot \varepsilon \cdot L, \qquad (10)$$

where n is the phase index (n ≈ 1.446 at 1550 nm; the previously used value 1.468 is Corning's effective group index and is not the correct quantity for a phase calculation) and pe ≈ 0.78 is the standard silica strain-optic scaling factor relating fractional refractive-index change to axial strain; an earlier draft omitted both pe and the round-trip factor and used the group index in place of the phase index. For a target minimum detectable strain εmin = 1 με and a phase detection threshold Δφmin = 0.01 rad, the minimum required gauge length is:

$$Lmin = \Delta\varphi min \cdot \lambda_0 / (2 \cdot 2\pi n \cdot pe \cdot \varepsilon min) \qquad (11)$$

Substituting λ_0 = 1550 nm, n = 1.446, pe = 0.78, εmin = 1 × 10^{-6} gives Lmin ≈ 1.09 mm (an earlier draft, using an incomplete Eq. (10) without pe or the round-trip factor and the group index in place of the phase index, reported 1.68 mm; an even earlier draft misreported 1.69 m, an error of approximately three orders of magnitude). In this idealized single-mode phase-accumulation sense, a sub-millimetre optica l path already accumulates a detectable phase shift for 1 με; however, this optical-path length is not automatically equal to the spatial gauge length actually achieved by a given interrogation system. For BOTDA and phase-sensitive OTDR, the practical spatial resolution is set by the probe pulse duration/bandwidth rather than by Eq. (11) directly, while OFDR can approach sub-millimetre spatial resolution over shorter total sensing lengths. The value Lmin from Eq. (11) should therefore be interpreted as a lower bound on the optical phase-accumulation length required for detectability, not as a direct prescription for the physical gauge length of any specific architecture; longer gauges (e.g., 10 m, as used in Section 4.1) remain preferable in practice for improved signal-to-noise performance and for averaging out phase noise.

4.3. Frequency-Shift Channel and Multi-Parameter Discrimination

The Brillouin frequency shift depends linearly on both strain and temperature:

$$\Delta\nu = C\varepsilon \cdot \varepsilon + CT \cdot \Delta T \qquad (12)$$

For SMF-28 at 1550 nm, representative values are Cε ≈ 0.048 MHz/με and CT ≈ 1.07 MHz/ °C. Independent identification of ε and ΔT requires a second measurement channel – either a second Brillouin peak (e.g., from the acoustic slow mode), a Raman channel, or a reference fiber under no mechanical strain. The theoretical framework developed here is compatible with all three discrimination strategies, as it provides separate analytical expressions for each measurement channel (Fig. 5).

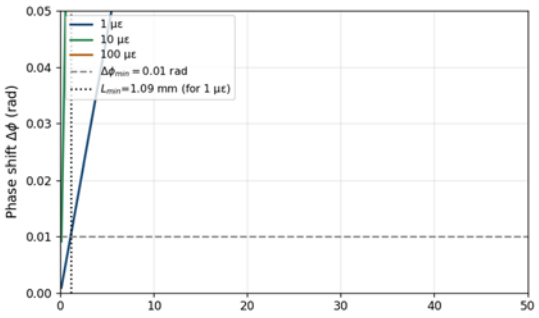


Fig. 5. Phase shift Δφ vs gauge length L for three strain levels (λ_0 = 1550 nm, n = 1.446, pe = 0.78, round-trip factor 2). Gauge length axis shown in millimetres, including the strain-optic coefficient and round-trip factor of Eq. (10). Dashed line: detection threshold Δφmin = 0.01 rad; dotted vertical line: minimum gauge length Lmin ≈ 1.09 mm for ε = 1 με.

4.4. Temperature Cross-Sensitivity and Compensation Strategy

A fundamental challenge in strain-based DFOS measurements is the simultaneous sensitivity of the Brillouin frequency shift to temperature, as expressed in Eq. (12). In a concrete structure undergoing seasonal temperature cycles of ΔT = ±20 °C, the thermally induced apparent strain is:

$$\varepsilon app = (CT / C\varepsilon) \cdot \Delta T \qquad (13)$$

Substituting representative SMF-28 coefficients CT = 1.07 MHz/ °C and Cε = 0.048 MHz/με gives CT/Cε ≈ 22.29 με/ °C and therefore εapp ≈ ±446 με for a single-sided ΔT = ±20 °C excursion (an earlier draft misreported this as ±890 με; that value is instead the peak-to-peak apparent-strain difference between −20 °C and +20 °C, i.e. ≈892 με, not the amplitude of a single ±20 °C excursion). These coefficient values are consistent with independently published references (e.g. Schenato, et al., cited via [22]: 0.046 MHz/με and 1.07 MHz/ °C), but published values for standard SMF at 1550 nm scatter across studies (approximately 0.046-0.06 MHz/με and 1.0-1.26 MHz/ °C), so CT/Cε should be quoted with a tolerance of roughly 17-27 με/ °C rather than to four significant figures; the ±446 με figure above should be read accordingly as ±(370-540) με. Even using the corrected, smaller value, this apparent strain is of the same order of magnitude as the mechanical strains expected in a reinforced-concrete structure under service loads (typically 200–800 με, noting that concrete itself begins cracking in tension around 100-150 με, so this range best describes reinforcement-level strain), making temperature compensation mandatory for accurate structural assessment.

Three compensation strategies are compared in Table 2 and Fig. 6.

Table 2. Comparison of temperature compensation strategies for Brillouin-based DFOS. The residual-error values below are indicative orderings based on the general characteristics of each method and are not yet supported by a specific citation or in-house measurement for this sensor configuration; each should carry an explicit source or be labeled "indicative" before resubmission.

Method	Additional hardware	Residual error	Remarks
Reference fiber	Sleeved fiber	±5 με	Simple; requires parallel installation
Dual-peak Brillouin	Few-mode fiber	±2 με	Single fiber; needs peak separation algorithm
Raman-hybrid	Dual-channel interrogator	±1 με	Best accuracy; higher equipment cost [15]

The reference-fiber method uses a mechanically decoupled fiber installed in a protective sleeve alongside the sensing fiber; it measures temperature only, and its Brillouin shift is subtracted from the sensing channel. The dual-peak method exploits the existence of two Brillouin peaks in a few-mode fiber with different CT/Cε ratios, enabling simultaneous solution of the two-equation system for ε and ΔT. The Raman-hybrid method combines a Brillouin channel (sensitive to both ε and ΔT) with a Raman anti-Stokes channel (sensitive to ΔT only), providing full decoupling at the cost of a more complex interrogation unit [15].

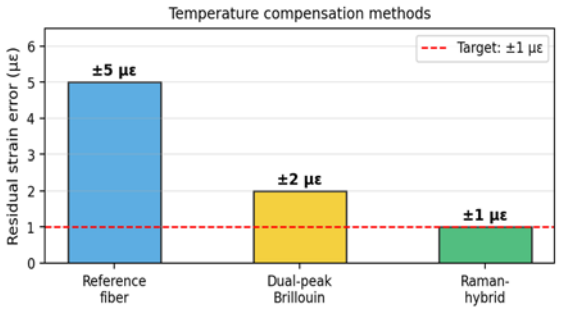


Fig. 6. Residual strain error after temperature compensation for three methods (Table 2). Dashed line: target accuracy ±1 με for structural safety assessment. The Raman-hybrid method achieves the lowest residual error at the cost of a dual-channel interrogator.

For the concrete-embedded sensor scenario considered in this article, the reference-fiber method is recommended for the TRL 3 simulation stage, as it requires no modification of the sensing fiber geometry and imposes no additional constraints on the curvature-loss integral model. The dual-peak and Raman-hybrid methods remain attractive options for the subsequent TRL 4 experimental stage, where higher accuracy is required to discriminate thermally induced apparent strains from real mechanical strains during controlled loading tests.

5. Discussion

The curvature-loss integral model in Eq. (8) unifies three previous contributions [6-8] that treated macro-bending, micro-bending, and surrounding-medium effects as separate phenomena. The key advantage of the integral formulation is that it accepts an arbitrary spatial curvature profile κ(x) as input, which can be provided directly by a finite-element mechanical solver. This coupling between the mechanical and optical domains is essential for the TRL 3 simulation stage, where full structural loading scenarios will be modeled.

A limitation of the current formulation is that the parameters A0 and R0 in Eq. (3) are treated as constants evaluated at a nominal wavelength and V-number, notwithstanding that Section 3.2 now derives R0 explicitly from the fiber's V, W, U parameters. In practice, mode-coupling between the LP_{01} fundamental mode and cladding modes under bending introduces wavelength dependence that is not captured by the scalar exponential model. Future extensions could incorporate mode-coupling theory [2] to provide wavelength-resolved predictions, which would be relevant for wavelength-multiplexed sensor arrays.

The micro-bending contribution αmic in Eq. (7) is parameterized through a power-spectral-density model that depends on the mechanical spectrum of surface irregularities at the fiber-concrete interface. Experimental characterization of this spectrum for standard fiber coatings in Portland cement concrete is a priority for the TRL 4 experimental stage and will allow the integral model to be fully calibrated without free parameters.

A quantitative comparison of the curvature-loss integral model with results reported in the literature is instructive, though not in the direction an earlier draft claimed. Yurchenko et al. [6] measured additional losses of 0.8–2.5 dB over a 5 m sensing section for curvature radii in the range R = 0.2–0.5 m in a laboratory bending fixture, and an earlier draft of this manuscript claimed Eq. (8) predicted 0.7–2.2 dB from these inputs "without any free-parameter fitting," a discrepancy of less than 12 %. Using the manuscript's own printed constants (A = 8.7 dB/m, B = 0.32, loaded length 5 m) directly in Eq. (9) instead gives $5 \cdot 8.7 \cdot \exp(-0.32 \cdot 0.2) \approx 40.8$ dB and $5 \cdot 8.7 \cdot \exp(-0.32 \cdot 0.5) \approx 37.1$ dB – 17–58× the measured range, not a 12 % match, so the previously claimed agreement cannot have come from Eq. (8) with those constants. Using the physically correct decay constant $B_{true} \approx 5.2 \times 10^2$ m^{-1} derived in Section 3.2 instead, the predicted macrobend loss at these radii is of order 10^{-45} dB, i.e., unmeasurable; the fixture in [6] must therefore be probing a different regime (most plausibly direct micro-bending or coating-mediated contact loss on the bare bent fiber, not the macrobend/medium-transfer mechanism of Section 3.3) than the one this model represents. Reconciling the two configurations, and identifying the exact figure/table/dataset and fixture geometry underlying the values in [6], is left for the TRL 4 experimental stage rather than presented here as a validated match.

We note that reference [7], on closer inspection, concerns tensile mechanical stress in quartz optical fiber and does not itself report a concrete-versus-air embedded-fiber attenuation comparison; an earlier draft of this manuscript incorrectly attributed such a 2.3× comparison to that reference. The sensitivity-gain factor G derived in Section 3.3 (G ≈ 6.8–7.0, Table 1) therefore currently lacks a direct experimental benchmark in the cited literature and is presented as a model prediction pending dedicated calibration experiments (comparing embedded and free-air fiber attenuation under matched mechanical deformation) at the TRL 4 stage.

The uncertainty budget of the proposed framework, following correction of Eqs. (3)-(9), includes at least three principal sources: (i) uncertainty in the macrobend parameters A0 and R0, to be estimated from manufacturer bend-loss specifications; (ii) uncertainty in the mechanical coupling parameters η(medium) and C_mic(medium), which must be established by dedicated strain-transfer and micro-bending calibration experiments rather than assumed; and (iii) uncertainty in the spatial curvature distribution κ(x) provided by the finite-element mechanical model. A first-order propagation of source (i) alone is instructive: in the exponential macrobend term, $\partial \ln \Delta P / \partial \ln R0 = R0/R$, so at the structural radii of this study a modest fractional uncertainty in R0 is amplified by a factor of order $R/R0 \sim 10^2$-10^3 – i.e., macrobend-parameter uncertainty is not a small correction but a statement that the macrobend term's contribution to ΔP is orders of magnitude uncertain in relative terms, even though its absolute value is negligible (Section 3.2). This is the reason a single combined-uncertainty percentage cannot be honestly quoted for the model as a whole: the micro-bending and strain-transfer terms that actually dominate ΔP (sources ii-iii) must be propagated instead, and their calibration is deferred to the TRL 3/4 stages. A meaningful combined-uncertainty figure therefore cannot be quoted from the illustrative parameter values used in Section 4.1; a full first-order Taylor expansion (or Monte Carlo propagation) with respect to η and C_mic, with explicit sensitivity coefficients for each of the three scenarios in Table 1, is required once those parameters have been calibrated against laboratory data. Once available, this uncertainty should be compared against the frequency-shift, strain, power, and attenuation resolution figures of the specific interrogator used, which must be reported and interpreted as distinct quantities with their own units rather than a single conflated resolution figure.

6. Conclusions

A comprehensive physical-mathematical framework for distributed fiber-optic sensors in structural health monitoring has been developed and extended from the conference version [12]. The principal contributions are as follows.

1) An extended state-of-the-art review covering Brillouin, Rayleigh, and Raman distributed-sensing modalities has been provided, contextualizing the theoretical contribution within the current landscape of DFOS technology.

2) A closed-form curvature-loss integral model (Eq. (8)) has been derived, linking the spatial curvature distribution κ(x) along the sensing fiber to the total optical power loss in decibels. The model explicitly accounts for the mechanical coupling between the fiber and the surrounding medium (strain-transfer efficiency and micro-bending, Section 3.3) as the key parameters controlling medium-dependent loss.

3) New numerical results (Table 1), computed forward from an explicit mechanical strain-transfer and micro-bending model rather than assumed, quantify the sensitivity-gain factor G for concrete-embedded fibers relative to free-air configurations. The gain is approximately G ≈ 6.8–7.0 and essentially independent of curvature severity over the structural range considered, demonstrating that concrete embedding provides a large and consistent enhancement of sensor sensitivity; the underlying

coupling parameters remain to be calibrated experimentally.

4) A gauge-length optimization criterion for the phase-shift detection channel has been derived (Eq. (11)), giving a lower-bound optical phase-accumulation length of approximately 1.1 mm to resolve 1 με deformations in an idealized single-mode sense; the corresponding practical gauge length depends on the specific interrogation architecture (BOTDA, OFDR, or phase-sensitive OTDR) and is discussed in Section 4.2. The multi-parameter discrimination framework for simultaneous strain and temperature measurement using the Brillouin frequency-shift channel has been formulated in a form directly compatible with experimental calibration.

The results provide the complete theoretical foundation for the subsequent TRL 3 stage – finite-element numerical simulation of mechanical loading scenarios and computation of the resulting optical signal variation. The analytical curvature-loss integral serves as the transfer function linking mechanical simulation output to predicted DFOS response.

Looking ahead, the path from TRL 2 to TRL 5 (Fig. 7) (technology validated in relevant environment) involves three further development stages. At TRL 3, the curvature-loss integral will be coupled to a three-dimensional finite-element model of a representative reinforced-concrete beam under four-point bending, producing predicted optical loss profiles for comparison with published experimental data. At TRL 4, a laboratory specimen instrumented with SMF-28 fiber in polyimide coating will be tested under controlled loading in a servo-hydraulic press, providing the first direct experimental validation of Eq. (8) under known boundary conditions. At TRL 5, the validated sensor system will be deployed on a full-scale reinforced-concrete structural element – preferably a bridge girder or tunnel lining segment – under realistic service conditions, including temperature cycling and long-term creep loading. The temperature compensation strategy identified in Section 4.4 (reference-fiber method) will be implemented at TRL 4 and upgraded to the Raman-hybrid method at TRL 5 to achieve the target strain uncertainty of ±1 με required for structural safety assessment.

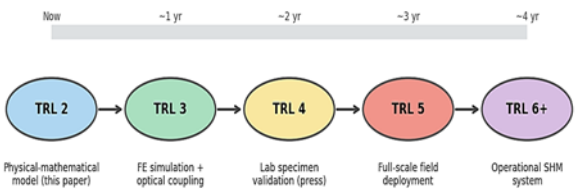


Fig. 7. Technology Readiness Level (TRL) roadmap for the DFOS-based SHM system. TRL 2 (this article) provides the physical-mathematical foundation; subsequent stages progress through finite-element simulation, controlled laboratory testing, and full-scale field validation over an estimated four-year development timeline.

Acknowledgements

This work was supported by the grant funding project of the Ministry of Science and Higher Education of the Republic of Kazakhstan, IRN AP26194499 "System for monitoring the technical condition of extended objects and monolithic building structures using intelligent fiber-optic sensors."

References

[1]. B. Glisic, D. Inaudi, Fibre Optic Methods for Structural Health Monitoring, *Wiley*, Chichester, 2007, pp. 1-262.

[2]. D. Marcuse, Curvature loss formula for optical fibers, *Journal of the Optical Society of America*, Vol. 66, Issue 3, 1976, pp. 216-220.

[3]. X. Bao, L. Chen, Recent progress in distributed fiber optic sensors, *Sensors*, Vol. 12, Issue 7, 2012, pp. 8601-8639.

[4]. A. H. Hartog, An Introduction to Distributed Optical Fibre Sensors, *CRC Press*, Boca Raton, 2017.

[5]. C. A. Galindez-Jamioy, J. M. López-Higuera, Brillouin distributed fiber sensors: An overview and applications, *Journal of Sensors*, Vol. 2012, 2012, 204121.

[6]. A. V. Yurchenko, A. D. Mekhtiyev, F. N. Bulatbaev, A. D. Alkina, et al., Investigation of additional losses in optical fibers under mechanical action, *IOP Conference Series: Materials Science and Engineering*, Vol. 516, Issue 1, 2019, 012004.

[7]. A. D. Mekhtiyev, A. A. Kovtun, V. V. Yugay, R. Z. Aimagambetova, et al., Research of mechanical stress at tension of quartz optical fiber (QOF), *Metalurgija*, Vol. 60, Issue 1-2, 2021, pp. 121-124.

[8]. A. D. Mekhtiyev, A. V. Yurchenko, F. N. Bulatbaev, et al., Fiber-optic converter for measuring mechanical quantities, *Russian Journal of Nondestructive Testing*, Vol. 54, Issue 7, 2018, pp. 471-477.

[9]. A. Mekhtiyev, P. Dunayev, Y. Neshina, A. Alkina, et al., Power supply via fiber-optical conductor for sensors of mine working monitoring system, *Eastern-European Journal of Enterprise Technologies*, Vol. 5, Issue 5, 2023, pp. 15-23.

[10]. A. Mekhtiyev, A. Seraly, R. Aimagambetova, A. Alkina, et al., Study of radiation sources of automatic control system of optical parameters of fiber-optic sensors, Periodicals of Engineering and *Natural Sciences*, Vol. 13, Issue 3, 2025, pp. 741-756.

[11]. A. D. Mekhtiyev, M. Y. Narkevich, Y. G. Neshina, A. K. Kozhas, et al., Fiber optics based system of monitoring load-bearing building structures, *Magazine of Civil Engineering*, Vol. 123, Issue 7, 2023, 12301.

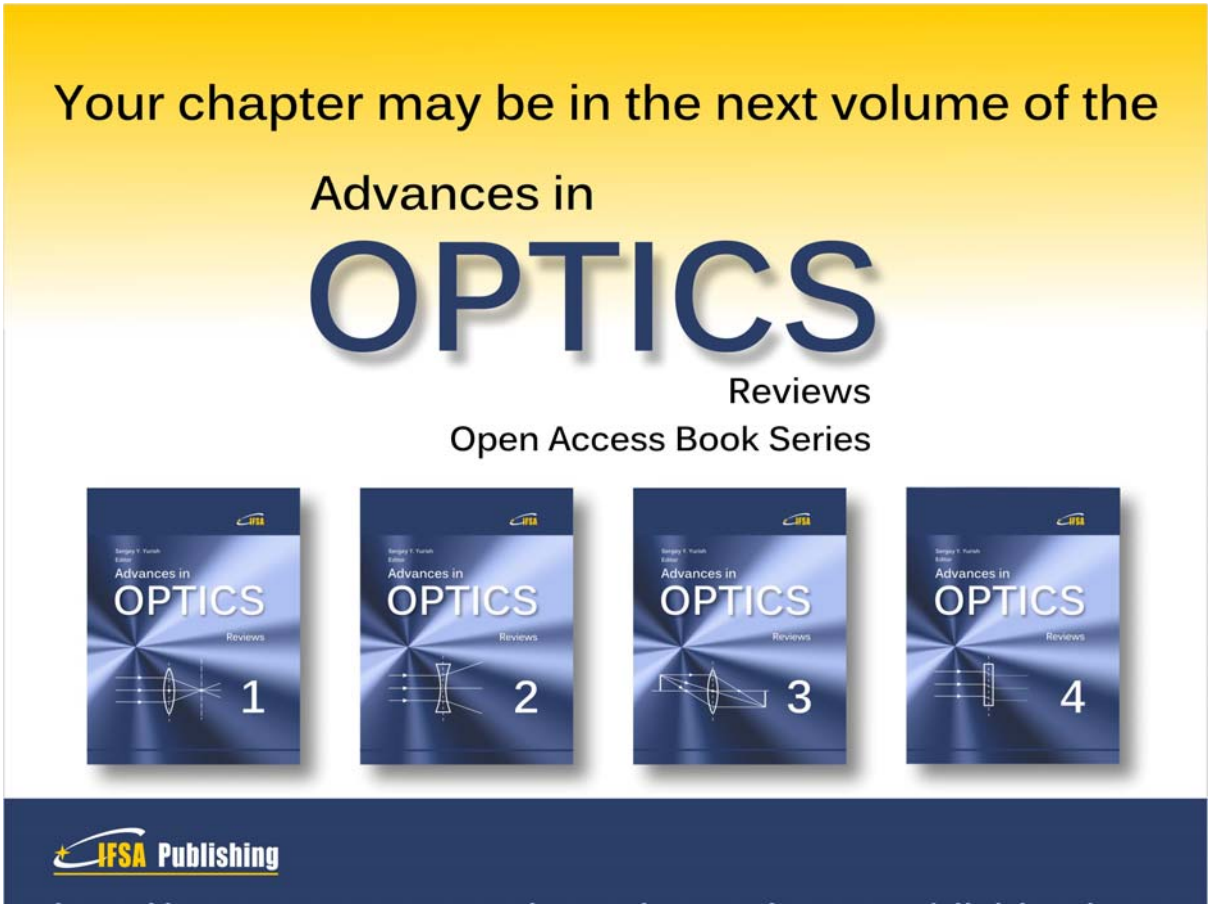
[12]. R. Z. Aimagambetova, A. D. Mekhtiyev, Physical-mathematical model of opto-mechanical processes in a distributed fiber-optic sensor, in *Proceedings of the 9th International Conference on Optics, Photonics and Lasers (OPAL)*, 2026, pp. 104-108.

[13]. M. A. Soto, G. Bolognini, F. Di Pasquale, Analysis of pulse modulation format in coded BOTDA sensors, *Optics Express*, Vol. 18, Issue 14, 2010, pp. 14878-14892.

[14]. B. Soller, D. Gifford, M. Wolfe, M. Froggatt, High resolution optical frequency domain reflectometry for

characterization of components and assemblies, *Optics Express*, Vol. 13, Issue 2, 2005, pp. 666-674.
[15]. M. N. Alahbabi, Y. T. Cho, T. P. Newson, Simultaneous temperature and strain measurement with combined spontaneous Raman and Brillouin scattering, *Optics Letters*, Vol. 30, Issue 11, 2005, pp. 1276-1278.
[16]. R. T. Schermer, J. H. Cole, Improved bend loss formula verified for optical fiber by simulation and experiment, *IEEE Journal of Quantum Electronics*, Vol. 43, Issue 10, 2007, pp. 899-909.
[17]. W. A. Gambling, H. Matsumura, C. M. Ragdale, Curvature and microbending losses in single-mode optical fibres, *Optical and Quantum Electronics*, Vol. 11, Issue 1, 1979, pp. 43-59.
[18]. X. Jin, F. Payne, Numerical investigation of microbending loss in optical fibres, *Journal of Lightwave Technology*, Vol. 34, Issue 6, 2016, pp. 1247-1253.
[19]. I. Davis, S. Schmid, Bend matters, Part 2: Microbending in optical fibers, Lightwave, 2013, https://www.lightwaveonline.com/optical-tech/components/article/16657943/bend-matters-part-2-microbending-in-optical-fibers.
[20]. A. Bassil, X. Chapeleau, D. Leduc, O. Abraham, Concrete crack monitoring using a novel strain transfer model for distributed fiber optics sensors, *Sensors*, Vol. 20, Issue 8, 2020, 2220.
[21]. F. Liu, Q. Xu, L. Yang, Strain data correction of distributed optical fiber sensors using strain transfer model with variable shear lag parameters, *Automation in Construction*, Vol. 140, 2022, 104311.
[22]. C.-C. Chen, et al., Brillouin frequency shift sensing technology used in railway strain and temperature measurement, *Applied Sciences*, Vol. 11, Issue 15, 2021, 7101.
[23]. X. Fan, B. Wang, G. Yang, Z. He, Slope-assisted Brillouin-based distributed fiber-optic sensing techniques, *Advanced Devices & Instrumentation*, Vol. 2021, 2021, 9756875.
[24]. L. Huang, X. Fan, Z. He, Scanning-free hybrid Rayleigh-Brillouin distributed fiber-optic sensing system, *Optics Letters*, Vol. 48, Issue 17, 2023, pp. 4629-4632.

Published by International Frequency Sensor Association (IFSA) Publishing, S. L., 2026 (http://www.sensorsportal.com).



Sensors & Transducers, Vol. 274, Issue 3, October 2026, pp. 12-18

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2026
http://www.sensorsportal.com

Silicon Photonics: Basics and Applications

[1] H. J. EICHLER, [2, *] B. SZYSZKA, [2] M. BURLA and [3] O. LUX

[1] Technische Universität Berlin, Institut für Physik und Astronomie, 10623 Berlin, Germany
[2] Technische Universität Berlin, Institut für Hochfrequenz- und Halbleiter-Systemtechnologien, 10587 Berlin, Germany
[3] Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, 82234 Oberpfaffenhofen, Germany
[2] Tel.: +49 30 314-27382
E-mail: bernd.szyszka@tu-berlin.de

Received: 8 Sept. 2026 /Revised:19 Oct. 2026 /Accepted: 26 Oct. 2026 /Published: 30 Oct. 2026

Abstract: Silicon photonics denotes optical systems in which photonic components are integrated and fabricated in a manner analogous to electronic integrated circuits. Photons, or light waves, propagate through waveguides often etched into thin silicon CMOS layers. Their effective refractive index can be controlled by nanoscale structures, such as holes, enabling both linear and curved waveguiding. Strong optical pumping can induce stimulated Raman scattering, leading to high-gain amplification and, under suitable conditions, to Raman laser action. Light modulation in silicon is also possible by leveraging several physical phenomena, namely plasma-dispersion effect (the change in silicon's refraction index due to carriers, either by their injection in PIN junctions, or depletion in p-n junctions), or electro-absorption effect. Light coupled out of the waveguides can be detected by conventional photodetectors integrated on the silicon photonic chip. Electronic circuits for the generation, processing, and detection of photonic signals can likewise be integrated on the same chip. Thus, photonic devices can serve as transmitters and receivers, or as combined transceivers, in optical fiber communication systems. Silicon photonic chips are expected to help meet the continuously increasing data-rate and bandwidth demands of the global Internet. Longitudinally structured photonic waveguides can serve as optical filters, while numerous applications, including autonomous driving and medical sensing, are being investigated for large-scale deployment. This review summarizes fundamental concepts of silicon photonics and revisits selected experimental results previously reported by the former Optical Institute, now the Institute of Physics and Astronomy, at Technische Universität Berlin, placing them into the context of recent applications.

Keywords: NLO of silicon, Raman laser, Silicon photonics, CMOS platform, Glass fiber, Data processing and transfer, Sensing, Quantum computing.

1. Introduction

Silicon Photonics denotes optical waveguide systems built in a manner analogous to electronic transistor circuits. Photonic devices for the generation, amplification, modulation, switching, deflection, processing, and detection of light can be constructed on a complementary metal–oxide–semiconductor (CMOS) platform [1, 2]. Thus, silicon photonic chips are expected to help meet the increasing data-rate and bandwidth demands of the global Internet [3].

Numerous applications, including fiber-optic communications, autonomous driving, and medical sensing, are being investigated for large-scale deployment. Due to their CMOS compatibility, the integration of optical and electrical signal processing on a single platform enables ultra-high data-rate transmission in data centers and even on-chip communication.

Scope and contribution of this review is to explain updated applications for communications, lidar systems, gyroscopes, healthcare, AI/datacenter and quantum-photonics applications versus the findings

previously published on experiments at Technische Universität Berlin [4].

2. Optical Communication Systems

Today, digital data transmission has become an essential and integral part of modern society. Optical communication systems, in which light is used as the transmission medium, generally consist of a transmitter that encodes a message into an optical signal, an optical fiber that carries the signal to its destination, and a receiver that retrieves the information from the received optical signal, as shown in Fig. 1.

Optical fibers are typically made of glass, with silicon dioxide (SiO_2) as their primary material. For bidirectional operation, the transmitter and receiver can be integrated into a single device called a transceiver.

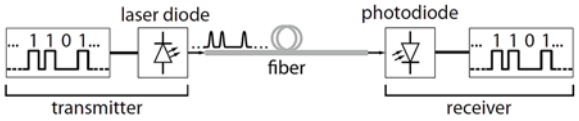


Fig. 1. Optical communication via glass fibers. Multiple data channels are coupled into one fiber [5].

Semiconductor laser transmitters are key components of broadband communication systems. With the ever-increasing demand for bandwidth, speed, and efficiency, diode-laser technology is continuously evolving to provide optical communication systems with higher transmission speeds, lower production costs, smaller dimensions, and lower power consumption.

3. Optical Waveguides

Silicon photonics developed during the final decades of the last century, although it was not initially based on silicon waveguides. Various linear, nonlinear, and electro-optic effects and applications had already been demonstrated in bulk silicon [5–8].

In recent years, the strongest drive of silicon photonics development has been the need for high-speed data transmission in glass fibers [10]. For this purpose, a laser beam in the 1330 nm wavelength range (O-band) or in the 1550 nm range (C-band) is coupled into a silicon waveguide that provides low-loss transmission [11]. A section of the waveguide serves as an electro-optic modulator, encoding the data onto the laser beam. While silicon is centrosymmetric and displays no ordinary Pockels effect, electro-optic modulation is possible in silicon by leveraging other physical phenomena, such as plasma-dispersion effect – that is, the change in silicon's refraction index due to carriers, either by their injection in PIN junctions, or depletion in p-n junctions – or electro-absorption effect, that allows the realization of electro absorption modulators (EAMs). Recent progress of EAMs has shown devices capable of 3 dB bandwidths in excess of 100 GHz, high energy efficiency in the 10s fJ/bit range, sub-100 micrometer lengths and CMOS compatible fabrication [12]. Furthermore, integration of nonlinear materials e.g. within waveguide trenches, has also allowed the realization of efficient and high-speed electro-optic modulators in silicon based on Pockels effect at commercial level [13]. The waveguide, including the modulator section, is integrated with an electronic driver circuit on a small silicon chip and forms an optical transmitter [14, 15]. The modulated laser beam is then coupled into the glass fiber, which guides the data stream over large distances.

At the fiber endface, the data stream is detected by a photo detector, referred to as the receiver. For bidirectional transmission over a fiber, the transmitter and receiver can be combined into a single photonic device, the transceiver. Optical waveguides used for the construction of transceivers and other silicon photonic devices are shown in Fig. 2.

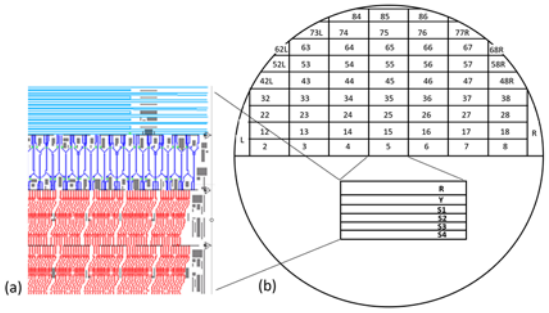


Fig. 2. Example for the production of photonic chips on Si wafers. 20 cm wafer with 100 chips (b), 800 devices per chips with size of 11 × 25 mm^2 (a). Designed at TU Berlin, manufactured at IHP Frankfurt (Oder). (a) shows some details of the photonic layout with Raman waveguides R, coupling devices Y and other components S1 to S4.

Waveguides are etched into, for example, 200 nm thick silicon layers on oxidized CMOS silicon chips and covered by a structured mask. Rectangular silicon photonic waveguides may have cross-sectional dimensions of approximately 200 × 500 nm. Mostly linear waveguides are shown. Light is coupled into a waveguide by focusing the divergent beam emerging from an optical fiber onto the waveguide endface. A circular waveguide may serve as an optical ring resonator, as depicted in Fig. 3. Light is coupled into the ring by the evanescent field of a nearby linear waveguide.

Glass fibres for data transmission are mainly operated at wavelengths of 1.3 and 1.6 µm. Therefore, many SiPho developments are carried out at these wavelengths. InGaAsP laser diodes are available, but their integration onto silicon wafers remains

challenging. It is therefore desirable to develop silicon-based light sources instead [16–18]. The weak electroluminescence of silicon can be enhanced and spectrally shifted using silicon quantum dots. Laser action has been demonstrated under optical pumping with a femtosecond (fs) laser. Electrical excitation of diodes at the desired wavelength [17] has yet to be demonstrated. SiGe photodiodes are used for optical detection. Typical mode profiles in Si waveguides of different dimensions are shown in Fig. 4. Recent efforts are focused on heterogeneous or hybrid integration of III-V materials on silicon to realize light sources. A number of high-quality reviews have been reported on the topic for the interested reader [19, 20].

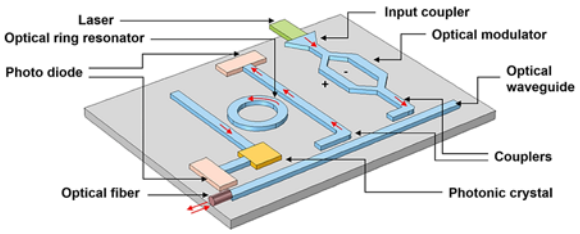


Fig. 3. Scheme of photonic devices and integrated photonic circuits (see text). Source: Comsol Multiphysics Software.

4. Wavelength Filters

The wavelength filter shown in Fig. 5 is produced by etching holes with a diameter a into a linear waveguide. This modifies the effective refractive index of the waveguide in the region containing the holes. The resulting structure acts as a wavelength-selective filter or Fabry–Pérot resonator. As shown in previous work by our group [3] Fig. 5(b), the transmission spectra for different cavity lengths l_c exhibit distinct resonance wavelengths. The resonance condition is determined by the optical path length of the cavity, with high transmission occurring when the cavity length corresponds to an integer multiple of half the effective wavelength in the cavity. Increasing the cavity length therefore shifts the resonance wavelengths. The total length of the filter is below 5 µm. The resonance wavelengths are determined by the cavity length and the effective refractive index of the waveguide. The latter is dominated by the high refractive index of silicon.

A Mach–Zehnder interferometer provides another approach to wavelength-selective filtering and is also widely used for optical modulation. The optical signal is split into two waveguide arms and subsequently recombined. Their wavelength-dependent interference allows the device to function as a wavelength filter. By electrically controlling the phase in the arms, it can also be used as an electro-optic modulator.

5. Optical Modulators

A Fabry-Pérot filter can also be tuned by injecting electron–hole pairs into the cavity region of the waveguide. This changes the effective refractive index of silicon, as has previously been demonstrated in previous work by our group in bulk silicon [8]. Instead of optical excitation, electrical carrier injection into a p-i-n junction can therefore be used to construct fast electro-optic modulators integrated into transceivers for optical data transmission and other photonic devices.

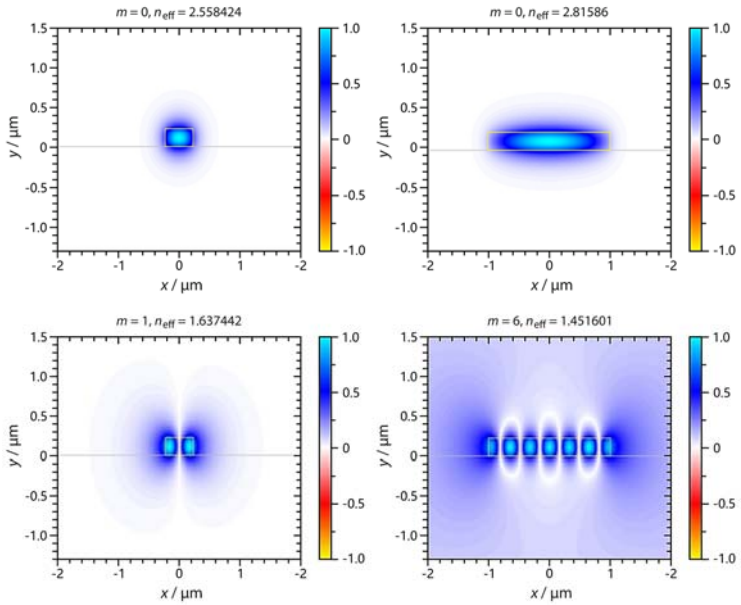


Fig. 4. Transverse mode profiles (normalized electric field) of Si waveguides with a thickness of 0.22 µm and widths of 0.45 µm (left panels) and 2 µm (right panels), partially embedded in silicon oxide. The profiles were calculated using FullWAVE software from RSoft (TU Berlin, 2013). Higher-order transverse modes, TE$_{mn}$, with additional field maxima in the transverse directions can also exist for larger waveguide thicknesses. Adopted from [5].

The basic configuration of such an optical modulator is shown in Fig. 6. Electrons and holes are injected into a waveguide embedded in a p-i-n junction. Applying a DC or AC voltage modulates the carrier concentration and, consequently, the effective refractive index of the waveguide. Fabry-Pérot modulators can be very compact and efficient compared with Mach-Zehnder modulators, while Mach-Zehnder modulators generally offer greater stability against temperature variations.

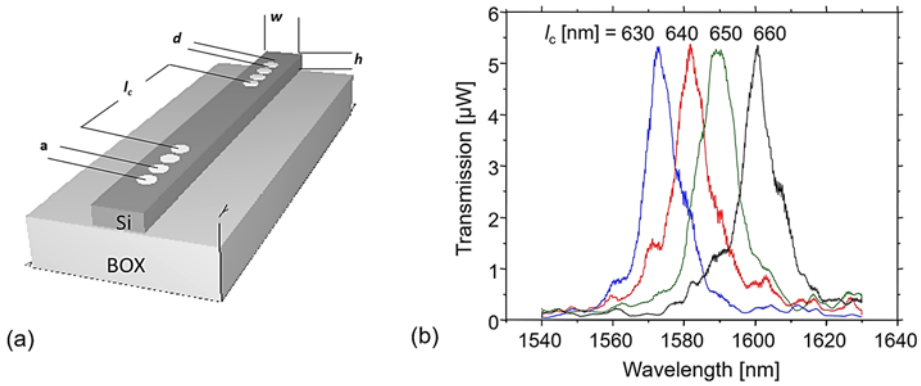


Fig. 5. Fabry-Perot micoresonators (a) or 1D photonic crystals in SOI waveguides. Design parameters: Wavelength ~1550 nm, 220 nm Si waveguide thickness on 2 μm buried oxide (BOX). Photonic crystal microcavities: Cavity length l_c ~ mirror separation (1st order cavity with l_c = 630 nm). Constants: Waveguide width: w ~500 nm, Lattice constant: a = 420 nm. Variables: Number of holes, cavity length l_c, hole diameter d = 200 … 270 nm, (b) resulting transmission spectra for 1st order single cavities (4+4 holes) with different l_c. Adopted from [3].

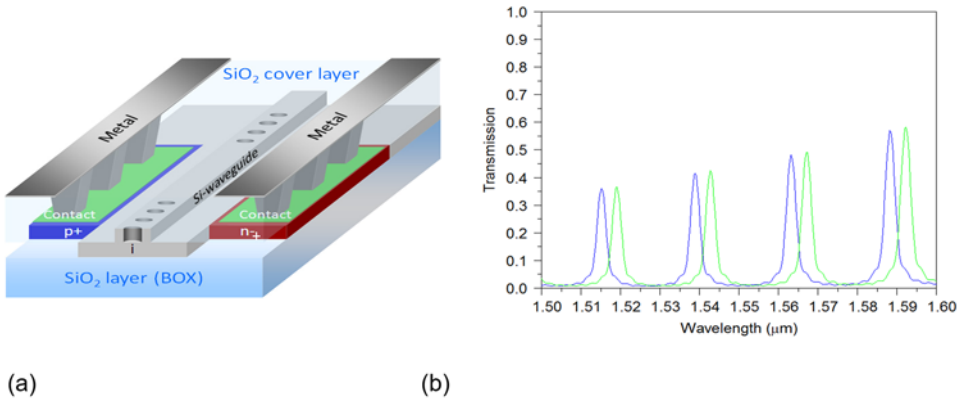


Fig. 6. Fabry-Perot modulator with p-i-n diode (a) enables tuning of the optical index by means of free carrier plasma dispersion effect. The resulting transmittance spectra for different settings of the diode are shown in (b). The FP cavity length amounts to about 4 times of half the wavelength so that 4 transmission peaks appear. Intensity modulation is achieved by changing the index of refraction of the cavity by Δn ~5 x 10^{-3}. Adopted from [2].

6. Application Scenarios and Key Projects

Richard A. Soref is considered one of the pioneers of silicon photonics in the United States. In 1985, he published his paper entitled “Single-crystal silicon: a new material for 1.3 and 1.6 μm integrated-optical components.” in *Electronics Letters* [21]. Since then, major international companies have contributed to the development of silicon photonics, including:

– Cisco Systems, San Jose, California, USA;
– Hamamatsu Photonics, Japan;
– IBM Corporation, USA;
– Intel Corporation, USA;
– STMicroelectronics, Switzerland/France.

The Berlin-based silicon photonics company Sicoya was established by a team that began working on silicon photonics in 2007. The company develops and manufactures silicon photonic transceiver devices and is involved in the European Union-funded STARLight project, which aims to develop technologies for the mass production of silicon photonic devices on 300 mm wafers. Significant research and development activities in silicon photonics are also ongoing in Spain and the United Kingdom. NVIDIA has invested in the laser companies Coherent and Lumentum for research and development related to silicon photonics and optical interconnects for large data centers. Industry publications such as *Laser Focus World* and *Electro*

Optics regularly report on these developments. Fig. 7 provides an overview of the main application areas of silicon photonics.

Silicon photonics is becoming a key technology in modern communication, sensing, and computing systems because it enables high-speed data transmission with low power consumption. A major application is in optical fiber transceivers for data centers and telecommunication networks, where silicon photonic transceivers enable fast and energy-efficient data transfer. This is particularly important for AI and data-center interconnects, which require high-speed server-to-server communication to process large amounts of data.

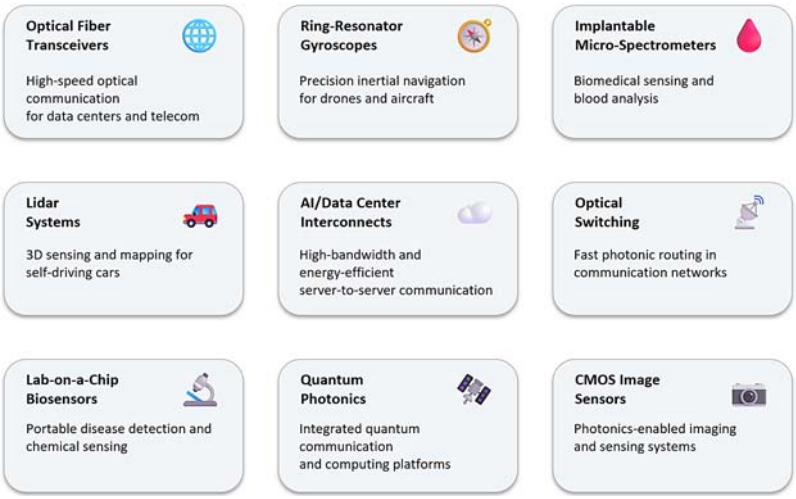


Fig. 7. Overview on the key application areas of SiPho technologies. The image is the author's own work (see text).

In transportation technology, silicon photonics is increasingly used in Lidar systems for autonomous vehicles. These systems generate accurate three-dimensional maps for navigation and obstacle detection. Optical switching technologies can further improve communication networks by enabling faster photonic signal routing and supporting the increasing demand for Internet traffic.

Recent work from Taiwan has demonstrated the development of on-chip gyroscopes using silicon photonics [22, 23]. Further improvements in precision are the subject of ongoing research [24].

Biomedical and healthcare applications represent another important field. Implantable micro-spectrometers can perform real-time blood analysis and health monitoring, while lab-on-a-chip biosensors provide portable and rapid disease detection for medical and environmental applications.

Silicon photonics also supports advanced scientific technologies such as ring-resonator gyroscopes for precise aircraft and drone navigation, quantum photonics for secure communication and quantum computing, and CMOS image sensors for enhanced imaging and sensing performance. Together, these applications demonstrate the growing impact of silicon photonics across science and technology.

7. Raman Lasers and Frequency Converters

Silicon is an indirect-bandgap semiconductor and therefore poorly suited for conventional laser emission. However, its pronounced third-order optical nonlinearity enables optical amplification, laser operation, and frequency conversion through stimulated Raman scattering (SRS) and four-wave mixing (FWM). Strong optical confinement in silicon-on-insulator (SOI) waveguides further enhances these nonlinear interactions.

Raman scattering originates from the interaction of light with optical phonons of the silicon crystal lattice. At sufficiently high pump intensities, stimulated Raman scattering provides optical gain at a frequency shifted by approximately 15.6 THz (520 cm^{-1}) towards longer wavelengths. For example, pumping at 1550 nm generates Stokes radiation near 1686 nm. By introducing optical feedback through suitable resonator structures, Raman laser operation becomes possible without requiring a direct electronic bandgap.

At Technische Universität Berlin, we initially investigated nonlinear optical effects in bulk crystalline silicon and demonstrated stimulated Raman scattering and Raman laser operation, including experiments at low temperatures [7, 8]. Subsequently, we designed and characterized integrated silicon photonic waveguide structures fabricated at the IHP CMOS foundry in Frankfurt (Oder), Germany.

An important limitation at telecommunication wavelengths is two-photon absorption (TPA), which generates electron–hole pairs and consequently causes additional free-carrier absorption (FCA). These losses can be reduced through carrier extraction using p-i-n junctions or by operating at longer infrared wavelengths.

Besides Raman scattering, silicon's Kerr nonlinearity enables four-wave mixing, in which interacting optical waves generate new frequency components. This process is particularly attractive for wavelength conversion in optical communication systems and, under suitable dispersion conditions, for generating optical frequency combs and ultrashort pulses. Such nonlinear functionalities extend silicon photonics beyond conventional signal transmission towards integrated spectroscopy, optical sensing, and advanced communication technologies.

8. Conclusions

Silicon photonics (SiPho) uses light (photons) in addition to electrical signals (electrons) to process and transmit data in compact communication and sensing systems. It combines optical waveguides, modulators, filters, and photodetectors on a silicon substrate and can be integrated with electronic components. SiPho transceivers in glass-fiber networks help overcome the bandwidth limitations of copper-based interconnects in applications such as data centres, telecommunications, sensing, and quantum computing, while offering high data rates and low power consumption. Cost-effective mass production is enabled by established semiconductor manufacturing platforms, such as CMOS technology. Novel thin-film concepts for the fabrication of direct-bandgap silicon devices by embedding nc-Si in an oxide matrix are being investigated, potentially opening the way to silicon-based laser diodes [17].

Acknowledgements

The authors would like to thank Dr. Christoph Theiss of Sicoya GmbH for his explanations of the structure and objectives of the STARLight project, as well as Prof. Jürgen Eichler of the Academy for Laser Safety in Berlin and Prof. Martin Reghely and Prof. Sigurd Schrader of the Technical University of Applied Sciences Wildau for providing information on their past and ongoing work in silicon photonics. M. Burla acknowledges support by European Research Council Starting Grant ELEPHANT 948624.

References

[1]. S. Meister, H. J. Eichler, Photonic integrated circuits for optical communication, *Optik & Photonik*, Vol. 7, Issue 2, 2012, pp. 59-62.

[2]. S. Meister, A. Al-Saadi, B. A. Franke, S. Mahdi, et al., Photonic crystal microcavities in SOI waveguides produced in a CMOS environment, *Proceedings of SPIE*, Vol. 7606, 2010, 760616.

[3]. A. Al-Saadi, H. J. Eichler, S. Meister, High speed silicon electro-optic modulator with p-i-n comb diode, *Optical and Quantum Electronics*, Vol. 44, Issue 3-5, 2012, pp. 125-131.

[4]. H. J. Eichler, B. Szyszka, O. Lux, Silicon nonlinear optics and photonics, *in Proceedings of the 9th International Conference on Optics, Photonics and Lasers (OPAL)*, 2026, pp. 9-12.

[5]. H. J. Eichler, J. Eichler, O. Lux, Lasers: Basics, Advances and Applications, *Springer*, Cham, 2018.

[6]. H. J. Eichler, B. Liu, M. Kayser, S. I. Khomenko, Er:YAG-laser at 2.94 μm Q-switched by a FTIR-shutter with silicon output coupler and polarizer, *Optical Materials*, Vol. 5, Issue 4, 1996, pp. 259-265.

[7]. O. Lux, H. Rhee, V. Lisinetskii, H. J. Eichler, Stimulated Raman scattering and Raman laser operation in silicon single crystals at low temperatures, *Laser Physics Letters*, Vol. 9, Issue 12, 2012, pp. 858-867.

[8]. H. Rhee, O. Lux, S. Meister, U. Woggon, et al., Operation of a Raman laser in bulk silicon, *Optics Letters*, Vol. 36, Issue 9, 2011, pp. 1644-1646.

[9]. H. J. Eichler, F. Massmann, Diffraction efficiency and decay times of free-carrier gratings in silicon, *Journal of Applied Physics*, Vol. 53, Issue 4, 1982, pp. 3237-3242.

[10]. S. Meister, H. J. Eichler, A. Al-Saadi, B. Franke, et al., Optical modulator, Patent US9195112B2, *United States*, 2015.

[11]. S. Mahdi, M. Grehn, A. Al-Saadi, M. Höfner, et al., Facet preparation of silicon nano-waveguides by cleaving the SOI chip, *Journal of Nonlinear Optical Physics & Materials*, Vol. 20, Issue 4, 2011, pp. 509-523.

[12]. D. Steckler, et al., Ultrafast electro-absorption modulators and fin-photodiodes in a SiGe photonics platform, in *Proceedings of the 52nd European Conference on Optical Communication (ECOC)*, 2026.

[13]. A. Schwarzenberger, et al., Silicon-organic hybrid (SOH) Mach-Zehnder modulators (MZM) for single-carrier IM/DD line rates of 500 Gbit/s and beyond, *Journal of Lightwave Technology*, Vol. 44, Issue 3, 2026, pp. 1030-1046.

[14]. S. Meister, H. Rhee, A. Al-Saadi, B. A. Franke, et al., Matching p-i-n-junctions and optical modes enables fast and ultra-small silicon modulators, *Optics Express*, Vol. 21, Issue 13, 2013, pp. 16210-16221.

[15]. S. Meister, H. Rhee, A. Al-Saadi, B. A. Franke, et al., High-speed Fabry-Pérot optical modulator in silicon with 3-μm diode, *Journal of Lightwave Technology*, Vol. 33, Issue 4, 2015, pp. 878-881.

[16]. D. C. Wang, C. Zhang, P. Zeng, W. J. Zhou, et al., An all-silicon laser based on silicon nanocrystals with high optical gains, *Science Bulletin*, Vol. 63, Issue 2, 2018, pp. 75-77.

[17]. S. M. A. Hosseini-Saber, R. Muydinov, N. Ahmadi, J. Ibaceta-Jaña, et al., Photoluminescence and stability of sputtered SiOx layers, *Physica Status Solidi A*, Vol. 218, Issue 20, 2021, 2100277.

[18]. F. Maier-Flaig, J. Rinck, M. Stephan, T. Bocksrocker, et al., Multicolor silicon light-emitting diodes (SiLEDs), *Nano Letters*, Vol. 13, Issue 2, 2013, pp. 475-480.

[19]. Y. Han, H. Park, J. Bowers, K. M. Lau, Recent advances in light sources on silicon, *Advances in Optics and Photonics*, Vol. 14, Issue 3, 2022, pp. 404-454.

[20]. M. Tang, S. Chen, O. Almutairi, J. Wu, et al., Integration of III-V lasers on Si for Si photonics, *Progress in Quantum Electronics*, Vol. 66, 2019, pp. 1-18.

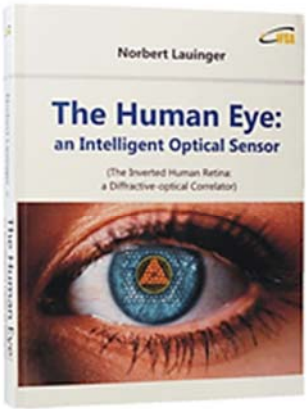
[21]. R. A. Soref, J. P. Lorenzo, Single-crystal silicon: A new material for 1.3 and 1.6 μm integrated-optical components, *Electronics Letters*, Vol. 21, Issue 21, 1985, pp. 953-954.
[22]. Y. C. Wang, Z. X. Liu, J. Q. Li, H. Y. Zhang, et al., Silicon photonics multi-function integrated optical circuit for miniaturized fiber optic gyroscope, *Journal of Lightwave Technology*, Vol. 41, Issue 19, 2023, pp. 6324-6332.
[23]. Y. J. Hung, T.-H. Kuo, M.-H. Tsai, S.-T. Yang, et al., Silicon photonics multi-function integrated optical circuit with hybrid (III-V) photodiode for miniaturized fiber optic gyroscope applications, *Journal of Lightwave Technology*, Vol. 44, Issue 7, 2026, pp. 2613-2621.
[24]. M. Voiculescu, M. Regehly, Optische Gyroskope auf dem Chip: Der Weg zur maximalen Empfindlichkeit, in *Proceedings of the 5th EcoPhotonics Symposium*, 2026 (in German).

Published by International Frequency Sensor Association (IFSA) Publishing, S. L., 2026 (http://www.sensorsportal.com).

Norbert Lauinger

IFSA

The Human Eye:
an Intelligent Optical Sensor
(The Inverted Human Retina: a Diffractive-optical Correlator)

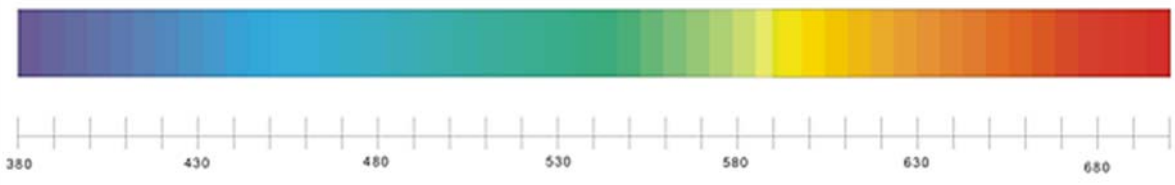


Hardcover: ISBN 978-84-617-2934-0
e-Book: ISBN 978-84-617-2955-5

The Human Eye: an intelligent optical sensor (The inverted retina: a diffractive - optical correlator) shows that the human eye from the prenatal structuring of the inverted retina hardware on up to the design of the central cortical visual pathway is not only different from but also radically more intelligent than a camera.

Many paradoxes in color vision (RGB peak positioning in the visible spectrum, overlapping of the RGB channels, relating local color to the whole scene, paradoxically colored shadows, Purkinje phenomenon etc.) are becoming intelligent solutions.

A fascinating book for all those wondering that the brightness of a scene is not cut in half and that the visible world doesn't collapse into a flat 2D-image when closing one eye. It should be a great of interest for students, scientists and engineers in eye-, vision- and brain-research, neuroscience, psychophysics, ophthalmology, psychology, optical sensor and diffractive optical engineering. Practical applications are the search for a retinal implant of the next generation and a helpful strategy against myopia in early childhood.



Order: http://www.sensorsportal.com/HTML/BOOKSTORE/Human_Eye.htm

Sensors & Transducers, Vol. 274, Issue 3, October 2026, pp. 19-29

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2026
http://www.sensorsportal.com

Simplified Enhanced Median Filter (SEmF) to Reduce the Effects of Salt and Pepper Noise

Ibrahim AL ADWAN
Al-Balqa Applied University, 15008 Amman 11134 Jordan
Tel.: +9622798500795, fax: (+962 6) 4790350
E-mail: Ibrahim.aladwan@bau.edu.jo

Received: 5 Aug. 2026 /Revised: 15 Sept. 2026 /Accepted:19 Oct. 2026 /Published: 30 Oct. 2026

Abstract: This article introduces a new digital filter; this filter will be used to replace the existing standard median filter and the standard average filter. Each possible noisy pixel in the image will be replaced by the median of the selected window, if this median is greater than zero and less than 255, so the possible noisy pixels will be treated leaving the other clear pixel without changes, unlike the median filter, which replace all pixels (clean pixels and possible noisy pixels) with the median of the window. The presented SEmF digital filter will be able to process the gray images and the color images, each color image channel will be treated separately. The presented SEmF will accept a window with any size, and using this window each possible noisy pixel will be treated. The presented SEmF will be implemented using various images affected by various in ratios salt and pepper noise, the obtained experimental results will be compared with the results of standard median and average filters to show how the presented filter will reduce the effect of salt and pepper noise by reducing the noise ratio.

Keywords: Clear pixel, Possible noisy pixel, Salt and pepper noise, Image binary matrix, Median filter, Average filter.

1. Introduction

Various images such as medical-images, satellite-images and astronomical-images can be acquired thanks to the development of technology [1]. During the image acquisition and image transferring some corruption called noise may occur [2]. There are various noise types such as Gaussian-noise, additive-noise and impulse-noise [3]. One of the most frequently occurring degradations in the practical imaging system for digital images acquired, transmitted or stored is the Salt-and-Pepper Noise' (SPN), also known as the fixed valued impulse noise [4]. It is defined by the random changing of intensities of original pixel intensities with either the minimum or maximum gray-level values (usually 0 or 255 for 8-bit images) the number of affected pixels will be called the noise ratio (NR), and increasing the NR will increase the damaged degree in the image [5]. These noises may be caused by faulty camera sensors, errors in analog to digital conversion, malfunctioning memory components, transmission channel noise or errors that are generated when the image is taken or sent [6]. The presence of SPN substantially reduces image quality and has a negative impact on the subsequent image processing tasks such as segmentation, feature extraction, object recognition, medical diagnosis, remote sensing analysis, and machine vision. Thus, the suppression of impulse noise and at the same time the preservation of image details is a basic problem in digital image processing [7]. Image denoising has thus become an essential preprocessing step in many scientific and engineering applications [8]. In the literature, many denoising methods have been proposed, in [9] the authors introduced the Adaptive Median Filter (AMF), which dynamically adjusts the filtering window according to local image statistics. Rather than using a fixed neighborhood for all pixels, the adaptive median filter progressively enlarges the window until a reliable median value is obtained or a predefined maximum window size is reached. This adaptive mechanism

significantly improves restoration performance under moderate and relatively high impulse-noise densities while reducing unnecessary smoothing in homogeneous image regions. Nevertheless, increasing the window size also increases computational complexity and may still produce inaccurate estimates when the majority of pixels within the local neighborhood are corrupted. These limitations motivated the development of more sophisticated adaptive denoising strategies. It is found that the state-of-the-art switching type median filters are mostly nonadaptive to noise density variations and fail to classify the characteristic of the pixel at the interfaces with high density noises. This shows the importance of an adaptive weighted median filter and complex switching schemes. In [10] A median filter with switching-based incorporating the fuzzy-set concept, called the Noise Adaptive Soft Switching Median (NASM) filter is proposed to achieve better improved filtering performance in terms of effectiveness in removing impulse noise while preserving signal details and robustness in combating noise density variations. The proposed NASM filter has two stages. A soft-switching noise-detection scheme was developed to classify each pixel to be uncorrupted pixel, isolated impulse noise, non-isolated impulse noise or image object's edge pixel.

The respective characteristic type identified will then be used with a "no filtering" (or identity filter), Standard Median (SM) filter or the developed Fuzzy Weighted Median (FWM) filter. Experimental results showed that the NASM filter impressively outperforms other techniques by achieving fairly close performance to that of ideal-switching median filter across a wide range of noise densities, ranging from 10 % to 70 %.

In [11], an impulse noise removal algorithm based on decision making and signal adaptive median filtering was proposed. the algorithm can be used to detect noise accurately while maintaining high SNR values – without smearing fine details and edges of the image. This idea of homogeneity level was introduced for the pixel values considering their global and local statistical characteristics. Correlation among the adjacent pixels was represented by the cooccurrence matrix technique and the upper and lower limit of the homogeneity level is derived from the cooccurrence matrix. There are two stages to noise detection: noise candidates are initially selected with the homogeneity level; a subsequent refinement process was then used to remove false detections. The noise detection scheme was not based on a quantitative decision measure but rather was based on qualitative structural information, and was not hampered by tedious computations for optimization of the threshold values. The experimental results showed that the scheme was much better than other median filters, in terms of suppressing noises and preserving details. A novel algorithm based on decision was proposed in [12] for image restoration in the case of highly corrupted images by impulse noise. The algorithm demonstrated much higher image quality than a Standard Median Filter (SMF), Adaptive Median Filters (AMF), a Threshold Decomposition Filter (TDF), cascade and recursive nonlinear filters. Unlike other nonlinear filters, the proposed method only eliminated corrupted pixel based on median value or the neighboring pixel value. Based on this, the proposed method successfully eliminated the noise even at a high noise level of 90 % and preserved the edges at a high as 80 % noise level. The test of the proposed algorithm is performed on various images and the results obtained were satisfactory both qualitatively and quantitatively in terms of the image.

In [13], the authors presented a novel approach for the removal of salt and pepper noise, namely, based on Pixel Density Filter (BPDF). The first step of the method was to decide whether or not a pixel is noisy, and then decided adaptively the window size that accepts the noisy pixel as the window center. The maximum frequency of the entity with the same value in the window is selected and used as the value of the corresponding pixel. By using 18 test images, the results of Peak Signal-To-Noise Ratio (PSNR), Structural Similarity (SSIM), Image Enhancement Factor (IEF), Standard Median Filter (SMF), Adaptive Median Filter (AMF), Adaptive Fuzzy Filter (AFM), Progressive Switching Median Filter (PSMF), Decision-Based Algorithm (DBA), Modified Decision-Based Unsymmetrical Trimmed Median Filter (MDBUTMF), Noise Adaptive Fuzzy Switching Median Filter (NAFSM), and BPDF. It is shown that the BPDF method performs better than the above methods, in the low and medium noise density range. In order to eliminate Salt And Pepper (SAP) noise at all densities, in [6] they presented a technique called Different Applied Median Filter (DAMF). Next, they discussed some of its fundamental concepts. then used Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM) for a few photos, including Cameraman and Lena, to compare the outcomes of the DAMF method with a few other approaches. For instance, the PSNR and SSIM results of the PSMF, DBA, MDBUTMF, and NAFSM techniques for a Cameraman image with an SAP noise ratio of 30 % are 28.27/29.28/29.44/32.09 and 0.9044/0.9324/ 0.7740/0.9494, respectively, but the PSNR and SSIM results of the DAMF method are 36.83 and 0.9844, respectively. Finally, they demonstrated that DAMF could effectively eliminate SAP noise across all densities.

The authors in [15] suggested using an Adaptive Frequency Median Filter (AFMF) to eliminate the noise caused by salt and pepper. The adaptive condition of the Adaptive Median Filter (AMF) is also used by AFMF. Instead of using the AMF median, AFMF uses the frequency median to restore the grey values of the corrupted pixels. The frequency median emphasizes the uniqueness of grey values and can filter out noisy pixels when assessing a grey value of the window's center pixel. As a result, the frequency median yields a grey value that is more similar to the initial grey value than the AMF median. AFMF hence

performs better than AMF. The authors conducted experiments using the TESTIMAGES data set and a range of natural photographs from the MATLAB library. They also contrasted the denoising outcomes of AFMF with those of other cutting-edge denoising techniques. The outcomes demonstrated that AFMF denoises more successfully than alternative techniques. In [16] they proposed a modified decision-based unsymmetrical trimmed median filter algorithm for the restoration of color and grayscale images severely distorted by salt and pepper noise. When other pixel values, 0s, and 255s are present in the chosen window, the suggested approach replaces the noisy pixel with the trimmed median value; when all of the pixel values are 0s and 255s, the noise pixel is replaced with the mean value of all the elements in the chosen window. The suggested approach outperformed the Progressive Switched Median Filter (PSMF), Standard Median Filter (MF), Decision Based approach (DBA), and Modified Decision Based Algorithm (MDBA).

To address the problem of salt-and-pepper noise in digital images, [17] introduced a Nonlocal Adaptive Mean Filter (NAMF) in literature. Their approach is based on the adaptive window detector and nonlocal mean filtering strategy. This adaptive detector first detects the corrupted pixels by dynamically determining the neighborhood size based on local image properties and then a nonlocal adaptive mean filter is used to estimate the intensity of pixels with noise by referring to similar image regions. A number of experiments showed that NAMF yielded superior restoration results compared to several of the common impulse-noise removal techniques for a broad range of noise densities and retained the edges and fine structures in the image. It demonstrates the increasing interest in using adaptive noise detection as well as nonlocal image statistics for denoising. In [18] they proposed an algorithm for the salt and pepper noise removal, which introduces detail-aware noise removal to overcome the drawback of texture loss while using median-type filters for salt and pepper noise removal. They start by using median-type filter to remove the impulse noise, and then use an adaptive nonlocal bilateral filter to minimize the restoration error caused by the first filter. The authors showed that the adaptive bilateral filtering method combined with median based restoration method significantly enhances the preservation of edges and recovery of texture to improve the effectiveness of the noise suppression property, by statistically analyzing the residual errors. They emphasize the need for a combination of filtering methods to be implemented rather than just median filtering.

The authors in [19] discussed the theoretical behavior of adaptive morphological filters in the suppression of salt-and-pepper noise. Unlike only median statistics, the proposed framework is based on adaptive mathematical morphology, which helps to preserve the boundary of the objects while removing impulse noise. Under various contaminations, the authors studied the filtering behavior and showed that the adaptive morphological processing has better edge preservation than conventional nonlinear filters. Their study helps to understand how adaptive neighborhood operations could be deployed to achieve reduction of impulse noise while retaining the structural image information. One of the most crucial steps in digital image processing for restoring an image's structural integrity and visual quality is denoising. Conventional techniques frequently have drawbacks such as over-smoothing, processing complexity, and the inability to maintain important features, especially edges. To overcome these difficulties, [20] presents a hybrid denoising technique that combines Adaptive Median Filter (AMF) and Modified Decision-Based Median Filter (MDBMF). The MDBMF significantly reduces noise while maintaining edges by selectively recovering damaged pixels without impacting intact regions, whereas the AMF dynamically modifies window sizes to accurately detect noisy pixels. The objective assessments that support the subjective analysis demonstrate that the hybrid approach's performance far surpasses that of current state-of-the-art techniques. Nine benchmark photos from a standard and medical dataset—specifically, images of the liver and chest with varying noise densities ranging from 10 to 90 %—are used for the test. The hybrid technique outperforms the state-of-the-art approaches in terms of performance, as demonstrated by quantitative evaluations of PSNR, MSE, IEF, SSIM, FOM, and VIF. Compared to other techniques like BPDF, AT2FF, and SVMMF, the PSNR improved by up to 2.34 dB, the IEF improved by more than 20 %, and the MSE improved by up to 15 %. Improved structural similarity is confirmed by improvements in SSIM values of up to 0.07. Additionally, the hybrid approach's outstanding performance is demonstrated by the FOM and VIF measures, which both surpassed all other denoising algorithms assessed, reaching 0.68 and 0.61, respectively. To remove salt-and-pepper noise while preserving the significant image boundaries, [21] presented an edge-adaptive total variation model. It uses edge detection and median filtering in homogeneous regions, and adaptive total variation model in edge regions to separate the edge regions from the non-edge regions. The method effectively attenuates impulse noise while preserving sufficient edge details of the image, thanks to the use of these two restoration techniques. The experimental results well justified the advantages of combining adaptive variational model with classical median filter based denoising method in terms of edge preservation capability.

In order to eliminate high-density salt and pepper noise, [22] introduced Nearest Value Based Mean Filter (NVBMF) which is a two-stage adaptive filtering algorithm. In the first step, the intensity of each noisy pixel is substituted by the intensity of the closer intensity in its neighboring pixels or by the average of the equally close intensity if there are more than one. The restored image is further enhanced in the second stage by an adaptive mean filtering operation based on the estimated noise density. For the

evaluation, the proposed method was applied to a number of benchmark image datasets, and compared with many state-of-the-art impulse-noise removal methods such as the Adaptive Frequency Median Filter (AFMF), Adaptive Weight Mean Filter (AWMF), Adaptive Switch Weight Mean Filter (ASWMF), and several Riesz- and Cesàro-based filters. The results of the experiments showed that NVBMF provided competitive and, in certain cases, even better PSNR, SSIM and Image Enhancement Factor (IEF), especially under high noise density conditions which revealed a capable performance in preserving information content of the image while reducing the impulse noise. In [23] presented an optimization-based adaptive switching filter using Modified Decision Based Unsymmetric Trimmed Median Filter (MDBUTMF) along with an optimization technique Hybrid Bee Colony Optimization-Firefly Bat Intelligence Algorithm (Hyb-BCO-FBIA). The optimization program adapts the filtering parameters in an automatic manner for better restoration quality in the case of a strong salt-and-pepper noise contamination. Experimental comparisons revealed that the proposed optimization framework outperform two or more several conventional techniques of adaptive median filtering in terms of PSNR and MSE at the same time maintaining the edges of images more effectively. In this work the authors illustrate the ever-growing trend of using intelligent optimization techniques to improve traditional impulse-noise removal algorithms. [24] outlined a two-stage filtering scheme consisting of an impulse-noise detector based on the characteristics differences and an adaptive mean filter based on the direction filter. The identification of noisy pixels is performed in both stages: In the first stage gray-level extrema and local characteristic differences are used to identify noisy pixels; and in the second stage, only pixels identified as noisy are restored using directional averaging. It was found through experiments that the proposed detector was effective in reducing the false classifications of noise-free pixels and enhanced the edge preservation from conventional median based detectors, especially at high density of impulses. This work confirms that dividing noise detection from pixel restoration is an effective approach which is followed by many contemporary adaptive algorithms for the extraction of noise from images.

Based on these observations, a Simplified Enhanced Median Filter (SEmF) is introduced in this article to suppress the salt-and-pepper noise in grayscale and color images. The proposed method differs from the traditional median and average filters that process all pixels without distinction, by only taking care of the possible noisy pixels whose intensities are related to the extreme gray levels of impulse noise. The pixels which are not changed during the filtering process are noise free, thus eliminating the need for unnecessary changes to the image information. Furthermore, the proposed algorithm uses an adaptive multi-window approach that uses a selectable padding factor to enable the use of larger neighborhoods as noise density increases so that more contextual information is necessary to ensure the correct median estimation. This design makes it possible to keep the proposed filter having enough restoration performance in a large range of impulse-noise levels, while more effectively preserving image details than is provided by conventional filter designs.

This work makes the following major contributions. A simplified selective median filtering framework is first introduced, which is applied to only some noisy pixels in the image and leaves all of the other pixels unchanged. Second, an adaptive multi-window median estimation strategy is introduced to enhance the restoration result under low and high intensity salt-and-pepper noise. Thirdly, the proposed algorithm works for both grayscale & color images without any modifications to the structure, as it degrades color channels separately. Finally, experimental results have been compared by the Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) and filtering time in order to prove the effectiveness of the proposed method when compared with the conventional median and average filtering methods.

2. Methodology

Recently, various digital filters are used to remove SPN [1-5], a lot of these filters are based on standard Median Filter (SMF) And Standard Average Filter (SAF) A median based filtering technique has been one of the most successful techniques for fixed valued impulse noise filtering, as the median value is not as susceptible to extreme value outlier noise. The main drawback of the standard Median Filter (SMF) is that it performs the filtering process on all of the image pixels, both corrupted and uncorrupted ones. Many pixels that do not produce noise are thus replaced, losing the fine details of the image, making edges round off and degrading the texture of images, especially with repeated filtering or by using large filtering windows. The SMF and SAF have the following disadvantage [25-29]:

- SMF treats all pixels (clear pixels (CP) and possible noisy pixels (PNP), it uses a window and the pixel value will be replaced by the window median, while SAF also treats all pixels by using a window and replacing the pixel value with the average of this window. SMF and SAF will replace the image pixels even if the image is clear as shown in Fig. 1. it is shown in Fig. 1 SMF and SAV produced damaged images, while the presented filter kept the clear image without any changes;
- SMF and SAF filters do not differentiate between CPs and PNPs, and they are processed in the same manner, Fig. 2 shows how these filters deal with CPs and PNPs.

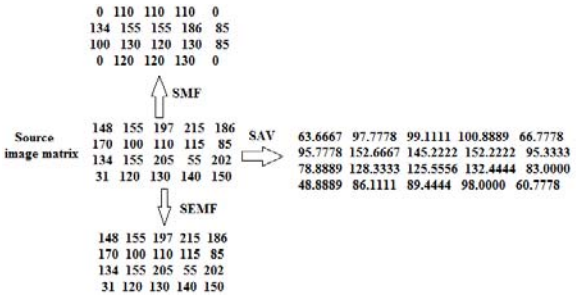


Fig. 1. Clear image filtering.

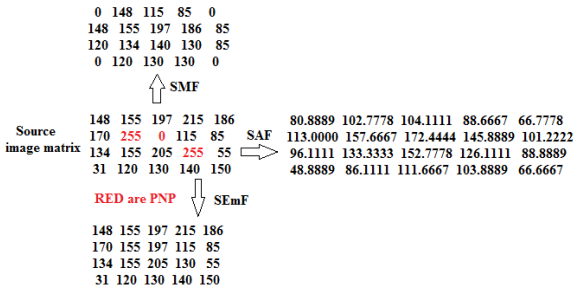


Fig. 2. Reducing SPN.

Proposed Simplified Enhanced Median Filter (SEmF) is based on the well-known selective processing philosophy, but with a simplified implementation. The pixels with gray level values at the extremes (0 or 255 in 8-bit images) are considered as potentially noisy pixels, while all the other pixels are treated as uncorrupted and hence are not modified. Because the detection strategy is based on pixel values being replaced by the minimum or maximum permissible gray levels, it is compatible with the statistical model of salt-and-pepper noise, and does not require unnecessary processing of the majority of pixels in the image. Limiting the restoration to noisy pixels that they consider "candidates" helps to avoid the loss of detail that is often found when the entire image is modified by the conventional median filter [9]. The proposed SEmF is different from the widely used median filter where a fixed neighborhood is applied throughout the image, and it introduces an adaptive multi-window restoration mechanism. The filtering process is user-controlled by the maximum neighborhood size (padding factor or PF). In this situation with relatively low impulse-noise density, a small neighborhood can be used, since most neighbouring pixels are not corrupted. However, under higher noise densities, a larger neighborhood will ensure that there will be enough valid pixels in the neighborhood to ensure that a good median estimate is calculated. Restoration has been demonstrated to be improved under severe impulse-noise contamination with similar adaptive window-expansion strategies that make it more likely to obtain representative neighborhood statistics [10].

The new multi-window strategy is a different implementation of classical adaptive median filters. In classical adaptive median filtering, the neighborhood size is increased until some prescribed statistical criteria are met, while in SEmF, the size of the neighborhood is evaluated in a predefined sequence of neighborhood size based on the padding factor. This design maintains the main benefit of adaptive neighborhood selection, and it is more easily implemented. Moreover, the maximum neighborhood size is given beforehand and thus the computational behavior of the proposed algorithm is known and can be tailored to the anticipated level of impulse noise and computational resources. Another special feature of the proposed algorithm is the retention of the image information. As only the noisy pixels are replaced, the original values of all the other pixels will be preserved. This is also important for maintaining edges and fine textures as these image structures are often lost by multiple corrections of already correct pixels. Many research works on switching and decision-based filters have shown that the visual quality and quantitative performance metrics like Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) are improved significantly if some of the pixels are not replaced unnecessarily.

The replacement value in SEmF is calculated with the median of the neighborhood that has been selected if the computed median value is not in the minimal and maximal gray-levels. With this rule, one impulse value may not be replaced by another impulse value, which can happen if there is a lot of noise pollution in the local neighborhood. The same ideas have been used in adaptive and decision-based impulse-noise filters to achieve better restored accuracy in high noise density, by choosing replacement values from the "reliable neighborhood statistics" instead of "corrupted observations. The algorithm proposed is also directly available for both gray-level and color images. When processing color pictures, each color channel is treated separately, and the same selective filtering strategy is applied to each channel. Channel-wise processing is commonly used in image restoration based on median filtering, since it ensures the simplicity of the algorithm used for each individual image channel while also effectively minimizing impulse noise in each component image. The proposed filter design allows the filter to function without any changes to the detection and restoration methods.

The complexity of SEmF is mainly determined by the padding factor and the neighborhood sizes used in the computation, from a computational point of view. If the impulse-noise density is low, the smallest neighborhood is enough, and the computational complexity is similar to that of conventional median filtering. Additional neighborhood evaluations provide more computational cost under higher noise densities, but yield better probability of obtaining a reliable median estimate which in turn results in better restoration quality. Furthermore, if the image is noise-free in some regions, many of the neighborhood evaluations will not be computed since only the candidate noisy pixels are evaluated, which reduces the processing time compared to algorithms that

process all pixels. This selective computation is an attractive compromise in terms of denoising and computational efficiency.

The presented SEmF will use the following input matrices:

- Noisy input image (IM), this matrix is a 2D matrix [26-30], which requires processing and which contains the PNP, the number of these pixels will depend on the noise ratio;
- Padded matrix [31-35], this matrix can be obtained from IM by padding it using a padding factor (PF), this factor will determine the maximum window which can be used to treat the image pixels, (PF = 1: 3 by 3 window, PF = 2: 5 by 5 window, PF = 3: 7 by 7 window and so on). In the presented SEmF the maximum value PF was selected to be equal 5. This will give the filter the ability to use window with size 11 by 11 down to window with size 3 by 3, Fig. 3 shows an example of various padding of a given image;
- Binary image matrix (BIM): this image is a 2D image with size equal the padded IM size, and it will contain zeros and ones, zero means that the associated pixel is a PNP, while 1 means that the associated pixel is CP, Fig. 4 show the BIM obtained from the source noisy image matrix;
- Pixel window, this window is a 2D matrix; it will be used to replace the associated pixel with the median of this window. The presented SEmF will use various windows depending on the selected value of the PF, the maximum window size used here in this filter is 11 by 11 with PF equal 5. The median calculation can be repeated by repeating windows calculation starting from highest PF to lowest PF, Fig. 5 shows some possible used pixel windows.

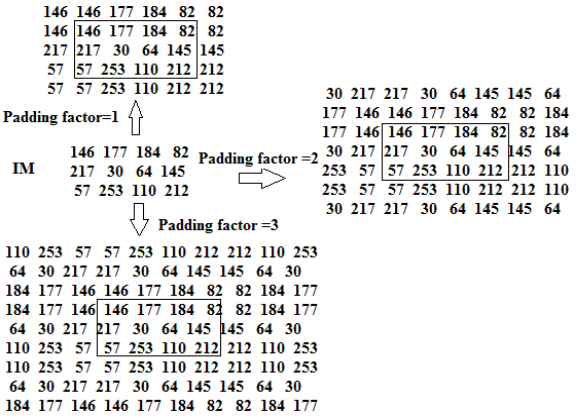


Fig. 3. Image matrix padding example.

The filter version can be selected by using the value of the selected padding factor (PF), version 1 will be selected when PF equal one, while the second version will be selected when PF is greater than 1. Version 1 will be used to remove SPN with low ratio, while version 2 will be used to remove SPN with higher noise ratio. The number of used windows to calculate the median value, which will be used to replace the value of PNP depends on the selected value of PF, Table 1 shows possible values of the PF and the associated windows to be used:

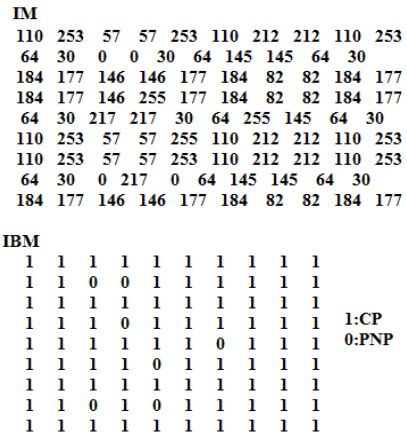


Fig. 4. Binary image matrix (BIM) example.

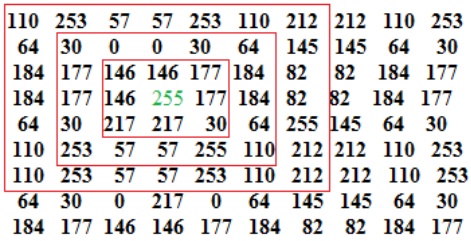


Fig. 5. Possible used pixel windows.

The presented SEmF will be implemented using one of the following algorithms:

- One window calculation algorithm (SEmF_1w);
- Multiple windows calculation algorithm (SEmF_mw).

Table 1. PF and the used windows.

PF	Algorithm version	Highest window	Used windows to calculate the median
1	SEmF_1w	W1 = 3 by 3	W1
2	SEmF_mw	W2 = 5 by 5	W2 down to W1
3	SEmF_mw	W3 = 7 by 7	W3 down to W1
4	SEmF_mw	W4 = 9 by 9	W4 down to W1
5	SEmF_mw	W5 = 11 by 11	W5 down to W1

Increasing the number of used windows will increase the time required to remove the SPN. The presented filter versions can be combined in one algorithm, and the selected value of the PF will determine the used version. The presented filter will be able to remove the noise from gray images, if the noisy image is a color one, then each 2D color-channel matrix must be treated separately.

SEmF Algorithm

Input: Noisy image a1, padding factor PF.
Output: Denoised image x.
Read a1.
Initialize PK
If the image is color call the filter 3 times, one time for each color 2D matrix.
Presented filter:
For each PF starting from the highest value down to 1 do:
Pad the image matrix.
Calculate the BIM.
For each pixel do
If the pixel is a CP then move to the next pixel.
Apply windows calculation:
For each window replace the PNP with the window median if the window median is greater than zero and less than 255.

Fig. 6 illustrates a MATLAB function, which can be used to implement the presented filter. The presented SEmF method is designed for the salt and pepper impulse noise model, in that, the corrupted pixels are commonly represented by the extreme intensity values of 0 and 255. Therefore, the proposed method identifies pixels with intensity values either 0 or 255 as corrupted pixels. However, pixels with values other than 0 or 255 assumed to be uncorrupted and are preserved.

This detection strategy provides a simple and computationally efficient technique for identifying the impulse noise pixels. Nevertheless, it depends on the assumption that the pixels with extreme values in the observed image are an impulse noise process generation. Indeed, in natural images, intensity values of 0 and 255 may also occur legitimately, as in dark region, saturated areas, strong image boundaries and dark regions. Based on that, such pixels may potentially be classified as corrupted and modified by SEmF. Based on that, the applicability of the current SEmF is primarily considered within the assumed salt and pepper impulse noise model. The possibility of false detection of naturally occurring 0- and 255-valued pixels is recognized as a limitation of the proposed approach.

```
function X=ZQF(A,PF)
%A: input noisy image; PF:padding factor
%X: Denoised image
A=double(A);
 %%For each PF staring from highest to lowest:
for p=PF: 1:1
%%image padding
 pA=padarray(A,[p p],'symmetric');
%%calculating BIM
 pB=(pA.~=0 & pA.~=255);
 %%get the image size
[m,n]=size(pB);
%%for each pixel do
 for i=1+p:m-p
  for j=1+p:n-p
%%if the pixel is PNP   (condition 1)
    if (pB(i,j)==0)
%%Apply windows calculations for the pixel
      for k=1:p
%%the window median must be greater than 0 and less than 255 to apply PNP replacement  (condition 2)
        if ((0<median(median(pA(i-k:i+k,j-k:j+k)))) & (median(median(pA(i-k:i+k,j-k:j+k)))<255) & (pA(i,j)==0| pA(i,j)==255))
          %%get the window
           Wk=pA(i-k:i+k,j-k:j+k);
          %%find the window median and replace the PNP with the median
           A(i-p,j-p)=median(median(Wk));
        break;
        end
      end
    end
  end
 end
end
 X=uint8(A);
end
```

Fig. 6. Presented filter MATLAB function.

The SEmF algorithm operates according to the following sequence:

1. Input and initialization: The padding factor (PF) and the noisy image α1 are given as an input. The filtering procedure is applied independently to the color image to each of the three color channels;
2. PF processing order: In this step, the padding factor is processed from the highest value down to PF = 1. Based on this, the filtering procedure starts with the largest predefined processing configuration and subsequently proceeds toward the smaller configurations;
3. Padding: For each PF, the image matrix is padded according to the padding factor required to construct the corresponding processing windows. The padding operation is performed before the BIM and window calculations for the PF;
4. BIM calculation: For each PF, the Binary Image Matrix (BIM) is calculated from the image being processed. This value used to identify the pixels corrupted or uncorrupted;
5. Pixel by pixel processing: The image is then scanned pixel by pixel. If the current pixel is identified as a clean pixel by the BIM, it is preserved and the algorithm move on to the next pixel directly;
6. Processing of noisy pixels: If the pixel is identified as a potential noisy pixel, the predefined processing windows are evaluated sequentially. For each window, the median value of the valid pixels within that window is calculated;
7. Replacement criterion: A PNP is replaced by the calculated window median only when the median satisfies;
8. 0 < median < 255;
9. Completing process.

3. Results and Discussion

3.1. Results

The performance of the presented SEmF will be evaluated using the following parameters:

- The quality of the denoised image Y reference to a clear image X, this quality can be measured by calculating the mean square error (MSE) (see equation 1) between the two images, and the peak signal to noise ratio (PSNR) (see equation 2) calculated between the two images, the lower MSE and the higher PSNR point to a denoised image with good quality.

$$\mathrm{MSE}(X,Y) := \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(x_{ij}-y_{ij}\right)^2, \tag{1}$$

$$\mathrm{PSNR}(X,Y) := 10\log\left(\frac{255^2}{\mathrm{MSE}(X,Y)}\right) \tag{2}$$

- Filtering time (FT) in seconds.

A color image (lena.png) was affected with SPN with low ratio (NR is low when NR is less or equal 5 %), the PF was varied, the MSE, PSNR and FT were calculated and Table 2 shows the obtained results:

Table 2. Performance of the presented SEmF when varying PF using a low ratio noise (NR = 1 %).

PF	MSE	PSNR	FT
1	1.1192	**47.6349**	0.4840
2	1.1030	**47.6983**	0.7570
3	1.2303	**47.2355**	1.2620
4	1.2444	**47.1860**	1.3820
5	1.2553	**47.1481**	2.0230

From Table 2 it is shown that selecting PF with 1 will give the best performance, so for low SPN ratio it is recommended to use the first version of the presented filter (using only one window).

The previous experiment was repeated by using a noisy image with SPN ratio = 50 %, and table 3 shows the obtained performance parameters results:

Table 3. Performance of the presented SEmF when varying PF using a high ratio noise (NR = 50 %).

PF	MSE	PSNR	FT
1	1940.4	**15.2536**	6.7800
2	136.3988	**26.7824**	8.6500
3	108.5730	**27.7743**	8.7440
4	105.9412	**27.8810**	9.0070
5	107.3995	**27.8217**	9.3710

From Table 3, it can be observed that PF = 4 provides the best overall performance under the high salt-and-pepper noise condition (NR = 50 %). PF = 4 achieves the lowest MSE (105.9412), the highest PSNR (27.8810 dB), and a shorter filtering time (9.0070) compared with PF = 5, which has an MSE of 107.3995, a PSNR of 27.8217 dB, and a filtering time of 9.3710. Therefore, PF = 4 is selected as the optimal configuration for the high-noise condition considered in Table2.

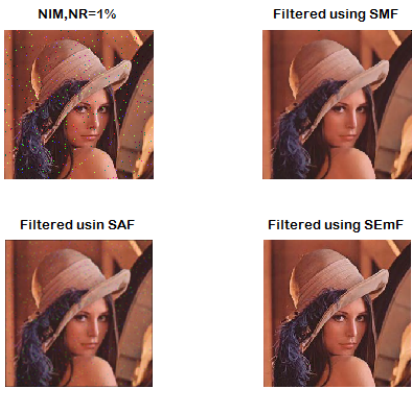


Fig. 7. Filtering NIM with low NR (NR = 1 %).



Fig. 8. Filtering NIM with high NR (NR = 30 %).

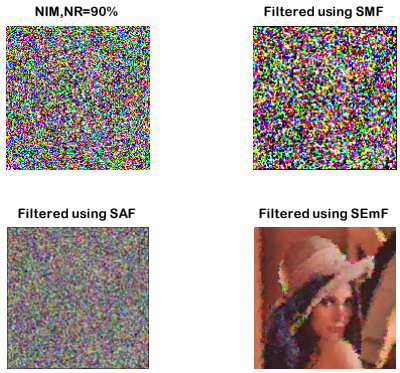


Fig. 9. Filtering NIM with high NR (NR = 90 %).

The image “lena.png” was affected by SPN noise using various NRs, the noisy image (NIM) was filtered using each of the filters: SMF, SAV and the presented SEmF, the quality parameters of the denoised images were calculated and Table 4 shows the obtained results.

Table 4. Quality results.

NR (%)	SAF		SMV		SEmF	
	MSE	PSNR	MSE	PSNR	MSE	PSNR
1 (low)	105.3041	**27.9072**	39.6728	**32.1467**	1.1561	**47.4936**
2 (low)	129.6134	**27.0019**	40.9721	**32.0067**	2.0895	**44.9285**
3 (low)	152.0068	**26.3128**	42.1428	**31.8830**	3.2743	**42.9818**
4 (low)	184.8466	**25.4618**	45.3452	**31.5654**	5.2566	**40.9245**
5 (low)	203.3728	**25.0469**	46.1149	**31.4924**	6.0648	**40.3038**
10 (high)	348.4787	**22.7076**	60.6322	**30.3042**	11.3460	**37.5816**
20 (high)	673.4300	**19.8492**	132.5255	**26.9083**	25.5604	**34.0567**
30 (high)	1057.1	**17.8899**	356.3827	**22.6101**	46.2195	**31.4836**
40 (high)	1521.3	**16.3085**	973.4563	**18.2473**	70.2549	**29.6635**
50 (high)	2066.4	**14.9787**	2124.6	**14.8588**	111.3595	**27.6625**
60 (high)	2673.8	**13.8604**	4226.8	**11.8692**	159.5115	**26.1017**
70 (high)	3343.6	**12.8897**	7148.2	**9.5882**	228.1709	**24.5473**
80 (high)	4099.5	**12.0001**	11051	**7.6966**	331.3728	**22.9299**
90 (high)	4803.9	**11.3152**	15239	**6.3010**	487.4655	**21.2533**

From Table 4 it is shown that the presented SEmF enhanced the quality of the denoised image when this image is affected with SPN with low or high NRs, SEmF gave better results for the values of MSE and PSNR as shown in Fig. 10.

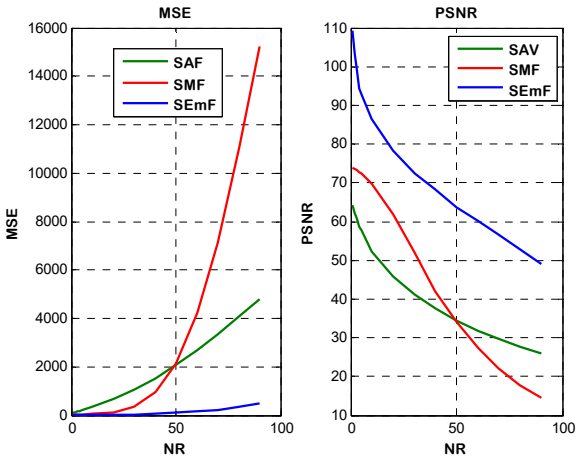


Fig. 10. Quality parameters comparisons.

3.2. Discussion

The experimental results show that the proposed Simplified Enhanced Median Filter (SEmF) can always restore better quality than the Standard Median Filter (SMF) and the Standard Average Filter (SAF) for both low- and high-density salt and pepper noise. The quantitative evaluations of Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR) and filtering time (FT) are widely used in image denoising algorithms and were used to assess the performance. In the first experiment, the effect of the padding factor (PF) on restoration performance was studied. The results show that when the ratio of noise to impulse (NR = 1 %) is low, the single-window version of SEmF (PF = 1) yields the minimum MSE, maximum PSNR and the fastest execution time for the images corrupted with low NR value. This is in line to the properties of low density impulse noise, which has the following properties: Most of the neighboring pixels are uncorrupted and only a small local neighborhood is enough to estimate the median value reliably. Under these conditions, increasing the size of the window adds needless calculations but does not boost restoration accuracy. The same conclusion can be drawn for adaptive median filtering, in which the use of larger neighborhoods is beneficial only if there is a large number of corrupted pixels in the neighborhood.

A different behavior is observed for heavily corrupted images. The results of the experiments demonstrate that the quality of the restoration is significantly enhanced with the multi-window implementation when the impulse-noise ratio becomes 50 %, because the MSE is much lower and the PSNR is much higher compared with the single-window implementation. The larger neighborhood makes it more likely that there will be enough uncontaminated (or uncorrupted) pixels in it to obtain a more representative median value, which will help to restore heavily contaminated areas. In high noise conditions the gain in restoration quality is greater than the time it takes to filter with larger window sizes. This relationship between the computational complexity and the accuracy of the restoration has also been noted in the adaptive and switching median filters reported in the literature. The proposed selective restoration strategy is further highlighted by the comparison with SMF and SAF. The SAF takes linear averaging in the local neighborhood and the conventional SMF replaces all image pixels, including corrupted ones. Both methods, therefore, change an excessive number of correct pixels, resulting in blurring of the edges and loss of fine image details. SEmF restores only the pixels that are detected as INC (impulse-noise candidates) and leaves the other pixels unchanged. This selective processing greatly minimizes the unwanted changes of valid image data and helps to improve PSNR and decrease MSE as seen in the results. This behavior is in keeping with the design philosophy of decision based and switching filters, whose performance it has consistently surpassed by decoupling the noise detection and restoration processes, as is conventional median filtering. It is worth noting that these results were obtained using MATLAB R2024b, using machine with (CPU model, RAM, GPU if applicable, e.g., Intel i7-8650U, 24 GB RAM, RTX 3060Ti 12 GB), image data characteristics 20×200×1, unit 8, it takes as an average overall running time running time 50 ms.

Table 4 shows that SEmF achieves lower MSEs and higher PSNRs than the other two cases (SMF and SAF) over all the noise densities studied. The performance appears to be more noticeable as the impulse-noise density becomes stronger, suggesting the proposed adaptive multi-window approach is more effective under severe degradation. The visual comparisons shown in Fig. 7-10 are also in line with the quantitative analysis results, as they demonstrate the improved preservation of edges, image structures and fine details following filtering with SEmF in comparison to the conventional methods. The other thing to note is computational efficiency. The multi-window implementation is more complex in the case of higher padding factors, but by not applying restoration techniques to pixels already classified as uncorrupted, the algorithm is able to avoid unnecessary computations. This means that the number of computations is significantly less than it would be if each of the image's pixels were processed many times with different neighborhood sizes. The proposed algorithm is an effective compromise between the quality of image restoration and the time taken to execute the algorithm and is thus suitable for practical applications of image processing, where the accuracy and execution time are important. It is worth noting that the limitation of the proposed SEmF is its reliance on the identification of 0 and 255 valued pixels as candidate impulse-corrupted pixels. Even this assumption is good for the salt pepper noise models used in the given experiments, however, natural values of image pixels may also have these extreme intensity values. As a result, a genuine saturated white or black pixel may be incorrectly classified as noisy and then modified. This limitation is more significant in images containing bright

regions, high contrast structures, saturated areas or large dark areas. Consequently, the current structure does not explicitly distinguish between a naturally occurring extreme valued pixels and the corrupted ones.

Recent WMF [35] and MWMF methods [36] share some concepts with SEmF, especially selective processing, padding operation and neighborhood-based filtering. These common concepts are therefore not claimed as individual novel contributions of SEmF. Instead, the novelty of SEmF lies in the specific combination and sequence of its processing steps. However, the novelty of SEmF lies in the specific combination and sequence of its processing steps.

In summary, the experimental results demonstrate that the proposed SEmF can efficiently achieve a selective noisy-pixel detection, as well as an adaptive median estimation. The algorithm avoids discarding unchanged pixels, reduces the loss of details and has better restoration performance than the traditional median and average filters, especially in the presence of moderate and high salt-and-pepper noise. The results are in line with previous research that demonstrated that selective switching strategies and adaptive neighborhood processing methods can offer substantial benefits when compared to indiscriminate filtering.

4. Conclusion

A simple and efficient SEmF was presented, the used filter was simple and easy to implement, and it can be efficiently used to remove salt and pepper noises with low and high noise ratios. The presented method was capable of denoising grayscale and color images, changing the image did not require any changes in the presented filter algorithm. The presented filter can be implemented in two versions: the first version used one window to calculate the median and it was recommended to be used to eliminate salt and pepper noise with low noise ratio, while the second version used multiple windows to calculate the median and it was recommended to be used to eliminate salt and pepper noise with high noise ratio. The presented filter processed only the possible noisy pixels keeping the clear pixels without changing. The possible noisy pixels were filtered base on the pixel value and the calculated window median, this pixel was replaced when the pixel value was equal 0 or 255 and the calculated window median was greater than 0 and less than 255.The presented SEmF enhance the performance of standard median and standard average filters by improving the quality of the denoised images, this filter optimized the values of MSE and PSNR for the denoised image. Future work will investigate contextual and neighborhood-based decision criteria that can distinguish genuine extreme valued image structure from impulse corrupted pixels, thereby reducing the possibility of false detection while preserving important image details.

References

[1]. R. H. Chan, C.-W. Ho, M. Nikolova, Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization, *IEEE Transactions on Image Processing*, Vol. 14, Issue 10, 2005, pp. 1479-1485.

[2]. T. Chen, H. R. Wu, Adaptive impulse detection using center-weighted median filters, *IEEE Signal Processing Letters*, Vol. 8, Issue 1, 2001, pp. 1-3.

[3]. T. Chen, K.-K. Ma, L.-H. Chen, Tri-state median filter for image denoising, *IEEE Transactions on Image Processing*, Vol. 8, Issue 12, 1999, pp. 1834-1838.

[4]. S. Enginoğlu, U. Erkan, S. Memiş, Pixel similarity-based adaptive Riesz mean filter for salt-and-pepper noise removal, *Multimedia Tools and Applications*, Vol. 78, Issue 24, 2019, pp. 35401-35418.

[5]. S. Enginoğlu, U. Erkan, S. Memiş, Adaptive Cesàro mean filter for salt-and-pepper noise removal, *El-Cezeri Journal of Science and Engineering*, Vol. 7, Issue 1, 2020, pp. 304-314.

[6]. U. Erkan, L. Gökrem, S. Enginoğlu, Different applied median filter in salt and pepper noise, *Computer and Electrical Engineering*, Vol. 70, 2018, pp. 789-798.

[7]. U. Erkan, D. N. H. Thanh, S. Enginoğlu, S. Memiş, Improved adaptive weighted mean filter for salt-and-pepper noise removal, in *Proceedings of the International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, 2020, pp. 1-5.

[8]. R. C. Gonzalez, R. E. Woods, Digital Image Processing, *Pearson*, New York, 2018.

[9]. H. Hwang, R. A. Haddad, Adaptive median filters: New algorithms and results, *IEEE Transactions on Image Processing*, Vol. 4, Issue 4, 1995, pp. 499-502.

[10]. H.-L. Eng, K.-K. Ma, Noise adaptive soft-switching median filter, *IEEE Transactions on Image Processing*, Vol. 10, Issue 2, 2001, pp. 242-251.

[11]. G. Pok, J.-C. Liu, A. S. Nair, Selective removal of impulse noise based on homogeneity level information, *IEEE Transactions on Image Processing*, Vol. 12, Issue 1, 2003, pp. 85-92.

[12]. K. S. Srinivasan, D. Ebenezer, A new fast and efficient decision-based algorithm for removal of high-density impulse noises, *IEEE Signal Processing Letters*, Vol. 14, Issue 3, 2007, pp. 189-192.

[13]. U. Erkan, L. Gökrem, A new method based on pixel density in salt and pepper noise removal, *Turkish Journal of Electrical Engineering & Computer Sciences*, Vol. 26, Issue 1, 2018, pp. 162-171.

[14]. U. Erkan, S. Enginoğlu, D. N. H. Thanh, L. M. Hieu, Adaptive frequency median filter for the salt-and-pepper denoising problem, *IET Image Processing*, Vol. 14, Issue 7, 2020, pp. 1291-1302.

[15]. S. Esakkirajan, T. Veerakumar, A. N. Subramanyam, C. H. P. Chand, Removal of high-density salt and pepper noise through modified decision-based unsymmetric trimmed median filter, *IEEE Signal Processing Letters*, Vol. 18, Issue 5, 2011, pp. 287-290.

[16]. H. Zhang, Y. Zhu, H. Zheng, NAMF: A nonlocal adaptive mean filter for removal of salt-and-pepper noise, *Mathematical Problems in Engineering*, Vol. 2021, 2021, 4127679.

[17]. Y. Liu, X. Zhang, et al., Salt and pepper noise removal method based on a detail-aware filter, *Symmetry*, Vol. 13, Issue 3, 2021, 515.

[18]. M. Mares-Javier, C. Guillén-Galván, R. Lemuz-López, J. Debayle, On the properties of some adaptive morphological filters for salt and pepper noise removal, *Image Analysis & Stereology*, Vol. 40, Issue 1, 2021, pp. 29-38.

[19]. F. Ullah, K. Kumar, T. Rahim, et al., A new hybrid image denoising algorithm using adaptive and modified decision-based filters for enhanced image quality, *Scientific Reports*, Vol. 15, Issue 1, 2025, 8971.

[20]. Y. Jiang, H. Wang, Y. Cai, B. Fu, Salt and pepper noise removal method based on the edge-adaptive total variation model, *Frontiers in Applied Mathematics and Statistics*, Vol. 8, 2022, 918357.

[21]. M. Turan, A new approach for SPN removal: Nearest value based mean filter, *PeerJ Computer Science*, Vol. 8, 2022, e1160.

[22]. S. Mohan, B. Paulchamy, Removal of salt and pepper noise using adaptive switching modified decision-based unsymmetric trimmed median filter optimized with Hyb-BCO-FBIA, *Automatika*, Vol. 65, Issue 3, 2024, pp. 852-865.

[23]. H. Ma, Y. Nie, A two-stage filter for removing salt-and-pepper noise using noise detector based on characteristic difference parameter and adaptive directional mean filter, *PLOS ONE*, Vol. 13, Issue 10, 2018, e0205736.

[24]. M. O. Al-Dwairi, A. Y. Hendi, M. S. Soliman, Z. A. A. Alqadi, A new method for voice signal features creation, *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 9, Issue 5, 2018, pp. 4092-4098.

[25]. A. A. Moustafa, Z. A. Alqadi, A practical approach of selecting the edge detector parameters to achieve a good edge map of the gray image, *Journal of Computer Science*, Vol. 5, Issue 5, 2009, pp. 355-362.

[26]. A. Al-Rawashdeh, Z. Al-Qadi, Using wave equation to extract digital signal features, *Engineering, Technology & Applied Science Research*, Vol. 8, Issue 4, 2018, pp. 1356-1359.

[27]. Z. Alqadi, B. Zahran, Q. Jaber, B. Ayyoub, et al., Enhancing the capacity of LSB method by introducing LSB2Z method, *International Journal of Computer Science and Mobile Computing*, Vol. 8, Issue 3, 2019, pp. 76-90.

[28]. S. Sivankutty, et al., Supercritical angle fluorescence for enhanced axial sectioning in STED microscopy, *Methods*, Vol. 174, 2020, pp. 20-26.

[29]. V. J. Rani, K. K. Thanammal, An extended review on impulse noise reduction for corrupted images, *Turkish Online Journal of Qualitative Inquiry*, Vol. 12, Issue 7, 2021, pp. 3053-3064.

[30]. D. Sauter, L. Parson, Spatial filtering for speckle reduction, contrast enhancement, and texture analysis of GLORIA images, *IEEE Journal of Oceanic Engineering*, Vol. 19, Issue 4, 1994, pp. 563-576.

[31]. I. A. Pardosi, A. A. Lubis, Grayscale image quality analysis result of noises reduction using adaptive fuzzy filter (AFF) and spatial median filter (SMF) against image depth variations, *Journal of Physics: Conference Series*, Vol. 1361, Issue 1, 2019, 012028.

[32]. T. Zhou, Y. Wang, A survey on domain adaptive object detection, *The Computer Journal*, Vol. 69, Issue 2, 2026, pp. 357-375.

[33]. Z. A. A. Alqadi, M. K. A. Zalata, G. M. Qaryouti, Comparative analysis of color image steganography, *JCSMC*, Vol. 5, Issue 11, 2016, pp. 37-43.

[34]. Z. A. Alqadi, M. Aqel, I. M. M. El Emary, Performance analysis and evaluation of parallel matrix multiplication algorithms, *World Applied Sciences Journal*, Vol. 5, Issue 2, 2008, pp. 211-214.

[35]. Q. M. Jaber, Z. A. A. Alqadi, H. F. Taha, Multiple_Window Median Filter (MWMF) for efficient removing salt and pepper noise, *International Journal of Computer Science and Mobile Computing*, Vol. 11, Issue 6, 2022, pp. 169-189.

[36]. H. G. Zaini, Window Median Filter (WMF) to eliminate salt and pepper noise from digital color image, *International Journal of Computer Science and Mobile Computing*, Vol. 10, Issue 9, 2021, pp. 23-37.

Published by International Frequency Sensor Association (IFSA) Publishing, S. L., 2026 (http://www.sensorsportal.com).



Sensors & Transducers, Vol. 274, Issue 3, October 2026, pp. 30-41

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2026
http://www.sensorsportal.com

A Polarized Image Multiplexor for In-Line Frame Calculation in a Holographic Parallel Video Processor for 3D Dynamic Display

Branimir IVANOV, * Ginka IVANOVA and Deyan DIMOV
Institute of Optical Materials and Technologies, Laboratory 'Holography',
109, Acad. G. Bontchev Str., 1113 Sofia, Bulgaria
Tel.: + 33 617446640
E-mail: givanova@iomt.bas.bg

Received: 31 Aug. 2026 /Revised: 26 Oct. 2026 /Accepted: 29 Oct. 2026 /Published:30 Oct. 2026

Abstract: This article extends an existing eight-channel, non-polarized holographic optical trigger intended for real-time three-dimensional display by adding a polarizing beam-splitter layer. At each of the trigger's eight branches, a polarizing beam-splitter cube separates the arriving light into orthogonal S and P components, which two parallel encoder stages then address independently, doubling the trigger into 16 polarization-addressable channels. A fixed convention assigns the S- and P-branch channels to opposite depth-plane directions of a 49-plane display, with a polychromatic red–green–blue reference channel. The result is architectural economy: this construction reaches the same three-trigger-equivalent target achieved elsewhere with three physical triggers, using only two and doubling one of them through polarization cloning, the functional duplication of its branches, rather than physical duplication. The article also presents a layered numerical model of the recording medium, its recording energy budget, an optically recorded zoom transformation, and polarimetric measurements of the polarization states of a physically constructed trigger and its polarization-cloned counterpart. The emphasis throughout is methodological, prioritizing the cloning technique over further channel-count scaling.

Keywords: Polarization multiplexing, Holographic optical trigger, Volume holography, Polarizing beam-splitter cube, Silver-halide recording medium, Three-dimensional holographic display.

1. Introduction

Real-time three-dimensional holographic display places a substantial computational burden on any system that must generate a distinct wavefront for every depth plane, color channel, and frame of a moving scene. Digital computer-generated holography addresses this burden through dedicated processing hardware, algorithmic acceleration, or trained neural networks, and recent work has demonstrated real progress along each of these routes. One single-chip field-programmable gate array processor computes full-color, ultra-high-definition holograms in real time, expanding the viewing angle through a dedicated steering-backlight unit, not through the native pixel count of the spatial light modulator (SLM) alone [1]. A single-step Fresnel calculation based on the split-Lohmann method computes an entire multi-depth scene in constant time regardless of the number of depth planes represented; conventional layer-based computation does not share this property, since its calculation time grows in direct proportion to depth-layer count [2]. Even so, both results depend on continuous, per-frame computation that dedicated processing hardware carries out for as long as the display operates. The display's space–bandwidth product imposes a direct trade-off between image size and viewing angle: with a fixed pixel count, increasing one necessarily decreases the other [1]. A recorded or multiplexed holographic

medium is not constrained in the same way, since the structure of the recording medium itself, rather than a discrete, fixed pixel array, sets its spatial resolution [3].

The present work continues a line of research into an all-optical holographic video graphics processing unit, referred to throughout as the HVGPU, in which channel and depth-plane multiplication are properties of the recording and of a small number of passive optical elements, established once during fabrication rather than recomputed for every frame [4, 5].

In-line frame calculation in the title refers to this multiplication of each input frame into depth-plane frames while the frame passes through the trigger, and not to a measured calculation speed.

1.1. Multiplexing as a General Principle in Holography

Multiplying the number of independently addressable channels obtainable from a single optical element, by exploiting some physical property of the illuminating light rather than by duplicating hardware, is an established and currently active strategy across several branches of optics. Two general routes toward that end are considered below: polarization specifically, since it forms the basis of the present architecture, and other physical degrees of freedom pursued for the same purpose elsewhere in the literature.

Polarization has itself been the subject of formal multiplexing frameworks in holography. A recently proposed orthogonal matrix of polarization combinations enables four or more independently addressable holographic channels recorded at a single position within a polarization-sensitive material, selectively reconstructed according to the polarization combination of the illuminating reference wave [6]. This particular doubling mechanism has already been explored in our earlier work: an infrared holographic optical element reported over a decade earlier used the same principle at smaller scale, separating two wavelength channels into four by their S- and P-polarized components [7]. Nor is orthogonal polarization multiplexing confined to linear states; a recent all-dielectric metasurface superimposes two independent holographic phase patterns addressed by orthogonal circular polarization, generating different structured-light outputs depending on whether the illumination is right- or left-circularly polarized, with an extension to spatially multiplexed channels at independently controlled focal planes [8]. Polarizing beam-splitter cubes performing a similar selective folding also appear in current commercial-grade near-eye optics, relaying holographic content through a compact eyepiece path [9]. Polarization can also change dynamically rather than remaining fixed at fabrication: fast-response liquid-crystal phase shifters, paired with polarization gratings, steer a single SLM's holographic output between several discrete positions in rapid temporal sequence, exploiting human visual persistence to create the impression of an enlarged display area [10].

More elaborate schemes push polarization multiplexing further still. Because the number of independently addressable polarization channels achievable from a single metasurface layer is fundamentally limited by the degrees of freedom of the Jones matrix – a ceiling of three channels under conventional orthogonal-polarization design [11, 12] – recent work has turned to deliberately engineered noise in the metasurface response to exceed that limit, reaching 55 distinguishable channels when combined with an additional axial focal-plane offset and neural-network optimization [11]. Researchers have since extended the same noise-engineered approach to combine polarization with position, wavelength, and orbital angular momentum, individually and simultaneously, within a single device, each additional physical degree of freedom multiplying the addressable channel count further [12]. The present architecture applies polarization as an orthogonal, physically realizable selecting variable through a dedicated polarizing beam-splitter-cube layer added ahead of an existing non-polarized channel set, rather than through engineered metasurface response.

1.2. Other Physical Degrees of Freedom Used for Holographic Multiplexing

Polarization is one instance of a broader family of techniques that recover multiple independent optical functions from a single element by varying some controllable property of the illuminating beam. Orbital angular momentum offers another such property. Recent work encodes independent image channels into distinct orbital-angular-momentum modes for large-field-of-view multiview display [13]. A further extension introduces a synthetic, mathematically defined operator dimension beyond orbital angular momentum's own physical mode index, addressing a broader constraint: the intrinsic physical degrees of freedom conventionally exploited for holographic multiplexing were approaching the limits of further scaling. Subdividing a single orbital-angular-momentum channel by an engineered coordinate-transformation operator yielded a ninefold capacity increase over conventional orbital-angular-momentum holography [14].

Angle of incidence under Bragg selectivity is a third such property: a multiplexed holographic optical element can hold two entirely different optical functions in superposition; the angle of a variable-focus illumination lens, rather than polarization, selects between them [15]. Beyond these, researchers have developed a wider taxonomy of multiplexing techniques – including shift, peristrophic, phase-coded, and correlation-based schemes – principally for holographic data storage rather than display; the present article does not develop these further [16].

1.3. Scope and Relation to Prior Work

An earlier study [4] describes and characterizes the eight-channel, non-polarized beam-splitter unit that anchors the present architecture: a splitter delivering eight physical output branches, a downstream encoder of half-wave plate rotators, and a beam combiner, with each branch carrying a distinct linear polarization angle generated from a single input beam. The present article, a continuation of the OPAL' 2026 design and fabrication study [5], takes that unit as a fixed starting point and does not re-derive or re-characterize it.

Relative to our previously reported eight-branch holographic trigger [4, 5], the present work introduces one specific architectural extension: each of the eight inherited branches is separated into orthogonal S and P components by a dedicated polarizing beam-splitter cube and processed by two parallel encoder stages, which produces 16 simultaneously accessible polarization-resolved optical paths from the original eight (Section 3). The contribution introduced in this study is therefore the polarization-branch duplication architecture based on the polarizing beam-splitter and its experimental characterization, and not the underlying eight-channel trigger, polarization multiplexing in general, or the previously reported 49-plane display concept. The separation into S and P components was already applied to a two-wavelength input in an earlier four-channel element [7]. The extension beyond the prior OPAL' 2026 conference publications [4, 5] consists of two new experimental results: the exposure dependence of the diffraction efficiency of the three U04 color holograms (Section 5) and the polarimetric measurements of the polarization states of the physical and polarization-cloned branches (Section 6). It also includes three design results: the architecture and channel accounting presented in Section 3, the recorded zoom transformation described in Section 4, and the layered numerical model with the recording energy budget presented in Section 5. The treatment of the recording material is limited to those properties required to interpret these results.

2. Theoretical Background

2.1. Volume Holography and Bragg Selectivity

A hologram is classified as thin or as thick (volume) according to whether the reconstruction beam interacts with a single recorded interference plane or with many recorded planes distributed through the thickness of the recording layer. The Klein–Cook parameter formalizes this distinction:

$$Q = 2\pi\lambda_0 d/(n_0\Lambda^2), \quad (1)$$

where λ_0 is the illumination wavelength in air, d is the recording-layer thickness, n_0 is the medium's average refractive index, and Λ is the interference-fringe period. Values of Q greater than unity place a grating in the volume, or Bragg-diffracting, regime, requiring a coupled-wave rather than a simple grating-diffraction analysis [16]. In this regime, Kogelnik's coupled-wave theory governs reconstruction, relating diffraction efficiency to the recording layer's thickness and its induced refractive-index or absorption modulation. This theory remains the standard theoretical treatment of thick-hologram diffraction efficiency across the volume-holography literature [17]. The recording medium used in the present work, described in Section 5, operates in this volume regime.

Depth distortion in a thick hologram does not occur when the geometry of the illumination beam at reconstruction exactly matches the geometry of the reference beam used at recording; this geometric correspondence is the governing condition for distortion-free reconstruction [17]. Distortion specifically arises when the reconstruction illumination is not aligned with the optical axis of the original reference beam; when that alignment is exact, no such distortion occurs. This criterion underlies the zoom transformation described in Section 4, achieved by varying the convergence or divergence angle of the illuminating beam rather than by digitally resampling an already-encoded pattern.

2.2. Wavefront Cloning

A hologram that records the wavefront produced by a physical optical element, whether a lens, a prism, or a more complex objective, reproduces the transformation that element applies to the illuminating wave, within the angular and spectral selectivity of the recording, for a subsequent wave passing through the recorded copy [18, 19]. At reconstruction in the recording geometry, the reconstructed wavefront approximates the wavefront that the original element would have produced, within the angular and spectral selectivity and the resolution of the recording. It follows that a SLM, displaying a computed interference pattern rather than a fixed recorded one, can synthesize the equivalent of a physical optical element without that element existing in bulk form. Switching the displayed pattern between frames is, by the same logic, functionally equivalent to physically exchanging optical elements at frame rate.

Both analog and digital contexts demonstrate this same principle: an interference pattern, whether recorded on a fixed medium or generated dynamically by a SLM, can carry an intended optical transformation while preserving the phase information of the reconstructed wavefront. Recent digital work, for instance, encodes a comparable rescaling transformation directly into a computed hologram's phase structure while explicitly tracking its effect on reconstruction depth [20].

The zoom transformation described in Section 4 applies the same underlying principle by physical

rather than computational means. The recording medium described in Section 5 records a real optical system's convergence and divergence behavior as an interference pattern, so that reconstructing that recorded pattern provides the convergence and divergence behavior of the zoom optics, for the recorded geometry, without requiring them to remain part of the operating system.

As in elemental-image-array approaches to light-field imaging, many elementary recordings construct a multichannel hologram of the kind described in Section 3, each differing from the others by its parallax angle relative to the recorded object; each elementary hologram, in effect, views the object from its own distinct angle [21].

3. Architecture of the Polarized HVGPU

3.1. Inherited Non-Polarized Splitter Architecture

The present architecture takes as its starting point a non-polarized, eight-channel beam-splitter unit and its associated encoder and combiner stages, described in full in our previous papers [4, 5]. The present article requires only two properties from that inherited architecture. The splitter delivers eight physical output branches from a single input, preserving the input beam's phase, polarization, and wavelength unaltered across all eight branches.

A downstream encoder stage then assigns each branch a distinct, though not necessarily orthogonal, linear polarization angle by setting a half-wave plate so that the polarization angle takes one of eight values from 0° to 70° in 10° steps. A half-wave plate whose fast axis is at an angle α to the input polarization rotates linear polarization by 2α, so these settings correspond to plate orientations from 0° to 35° in 5° steps.

In this base architecture, angle rather than orthogonality separates adjacent channels; two states are distinct as long as they differ, whether or not they are perpendicular to one another. The addition described in the remainder of this section builds on that base without altering it and treats the base splitter and its eight branches as a given configuration.

Throughout, a branch is one of the eight outputs of the splitter; a channel is one physical optical path through the trigger, so that each branch gives an S channel and a P channel; a color component is one wavelength carried by a channel; a virtual color channel is one color component in one channel; a depth plane is one image plane of the display; and a trigger is the complete multichannel device.

3.2. The Polarizing Beam-Splitter Layer

A polarizing beam-splitter cube is introduced ahead of the encoder stage at each of the eight branches produced by the inherited splitter. Each cube divides the light arriving at its branch into two orthogonally polarized components, S and P, which propagate onward as two physically separate beams. Applied across all eight branches, this yields 16 output beams: eight carrying the S-polarized component of their respective branch (S_1 through S_8) and eight carrying the P-polarized component (P_1 through P_8). S and P are, by construction, mutually orthogonal. This orthogonality is a stronger property than the mere distinctness the inherited eight-channel angle encoding relies upon, and the present architecture applies it deliberately, as an additional, independent multiplication of the channel count rather than as a replacement for the angle-based distinctness already present in the inherited base. S and P are not the only pair of states with this property; any pair of mutually orthogonal polarization states could in principle serve the same role, circular polarization included, but S and P are the simplest and most directly obtainable pair from a single passive cube.

Of the two orthogonal outputs at each branch, we arbitrarily designate one to continue propagating in the same direction as the branch's original beam – here, the S-polarized output. The P-polarized output, newly generated by the cube, travels onward to a second, physically separate encoder stage.

In this article, “polarization cloning” is an engineering term for the functional duplication of a processing branch: the cube decomposes the light of each branch into orthogonal S and P components, and each component is processed by its own encoder, so that the P branch carries the same eight channels of image information as the S branch in the orthogonal polarization state, which is the sense in which it performs the function of a second, separately built trigger. No claim of equal diffraction efficiency or power is made. The term does not refer to quantum cloning or to cloning of an arbitrary polarization state, which a passive polarizing beam-splitter cube does not perform.

3.3. Dual Encoder Architecture: S-Branch and P-Branch

Two encoder stages operating in parallel process the 16 channels the polarizing beam-splitter layer produces, each stage a direct copy of the encoder already established in the inherited base architecture, each built from eight half-wave plates. The first encoder receives the eight S-polarized channels and applies the base architecture’s own angle assignment unmodified: polarization angles from 0° to 70° in 10° steps, one angle per channel. The second encoder receives the eight P-polarized channels and otherwise matches the first exactly; only its polarization angles differ, offset by exactly 90°; its plates are mounted at the S-branch orientations plus 90°. This offset reflects the 90° separation between S and P at the point where the two branches diverge; the second encoder is

otherwise an unmodified reproduction of the first rather than an independently designed unit.

The encoder stage performs continuous checks on the calculated polarization angle assigned to each channel during operation. Polarization angles are normalized modulo 180°, since orientations of linear polarization that differ by 180° describe the same polarization axis, and duplicate orientations are rejected to preserve channel distinguishability. The encoder checks and corrects this condition during operation, in the same manner the base architecture's own angle-assignment function already requires; extending this checking to 16 channels across two encoders, rather than eight channels across one, adds negligible additional computational burden.

3.4. Mapping Polarization Branches to Depth Planes

A fixed convention assigns the 16 channels the S- and P-branch encoders produce to the display's depth planes. The eight S-branch channels, S_1 through S_8, go to the 24 depth planes on one side of the reference plane (8 × 3 colors), and the eight P-branch channels, P_1 through P_8, go to the 24 depth planes on the other, each channel carrying all three color components.

This mapping is a deliberate representational choice adopted for clarity of exposition, not a physical necessity imposed by the polarization mechanism itself; S and P carry no inherent association with depth sign, and the assignment could in principle be made differently. The convention is nonetheless useful, since it gives the architecture's two most naturally distinguished quantities, the sign of depth displacement and the S/P polarization axis, a single, consistent correspondence throughout the remainder of this article.

Fig. 1 summarizes this dual-branch architecture. The inherited splitter's eight outputs each feed a polarizing beam-splitter cube, producing the S- and P-polarized branches described above. The eight channels in each branch pass through their respective half-wave plate sets and encoders – identical in structure, offset by 90° in angle between the two – before recombining through eight-input beam combiners at the combiner stage, reproducing the base architecture's splitter-to-combiner path twice over rather than once.

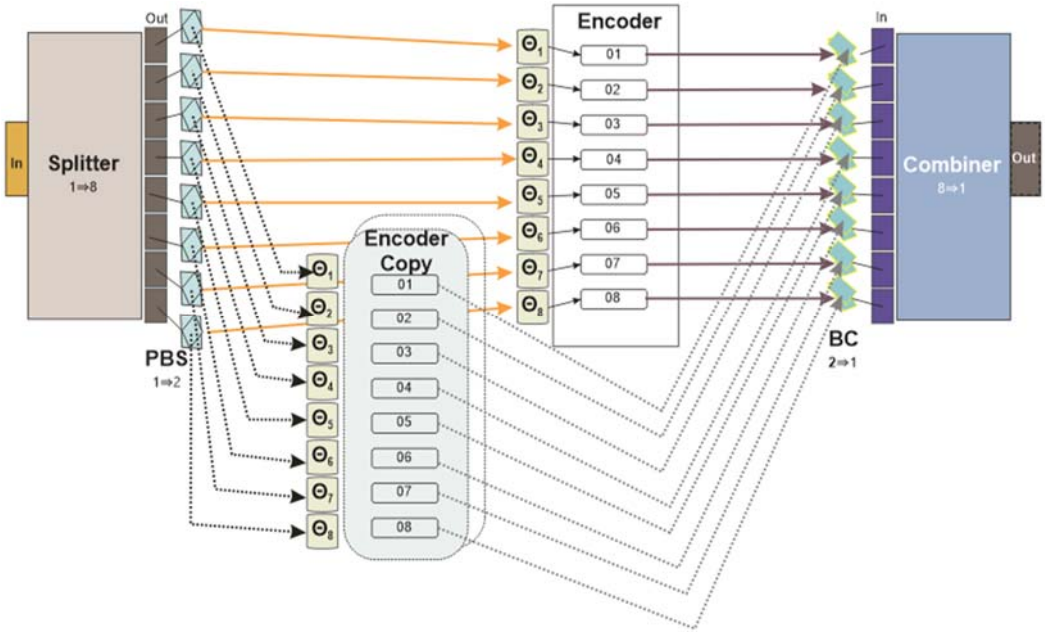


Fig. 1. Polarized HVGPU architecture: an eight-channel splitter (1→8) feeds eight polarizing beam-splitter cubes PBS (1→2), each separating its branch into orthogonal S and P components. The physical encoder and its copy – with the polarization angles Θ_1–Θ_8 set by the half-wave plates and transformation functions 01–08 in each – process the S and P branches independently before both branches recombine through a beam combiner BC (2→1) and enter the eight-input channels of the combiner unit (8→1).

3.5. On-Axis Propagation

At every channel, the image propagates along the optical axis rather than off-axis. This choice avoids the need for compensation of the geometric distortions that off-axis propagation would otherwise introduce, and follows directly from the reconstruction-geometry criterion given in Section 2.1: keeping both recording and reconstruction beam geometry on the optical axis most straightforwardly maintains their exact correspondence. This article makes no claim regarding the relative merits of off-axis architectures generally; it describes and characterizes the on-axis

configuration actually built, and leaves treatment of off-axis alternatives to future work.

3.6. The Reference Channel and Whole-Assembly Positioning

A 49th channel, sourced from a single physical output of a separate, ordinary eight-channel non-polarized trigger, serves as the reference plane against which the 24 positive and 24 negative depth planes are measured. Like each of the 16 polarization-derived channels, the reference channel carries the red, green, and blue components. The three components co-propagate spatially in the reference channel and remain spectrally distinct, together forming a polychromatic RGB reference beam. No component appears at an intermediate wavelength: a spectrometer in this channel records three lines at 632.8, 532, and 488 nm, and a detector or the eye receives the sum of the three intensities, which is perceived as a mixed color by additive mixing. These components all carry the unmodified reference image, so they share one channel number and constitute a single reference plane. Counted by color, they are three virtual color channels. The reference channel is controlled in the same manner as the other 48 channels, rather than treating it only as a passive positional marker.

Because the physical spacing between the reference channel and each of the 48 depth-plane channels is fixed by construction, a single translation applied to the reference channel alone repositions the entire 49-plane assembly along the propagation axis as a rigid whole. The assembly's overall standoff distance from the source thereby becomes one adjustable quantity, governed at the reference channel, rather than 48 separately tracked positions across both branches of the polarization-cloned trigger.

3.7. Physical-to-Virtual Channel Accounting

The architecture described above amounts to an equivalence, presented here in those terms: a single physical trigger, augmented with the polarizing beam-splitter layer described in Section 3.2, is designed to reproduce the function of two separately constructed physical triggers. Building the 16-channel polarized trigger described here, together with the single reference channel drawn from a separate ordinary trigger, therefore reaches the same three-trigger-equivalent target architecture as an all-non-polarized, three-physical-trigger implementation, using two physical triggers rather than three. This article offers it as a description of the method's structural economy – reaching a given target architecture by a different, more economical route – rather than as an argument for the value of channel-count scaling as such.

The economy is one of hardware and not of optical power: the polarizing beam-splitter cube divides the power of each branch between its S and P outputs, so that, for an ideal lossless cube, each of the 16 polarization-separated channels receives half of the power of the corresponding channel of a non-polarized trigger illuminated by its own full beam. This halving is the price of saving the second trigger; it can be recovered through the source power and, in recording, through exposure (Section 5.3), and a complete power budget is not characterized in this article.

The present article's emphasis is on the polarization technique itself, and on the design of a polarization-cloned trigger intended to reproduce the function of a physically constructed one, not on the raw number of channels obtainable by extending the technique further.

In total, the architecture described in this section comprises 17 active physical channels: the 16 polarization-separated channels that the polarizing beam-splitter layer produces from the eight paths of the inherited splitter (eight S and eight P), plus one reference path taken from a single output of a separate, ordinary eight-channel non-polarized trigger, with the other seven outputs left unused. Each of the 16 polarization-separated channels carries the red, green, and blue components, giving $16 \times 3 = 48$ virtual color channels, 24 in each branch. Together with the single reference plane, this realizes 49 depth planes; counted by color, the reference path carries three components, so the total is 51 virtual color channels.

4. Optical Implementation and Recorded Realization of a Function

Eliminating translating bulk optics from a zoom function is a design goal pursued across imaging systems generally, independent of any holographic context. Deformable-mirror-based zoom systems, for example, achieve variable focal length by reshaping a mirror surface under microactuator control, avoiding lens-group translation altogether; such systems report zoom ratios on the order of 2×, with each new target focal length requiring fresh calculation of the required surface profile and per-actuator voltage control to manage the aberrations the reshaped surface itself introduces [22]. The present work shares the goal of a zoom function without translating bulk elements, but pursues it by a different route.

In the encoder stage described in Section 3.3, between the half-wave plate rotators and the combiner, an optical teletransfer system gives each channel independently controllable beam convergence or divergence. This sets the effective distance of the channel's reconstructed depth plane by the mechanism of Section 2.1: varying the convergence or divergence angle of the illuminating beam relative to the geometry used at recording. The axis of the illuminating beam stays coincident with that of the recording reference beam, so the reconstruction remains free of depth distortion. Here the optical function of this zoom teletransfer system is recorded as an interference

pattern directly in the recording medium of Section 5, forming a thick hologram of the teletransfer optics and not a digitally computed or resampled pattern.

Reconstructing this recorded pattern provides the convergence and divergence behavior of the teletransfer system for the recorded geometry, by the wavefront-cloning principle of Section 2.2, so that the physical zoom optics need not remain part of the operating system.

A digital reproduction of the same function with a spatial light modulator, for instance a computed phase pattern equivalent to a variable-focal-length Fresnel lens, is a natural extension of this principle. Such a realization would make the zoom state reconfigurable, in contrast to the fixed recorded pattern of the present implementation, and is left for future work.

5. Recording Medium and Energy Budget

5.1. Recording Medium

The present work employs U04 silver-halide holographic emulsion as the recording medium, in accordance with the original architecture [4]. U04 is an iso-panchromatic silver-halide emulsion with a 4 nm average grain size, a resolution of 18000 lines/mm, and spectral sensitivity spanning 440-700 nm, sufficient for recording all common holographic laser wavelengths without diffusion-related blur [23]. Whereas the design and fabrication study [5] used the LiMo silver-halide formulation, the present work returns to U04.

Each holographic plate consists of a photosensitive layer approximately 4 μm thick, deposited over an underlayer approximately 0.4 μm thick, both supported on a glass substrate. For monochrome recording of a reflective diffuse object, [23] reports a diffraction efficiency of 98 % for this medium. This article limits its materials treatment to the properties needed to interpret the layered model and energy budget below, and does not attempt a comprehensive characterization of U04. It makes no systematic comparison between recording emulsions, since the aim is to characterize what the present technique achieves with this medium.

5.2. Layered Numerical Model of the Recording Medium

Rather than treat the active layer of U04 as a single, uniformly responding medium, this article models it as a stack of thin slices, each acting as a thin diffraction grating, which must operate in phase with one another for the reconstructed image to build correctly. This is a numerical discretization only: a layer of about 4 μm sampled at a nominal spatial step of 4 nm, the grain size reported for U04 [23], contains about 1000 slices, which should not be interpreted as 1000 physically distinct diffraction gratings. The recorded grating period is determined by the interference geometry and the wavelength and not by the grain size; for a reflection recording at 532 nm in a medium of refractive index near 1.5, the fringe planes are about 177 nm apart, which gives of the order of 20 Bragg planes across the layer.

5.3. Recording Energy Budget

In the present work, ordinary exposure-energy management – rather than any active optical gain mechanism – recovers the diffraction efficiency lost to the additional absorption that the added polarizing-beam-splitter layer and its associated second encoder introduce. Recording quality depends on matching the object-beam and reference-beam intensities for maximum interference contrast, and extending exposure time for a given channel, within the linear part of the recording material's exposure characteristic, compensates for any reduction in available signal at that channel (Fig. 2).

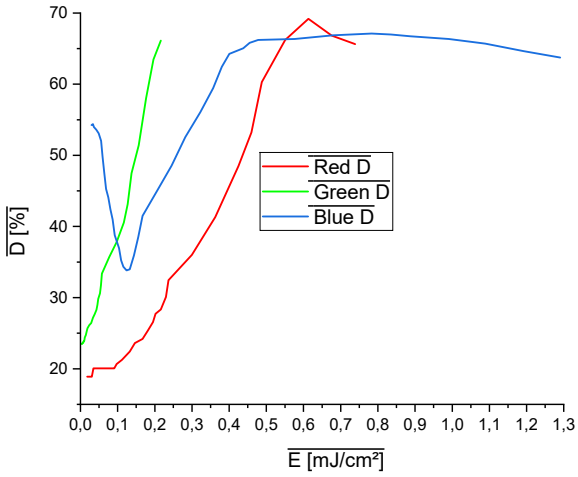


Fig. 2. Dynamic Range of the R–G–B Diffraction-Efficiency Response in the U04 Emulsion.

The developers of the emulsion recommend a minimum exposure of 320 μJ/cm² per color component for full-color recording, rising to roughly 600 μJ/cm² for monochrome [23]. In this work, exposures ranging from 320 μJ/cm² up to 1.2 mJ/cm² were investigated.

In the present measurements (Fig. 2), the maximum diffraction efficiencies of the U04 plates are about 69 % in the red, 66 % in the green, and 67 % in the blue, reached at exposures of about 0.6, 0.2, and 0.8 mJ/cm², respectively. For comparison, [24] reports values above 50 % for an emulsion with a 10 nm grain under different exposure conditions, and [23] reports 98 % for monochrome recording of a reflective diffuse object. For a color holographic optical element such as the HVGPU, however, the quantity that matters is not the maximum efficiency of each color but the balance

among the three, since the relative output powers of the red, green, and blue components determine the white balance perceived by the eye. This balance is reached at operating efficiencies well below the maximum attainable for one or two of the colors, and it is set by the exposure of each color hologram and by the source powers. The lower maxima of the present plates, which we attribute to a recording and swelling procedure not optimized for peak efficiency, therefore do not by themselves impair the operation of the element.

The efficiency cost of recording several wavelengths together in one shared layer, instead of in separate per-color recordings as here, has been measured elsewhere: a photopolymer volume hologram recorded with red, green, and blue exposed together reached per-channel diffraction efficiencies of 34 % to 67 % (mean 53 %), the shortest-wavelength channel receiving the least favorable exposure in that configuration [3]. This is consistent with our choice of a separate recording medium per color, which allows independent control of exposure and phase for each color component.

6. Experimental Characterization

The following measurements characterize the individual color holograms comprising the present architecture, and the polarization states of the physical and cloned branches of the trigger. We obtained the diffraction-efficiency values reported in Fig. 2 following the exposure procedure described in Section 5. We characterized each of the three color components – recorded independently at 632.8 nm, 532 nm, and 488 nm on separate U04 plates – individually before assembly into the S/P-cloned trigger, to obtain a per-color baseline against which we could later judge the combined trigger's behavior.

6.1. Polarization-State Verification

The polarization states at the output of the optical processor were verified with a polarimeter placed at its end, that is, after the beam-splitter unit, the encoder unit, and the beam-combiner unit, so that each measurement includes the combined effect of all three units. The polarimeter is a Thorlabs PAX1000VIS/M, which operates in the visible spectral range and was set to the wavelength of each color component (632.8, 532, and 488 nm).

The measurements confirm that the assembled processor – the splitter units that form the S and P branches (Section 3.2), followed by their encoder units (Section 3.3) – produces the intended polarization angles and not only the nominal design values. Fig. 3 presents the results for the three color components – red (Fig. 3a), green (Fig. 3b), and blue (Fig. 3c) – sampled at representative polarization angles from the full 16-channel set: 0°, 20°, and 70° for the physical S-branch trigger, and 90° and 160° for the polarization-cloned P-branch trigger.

Each panel shows the measured polarization ellipse together with its azimuth and ellipticity angle; an ellipticity angle of 0° corresponds to ideal linear polarization. The polarimeter reports azimuths in the interval from −90° to +90°, so the readouts of about −20° in the 160° panels are equivalent to 160° modulo 180°, in agreement with the periodicity of linear polarization (Section 3.3).

The azimuths printed in the panels agree with the nominal polarization angles within 0.7° in 13 of the 15 measurements and within 2.4° in the remaining two (red at 70° and 160°), and the ellipticity angles are at most 3.8°, so all measured states are close to linear. An ellipticity angle of 3.8° corresponds to a minor-to-major axis ratio of 0.07 in amplitude, that is, an orthogonal component of about 0.4 % of the intensity. The largest azimuth deviation, 2.4°, is less than half of the 10° spacing between adjacent channels, so each measured state remains closer to its own nominal angle than to that of a neighboring channel.

The polarization-cloned P branch reproduces the polarization quality of the physical S branch: its largest azimuth and ellipticity deviations are 1.8° and 3.2°, against 2.4° and 3.8° for the S branch. The panel labels are the azimuths of the linear polarization read on the polarimeter, referred to the polarization axis of the S branch at the encoder input, and not mechanical positions of the half-wave plate axes.

A plate whose fast axis is at an angle α from this axis maps a linear polarization at an angle ψ to $2\alpha - \psi$. The S-branch plates receive $\psi = 0°$ and, at α from 0° to 35°, produce 0° to 70°. The P-branch plates receive the orthogonal polarization $\psi = 90°$ and are mounted at the S-branch orientations plus 90°, so that they produce 90° to 160°.

Crosstalk between adjacent channels is not evaluated here, since no pair of successive channels was measured in this article; it is reserved for the companion paper on the base architecture.

6.2. Scope of the Comparison between the Physical and the Cloned Trigger

The polarization states of the physical S-branch and the polarization-cloned P-branch are compared in Fig. 3, and the per-color diffraction-efficiency baselines of Fig. 2 were measured before assembly. A direct comparison of the diffraction efficiencies of the physical and the cloned trigger, with uncertainties, is not carried out in this article and is reserved for a subsequent publication.

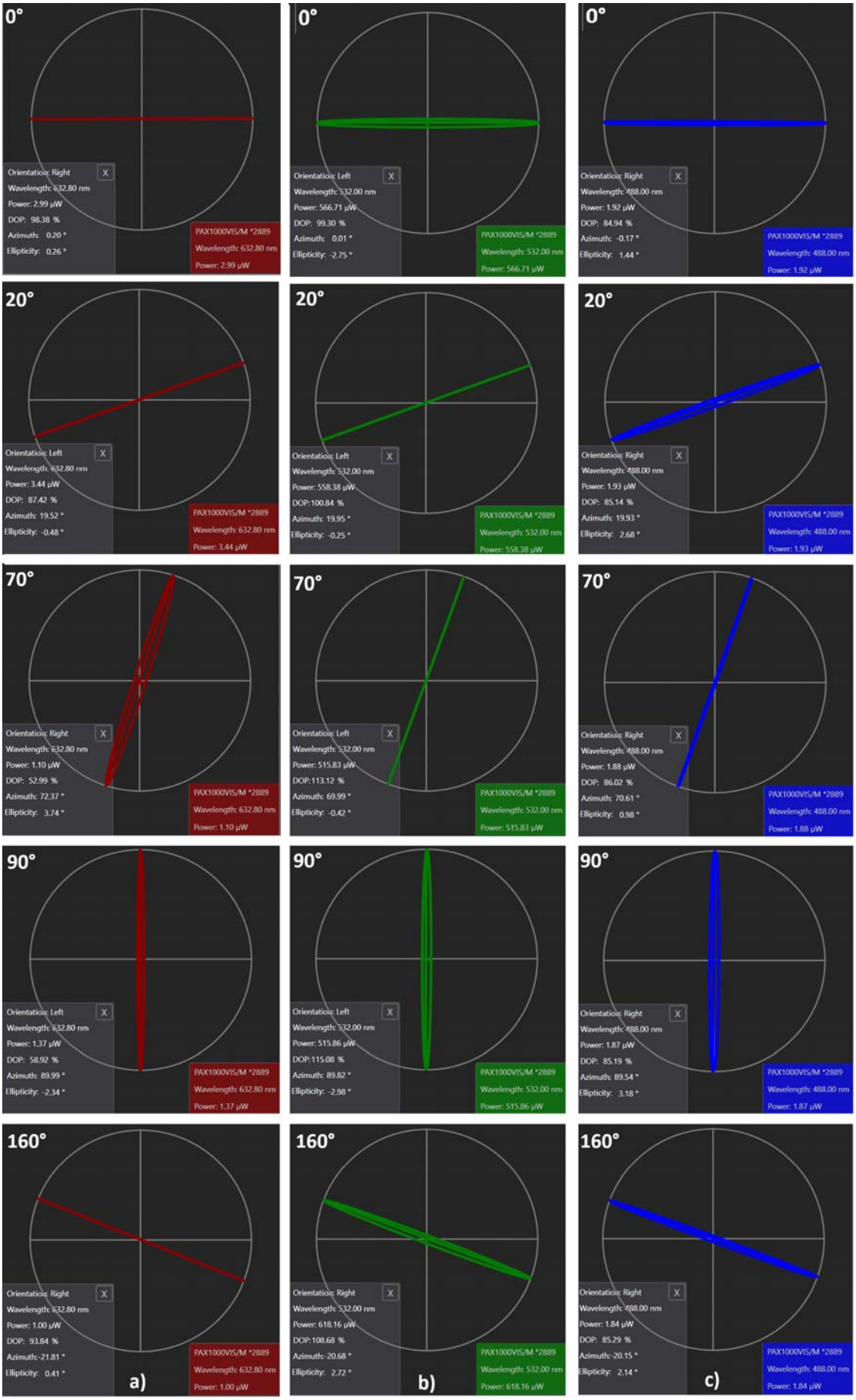


Fig. 3. Polarization azimuths of physical trigger (0°, 20°, and 70°) and polarization-cloned trigger (90° and 160°).

7. Discussion

7.1. Efficiency and Crosstalk across Polarization-Multiplexing Mechanisms

The efficiency cost of polarization-based multiplexing depends strongly on the physical mechanism performing the channel separation. Where multiple holograms are recorded as overlapping interference patterns at a single position in an anisotropic material and distinguished at readout by reference-wave polarization, diffraction efficiency falls to a fraction of a percent as channel count increases, even as crosstalk remains low: approximately 6×10^{-4} for a four-channel polarization-holographic recording in a photoinduced-anisotropic material [6]. Metasurface-based approaches that engineer a departure from strict polarization orthogonality to exceed the three-channel Jones-matrix limit report substantially higher per-channel efficiencies: 19 % to 28 % for polarization multiplexing alone, and 10 % to 16 % when combining polarization with wavelength multiplexing across 10 channels. Both approaches confirm that crosstalk grows as channel count increases, and at least one configuration requires active optimization to correct uneven energy distribution among channels [11, 12]. Researchers have measured polarization-based folding of the optical path in near-eye relay optics – built from cascaded quarter-wave plates, polarizing beam-splitters, and half-mirrors – at 8.5 % to 17.5 % against a 25 % theoretical maximum. They attribute the shortfall to imperfect polarization states and component performance rather than to any flaw in the underlying principle [9]. Where polarization switches dynamically in time rather than remaining fixed at fabrication, researchers have measured cascaded liquid-crystal phase-shifter stages at approximately 80 % transmittance per stage, with total transmission falling approximately as 0.8 raised to the number of cascaded stages [10]. The present architecture separates its 16 channels spatially, at a single polarizing-beam-splitter layer positioned once ahead of the encoder stage, rather than through repeated cascaded stages, overlapping recordings at one location, or engineered departure from polarization orthogonality; Fig. 2 reports its own measured diffraction efficiency.

7.2. Parameter Sensitivity: Robustness versus Security

Design priorities for polarization-based multiplexing depend heavily on the intended application. Display designers typically treat extreme sensitivity to small parameter deviations as a limitation to engineer around; elsewhere, researchers have deliberately exploited comparable sensitivity as a security feature instead. One holographic encoding scheme requires exact alignment of several interdependent optical parameters for correct reconstruction, with a propagation-distance error of only 0.1 cm sufficient to destroy the reconstructed image entirely [14]. The present architecture pursues the opposite goal, favoring robust, stable channel separation from a small number of fixed, passive optical elements over fine parameter sensitivity in either direction.

Polarization can also select between distinct behaviors of a single fixed element: a nonlinear flat-optics edge detector reconstructs edges in different directions depending on the polarization of its illumination [25]. This differs from the channel multiplication described here, but it relies on the same underlying principle: a controllable property of the illuminating beam governs the behavior of a fixed passive element.

7.3. Active versus Passive Channel Generation

Polarization can also switch dynamically in time, instead of being fixed at fabrication. Fast-response liquid-crystal phase shifters paired with polarization gratings steer the output of a single SLM between discrete positions in rapid sequence, synchronized frame-for-frame with the refresh rate, so that viewers perceive positions occupied one at a time as simultaneous, through persistence of vision. The same work reports that the perceived brightness per position falls in proportion to the number of positions, because a fixed instantaneous optical power is divided across sequential time slots and not distributed once across permanently separate paths [10]. The present architecture divides the power spatially, once, at the polarizing-beam-splitter layer: its 16 channels exist simultaneously as separate polarization-resolved paths, without time multiplexing, so that no switching speed or synchronization is required to separate them. The two approaches serve different goals: enlarging the effective projection area of a single SLM in one case, multiplying the channel count of one recorded trigger in the other.

7.4. Digital Computational Alternatives for Multi-Depth Reconstruction

Real-time, full-color holographic video by digital means is possible, though at a substantial engineering cost. A slim-panel holographic video display computes full-color, ultra-high-definition holograms at 30 frames per second with a custom single-chip processor built for the task, and expands the viewing angle about 30-fold through a dedicated steering-backlight unit [1]. Not every digital approach scales poorly with the number of depth planes: a single-step Fresnel calculation based on the split-Lohmann method computes a multi-depth scene in constant time regardless of depth-layer count, up to

18 times faster than conventional layer-based computation, but it still depends on dedicated graphics-processing-unit hardware computing a fresh hologram every frame and has its own constraints, a minimum reconstructible depth and a residual lateral-shift artifact [2]. Other digital approaches to related multi-depth or multi-view problems report substantial training overhead, tens of hours of network training in more than one case [26], and typically represent scenes by a small number of discrete depth or view states, a number set mainly by computational cost: three depth planes in one recent multi-depth holographic display [9] and eight accommodation states in a near-eye light-field reconstruction system [26].

The distinction between these digital results and the present architecture is not that digital multi-depth computation cannot be made fast or cannot scale; recent results show that it can be made both. The distinction is that digital computation, however fast, must run continuously, once per frame, for as long as the display operates, whereas channel and depth-plane multiplication in the present architecture is a fixed property of the recording and of the polarizing-optics layer, established once during fabrication rather than recomputed indefinitely. The throughput of the present architecture is therefore bounded by the refresh rate of the source of the input frames, a HOLOEYE LETO-3-VIS-009 phase-only spatial light modulator [27], whose manufacturer-specified input frame rate is 60 Hz, or 180 Hz in color-field-sequential mode. In our setup it projects 90 monochrome frames per second, which corresponds to 4410 depth-plane sub-frames per second for 49 planes, as in the base architecture [4]. Source modulation, synchronization, detector response, scene updating, and end-to-end latency are not characterized in this article. This reflects a broader motivation shared with unrelated recent work in other areas of optics entirely: researchers have proposed all-optical analog computing as a remedy for the complexity, power consumption, and speed limitations inherent to digital signal processing generally [25].

7.5. Design Perimeter: Technique over Channel Count

The central claim of the present article is methodological rather than a claim about achievable scale. A polarization-cloned trigger, generated from a single physical trigger by the polarizing-beam-splitter layer described in Section 3, provides orthogonally polarized copies of the eight branches, and the measured polarization states of the physical and cloned branches (Fig. 3) are consistent with this design. Whether the cloned trigger matches an independently constructed one in diffraction efficiency remains to be shown.

Three further routes to multiplying channel count are noted as directions available to the present technique, rather than as content developed or characterized here. These are: additional physical degrees of freedom, such as a further polarizing-prism type beyond the polarizing beam-splitter cube used here (a Wollaston or Glan prism); further zoom-state, global-convergence, or spectral-shift dimensions layered onto the present architecture; and dynamic, spatial-light-modulator-based generation of additional virtual triggers along the lines discussed in Section 2.2. A full three-trigger, 49-plane system – assembled and characterized as a complete unit, rather than through the paired physical/cloned-trigger comparison reported in Section 6 – is likewise reserved for future work.

8. Conclusion and Future Work

This article has described the addition of a polarizing beam-splitter-cube layer to an existing eight-channel, non-polarized holographic trigger architecture, doubling it into 16 polarization-addressable channels through separation of each branch's light into orthogonal S and P components. Two encoder stages, each a direct reproduction of the base architecture's own encoder offset by the 90° separation between S and P, assign each of the 16 channels its own polarization angle. A fixed representational convention then maps the S-branch and P-branch channels to the positive and negative depth planes of a 49-plane display, with a 49th, polychromatic RGB reference channel drawn from a separate, ordinary trigger.

Future work includes assembling and characterizing the complete three-trigger, 49-plane system as a single unit, and extending the polarization-cloning principle to other orthogonal-polarization pairs and other polarizing-prism types. It also includes combining the present cloning mechanism with additional physical multiplexing dimensions – zoom-state, global-convergence-angle, and spectral-shift multiplication – and exploring a spatial-light-modulator-based, dynamically reconfigurable realization of the recorded zoom teletransfer system described in Section 4.

References

[1]. J. An, K. Won, Y. Kim, J.-Y. Hong, et al., Slim panel holographic video display, *Nature Communications*, Vol. 11, 2020, 5568.

[2]. A. Lagrange, A. Gilles, K. Heggarty, B. Fracasso, et al., Real-time calculation of Fresnel holograms using the split-Lohmann method, *Optics Letters*, Vol. 50, Issue 8, 2025, pp. 2554-2557.

[3]. T. Wu, J. Ma, C. Wang, H. Wang, et al., Full-color see-through three-dimensional display method based on volume holography, *Sensors*, Vol. 21, Issue 8, 2021, 2698.

[4]. G. Ivanova, B. Ivanov, D. Dimov, A 24-channel optical trigger for use in 3D holographic display systems, in *Proceedings of the 9th International Conference on Optics, Photonics and Lasers (OPAL)*, 2026, pp. 35-39.

[5]. B. Ivanov, G. Ivanova, D. Dimov, Design and fabrication of a 24-channel holographic optical trigger for high-speed optical computation, in *Proceedings of the 9th International Conference on Optics, Photonics and Lasers (OPAL)*, 2026, pp. 86-91.

[6]. S. J. Zheng, J. R. Tan, H. J. Liu, X. Lin, et al., Orthogonal matrix of polarization combinations: Concept and application to multichannel holographic recording, *Opto-Electronic Advances*, Vol. 7, 2024, 230180.

[7]. B. Ivanov, M. Shopova, A. Baldjiev, E. Stoykova, et al., Four-channel holographic infrared optical element, *Proceedings of SPIE*, Vol. 7747, 2011, 77470K.

[8]. X. Xu, Z. Lu, H. Pan, S. Zhou, et al., Polarization-encoded switchable structured light generator based on all-dielectric holographic metasurfaces, *Coatings*, Vol. 16, Issue 7, 2026, 858.

[9]. Q. Hou, S. Xu, Z. Li, C. Chi, et al., Full-color multi-depth holographic pancake display, *Optics Express*, Vol. 34, Issue 13, 2026, pp. 24805-24828.

[10]. Q. Han, T. Wang, G. Zhang, W. Kamal, et al., Holographic projection display enlargement via polarization grating beam steering with fast-response liquid crystal pi cells, *ACS Photonics*, Vol. 13, Issue 10, 2026, pp. 2943-2953.

[11]. J. Wang, J. Chen, F. Yu, R. Chen, et al., Unlocking ultra-high holographic information capacity through nonorthogonal polarization multiplexing, *Nature Communications*, Vol. 15, 2024, 6284.

[12]. Y. Liu, Y.-T. Xiao, X.-Y. Huo, Q. Cai, et al., Dynamic holographic display with multiplexed dielectric metasurfaces, *Nanophotonics*, Vol. 15, Issue 12, 2026, e70114.

[13]. Y. Ye, H. Su, Y. Jia, B. Li, et al., High-capacity multiview display with large viewing angle via orbital angular momentum encoded nanograting arrays, *Nanophotonics*, Vol. 14, Issue 24, 2025, pp. 4341-4350.

[14]. F. Shen, Z. Mao, W. Fan, J. Wu, et al., OAM multiplication operator enabled holographic multiplexing, *Light: Science & Applications*, Vol. 15, 2026, 18.

[15]. Q. Ji, H. Deng, H. Zhang, W. Jiang, et al., Optical see-through 2D/3D compatible display using variable focus lens and multiplexed holographic optical elements, *Photonics*, Vol. 8, Issue 8, 2021, 297.

[16]. A. Márquez, C. Li, A. Beléndez, S. A. Maier, et al., Information multiplexing from optical holography to multi-channel metaholography, *Nanophotonics*, Vol. 12, Issue 24, 2023, pp. 4415-4440.

[17]. H. Kogelnik, Coupled wave theory for thick hologram gratings, *Bell System Technical Journal*, Vol. 48, Issue 9, 1969, pp. 2909-2947.

[18]. D. Gabor, A new microscopic principle, *Nature*, Vol. 161, Issue 4098, 1948, pp. 777-778.

[19]. P. Hariharan, Optical Holography: Principles, Techniques and Applications (2nd ed.), *Cambridge University Press*, Cambridge, 1996.

[20]. F. Wang, J. Starobrat, Y. Wang, M. Goto, et al., Local zoom issue in holographic 3D display, *Journal of Physics: Photonics*, Vol. 8, Issue 3, 2026, 035019.

[21]. K. H. Kwon, M.-U. Erdenebat, N. Kim, A. Khuderchuluun, et al., High-quality 3D visualization system for light field microscopy with fine-scale shape measurement through accurate 3D surface data, *Sensors*, Vol. 23, Issue 4, 2023, 2173.

[22]. Y. Ning, S. Chen, D. Gao, Y. Zhou, et al., Deformation strategy of deformable mirrors for a typical zooming system, *Applied Optics*, Vol. 63, Issue 8, 2024, pp. 2311-2320.

[23]. P. Gentet, Y. Gentet, S.-H. Lee, Ultimate 04: The new reference for ultra-realistic color holography, in *Proceedings of the International Conference on Emerging Trends & Innovation in ICT (ICEI)*, 2017, pp. 162-166.

[24]. B. Ivanov, M. Shopova, E. Stoykova, V. Sainov, Reconstruction of silver halide volume reflection holograms at the wavelengths of recording, *Optica Applicata*, Vol. 42, Issue 3, 2012, pp. 643-650.

[25]. D. De Ceglia, A. Alù, D. N. Neshev, C. De Angelis, Analog image processing with nonlinear nonlocal flat optics, *Optical Materials Express*, Vol. 14, Issue 1, 2024, pp. 92-100.

[26]. U. Gudelek, E. Sahin, A. Gotchev, Perceptually optimized model for near eye light field reconstruction, in *Proceedings of the IEEE International Workshop on Multimedia Signal Processing (MMSP)*, 2023, pp. 1-6.

[27]. LETO-3 phase only spatial light modulator datasheet, HOLOEYE Photonics AG, https://holoeye.com/products/spatial-light-modulators/leto-3-phase-only/

Published by International Frequency Sensor Association (IFSA) Publishing, S. L., 2026 (http://www.sensorsportal.com).

Sensors & Transducers, Vol. 274, Issue 3, October 2026, pp. 42-54

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2026
http://www.sensorsportal.com

An All-Optical Holographic Video-Graphics Processor for Real-Time Volumetric 3D Display

Ginka IVANOVA, * Branimir IVANOV and Deyan DIMOV
Institute of Optical Materials and Technologies, Laboratory "Holography",
109, Acad. G. Bontchev Str., 1113 Sofia, Bulgaria
Tel.: + 33 601825261
E-mail: branimirivanov@iomt.bas.bg

Received: 31 Aug. 2026 /Revised:26 Oct. 2026 /Accepted: 29 Oct. 2026 /Published:30 Oct. 2026

Abstract: This article presents an extended treatment of an all-optical holographic trigger intended for real-time volumetric three-dimensional display. The architecture combines three holographic triggers: two of them, each generating eight parallel optical channels in three colors through phase and polarization multiplexing, supply 24 virtual color channels each, and the third supplies a single reference plane, yielding forty-nine simultaneously available depth planes from a single input frame. Unlike digital computer-generated holography, where computing depth- and color-resolved content anew for every frame costs more as resolution and depth-plane count grow, the present architecture performs channel multiplication within the optical domain itself, requiring no digital conversion at any stage. The article develops the underlying splitter–encoder–combiner architecture and the physical mechanism generating and assigning discrete polarization states to individual channels, and works through an explicit accounting of physical versus virtual optical channels across the full three-trigger assembly – eight physical channels per full trigger yielding 24 virtual color channels, plus one reference channel – and presents a spectral and polarization characterization methodology for verifying correct channel operation across the visible range, positioned throughout against the digital holography literature addressing the same underlying problem.

Keywords: Holographic optical trigger, Volumetric three-dimensional display, Multichannel volume holography, Splitter–encoder–combiner architecture, Phase and polarization multiplexing.

1. Introduction

Volumetric three-dimensional display has resisted a fully satisfactory solution across several decades of active research, despite its evident value across a broad range of practical applications. A holographic reconstruction reproducing motion parallax, ocular accommodation, and correct occlusion at video frame rates places demands on both the computation and the optics that few systems built to date have met simultaneously. This article is an extended version of two OPAL' 2026 conference papers [1, 2] and develops, in considerably greater depth, a candidate solution built around a holographic optical trigger (HOT): a device performing depth- and color-channel multiplication within the optical domain itself, dispensing with the digital computation step that dominates the cost of most competing approaches. We compare the architecture with digital hologram synthesis and with earlier multichannel optical hardware, and we test its three color-specific units across the visible range. Compared with [1, 2], this work contributes (i) a splitter–encoder–combiner formulation, (ii) an explicit accounting of physical and virtual channels, (iii) wavelength-resolved measurements of transmittance, reflectance, and diffraction efficiency of each color-specific holographic video-graphics processor unit (HVGPU), (iv) polarimetric verification of the input state and of the channel polarization states at the combiner output, and (v) experimental documentation of multi-color synchronization.

1.1. Computational Pathways to Three-Dimensional Hologram Synthesis

Several research efforts have addressed the computational cost of digital hologram synthesis through learned or algorithmically optimized methods, and their results are instructive about where the underlying difficulty actually resides. A deep-learning approach trained on layered-depth-image (LDI) scene representations reaches real-time rates on a desktop graphics card, yet falls to roughly 5 frames per second on a mobile graphics processing unit (GPU) [3]. This is one figure among several illustrating a trade-off that runs through this entire body of work: raising reconstruction quality tends to cost runtime, and the exchange rate is steep. A different class of method, built around overlap-add stereograms, encodes a light field into a single computed two-dimensional hologram per frame, treating depth as a property recovered mathematically through wavefront propagation rather than as a physically distinct output [4]. Reported computation times for this approach span tens of seconds on general-purpose hardware down to under 100 milliseconds with GPU acceleration, for a single hologram at one depth configuration [4]. Both figures describe near-eye display geometries – a virtual accommodation volume of a few millimeters synthesized behind a spatial light modulator (SLM) and viewed through magnifying eyepiece optics – a configuration distinct from the physically projected, free-space multi-plane output pursued here. These numbers are not interchangeable with the depth-plane count developed here.

A related strand extends digital synthesis to layered scene representations propagated toward the hologram plane through a chain of angular-spectrum computations, in which each depth layer's contribution depends recursively on the propagated result of the layer preceding it [5]. That dependency is serial by construction: producing N layers costs N sequential propagation steps, each awaiting the completion of the one before it. The architecture developed in this article generates its depth-resolved channels through physically parallel optical paths carrying no such dependency – no channel waits on another. The handling of specular, non-Lambertian surfaces shows a comparable variety of digital solutions: directionally weighted diffuser modeling in the frequency domain [5], phase-added stereogram techniques incorporating per-ray depth [4], and hybrid schemes switching between plane-wave and spherical-wave primitives according to local surface properties [6]. Three separately engineered answers to a problem that a physically reconstructed wavefront does not need to solve at all, since directional light behavior belongs to the propagating field itself rather than to a computation performed about it.

Lensless imaging offers a useful counterpoint from outside the CGH mainstream. Fourier transform holography (FTH) recovers a full complex object field from a single recorded intensity pattern via one computed Fourier transform, given a reference wave co-located with the object [7]. This demonstrates that full-field reconstruction from a single capture is tractable, yet the reconstruction itself remains computational: software executes this algorithmic solution to the phase problem after acquisition. Thick-medium, Bragg-selective volume holography of the kind employed throughout the present work and its precursors takes the opposite route: wavelength- and angle-selectivity are physical properties of the recording medium, and reconstruction proceeds by simple illumination, with no digital step involved at any point [8, 9]. The Fourier-transform-holography literature itself draws a related distinction between holography and coherent diffraction imaging as two structurally different answers to the same phase-recovery problem. The present architecture belongs to the physical-reconstruction camp, consistent with its broader aim of avoiding analog-to-digital conversion anywhere in the display pipeline.

1.2. Diffraction Accuracy and Off-Axis Wavefront Modeling

The accuracy with which off-axis diffraction can be modeled also bounds reconstructed image quality in holographic systems – a difficulty that arises whenever a diffractive or holographic element must reconstruct content away from its own optical axis. Numerical work on this problem has recently improved substantially through an adaptive sampling strategy built on the angular spectrum framework common to essentially every digital CGH method discussed above, converting an off-axis propagation problem into an equivalent, more efficiently sampled on-axis one [10]. The reported gains are considerable: roughly 35 times faster than a prior state-of-the-art method at 20° of incidence, and a 35° off-axis field computed in about one second on ordinary commercial hardware [10]. One detail in that comparison matters more than the headline speedup. Constrained to the same sampling budget as the proposed method rather than allowed extra samples to match its accuracy, the comparison method produces an incorrect result at an incidence angle as modest as 5° [10].

This consideration informs a design choice developed further in Section 2.5: every channel of the present trigger propagates along its own local optical axis rather than at an angle to it. The physical architecture itself thereby avoids the off-axis correction burden described above, rather than compensating for it after the fact – digitally or otherwise.

1.3. Antecedents of Multichannel Holographic Multiplexing

Multichannel holographic interferometry has a longer history than the display application pursued here. A four-channel holographic optical element built

for phase-stepping digital speckle-pattern interferometry [9] encoded two illumination planes – vertical and horizontal, needed for full three-dimensional displacement-vector recovery – at two near-infrared wavelengths, yielding four independently addressable channels within one compact device. That element originally served precision strain and displacement measurement rather than display; its sixfold scaling into the present 24-channel architecture [1, 2] represents a repurposing of an established multiplexing principle toward a new application area, not a straightforward continuation within the same field.

Demonstrations of coherent beam combination through angle multiplexing of several Bragg gratings recorded sequentially in a thick polymeric substrate separately support the broader feasibility of sustaining many channels within a single volume-holographic element [11], and related work on nanocomposite photoanisotropic materials further reinforces this feasibility in polarization-sensitive holographic media [12]. Such devices combine and split multiple beams with diffraction efficiency well predicted by coupled-wave theory, and the angle-multiplexing literature cited within that work indicates volume-holographic media of this class can in principle sustain well over a thousand independently multiplexed channels [11]. That precedent bears directly on the channel-scaling route outlined in Section 1.4 below and developed architecturally in Section 2, pointing toward substantially more than 24 channels using the same recording medium and without additional laser sources.

1.4. Recording Media: Lineage, Properties, and Scaling Potential

The present prototype records on a silver-halide emulsion (U04) reported to offer 18000 lines/mm resolution and a 4 nm grain, roughly two to three times finer, respectively, than the comparable panchromatic materials used for benchmarking (6000–10000 lines/mm; 8–10 nm grain) [13]. We record on U04, as in [1]. This is a change from [2], which used the LiMo emulsion that work describes as photoanisotropic. We chose U04 because it is initially isotropic, which makes it suitable for later study of anisotropy induced by the recording itself; that study lies outside the present article. The blue source is also changed: the prototype of [1] used 442 nm illumination, whereas the blue hologram of the present HVGPU was recorded and characterized with a 488 nm Coherent OBIS laser. The red and green sources are unchanged, and all blue-channel results reported here correspond to the 488 nm source.

That material sits at a later stage of a materials lineage running through our own earlier work. A four-channel interferometric element built previously [9] used an in-house silver-halide emulsion requiring post-development chemical swelling to shift the reconstruction wavelength from the visible recording wavelength into the near-infrared operating band. A later, systematic study applied the same swelling principle to the considerably smaller, color-balance-critical shifts required for accurate multi-color reflection holography. That study established that controlled post-development swelling can correct shrinkage-induced wavelength shifts of tens of nanometers, and that a residual shift of only a few nanometers is sufficient to visibly disturb color balance [8]. The same study found that recording several wavelengths sequentially, through successive exposures on one plate, yields lower diffraction efficiency than a simultaneous multi-wavelength exposure. This observation bears directly on the present architecture, where a sequential exposure at a distinct half-wave plate setting records each of the eight phase-multiplexed channels within a single-color hologram. That sequential-exposure effect offers a physically grounded reason for the channel-to-channel efficiency variation reported in the fabricated prototype [1, 2], addressed further in Section 3.2. Reported single-wavelength, single-exposure efficiencies of roughly 80 %, 75 %, and 50 % at the red, green, and blue recording wavelengths respectively give a materials-level baseline against which the lower efficiencies measured under repeated [8], multichannel exposure become explicable rather than an unexplained figure.

The present recording setup – three laser sources combined through a dichroic beam-combining stage feeding a shared spatial filter – follows the general optical arrangement of the earlier work discussed above [8], establishing a direct line of design descent between the two setups. Recording the three colors simultaneously within a single shared photosensitive layer, rather than on three separate plates, carries a trade-off. A single shared layer minimizes total optical absorption, since light passes through one holographic layer rather than three; against that, simultaneous multi-color recording demands precise alignment of the fringe systems of all participating wavelengths, which substantially complicates achieving accurate white balance in the reconstructed image. Given that the present stage of this research remains experimental, recording each color on its own holographic plate allows independent characterization and adjustment of amplitude and phase per color. Once this separated-hologram approach establishes amplitude balance and fringe alignment, a single-layer, simultaneous multi-color implementation becomes a reasonable next step, trading the synchronization complexity just described for lower total absorption. Each plate used here consists of a photosensitive layer approximately 4 μm thick over a thin underlayer approximately 0.4 μm thick, both supported on a BK7 glass substrate 3–6 mm thick. Three plates rather than one raise the total absorption of the assembled trigger accordingly, a consideration for the overall optical efficiency budget of the finished system.

Reaching high diffraction efficiency in U04 recording depends on maintaining comparable optical

power between the object and reference beams during exposure, since the two beams achieve maximum interference contrast only when they reach the plate at closely matched intensity. The developers of the emulsion recommend a minimum exposure of 320 μJ/cm² per laser for full-color recording, rising to roughly 600 μJ/cm² for a monochrome exposure [13]. When the available object-beam power falls short relative to the reference beam, extending the exposure time – rather than altering beam geometry – achieves the required energy level.

The same wavelength-shifting swelling technique offers an inexpensive path to a higher channel count without adding more laser sources. Starting from the three recording wavelengths already in use here (632.8 nm, 532 nm, 488 nm), controlled post-development swelling shifts the effective reconstruction wavelength of a given recording in either direction from its own recording wavelength. The available shift is up to roughly 50 nm toward longer wavelengths and, independently, up to roughly 50 nm toward shorter wavelengths, so that each of the three base wavelengths can, in principle, yield three distinct reconstruction wavelengths: for example, the original wavelength variant, a red-shifted variant, and a blue-shifted variant. Applied across all three lasers, this gives up to nine reconstruction wavelengths from three physical sources. With the present recording wavelengths and shifts of ±50 nm, three pairs of these wavelengths (582 and 582.8 nm, 532 and 538 nm, 482 and 488 nm) lie closer together than the spectral width $\Delta\lambda$ of the corresponding diffraction lobe (Section 3.1), so about six of the nine are resolvable. Choosing the shift of each color so that neighboring wavelengths differ by more than $\Delta\lambda$ is an objective for optimization in subsequent work. Paired with the eight-channel phase multiplexing already demonstrated per color, this points toward a trigger of up to 72 channels, nine wavelength states multiplied by eight phase states, as an upper bound reachable with the three lasers already on the bench once the nine wavelengths are spaced by more than $\Delta\lambda$; with the shifts considered here, the resolvable count is about 48 channels. That channel count sits comfortably within the material precedent noted in Section 1.3 [11], and matches the 72-channel scaling direction identified as a target for future development [1]. One caveat belongs here as laboratory practice rather than device design: color-dependent diffraction efficiency, dynamic range, and overall optical response are not fixed properties of an emulsion. They vary between production batches of nominally identical material and drift over a plate's working life, so calibration measurements from one recording session do not guarantee identical results in the next without renewed calibration. Section 5.3 returns to this point in the context of experimental reproducibility.

Beyond this lineage, a dedicated holographic printer has also produced full-color holographic stereograms from AI-generated content on the U04 platform [14], underscoring the material's continued relevance within the broader color-holography literature.

2. Conception and Architecture of the Holographic Video-Graphics Processor Unit

2.1. The Holographic Optical Trigger as a Multichannel Switching Element

A holographic optical trigger, in the sense used throughout this article and its precursor [1], is a holographic image of a real multichannel optical switch. Depending on the direction of use, the same recorded structure functions either as a splitter – one input beam replicated into several outputs – or as a combiner, the reverse optical operation. Its purpose in either direction is the replication or merging of light beams across a defined band of the electromagnetic spectrum; here that band spans roughly 400 nm to 900 nm, the visible range plus a short margin on either side, sufficient for real-space 3D video projection. This element's functions govern the wavelength-dependent behavior of different color beams within the video stream without digitization between input and output; the transit time through the element is negligible compared with the frame period and does not limit the frame rate.

Three such devices combine here into what we term a holographic video-graphics processor unit (HVGPU): a splitter, an encoder, and a combiner, assembled in sequence. Three independent HVGPUs – one per working color – operate in parallel, each producing eight holographic projection planes for its assigned color. We chose the resulting count of 24 virtual color channels per trigger deliberately – eight per color across three colors – to echo the eight-bit byte of conventional digital architectures, easing comparison and future software migration between digital and optical computing platforms. This section develops that architecture well beyond the level of detail the treatment of the HOT concept in [1] allowed. That earlier treatment established the internal structure of the eight-channel base unit – the frequency mode generator, the frequency synchronization and spatial filtering stage, and the trigger proper, comprising the multiplexer, half-wave plate rotators, and depth-focus correction system [1] – which is recast within the splitter-encoder-combiner framework developed in this work.

2.2. Splitter–Combiner Architecture

Each module–the splitter and the combiner–comprises seven non-polarizing beam-splitter cubes, four 90° R100 mirrors, three intermediate path-length equalizers, and seven optical correctors that equalize the total optical path length of each channel. The splitter's task is straightforward to state and

considerably less straightforward to build: one incoming beam, with a defined phase, polarization direction, and wavelength, must emerge as eight parallel outputs in the same output plane, equidistant along one axis, each carrying exactly the input phase, polarization, and wavelength unaltered. The splitter clones one input channel onto eight output channels without modification of any optical parameter. In the present unit, one cube of the cascade divides the S and P amplitudes unequally; the input polarization conditioning described in Section 4.1 compensates for this, so the unit behaves as designed.

The combiner is the same physical unit run in reverse – eight inputs, one output – performing the corresponding merge. Each of the eight input channels' phase, polarization, and wavelength must appear at the single output unchanged relative to its own input. Where two channels of the same color enter the combiner with the same polarization and phase, they superpose coherently at the shared output. Channels of different colors do not interfere in this sense: their relative phase advances at the optical difference frequency, so a red and a green channel with the same polarization co-propagate and superpose in the same geometric channel, and the time-averaged intensity at the output is the sum of the two intensities.

Placed splitter-first, encoder second, combiner third – with the splitter's eight outputs facing the combiner's eight inputs – the assembly defines a net of eight optical channels into which the encoder can insert arbitrary transformations of the wavefront between splitting and combining. The whole assembly preserves phase, polarization, and wavelength, except where a deliberately applied transformation changes them; the assembly itself introduces no error beyond whatever transformation the encoder carries out.

Fig. 1 summarizes this splitter–encoder–combiner path. The splitter's eight output branches each carry, within the encoding stage, a half-wave plate set to produce its own polarization angle (Θ_1 through Θ_8) and the channel's programmed transformation function (01 through 08) – the zoom function of Section 2.3, for instance. The combiner then merges all eight branches back into the single output plane, with every branch remaining on-axis throughout rather than diverging at an angle. This spatial recombination does not erase channel identity: each branch retains the distinct phase and polarization state its half-wave plate assigned it, so the eight channels remain individually distinguishable within the combined output rather than collapsing into a single indistinguishable beam.

2.3. Programmed Wavefunction Transformation and the Zoom Function

The encoder applies a programmed transformation to the wavefunction of the propagating image entirely within the optical domain – a zoom operation, a Fourier transform, a recognition function, or any comparable optical-domain algorithm, depending on what the operator loads into the stage. For converting a flat two-dimensional source into a real-space three-dimensional display, the zoom function is the relevant case. Rescaling the image differently across different channels changes the effective viewing perspective of each channel, so that a viewer perceives an object built from several flat 2D images, each rescaled and repositioned appropriately, as a single three-dimensional form.

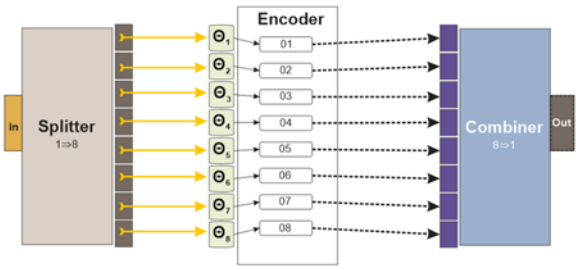


Fig. 1. Splitter–encoder–combiner architecture of the eight-channel HVGPU trigger, showing the polarization angles Θ_1–Θ_8 set by the half-wave plates and per-channel transformation functions 01–08 in the encoding stage. All eight branches propagate on-axis.

Achieving this rescaling optically, without introducing depth distortion, depends on a specific geometric condition. Distortion introduced by angular parallax appears in hologram reconstruction only when the illuminating beam's axis fails to coincide with that of the reference beam used at recording. Where that correspondence is exact, no such distortion occurs, regardless of how strongly the illuminating beam converges or diverges relative to the recording geometry. This is the criterion that makes the zoom function described above tractable directly through the optics. Because axis correspondence, rather than a fixed convergence angle, defines the reconstruction condition, varying the convergence or divergence of the illuminating beam changes the apparent depth of the reconstructed plane without disturbing the geometric correspondence that distortion-free reconstruction requires. This principle underlies the per-channel zoom function implemented in the encoder, and establishes why reconstruction must preserve the recording geometry of the U04 medium introduced in Section 1.4, for each of the eight channels.

2.4. Discrete Polarization States and the Placement of the Half-Wave Plates

The encoder's programmed transformation generates, among its other outputs, a set of distinct polarization states – one per channel. These are eight discrete polarization states, each corresponding to a fixed orientation of the linear polarization of the light in one channel. The number of such states equals the number of channels carrying a distinct polarization

angle. A linear polarization angle spans only the range between 0 and π. The half-wave plate's transversal plane possesses two orthogonal reference directions, the fast axis and the slow axis, perpendicular both to the optical axis of propagation and to the Poynting vector describing the direction of energy flow. Rotating the plate by an angle Θ/2 about the optical axis rotates the linear polarization of the transmitted light by Θ within that same transversal plane, so each channel is specified here by its polarization angle Θ. The eight polarization angles run from 0° to 70° in steps of 10°, which corresponds to plate orientations from 0° to 35° in steps of 5°; the half-wave plate angles quoted in [1, 2] are these polarization angles.

The half-wave plates that generate the eight polarization states sit at the output of the beam-splitter unit, as part of the encoder unit rather than of the splitter unit itself. The splitter unit transforms the single input channel into eight output channels (Section 2.2), and the half-wave plates, positioned immediately at that output, apply their rotation as the first operation the encoder unit performs on each of the eight channels. Because this set of plates belongs functionally to the encoder rather than to the splitter, the combiner remains a true mirror image of the splitter – the same optical design run in reverse – with no half-wave plates of its own at its input.

The rotation these plates apply must remain consistent with whatever other transformations – the zoom function of Section 2.3, for instance – the encoder applies simultaneously. Verification is therefore continuous, confirming that the combined effect of several transformations acting together does not drive any two channels to the same angle, and does not drive any channel past the π limit, beyond which a polarization angle repeats. The encoder already requires this real-time verification for its other programmed transformations regardless of polarization, so folding the rotation function's own consistency checks into that same process costs little in addition. The half-wave plate offers, in this role, the most direct available means of setting the electric field vector to a precise angle.

2.5. On-Axis Propagation within Each Channel

Each of the eight channels defined by the splitter–encoder–combiner assembly propagates along its own local optical axis rather than at an angle to it. The choice is deliberate. An off-axis channel geometry would introduce the distortion and aberration effects associated with off-axis propagation discussed in Section 1.2, and would require correction of the same general kind digital methods apply computationally when modeling off-axis diffraction [10]. Keeping every channel on-axis relative to its own local propagation direction avoids that correction burden at the architectural level, rather than compensating for it afterward.

2.6. Recording the HVGPU Function as Single-Color Holograms

Once assembled and characterized, the entire HVGPU function – splitter, encoder, and combiner together – can be recorded into a single holographic plate for a given color, condensing what is, in its discrete-component form, a fairly large optical bench into one compact element. Because each color is recorded separately (Section 1.4), the result is a set of three holograms: an R-HVGPU hologram, a G-HVGPU hologram, and a B-HVGPU hologram, each recorded with its own laser and each carrying the global wavelength function of the HVGPU encoded into that color's own spectral diffraction characteristic. The red hologram, for instance, records the global HVGPU function as expressed through the spectral diffraction efficiency curve obtained under red-laser recording; the green and blue holograms record the equivalent function at their own wavelengths. Additional holograms, recorded with additional coherent sources, extend color reproduction without altering the geometric parameters of the displayed image – a route to richer color rendering that leaves the zoom function and channel geometry established above untouched.

2.7. Physical and Virtual Channel Accounting Across the Trigger Assembly

The distinction between a physical channel and a color channel needs stating explicitly, since the two figures describing this system – 24 and 49 – arise from counting different things. A physical channel is one of the eight polarization-encoded paths within a single trigger; a color channel, or virtual channel, is one color propagating through one physical channel. Because all three colors travel through each of the eight physical channels of a given trigger, one trigger's eight physical channels yield 8 × 3 = 24 virtual color channels, consistent with the 24-channel figure already established for a single trigger [1].

Three triggers compose the full assembly. The first supplies the 24 positive-depth virtual channels; the second, an identical unit, supplies the 24 negative-depth virtual channels. Between them, this accounts for 48 virtual channels drawn from 16 physical channels (8 + 8). The third trigger is used differently: only one of its eight physical channels is active. Red, green, and blue propagate together through that channel as three spectral components of one beam and remain spectrally distinct. Because the three components perform the same optical function, namely transferring the reference image unmodified from the input to the output for comparison, we group them under one channel number and count them as a single reference plane. Counted by color, they are three virtual color channels. No component appears at an intermediate wavelength: a spectrometer in this channel records three lines at 632.8, 532, and 488 nm, and a detector or the eye receives the sum of the three

intensities, which is perceived as a mixed color by additive mixing.

Total physical channels: 8 (first trigger) + 8 (second trigger) + 1 (third trigger) = 17. Total focal planes: 24 + 24 + 1 = 49, the figure reported in [1], where the three color components of the reference plane share one plane number; counted by color, the assembly comprises 24 + 24 + 3 = 51 virtual color channels.

This reference plane serves as the calibration point for white balance across the other 48 virtual channels, and doubles as the plane-zero reference against which the ±24 depth-displaced planes are measured. It carries no zoom transformation and no color-specific processing of its own, passing the original frame from the source SLM through unmodified, and functions in that respect as an untransformed control against which the transformed planes can be checked. A single optical translation applied only to this reference plane moves the entire 49-plane package as a rigid unit relative to the source, since the spacing between the reference plane and each of the other 48 planes remains fixed. The whole assembly's position along the propagation axis thereby becomes a single adjustable parameter, rather than 48 independently tracked ones.

The finished HOT, in summary, comprises eight physical channels, corresponding to the eight polarization states of the light within one physical device, yielding 24 virtual color channels per device; three such devices, combined as described above, yield the complete 49-plane trigger assembly.

3. Spectral and Polarization Characterization Methodology

3.1. Measurement Configuration

Spectral and polarization characteristics of the assembled HVGPU system are measured across the 400–900 nm range using a three-unit Ocean Optics spectrometer suite (HR4000, HR2000+, FLMS). The measurement examines how the system distributes optical energy among the polarized channels in transmittance, reflectance, and diffraction. Because the encoder applies programmed instructions to the propagating wavefunction, the spectral and polarization response is expected to shift accordingly, tracking whatever function is currently loaded rather than remaining fixed.

The resulting data provide a direct check that each color-specific processor – R-HVGPU, G-HVGPU, and B-HVGPU – executes its programmed function correctly across the visible range. Each unit is, physically, a single hologram: the R-HVGPU, for example, is recorded with a coherent red laser on a U04 emulsion [13], operating through the spectral diffraction efficiency of that recording. Measuring its response across 400–900 nm therefore verifies both the recording and the correctness of the encoded wavefunction transformation.

We characterize the spectral selectivity of each color-specific HVGPU by the full width at half maximum, $\Delta\lambda$, of its diffraction-efficiency lobe in Figs. 2–4. Two functions executed at neighboring wavelengths remain resolvable when their center wavelengths differ by at least $\Delta\lambda$, so each optical "CPU" operates within its own spectral band. The values read from Figs. 2–4 are $\Delta\lambda_R = 15$ nm, $\Delta\lambda_G = 11$ nm, and $\Delta\lambda_B = 8$ nm for the red, green, and blue units.

The diffraction maxima of the three units lie at 627 nm (red), 522 nm (green), and 483 nm (blue), which is 5.8, 10.0, and 5.0 nm below the recording wavelengths of 632.8, 532, and 488 nm. At the recording lines the diffraction efficiency is therefore about 47 %, 18 %, and 29 %, compared with maxima of about 65–66 %. This shift is consistent with residual emulsion shrinkage after processing, which the swelling procedure of Section 1.4 is designed to compensate [8]. Because color balance is sensitive to shifts of a few nanometers [8], it is not optimized in the present prototype; aligning the diffraction maxima with the recording wavelengths is an objective for optimization in subsequent work.

For each of the three color holograms, transmittance, reflectance, and diffraction efficiency were measured as a function of reconstruction beam power, to characterize the operating range over which the recorded gratings respond linearly before saturation or measurable degradation sets in.

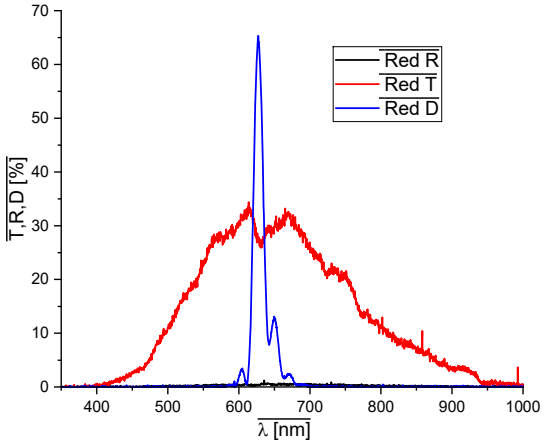


Fig. 2. Spectral dependence of transmittance, reflectance, and diffraction for the Red HVGPU (632.8 nm).

3.2. Anticipated Sources of Channel Non-Uniformity

The measurement campaign described above was designed, in part, to characterize numerically a set of non-uniformities already documented at the fabrication stage of the underlying trigger [1]. Diffraction efficiency across the eight physical channels of a single trigger was previously found to range between roughly 20 % and 50 %, averaging near 35 % – a spread consistent with the sequential-exposure mechanism discussed in Section 1.4. In the

wavelength-multiplexed recordings of [8], successive exposures on one plate lowered the diffraction efficiency relative to a simultaneous exposure; we expect the same mechanism when the eight channels are recorded through successive half-wave plate exposures. Where efficiency falls toward the lower end of the observed range, a proportional increase in source illumination – no more than a factor of 2.5, well within the range of standard diode laser drivers – compensates the hologram's contribution to channel non-uniformity without requiring any change to the recording itself [1].

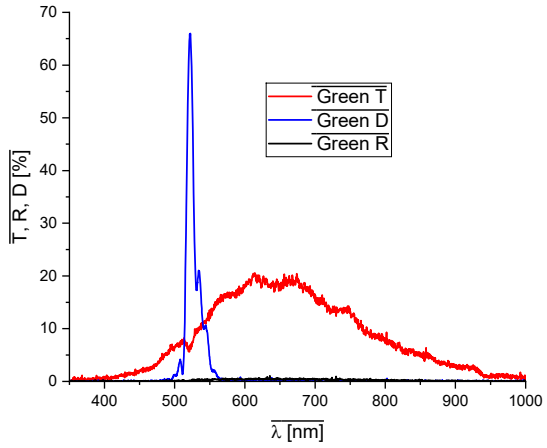


Fig. 3. Spectral dependence of transmittance, reflectance, and diffraction for the Green HVGPU (532 nm).

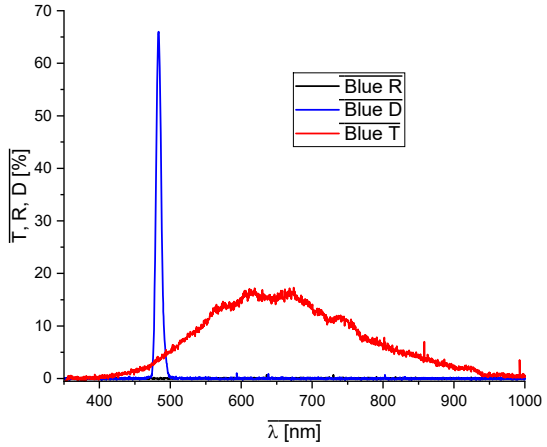


Fig. 4. Spectral dependence of transmittance, reflectance, and diffraction for the Blue HVGPU (488 nm).

We reproduce the measurements reported previously [1] only as the baseline against which the new HVGPU-level spectral and polarization measurements are evaluated. The new measurements extend the characterization to the assembled HVGPU, including the splitter and combiner stages, and to the response of the spectral energy distribution to the wavefunction transformations programmed into the encoder. We leave further functions and manipulations beyond the zoom operation described in Section 2.3 to subsequent work.

For the present experiments, we set a maximum crosstalk of 12.5 % between adjacent channels as the initial technical specification of the optical system. To meet it, we adjusted the polarization azimuth of each channel with a resolution of 0.2° and its ellipticity angle with a resolution of 0.4°, using a Thorlabs PAX1000VIS/M polarimeter, which also gave the degree of polarization. We measured each channel at the output of the beam combiner, where the optically manipulated color images of the eight physical channels of each color merge, with the other seven channels blocked. We define the overlap (crosstalk) as the fraction of the area of the narrower polarization ellipse that the ellipse of the physically adjacent channel of the same color covers, calculated from the azimuths and ellipticity angles read on the polarimeter. This measure describes the separability of the polarization states, not the transferred power. For example, the blue component had an azimuth (referenced to the horizontal axis) of 89.97° and an ellipticity angle of 0.25° in channel 1, and 79.91° and 0.71° in channel 2, a separation of 10.06°. The overlap of the two ellipses is approximately 9 %, below the specification, so these two channels needed no further adjustment. These values are approximate: they show that the overlap can be estimated with simple means, which is the purpose of this demonstration. All adjacent pairs of the three colors met the specification. The crosstalk values of [2] were obtained for different azimuth inclinations and polarization ellipses and therefore do not apply here.

4. Experimental Realization and Verification

This section presents direct experimental verification of the assembled trigger: confirmation that the light entering the HVGPU carries the correct input polarization state (Section 4.1), and photographic documentation of the physical trigger across successive stages of assembly and calibration (Section 4.2).

4.1. Input Polarization Conditioning and Verification

The polarization state at the HVGPU input is a controlled parameter, because the splitter cascade and the encoder both require a defined polarization state. In [2], the illumination was strictly linear, which that work identified as a critical condition. In the present cascade, one non-polarizing cube divides the amplitudes of the S and P components unequally, depending on the polarization direction of the incident light. To keep the division equal in both directions of the cube, we use left-circularly polarized light at the splitter input, which supplies equal S and P components. A quarter-wave plate at each cascade output then restores vertical linear polarization,

identical for all eight channels, as we verified with the polarimeter, so the encoder receives the same linear state as in [2]. These measures compensate for the cube within the splitter unit; the architecture of the trigger remains that of [1, 2]. Fig. 5 shows the input polarization state at the splitter entrance for each wavelength, upstream of the combiner output where we measure the crosstalk with the polarimeter. The polarization chain follows from Jones calculus in the cube-arm s–p basis. The left-circular input is

$$E_{\text{in}} = \left(1/\sqrt{2}\right)(1, i)^{\tau} \quad (1)$$

In arm k the cascade acts as:

$$C_k = diag(a_k, b_k\, e^{i\,\delta_k}), \quad (2)$$

where a_k and b_k are the amplitude factors for s and p and δ_k is their relative phase, so the arm output is:

$$E_k = \left(1/\sqrt{2}\right)\left(a_k, i\, b_k\, e^{i\,\delta_k}\right)^{\tau} \quad (3)$$

For an ideal non-polarizing cube a_k equals b_k and δ_k is equal to 0, and the arm output stays circular. A quarter-wave plate whose fast axis is aligned with a principal axis of the polarization ellipse converts an elliptical state to linear light, so the plate at the output of arm k is set with the polarimeter until the arm delivers linear polarization as close as possible to the vertical reference axis.

The residual azimuth error, determined by the amplitude imbalance, is absorbed in the setting of the half-wave plate. A half-wave plate whose fast axis is turned by $\Theta_k/2$ away from the input polarization then rotates this linear state by Θ_k, giving:

$$\mathrm{E}'_{\mathrm{k}} = (\sin\Theta_{\mathrm{k}}, \cos\Theta_{\mathrm{k}})^{\tau}, \quad (4)$$

up to a common phase factor, with

$$\Theta_{\mathrm{k}} = 0°, 10°, \ldots, 70°, \quad (5)$$

and a polarimeter azimuth of 90° minus Θ_k measured from the horizontal axis. The circular input therefore enters only through the splitting ratio of the cascade and the quarter-wave plate settings, and does not appear in the eight final states, which are linear. The states measured at the combiner output follow this prediction: the blue component had azimuths of 89.97° and 79.91° in channels 1 and 2 (Section 3.2), against 90° and 80° predicted.

4.2. Physical Assembly and Multi-Color Synchronization

Fig. 6 documents the trigger's physical assembly across three stages of adjustment. Fig. 6(a) shows the beam-splitter unit (left) and the beam combiner (right) positioned before adding the encoding stage between them – the configuration's starting point, prior to installing the half-wave-plate rotators and their associated transformation functions. Fig. 6(b) shows the multi-color output before channel synchronization: the red, green, and blue channels each appear as a separate, diffuse cloud of points rather than overlapping into a single spot. The combiner overlaps the arriving channels into a single spot (Section 2.2) only where they share spatial alignment, a condition the unsynchronized setup does not yet meet. Fig. 6(c) shows the trigger in its working, synchronized state, with the optical propagation paths of the blue and green channels visible along their designed trajectories through the assembly.

5. Discussion

5.1. Positioning Relative to Digital Computer-Generated Holography

The present approach and the digital CGH literature address the same problem through different means, and neither is presented here as a simple replacement for the other.

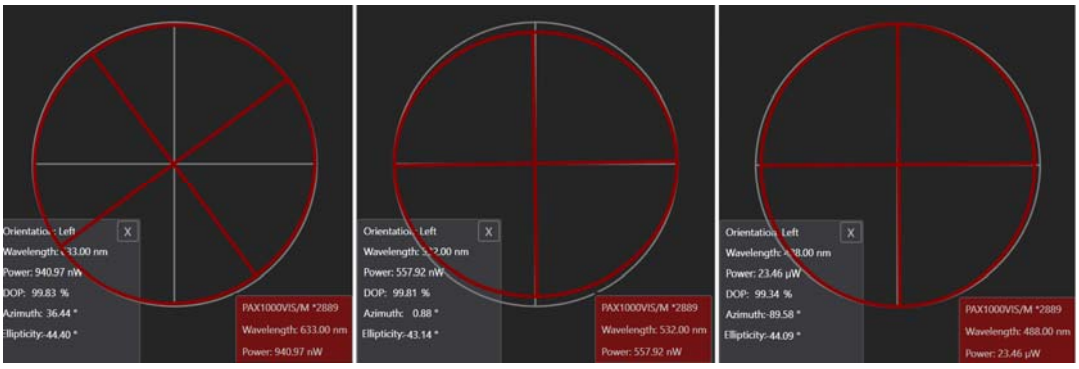


Fig. 5. Polarimeter measurements of the R–G–B input polarization states for the HVGPU at the splitter entrance for left-circularly polarized light.

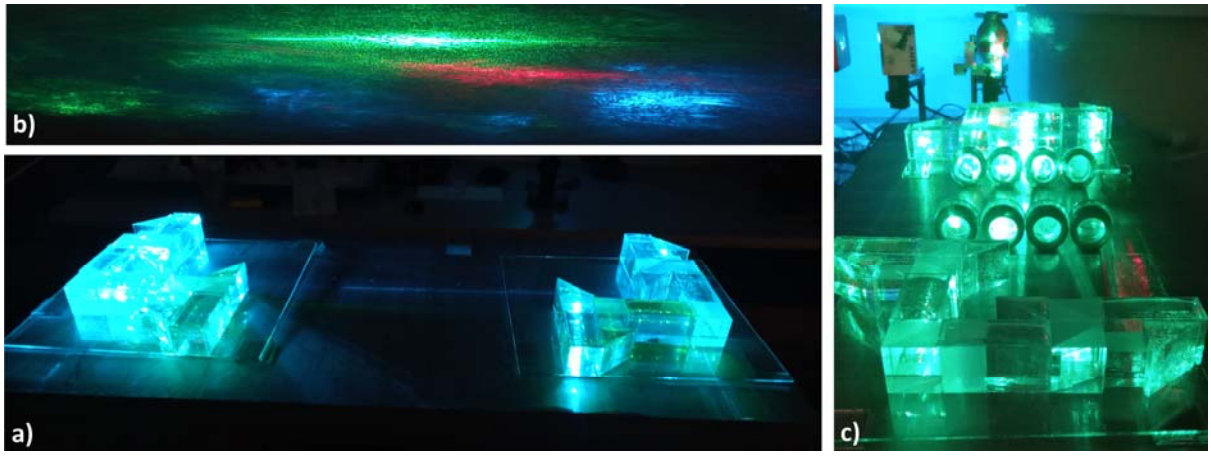


Fig. 6. Trigger assembly across three stages of adjustment: (a) beam-splitter (left) and beam combiner (right) positioned before encoder installation; (b) non-synchronized multi-color output, with red, green, and blue channels visible as separate point clouds; (c) working, synchronized state, showing the propagation paths of the blue and green channels.

A pattern recurs across the digital methods reviewed above: gains in accuracy tend to arrive with a disclosed, often substantial cost in computation time. A hybrid method combining planar and spherical wave primitives to more accurately model both diffuse and specular reflections features an analytically accurate configuration that runs roughly 60 times slower than a widely used overlap-add stereogram baseline [6]. Recovering real-time performance requires training a neural network, over several days of dedicated GPU time, to approximate that slower analytical method. A related line of work on reconstructing a dense light field from a sparse set of captured views – logically upstream of hologram synthesis itself – reports a comparable offline cost, on the order of 40 hours of training on a single high-end GPU before real-time inference becomes available [15]. A fully real-time digital pipeline for high-quality volumetric display, taken as a whole, consists of a chain of separately trained stages – sparse-view capture, learned dense reconstruction, hologram synthesis, phase-only conversion – each carrying its own accuracy–speed trade-off and its own offline training cost. Generating depth- and color-resolved channels within a single physical optical pass, as the present architecture does, requires no equivalent to any of these separately trained stages.

A formal evaluation grounded in wave-optics and sampling theory, rather than empirical benchmarking alone, demonstrates that several widely used image-based CGH methods – holographic stereograms, phase-added stereograms, and the overlap-add approach discussed in Section 1.1 – are different heuristic answers to the same mathematically posed problem. That problem is recovering a complex hologram from a light field by inverting a short-time Fourier transform, whose magnitude-squared formulation discards the phase information exact inversion needs [16]. Under that framework, the authors estimate that artifact-free, full-color reconstruction for a representative near-eye configuration requires on the order of several thousand sampled perspective views [16] – evidence that the accuracy–resolution trade-offs seen throughout this literature follow from the underlying computational formulation itself rather than from any one implementation's shortcomings. A separate method addressing computational cost through an adaptive combination of point-source and wave-field propagation primitives, chosen per depth layer according to local scene content, reports a computation-time reduction of up to 68 % relative to either primitive alone [17]. This reduction lowers the cost of digital synthesis, but it leaves the per-plane computation in place, and that computation remains the dominant cost for continuous volumetric video at high frame rates and high depth-plane counts.

The clearest single illustration of this trade-off, reported for one static two-dimensional hologram at one depth plane, comes from a formal gradient-based optimization framework for phase-only hologram computation. A common fast heuristic encoding technique runs in around 10 milliseconds but at markedly reduced quality, while recovering higher accuracy through iterative gradient-based optimization takes on the order of 30 seconds on comparable hardware [18]. The same authors note that extending their optimization to true multi-depth-plane holograms, by summing multiple computed layers, costs computation proportional to the number of planes, and state plainly that this does not scale well in practice. Rather than pursue simultaneous multi-plane reconstruction, they substitute gaze-contingent single-plane refocusing: eye tracking determines a single depth, and the system computes and displays only that depth at any given moment [18]. This is a considered and effective answer for the near-eye context that work addresses. It illustrates how differently the same multi-plane reconstruction challenge can be approached. A single dynamically refocused plane stands in for true simultaneous multi-plane content in that near-eye work, against

generating a larger number of physically distinct planes directly – the route the present architecture pursues, targeting 49 planes available simultaneously within a single optical pass. Motivating that target is the observation that a modest number of representation layers – five, in one representative neural CGH implementation [3] – may not by itself supply the depth resolution a free-space volumetric display requires. These five layers and the present 49 planes describe physically different quantities, however, a compact internal representation versus physically separated projection planes, and this distinction should be borne in mind when comparing the two figures numerically.

5.2. Structural Comparison to Other Multichannel and Multiplexing Strategies

The efficiency figures reported for the present prototype (Section 3.2) fit a pattern also seen in other reported multichannel volume-holographic devices. The coherent beam combiner discussed in Section 1.3 reports a measured combined-beam output roughly 20 % below its own coupled-wave-theory prediction, attributed to real beams departing from the idealized plane-wave assumptions the theory rests on, together with practically unavoidable imperfections in angular alignment [11]. That offers a documented precedent, from a different device, for the diffraction-efficiency shortfall relative to the recording material's specified theoretical maximum observed here, suggesting the shortfall may be a general feature of angle-, phase-, or wavelength-multiplexed volume-holographic devices rather than something specific to the present fabrication run.

Different lines of work on related multi-viewpoint or multichannel problems have arrived at different multiplexing strategies. One recent approach to eliminating the eyepiece in near-eye holographic displays generates several viewpoints within the eye's pupil through time-division multiplexing, displaying each viewpoint's hologram in sequence, fast enough to exploit visual persistence, to avoid the coherent interference that simultaneous display under shared illumination would otherwise produce [19]. The present architecture multiplexes its channels simultaneously instead – spatially, by polarization, and by wavelength, through physically parallel optical paths rather than a time sequence – a different resolution of the same underlying requirement to generate and separate multiple coherent channels without mutual interference. A smaller-scale precedent for solving a capacity limit through added physical hardware rather than algorithms alone appears in a cascaded dual-spatial-light-modulator configuration used to synthesize a higher-resolution hologram than either device produces alone [18]. On a more modest scale, this reflects the same design instinct – extending optical capability through added physical structure – underlying the present multichannel trigger.

Content-adaptive resource allocation represents a further, different strategy, applied within a different class of display technology. Multilayer, attenuation-based light-field displays, built from stacked LCD panels under incoherent illumination rather than coherent holographic recording media, benefit from weighting reconstruction fidelity according to human visual acuity across the field of view [20]. For displays with a hard physical ceiling on reproducible depth range, they benefit further from automatically positioning that limited budget around the most visually salient scene content rather than a fixed window [21]. These represent a different philosophy applied to a related resource problem: dynamically prioritizing a small, fixed budget by scene content, rather than physically increasing the number of available channels as pursued here. This article notes them for the breadth of strategy the wider display research community has brought to related constraints, given how different the underlying display physics remains between the two device classes.

5.3. On Experimental Reproducibility and Calibration

The efficiency and crosstalk figures reported for the present prototype (Section 3.2) are characteristic of the fabrication run and emulsion batch tested, because the response of the emulsion varies between batches and over the life of a plate (Section 1.4). Any replication or extension of the architecture, including the higher channel counts discussed in Section 1.4, should therefore include routine recalibration.

5.4. Summary

Across the digital CGH methods reviewed in this article, substantial engineering effort has gone into narrowing the trade-off between reconstruction accuracy, achievable frame rate, and the number of depth planes reconstructed simultaneously, and each method shows incremental progress along one or more of these axes. The present architecture addresses the same underlying multi-depth-plane, multi-color reconstruction problem by a fundamentally different route: physically parallel, all-optical multiplexing rather than sequential digital computation. This route removes the per-plane computation from the frame pipeline, at the cost of the fabrication complexity and calibration burden discussed in Sections 2 and 5.3. The throughput remains bounded by the refresh rate of the source of the input frames, a HOLOEYE LETO-3-VIS-009 phase-only spatial light modulator [22], whose manufacturer-specified input frame rate is 60 Hz, or 180 Hz in color-field-sequential mode. The prototype of [1] operates it at 90 monochrome frames per second, which corresponds to 4410 depth-plane sub-frames per second for 49 planes. A complete optical power budget, source modulation,

synchronization, detector response, and scene updating are not characterized in this article. Neither route displaces the other outright: digital CGH retains an advantage in flexibility, since a trained model or algorithm can synthesize arbitrary new scene content without new optical hardware, whereas the HVGPU requires physical fabrication for new channel configurations but delivers its full channel count and depth-plane set from each input frame without per-plane computation, at the frame rate of the source. Both directions read best as complementary parts of a shared effort toward practical, real-time volumetric holographic display.

6. Conclusions

The architecture developed in this work demonstrates that physically parallel, all-optical multiplexing can address the multi-depth-plane, multi-color reconstruction problem without the per-plane computation that constrains digital CGH approaches. By performing depth- and color-channel multiplication entirely within the optical domain, the holographic HVGPU offers a route toward real-time volumetric reconstruction that scales with fabrication rather than computation, with the frame rate bounded by the source of the input frames. Although this approach carries its own fabrication and calibration demands, it complements algorithmic CGH methods by providing a hardware pathway to high-speed, multi-plane holography. Two measured results define the next optimization steps: the diffraction maxima lie 5–10 nm below the recording wavelengths, and swelling shifts of ±50 nm place some of the nine reconstruction wavelengths closer together than the measured $\Delta\lambda$. Aligning the maxima and spacing the shifts are objectives for subsequent work. Together, these directions outline a convergent trajectory toward practical, real-time volumetric display systems capable of supporting dynamic, multi-color holographic scenes.

References

[1]. G. Ivanova, B. Ivanov, D. Dimov, A 24-channel optical trigger for use in 3D holographic display systems, in *Proceedings of the 9th International Conference on Optics, Photonics and Lasers (OPAL)*, 2026, pp. 35-39.

[2]. B. Ivanov, G. Ivanova, D. Dimov, Design and fabrication of a 24-channel holographic optical trigger for high-speed optical computation, in *Proceedings of the 9th International Conference on Optics, Photonics and Lasers (OPAL)*, 2026, pp. 86-91.

[3]. L. Shi, B. Li, C. Kim, P. Kellnhofer, et al., End-to-end learning of 3D phase-only holograms for holographic display, *Light: Science & Applications*, Vol. 11, Issue 1, 2022, 247.

[4]. N. Padmanaban, Y. Peng, G. Wetzstein, Holographic near-eye displays based on overlap-add stereograms, *ACM Transactions on Graphics*, Vol. 38, Issue 6, 2019, 214.

[5]. A. Gilles, P. Gioia, R. Cozot, L. Morin, Computer generated hologram from multiview-plus-depth data considering specular reflections, in *Proceedings of the IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, 2016, pp. 1-6.

[6]. J. Mäkinen, U. Gudelek, L. Varjo, E. Sahin, et al., Computer-generated holography using hybrid planar-spherical wave primitives, *Optics Express*, Vol. 34, Issue 8, 2026, pp. 14780-14800.

[7]. S. Mustafi, T. Latychevskaia, Fourier transform holography: A lensless imaging technique, its principles and applications, *Photonics*, Vol. 10, Issue 2, 2023, 153.

[8]. B. Ivanov, M. Shopova, E. Stoykova, V. Sainov, Reconstruction of silver halide volume reflection holograms at the wavelengths of recording, *Optica Applicata*, Vol. 42, Issue 3, 2012, pp. 643-650.

[9]. B. Ivanov, M. Shopova, A. Baldjiev, E. Stoykova, et al., Four-channel holographic infrared optical element, *Proceedings of SPIE*, Vol. 7747, 2011, 77470K.

[10]. H. Wei, X. Liu, X. Hao, E. Y. Lam, et al., Modeling off-axis diffraction with the least-sampling angular spectrum method, *Optica*, Vol. 10, Issue 7, 2023, pp. 959-962.

[11]. H. N. Yum, P. R. Hemmer, A. Heifetz, J. T. Shen, et al., Demonstration of a multiwave coherent holographic beam combiner in a polymeric substrate, *Optics Letters*, Vol. 30, Issue 22, 2005, pp. 3012-3014.

[12]. D. Nazarova, L. Nedelchev, N. Berberova-Buhova, G. Mateev, Nanocomposite photoanisotropic materials for applications in polarization holography and photonics, *Nanomaterials*, Vol. 13, Issue 22, 2023, 2946.

[13]. P. Gentet, Y. Gentet, S.-H. Lee, Ultimate 04: The new reference for ultra realistic color holography, in *Proceedings of the International Conference on Emerging Trends & Innovation in ICT (ICEI)*, 2017, pp. 162-166.

[14]. P. Gentet, M. Coffin, Y. Gentet, S.-H. Lee, From text to hologram: Creation of high-quality holographic stereograms using artificial intelligence, *Photonics*, Vol. 11, Issue 9, 2024, 787.

[15]. U. Gudelek, E. Sahin, A. Gotchev, Perceptually optimized model for near eye light field reconstruction, in *Proceedings of the IEEE International Workshop on Multimedia Signal Processing (MMSP)*, 2023, pp. 1-6.

[16]. U. Gudelek, E. Sahin, A. Gotchev, A simulation framework for perceptual characterization of holographic near-eye displays, *Proceedings of SPIE*, Vol. 13821, 2026, 1382105.

[17]. A. Gilles, P. Gioia, R. Cozot, L. Morin, Complex modulation computer-generated hologram by a fast hybrid point-source/wave-field approach, in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 2015, pp. 4962-4966.

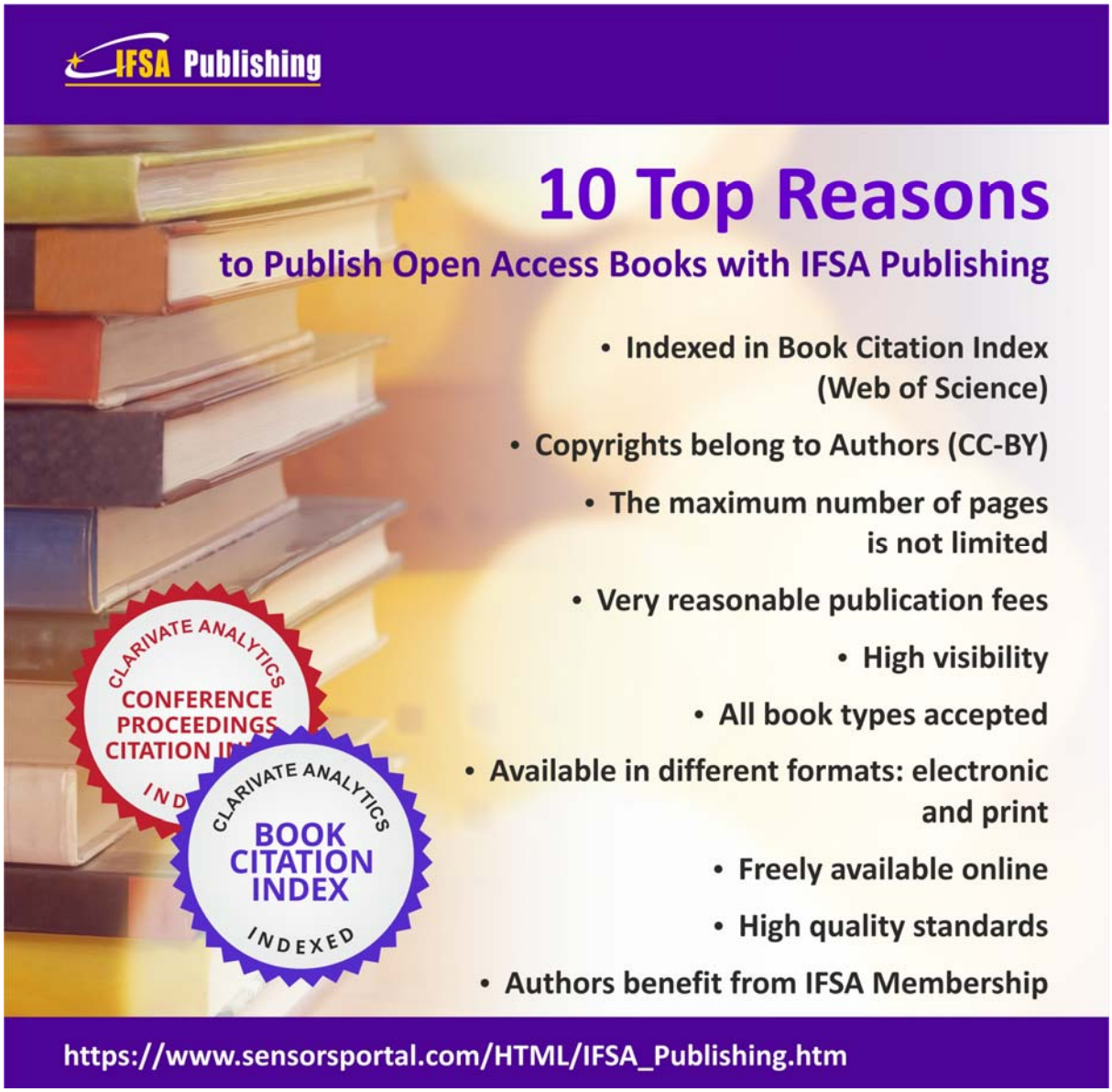
[18]. P. Chakravarthula, Y. Peng, J. Kollin, H. Fuchs, et al., Wirtinger holography for near-eye displays, *ACM Transactions on Graphics*, Vol. 38, Issue 6, 2019, 213.

[19]. J. Zhou, S. Xie, L. Jiang, Y. Luo, et al., High-quality, depth-cue-enhanced lensless holographic near-eye displays via pupil optimization, *Optics Letters*, Vol. 50, Issue 21, 2025, pp. 6811-6814.

[20]. J. Ding, M. Liu, Q. Zhong, H. Li, et al., Optimization algorithm of near-eye light field displays based on

human visual characteristics, *Chinese Optics Letters*, Vol. 14, Issue 4, 2016, 041101.
[21]. S. Wang, W. Liao, P. Surman, Z. Tu, et al., Salience guided depth calibration for perceptually optimized compressive light field 3D display, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 2031-2040.
[22]. LETO-3 phase only spatial light modulator, Datasheet, HOLOEYE Photonics AG, https://holoeye.com/products/spatial-light-modulators/leto-3-phase-only/

Published by International Frequency Sensor Association (IFSA) Publishing, S. L., 2026 (http://www.sensorsportal.com).



Sensors & Transducers, Vol. 274, Issue 3, October 2026, pp. 55-60

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2026
http://www.sensorsportal.com

Berry Phase and Statistical Properties of the Displaced Fock and Squeezed States

[1, *] Juan Carlos VEGA and [2] Didier OJEDA-GUILLÉN

[1] Escuela Superior de Física y Matemáticas, Instituto Politécnico Nacional, Ed. 9 U. P. Adolfo López Mateos, Alc. Gustavo A. Madero, C.P. 07738, Ciudad de México, México

[2] Escuela Superior de Cómputo, Instituto Politécnico Nacional, Av. Juan de Dios Bátiz esq. Av. Miguel Othón de Mendizábal, Col. Lindavista, Alc. Gustavo A. Madero, C.P. 07738, Ciudad de México, México

E-mail: jvegap1601@alumno.ipn.mx

Received: 1 Sept. 2026 /Revised: 25 Oct. 2026 /Accepted: 28 Oct. 2026 /Published: 30 Oct. 2026

Abstract: We study the displaced Fock states and squeezed displaced states by calculating their Berry phase when their parameters are explicitly time dependent. Furthermore, we analyze their statistical properties to identify the parameter regimes associated with classical and non-classical behavior. To this end, we assume that both the displaced Fock states and the squeezed displaced states are an explicit function of time where their parameters vary in time to obtain the geometrical phase. Then, we calculate the Mandel Q-parameter and the second-order correlation function $g^{(2)}(0)$ for each state using the similarity transformations of the ladder operators that allow us to obtain the expectation values needed.

Keywords: Displaced Fock states, Squeezed displaced states, Berry phase, Mandel Q-parameter, Second-order correlation function, Classical and non-classical behavior.

1. Introduction

In quantum optics, there is a growing interest in non-classical states of light. The harmonic oscillator is a key tool for these studies. By quantizing the electromagnetic field, we can map different optical systems to a harmonic oscillator [1-5]. This allows us to explore fundamental concepts in quantum mechanics.

We can define the states of light as eigenstates of Hamiltonians in a Fock space. Some of these states show classical properties, while others show non-classical properties [6]. Studying the quantum features of these states is very important in optics [2-6].

Another useful property of these quantum states is the Berry phase. This phase appears in many physical systems [3, 6, 7]. It is a geometric phase, it is a gauge invariant, and it connects closely with differential geometry [3, 6].

Furthermore, the study of these non-classical states of light is relevant for quantum metrology. Specifically, the reduction of photon number fluctuations helps to improve the signal-to-noise ratio in optical measurements because the noise is below the classical shot-noise limit. Therefore, understanding their geometrical phases and statistical properties provides a theoretical basis for future quantum transducers. The exact implementation of these states depends on the specific sensor architecture, which is beyond the scope of this fundamental analysis.

The aim of this work is to extend the study of the displaced Fock states and squeezed displaced states. The statistical results, including the Mandel Q-parameter and the second-order correlation function $g^{(2)}(0)$, were previously reported in a preliminary form in ref. [2]. The present article extends that conference paper by calculating the Berry phase when the state parameters are explicitly time-dependent.

While we have previously used the $su(1,1)$ algebra to calculate the Berry phase for the parametric amplifier [3], this manuscript applies the method to a different physical system. Furthermore, the geometric phase for generalized squeezed coherent states was established in ref. [6]. In Section 2, we derive our equations as special cases of their general expression. The new physical contribution of our specific treatment is that we use the algebraic approach to connect the exact geometric phase directly to the parameter boundaries that define classical and non-classical statistical behavior.

This work is organized as follows. In Section 2, we compute the Berry phase of both states. Section 3 reproduces the statistical properties from our previous work to identify the regimes where non-classical behavior appears. Finally, some concluding remarks are given.

2. Berry Phase

The geometric phase appears when we study a parameterized family of quantum states. We assume these states are written in terms of a set of parameters which are time-dependent and make up a closed loop C in the parameter space. After completing the cyclic evolution, the state returns to itself up to a phase factor. The geometric phase is defined as

$$\gamma_n(t) = i \oint_C \langle \varphi_n | \nabla_R | \varphi_n \rangle \, dR \tag{1}$$

This line integral, in general, is different to zero and depends on the path instead of how quickly it is traversed [3, 7, 8].

2.1. Berry Phase of the Displaced Fock States

Since we are mapping the states of light to eigenstates of Hamiltonians which can be defined in an appropriate Fock space, we need to use its basis states called Fock states, which are states with specific number of energy quanta, defined as [1, 2, 9, 10]

$$\hat{a}^\dagger \hat{a} |n\rangle = n|n\rangle, \tag{2}$$

where the number operator is $\hat{n} = \hat{a}^\dagger \hat{a}$ and the states $|n\rangle$ are the Fock states with $n = 0,1,2, \ldots$. Furthermore, the n photon state is defined as

$$|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}} |0\rangle, \tag{3}$$

with a factor that leads to the correct normalization of the state [1, 2, 10].

In particular, the harmonic oscillator states which minimizes the Heisenberg uncertainty relation, $\Delta p \Delta q \geq \hbar/2$, plays a significant role in the quantum theory of radiation and in quantum optics [2, 10-12].

These states which simultaneously minimize the variance in both the position and momentum operators, are called coherent states [2, 10, 11]. In [2-5, 12], the coherent states are defined as the action of the displacement operator on the vacuum state

$$|\alpha\rangle = \hat{D}(\alpha)|0\rangle, \tag{4}$$

where $\hat{D}(\alpha) = e^{\alpha \hat{a}^\dagger - \alpha^* \hat{a}}$ and $\alpha = |\alpha| e^{i\phi}$.

To compute the Berry phase, we assume that the parameter of the displaced Fock states is time dependent

$$|\alpha(R(t))\rangle = \hat{D}(\alpha(R(t)))|n\rangle, \tag{5}$$

where $\hat{D}(\alpha(R(t))) = e^{\alpha(R(t))\, \hat{a}^\dagger - \alpha^*(R(t))\, \hat{a}}$ and $\alpha\big(R(t)\big) = \alpha\big(\phi(t)\big) = |\alpha| e^{i\phi(t)}$. So, the geometrical phase is given by

$$\gamma = i \oint_C \langle \alpha\big(R(t)\big) | \nabla | \alpha\big(R(t)\big) \rangle \, dR, \tag{6}$$

Using Eq. (5), we obtain the following expression

$$\gamma = i \oint_C \langle n | \hat{D}^\dagger\big(\alpha(R)\big) \frac{d}{dR} \hat{D}\big(\alpha(R)\big) | n \rangle \, dR \tag{7}$$

Using the Hadamard identity which is defined as

$$\begin{aligned} e^A B e^{-A} = B + [A,B] + \tfrac{1}{2!}\big[A,[A,B]\big] + \\ + \tfrac{1}{3!}\Big[A,\big[A,[A,B]\big]\Big] + \cdots, \end{aligned} \tag{8}$$

we can calculate the similarity transformation of the operator $\frac{d}{dR}$ as follows

$$\begin{aligned} e^{\alpha^* \hat{a} - \alpha \hat{a}^\dagger} \frac{d}{dR} e^{\alpha \hat{a}^\dagger - \alpha^* \hat{a}} = \\ = \frac{d}{dR} - \dot{\alpha}^* \hat{a} + \dot{\alpha} \hat{a}^\dagger + \tfrac{1}{2}(\alpha^* \dot{\alpha} - \alpha \dot{\alpha}^*) \end{aligned}$$

Therefore, by substituting we obtain that the Berry phase is given as

$$\gamma = \tfrac{i}{2} \oint_C (\alpha^* \dot{\alpha} - \alpha \dot{\alpha}^*) \, dR, \tag{9}$$

where $\alpha^* \dot{\alpha} - \alpha \dot{\alpha}^* = 2i|\alpha|^2 \dot{\phi}$ and the displacement parameter satisfies ϕ: $[0,2\pi]$

$$\gamma = \tfrac{i}{2} \oint_0^{2\pi} 2i \, |\alpha|^2 d\phi = -2\pi |\alpha|^2 \tag{10}$$

Therefore, the Berry phase has been calculated. This result is a special case of the general expression presented in ref. [6], obtained here specifically for the displaced Fock states using the translation properties of the operators.

2.2. Berry Phase of the Squeezed Displaced States

Another harmonic oscillator state which minimizes the Heisenberg uncertainty relation are the squeezed states where the standard deviation of one observable

is squeezed, and the other conjugate observable is stretched exhibiting properties that closely resemble the classical harmonic oscillator [2, 10-12]. These states are defined by the combined action of the displacement and squeezing operators on the vacuum state

$$|\alpha,\xi\rangle = \hat{D}(\alpha)\hat{S}(\xi)|0\rangle, \tag{11}$$

where $\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}$ and $\hat{S}(\xi) = e^{\frac{1}{2}\left(\xi^*\hat{a}^2 - \xi\hat{a}^{\dagger 2}\right)}$ with $\xi = re^{i\theta}$ which satisfies $0 \le r < \infty$ and $0 \le \theta \le 2\pi$ [2-5,10-14].

Following a similar procedure to calculate the Berry phase of the displaced squeezed states, we need to assume that their parameters are time dependent

$$\begin{gathered}|\alpha(R(t)),\xi(R(t))\rangle = \\ = \hat{D}(\alpha(R(t)))\hat{S}(\xi\big(R(t)\big))|0\rangle,\end{gathered} \tag{12}$$

where $\hat{D}(\alpha(R(t))) = e^{\alpha(R(t))\,\hat{a}^\dagger - \alpha^*(R(t))\,\hat{a}}$ and $\hat{S}(\xi(R(t))) = e^{\frac{1}{2}(\xi^*(R(t))\hat{a}^2 - \xi\left(R(t)\right)\hat{a}^{\dagger 2})}$. So, the geometrical phase is given by

$$\gamma = i\oint_C \langle\alpha(R),\xi(R)|\nabla|\alpha(R),\xi(R)\rangle\, dR \tag{13}$$

Using Eq. (12), we obtain the following expression

$$\gamma = i\oint_C \langle 0|\,\hat{S}^\dagger(\xi)\hat{D}^\dagger(\alpha)\frac{d}{dR}\hat{D}(\alpha)\hat{S}(\xi)|0\rangle\, dR \tag{14}$$

Considering the generators of the $su(1,1)$ lie algebra defined as [3-5, 12]

$$\begin{gathered}K_0 = \frac{1}{2}\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right), K_+ = \frac{1}{2}\hat{a}^{\dagger 2}, \\ K_- = \frac{1}{2}\hat{a}^2,\end{gathered} \tag{15}$$

which satisfy $[K_0, K_\pm] = \pm K_\pm$ and $[K_-, K_+] = 2K_0$, we simplify the calculation of the similarity transformations

$$\begin{gathered}e^{\xi K_+ - \xi^* K_-}\frac{d}{dR}e^{\xi^* K_- - \xi K_+} = \\ = \frac{d}{dR} - \left(\dot{\xi}K_+ - \dot{\xi}^* K_-\right) + \left(\xi^*\dot{\xi} - \xi\dot{\xi}^*\right)K_0 - \\ -\frac{1}{3}\left(\xi^*\dot{\xi} - \xi\dot{\xi}^*\right)(\xi K_+ + \xi^* K_-) + \\ + \frac{1}{3}\left(\xi^*\dot{\xi} - \xi\dot{\xi}^*\right)|\xi|^2 K_0 + \\ -\frac{1}{15}\left(\xi^*\dot{\xi} - \xi\dot{\xi}^*\right)|\xi|^2(\xi K_+ + \xi^* K_-) + \\ + \frac{2}{45}\left(\xi^*\dot{\xi} - \xi\dot{\xi}^*\right)|\xi|^4 K_0 - \cdots,\end{gathered}$$

$$\begin{gathered}e^{\frac{1}{2}(\xi\hat{a}^{\dagger 2} - \xi^*\hat{a}^2)}(-\dot{\alpha}^*\hat{a} + \dot{\alpha}\hat{a}^\dagger)\, e^{\frac{1}{2}(\xi^*\hat{a}^2 - \xi\hat{a}^{\dagger 2})} = \\ = -\xi\dot{\alpha}^*\hat{a}^\dagger + \xi^*\dot{\alpha}\hat{a} + |\xi|^2\dot{\alpha}^*\hat{a} - |\xi|^2\dot{\alpha}\hat{a}^\dagger - \\ - |\xi|^2\xi\dot{\alpha}^*\hat{a}^\dagger + |\xi|^2\xi^*\dot{\alpha}\hat{a} + \cdots,\end{gathered}$$

$$\begin{gathered}e^{\frac{1}{2}\left(\xi\hat{a}^{\dagger 2} - \xi^*\hat{a}^2\right)}\left(\frac{1}{2}(\alpha^*\dot{\alpha} - \right. \\ \left. -\alpha\dot{\alpha}^*)\right)\, e^{\frac{1}{2}(\xi^*\hat{a}^2 - \xi\hat{a}^{\dagger 2})} = \frac{1}{2}(\alpha^*\dot{\alpha} - \alpha\dot{\alpha}^*)\end{gathered}$$

Since $\xi = |r|e^{i\theta}$, therefore $\left(\xi^*\dot{\xi} - \xi\dot{\xi}^*\right) = 2ir^2 d\theta$ and the similarity transformation can be written as

$$\begin{gathered}e^{\xi K_+ - \xi^* K_-}\frac{d}{dR}e^{\xi^* K_- - \xi K_+} = \\ \frac{d}{dR} + 2i\, sinh^2(r)\, d\theta K_0 - \left(\dot{\xi}K_+ - \dot{\xi}^* K_-\right) - \\ -\frac{1}{3}\left(\xi^*\dot{\xi} - \xi\dot{\xi}^*\right)(\xi K_+ + \xi^* K_-) - \\ -\frac{1}{15}\left(\xi^*\dot{\xi} - \xi\dot{\xi}^*\right)|\xi|^2(\xi K_+ + \xi^* K_-) - \cdots\end{gathered}$$

Now, using the action of these operators

$$\begin{gathered}K_+|k,n\rangle = \sqrt{(n+1)(2k+n)}|k,n+1\rangle, \\ K_-|k,n\rangle = \sqrt{n(2k+n-1)}|k,n-1\rangle, \\ K_0|k,n\rangle = (k+n)|k,n\rangle,\end{gathered}$$

the Berry phase is given as

$$\gamma = \mathrm{i}\oint_C \left(\frac{1}{2}(\alpha^*\dot{\alpha} - \alpha\dot{\alpha}^*) + \frac{1}{2}i\, sinh^2(r)\right)d\theta, \tag{16}$$

where the first term is the same as the Berry phase of the displaced Fock states. For the closed path C, we assume that the phase parameters ϕ and θ vary simultaneously from 0 to 2π. Thus, integrating over this cycle gives

$$\begin{gathered}\gamma = \frac{i}{2}\oint_0^{2\pi} 2i\,|\alpha|^2 d\theta + \frac{i}{2}\oint_0^{2\pi} i\, sinh^2(r)d\theta = \\ = -2\pi|\alpha|^2 - \pi sinh^2(r)\end{gathered} \tag{17}$$

The Berry phase has been calculated for this system. This equation is a special case of the general geometric phase established in [6], derived here through the explicit time-dependent evolution of the $su(1,1)$ generators.

3. Statistical Properties

Depending on the presence of some quantum features such as the relative fluctuation in photon number, the quantum states of light can be classified as classical and non-classical states [1, 2, 4].

Specifically, the Mandel Q-parameter allows us to measure the deviation from the Poisson distribution to distinguish quantum processes from classical ones because the photon number distribution for a coherent field is Poissonian, so, any distribution which is narrower than Poissonian must correspond to a non-classical field [1, 2, 4, 13]. This parameter is defined as

$$Q = \frac{\langle\hat{n}^2\rangle - \langle\hat{n}\rangle^2}{\langle\hat{n}\rangle} - 1 = \frac{\langle(\Delta n)^2\rangle - \langle\hat{n}\rangle}{\langle\hat{n}\rangle}, \tag{18}$$

where it describes a super-Poissonian distribution for $Q > 0$, a Poissonian distribution for $Q = 0$, and a sub-Poissonian distribution for $Q < 0$ [1, 2, 4, 13].

The second-order correlation function $g^2(0)$ quantifies the intensity fluctuations and provides

information on the bunching or the anti-bunching effects, is defined as

$$g^2(0) = \frac{\langle \hat{n}^2 \rangle - \langle \hat{n} \rangle}{\langle \hat{n} \rangle^2} = \frac{Q}{\langle \hat{n} \rangle} + 1, \quad (19)$$

where the bunching effect appears when $g^2(0) > 1$ and the anti-bunching effect when $g^2(0) < 1$ [1, 2, 4, 13-15].

3.1. Statistical Properties of the Displaced Fock State

By applying the Hadamard identity from Eq. (8), we find the translation properties for the annihilation and creation operators [2-5, 12-14]

$$\hat{D}^\dagger(\alpha)\hat{a}\hat{D}(\alpha) = \hat{a} + \alpha, \quad (20)$$

$$\hat{D}^\dagger(\alpha)\hat{a}^\dagger\hat{D}(\alpha) = \hat{a}^\dagger + \alpha^* \quad (21)$$

Next, we calculate the required expectation values to find the Mandel Q-parameter. We use the standard relations [2, 12]

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle, \quad (22)$$

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \quad (23)$$

Consequently, the expectation values are

$$\langle \hat{a}^\dagger\hat{a} \rangle = n + |\alpha|^2, \quad (24)$$

$$\langle (\hat{a}^\dagger\hat{a})^2 \rangle = n^2 + 4|\alpha|^2 n + |\alpha|^4 + |\alpha|^2 \quad (25)$$

This gives the following Mandel Q-parameter

$$Q = \frac{n(2|\alpha|^2-1)}{n+|\alpha|^2} \quad (26)$$

As illustrated in Fig. 1, a sub-Poissonian distribution occurs when $|\alpha|^2 < 1/2$. For larger values, we observe a super-Poissonian distribution. This transition applies strictly for $n > 0$. For the special case of $n = 0$, we obtain $Q = 0$ for all values of α. This is expected because the displaced vacuum state is exactly a coherent state. Fig. 2 displays the Mandel Q-parameter depending on the quantum number n and the displacement α.

Following this, we determine the second-order correlation function $g^2(0)$

$$g^2(0) = \frac{n(2|\alpha|^2-1)}{(n+|\alpha|^2)^2} + 1 \quad (27)$$

Antibunching effects are present when $|\alpha|^2 < 1/2$. This indicates that small displacements preserve the non-classical nature of the Fock state. However, for large displacements, classical behavior takes over. Fig. 3 shows that the exact transition happens at $|\alpha|^2 = 1/2$, where the system acts like a coherent state.

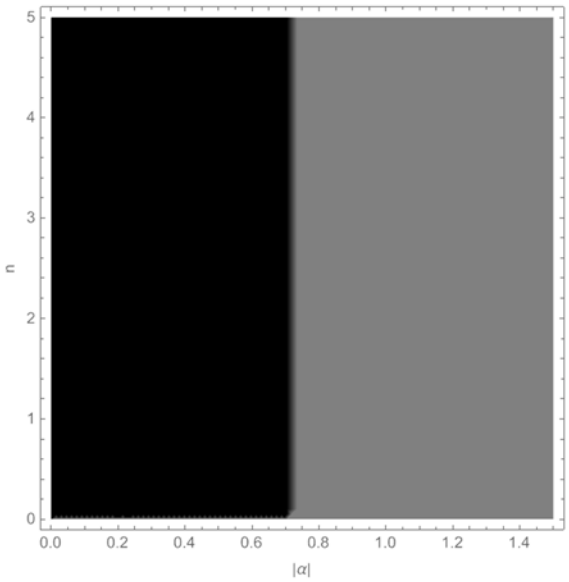


Fig. 1. Phase diagram of the Mandel Q-parameter for the displaced Fock state. Black region: $Q < 0$. Gray region: $Q > 0$.

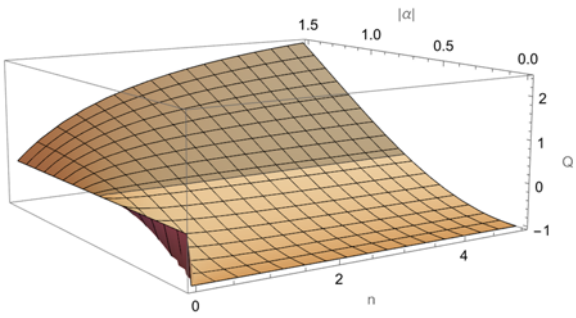


Fig. 2. Mandel Q-parameter for the displaced Fock state.

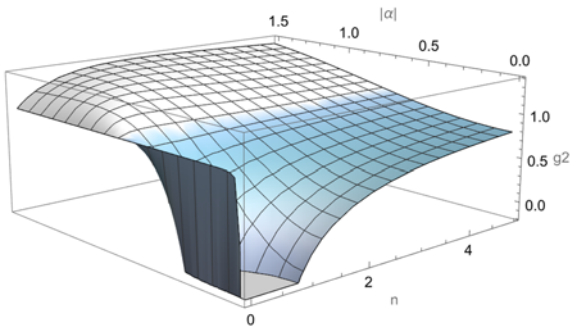


Fig. 3. Second-order correlation function $g^2(0)$ for the displaced Fock state.

3.2. Statistical Properties of the Squeezed Displaced State

With the help of Eqs. (8), (20) and (21), we derive the properties of the operators for this state [2, 12-14]

$$\hat{S}^{\dagger}(\xi)\hat{D}^{\dagger}(\alpha)\hat{a}\hat{D}(\alpha)\hat{S}(\xi) = \mu\hat{a} - v\hat{a}^{\dagger} + \alpha, \quad (28)$$

$$\hat{S}^{\dagger}(\xi)\hat{D}^{\dagger}(\alpha)\hat{a}^{\dagger}\hat{D}(\alpha)\hat{S}(\xi) = \mu\hat{a}^{\dagger} - v^{*}\hat{a} + \alpha^{*}, \quad (29)$$

where $\mu = \cosh(r)$ and $v = e^{i\theta}\sinh(r)$. Using these, the expectation values needed for the Mandel Q-parameter are

$$\langle \hat{a}^{\dagger}\hat{a} \rangle = |\alpha|^2 + sinh^2(r), \quad (30)$$

$$\langle (\Delta n)^2 \rangle = |\alpha|^2 \begin{pmatrix} e^{-2r}cos^2\left(\phi - \frac{\theta}{2}\right) + \\ +e^{2r}sin^2\left(\phi - \frac{\theta}{2}\right) \end{pmatrix} + \\ +2sinh^2(r)(sinh^2(r) + 1) \quad (31)$$

We define $\alpha = |\alpha|e^{i\phi}$. Then, the Mandel Q-parameter is

$$Q = \frac{|\alpha|^2(e^{-2r}-1)+2sinh^4(r)+sinh^2(r)}{|\alpha|^2+sinh^2(r)} \quad (32)$$

Here, we set $\phi = \theta/2$. This aligns the displacement with the squeezed quadrature to minimize the photon number fluctuations [2, 14].

In this case, it is possible to obtain both sub- and super-Poissonian distribution depending on the values of r and α, as shown in Fig. 4 and Fig. 5 [14, 15]. The exact boundary between these two regimes is determined analytically when the numerator of Eq. (32) is equal to zero: $|\alpha|^2(e^{-2r} - 1) + 2sinh^4(r) + sinh^2(r) = 0$. Supported by this condition, the super-Poissonian regime dominates for larger values of r. In contrast, the sub-Poissonian distribution appears for small values of r and $|\alpha|^2 \neq 0$.

Finally, we compute the second-order correlation function $g^2(0)$

$$g^2(0) = \frac{|\alpha|^2(e^{-2r}-1)+2sinh^4(r)+sinh^2(r)}{(|\alpha|^2+sinh^2(r))^2} + 1 \quad (33)$$

Depending on r and α, this state exhibits either bunching or antibunching effects, as seen in Fig. 6 [2, 13]. Bunching occurs exclusively when r is large. Conversely, antibunching is present for small r values if $|\alpha|^2 \neq 0$. This shows that adding a small amount of squeezing generates non-classical behavior.

4. Conclusions

In this article, we studied the geometric phase and the statistical properties of both displaced Fock states and squeezed displaced states. By assuming that the state parameters change over time and complete a closed path, we successfully calculated the exact Berry phase for both systems.

Through the Mandel Q-parameter and the second-order correlation function, we mapped the regions of classical and non-classical behavior. For the displaced Fock states, we found that non-classical features like sub-Poissonian statistics and antibunching dominate when $|\alpha|^2 < 1/2$. The transition to classical coherent behavior happens exactly at $|\alpha|^2 = 1/2$ for $n > 0$.

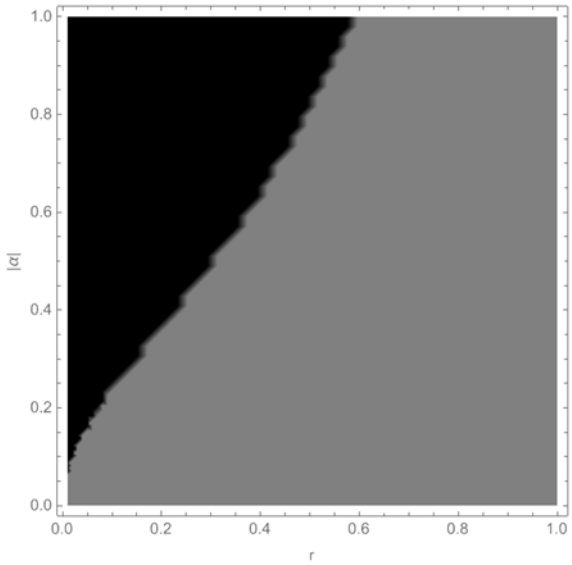


Fig. 4. Phase diagram of the Mandel Q-parameter for the squeezed displaced state. Blue region: $Q < 0$. Red region: $Q > 0$.

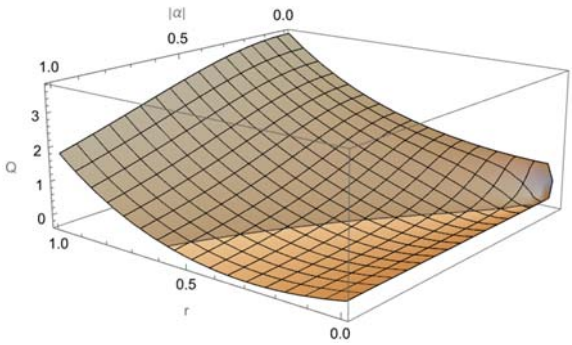


Fig. 5. Mandel Q-parameter for the squeezed displaced state.

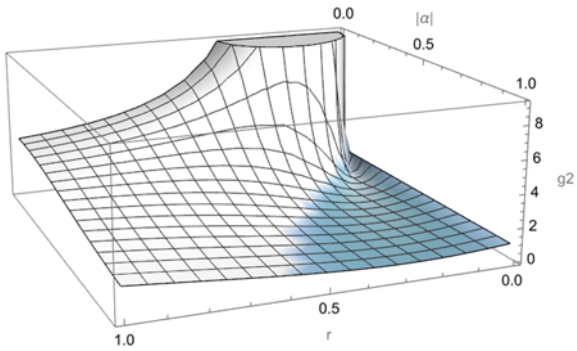


Fig. 6. Second-order correlation function $g^2(0)$ for the squeezed displaced state.

For the squeezed displaced states, we minimized the photon number fluctuations by setting $\phi = \theta/2$. Our analysis shows that this state can present both

sub-Poissonian and super-Poissonian distributions, as well as bunching and antibunching effects. A key finding is that a small squeezing parameter r is enough to produce non-classical effects, provided that the displacement is not zero.

The non-classical effects identified in this work, such as sub-Poissonian distributions and antibunching, are resources for quantum metrology. By finding the parameter regimes where the quantum noise is below the standard limit, our results show a theoretical advantage for improving the signal-to-noise ratio. The specific experimental implementation of these states depends on the transducer architecture, which is left for future applied research.

References

[1]. G. S. Agarwal, Quantum Optics, *Cambridge University Press, Cambridge*, 2013.

[2]. J. C. Vega, D. Ojeda-Guillén, Statistical properties of the displaced Fock and squeezed states, in *Proceedings of the 9th International Conference on Optics, Photonics and Lasers (OPAL)*, 2026, pp. 69-72.

[3]. J. C. Vega, E. Choreño, D. Ojeda-Guillén, R. D. Mota, Berry phase and the Mandel parameter of the non-degenerate parametric amplifier, *Journal of the Optical Society of America B*, Vol. 41, Issue 5, 2024, pp. 1084-1090.

[4]. J. C. Vega, D. Ojeda-Guillén, R. D. Mota, SU(1,1)×SU(2) approach and the Mandel parameter to the Hamiltonian of two oscillators with weak coupling, *Journal of the Optical Society of America B*, Vol. 42, Issue 3, 2025, pp. 683-692.

[5]. J. C. Vega, D. Ojeda-Guillén, R. D. Mota, Exact solution of the isotropic and anisotropic Hamiltonian of two coupled harmonic oscillators, *Indian Journal of Physics*, Vol. 99, Issue 5, 2025, pp. 5107-5116.

[6]. S. Seshadri, S. Lakshmibala, V. Balakrishnan, Geometric phases for generalized squeezed coherent states, *Physical Review A*, Vol. 55, Issue 2, 1997, pp. 869-875.

[7]. A. Shapere, F. Wilczek, Geometric Phases in Physics, *World Scientific*, Singapore, 1989.

[8]. D. J. Griffiths, D. F. Schroeter, Introduction to Quantum Mechanics (3rd ed.), *Cambridge University Press*, Cambridge, 2018.

[9]. M. Hofheinz, E. M. Weig, M. Ansmann, R. C. Bialczak, et al., Generation of Fock states in a superconducting quantum circuit, *Nature*, Vol. 454, Issue 7202, 2008, pp. 310-314.

[10]. C. Gerry, P. Knight, Introductory Quantum Optics, *Cambridge University Press*, Cambridge, 2005.

[11]. P. Meystre, Quantum Optics: Taming the Quantum, *Springer*, Cham, 2021.

[12]. D. Ojeda-Guillén, Mecánica cuántica de los estados coherentes y comprimidos, Master's Thesis, *Instituto Politécnico Nacional*, Mexico City, 2010 (in Spanish).

[13]. M. Orszag, Quantum Optics, Springer, Berlin, 2008.

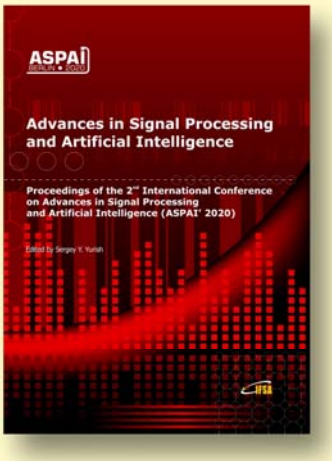
[14]. L. Mandel, E. Wolf, Optical Coherence and Quantum Optics, *Cambridge University Press*, Cambridge, 1995.

[15]. M. Fox, Quantum Optics: An Introduction, *Oxford University Press*, Oxford, 2006.

Published by International Frequency Sensor Association (IFSA) Publishing, S. L., 2026 (http://www.sensorsportal.com).

Advances in Signal Processing and Artificial Intelligence

Proceedings of the 2nd ASPAI' 2020 Conference



The proceedings contains all accepted and presented papers of both: oral and poster presentations at ASPAI' 2020 conference of authors from 23 countries. The coverage includes artificial neural networks, emerging trends in machine and deep learnings, knowledge-based soft measuring systems, artificial intelligence, signal, video and image processing.

Formats: hardcover (print book) and PDF (e-book), 264 pages
ISBN: 978-84-09-21931-5, e-ISBN: 978-84-09-21930-8
IFSA Publishing, 2020

https://www.sensorsportal.com/HTML/BOOKSTORE/ASPAI_2020_Proceedings.htm

Sensors & Transducers, Vol. 274, Issue 3, October 2026, pp. 61-75

www.sensorsportal.com

Sensors & Transducers

Published by IFSA Publishing, S. L., 2026
http://www.sensorsportal.com

Performance Analysis of Hybrid Beamforming Optimization Algorithms for mm-Wave Massive-MIMO Systems

[1] Girija SHANKAR SAHOO, [1] Anumoy GHOSH and [2,*] Abhishek SAHA

[1] Department of Electronics Communication and Engineering,
National Institute of Technology, Mizoram, India
[2] Department of Electronics and Communication Engineering,
Dream Institute of Technology, Kolkata, India
E-mail: abhishek.saha@dreaminstituteonline.edu.in

/Received: 9 May 2026 /Revised:21 June 2026 /Accepted: 19 Oct. 2026 /Published: 30 Oct. 2026

Abstract: The hybrid beamforming(HBF) in millimeter wave (mmWave) MIMO wireless communication system is an emerging technique used in modern communications system to fulfill the demandas of high data rate and low latency. This article investigates the best suited optimization technique to improve the beamforming gains thereby reducing the propagation loss. The investigation was conducted on the analysis and implementation of General Eigen Value Decomposition (GEVD), Orthogonal Matching Pursuit (OMP), and Manifold Optimization HBF techniques at precoder and combiner. Scattered channel model is used in this work. 2×2, 4×4, 8×8 and 16×16 massive MIMO antenna array structures are considered at the transmitter. Data streams ranging from 1 to 4 are considered. The performance is evaluated on the basis of spectral efficiency, mean square error (MSE) and Bit Error Rate (BER). Effects of massive MIMO array size and data streams on the performance metrics of the HBF algorithms are demonstrated with analysis. The performance of the HBF algorithms is compared with full digital beamforming technique to comprehend their relative advantages over the sophisticated and power hungry digital beamforming technique.

Keywords: Hybrid beamforming, Multi input multi output system, Manifold optimization, Orthogonal matching pursuit, Geeneralized eigenvalue decomposition.

1. Introduction

The next-generation mmWave wireless communication system is provide a promosing solution for modern communication systems that require extremely high data rates and bandwidth to accommodate large number of user and provide fast communication channels. These goals can be met if the communication spectrum is shifted to the mm-wave region using massive multi-input and multi output (MIMO) systems. The mm-wave massive MIMO system also increases the spectral efficiency and decreases latency by manifold [1]. Numerous numbers of antenna element use in massive MIMO system, generates significant interference that adversely affects the channel capacity and the performance of 5G communication systems [1]. Beamforming technique is employed to circumvent this problem which reduces the multipath fading and inter channel interference by concentrate the signal in desired direction [2]. This Beamforming realized by assembling the elements in the massive MIMO configuration in an organized array. Beam in a specific direction are added and subtracted in the other direction. Beamforming algorithm implemented at baseband and RF modules of the MIMO systems can also reduce path loss. Thus, beamforming is an effective technique which when used in conjunction

with massive MIMO system provides very high spectral efficiency and reduces the bit error rate [3].

1.1. MIMO

In the new in future of communication increasing demand of wireless communication services, there is an increased requirement of high data rate and energy efficiency. The massive MIMO wireless communication technology fulfills this requirement. Fig. 1, shows the massive MIMO cellular system with a station (BS) and multiple user equipment (UE) devices equipped with antenna array. Multiple UEs are scattered throughout the cell area. Each UE is equipped with multiple antennas, allowing it to receive signals with spatial diversity. Wireless communication link between BS and UEs with beamforming techniques are used to focus the transmitted signals towards the intended UE, optimizing signal strength and minimizing interference [4]. Successfully implement the mmWave massive MIMO technology support broad range of technology services including machine to machine (M2M), Internet on vehicles (IoV), device to devices (D2D) and backhaul access cellular network. Efficient beamforming techniques are to be employed to provide the capability of higher spectral and energy efficiencies.

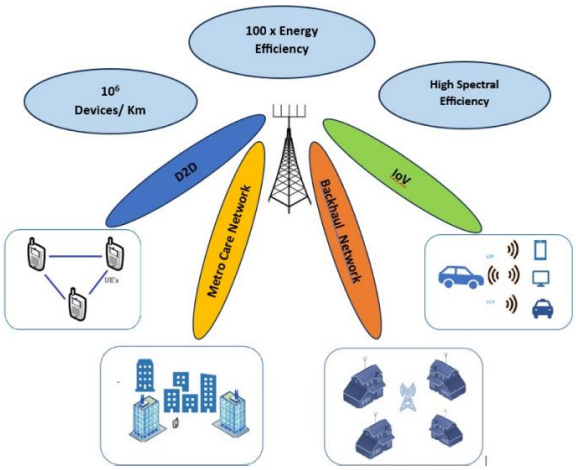


Fig. 1. An example of massive MIMO cellular communication system.

1.2. Hybrid Beamforming

Hybrid beamforming was achieved using initial stage the baseband precoder and second stage is RF precoder are used for pre signal processing and phase controls, respectively. The hybrid approach tradeoffs the advantages of both the configuration analog and digital beamformer. Based on the grouping of the antennas, there is cluster-based fully connected and partially connected antenna hybrid beamforming technique. In the cluster-based hybrid technique balance the hardware complexity and performance by reducing the number of RF chain [2, 5]. Fully connected hybrid beamforming increases the number of the RF chain, while it provides the highest level of the flexibility and adaptability [6]. In partially connected hybrid beamforming antennas are divided in small groups. Each group is connected to a shared RF chain and group have different RF beamforming weight to support spatial multiplexing. Partially connected beamforming provide the balance between performance and hardware complexity [5, 7].

Initially, analog beamforming was the preferred technique, but it gradually shifted to digital beamforming [8, 9]. In massive MIMO scenario, disadvantage of analog and digital beamforming with respect to hardware complexity and power requirement respectively [10] becomes significant. Thus, the hybrid beamforming (HBF) technique, where digital baseband precoder and analog RF precoder [10] avoiding the disadvantages of both. HBF as compared to full digital beamforming is difficult in term of optimizing four variables at baseband and RF module [11, 12]. Existing problems on HBF algorithms solve by breaking actual problem into subproblem for precoding and combining section in the transmitter and receiver sections [13]. One popularly used approach to hybrid beamforming design is singular value decomposition and Euclidean distance of hybrid beamformer with optimal beamforming [14]. Orthogonal matching pursuit (OMP) is used for spatially sparse precoding and combining algorithm to solve by matrix factorization problem [14]. In [15], a low complexity manifold optimization algorithm discussed, other than the matrix factorization approach. Iterative algorithm is designed for closed form solution of beamformer with original objective function [14, 16].

All the previous works in the development of HBF algorithms are based on maximizing the spectral efficiency. Other important performance metrics beside the spectral efficiency are Mean Square Error (MSE) and Bit Error Rate (BER) that can be controlled by the optimization of precoding and combining weights at precoder and RF sections respectively [17-19]. In practical systems, some modulation scheme is used for data transmission and thus MSE is an important metric to measure the transmission reliability [20]. Signal to interference noise ratio (SINR) and BER metrics are related to MSE in the beamforming design for multiple data. Thus the MSE and BER metrics as alternate objectives of spectral efficiency [18-20] for HBF algorithms to compare their performance with respect to full digital beamforming. Precoding matrices optimized by minimizing the MSE can also improve the spectral efficiency [21]. The traditional low complex Othogonal Matching Pursuit (OMP) dictionary based beamselction method to approximate the digital precoder. The four variables on the beamformer are optimized instead of factorizing the optimal full digital beamforming. HBF is optimized the MMSE enhance the performance [22]. In the HBF design approach, the minimization problem of MSE is decomposed into

subproblem hybrid precoding and combining. The selection of hybrid beamforming algorithm based on reduced complexity and better performance. Manifold Optimization (MO) Hybrid Beamforming approach to minimized the Euclidean distance between the optimal and hybrid beamformer to achieve the convergence [23]. The purpose of the Hybrid beamforming algorithm (HBF) is optimizing the MSE without any constraint of the space feasible solution and achieve the higher performance. We seek low complexity algorithm Generalized Eigen Value Decomposition (GEVD) beamforming matrix optimize column wise column to avoid the complexity and achieve the high data rate. Key issues and problem statement identified in previous work as follows:

- In traditional beamforming technique, large scale antenna array system encounters the high computational complexity and suboptimal performance. Hybrid beamforming, which combine digital and analog beamforming to strike balance between performance and complexity;
- Hybrid Beamforming algorithm optimize the beamforming process by incorporate the MMSE criteria to mitigate the interference and noise. While maximizing signal quality, thus improving overall system performance and robustness to the channel variation;
- Practical implementation of hybrid beamforming, complexity analysis is crucial parameter to tradeoff between sustainability of system and performance.

In this article several hybrid beamforming technique discussed for mmWave massive MIMO system. The Orthoganal Matching Pursuit (OMP) hybrid beamforming leverage the channel sparsity, where only few dominant path exist to approximate the optimal digital precoder by using dictionary based selection. The OMP hybrid beamforming has lower computational efficiency and its performance degarde with increasing data stream and channel correlation. The itrative based Manifold Optimization (MO) and Generalized Eigen Value Decomposition (GEVD) hybrid beamforming better performamnce over traditional beamforming technique. The GEVD Hybrid beamforming method gives the attractive trade-off between and complexity.

The most of the studies focus on either maximization of spectral efficiency or MMSE optimization under fixed antenna configuration and simulation assumption. Which makes difficult for direct comparison and to choose better hybrid beamformer configuration. Therefore, our work present unified framework to measure the performance of OMP, MO and GEVD hybrid beam forming algorithm under identical channel condition, RF chain constraint and power allocation. The contribution of our work summarized as follows:

- Design and analysis of 2×2, 4×4, 8×8 and 16×16 rectangular massive MIMO antenna array operating 28 GHz frequency;
- Implementation of Hybrid beamforming algorithm OMP, MO and GEVD under common MMSE base framework;
- Investigate the performance metric spectral efficiency, MSE, BER and Energy efficiency computational complexity varying antenna array size and data stream;
- Compare the performance and trade-off between performance and computation cost among the hybrid beamforming algorithm. The work identify selection of practical beamformer for future mmWave massive MIMO system.

This work extends our previous study in the persented in the IEEE Wireless Antenna and Microwave Propogation (WAMS 2022) [21], where antenna array design for hybrid beamforming investigated. In the persent work, we significantly extends that study by implementation of hybrid beamforming algorithms and evaluating the performance of these algorithm in term of spectral efficiency (SE), mean square error (MSE), bit error rate (BER) and energy efficiency (EE). This article is structure as follows: Section 2 elucidates the channel modeling and massive MIMO antenna array considered for the proposed work. Section 3 gives the detailed analysis of the three HBF algorithms and deriving the SINR, SE, MSE and BER thereof. Section 4 illustrates the simulation results relying on the analysis in Section 3 and presents a performance evaluation and comparison among the HBF algorithms based on SE, MSE and BER. Section 5 concludes this work with final comments. Throughout this article, uppercase letter and lowercase letters are represent the matrices, column vector and scalar value. The subscripts $()^T$,$()^*$ and $()^H$ represent transpose matrix, complex conjugate and hermitian matrix. tr() and ||.|| denote the trace of the matrix and Euclidean norm of the vector respectively. Conjugate gradient function is $\nabla()$ and E{} indicates expectation operator. $[A]_{i,j}$ denotes the $i-j^{th}$ entry of a matrix A.

2. Designing and Modelling of Massive MIMO System

The massive MIMO mmWave communication technology increases the spectral efficiency and channel capacity significantly by increase the antenna element. This technique encountered with certain challenges such as interference caused by deployment of large antenna array. Beamforming technique popularly used to mitigate the interference challenge in massive MIMO system [12, 14]. This articler discussed, the shape and size of antenna array play pivotal role in characteristics channel model. In MIMO system, spacing between antenna element in rectangular antenna array determine the mutual coupling and channel matrix. For scenario user equipment positioned very closely, correlation between user equipment higher as compared to independent identical distributed (i.i.d) Rayleigh

channel. This can be minimized and uncorrelated by large antenna array.

2.1. Channel Model

The correlation-based channel model used to analyze performance and the statistical properties of wireless communication. it is also used for the design theoretical model when UE employs single antenna further extended to UE with multiple antenna. A list of symbols is presented in Table 1.

Table 1. List of Symbols.

Symbol	Definition
G	Channel Matrix
D	Propagation Matrix
β_k	Large scale fading factor
ϕ	Gain constant
d_k	Distance between the k_{th} UE and the base station
ξ_k	Shadow fading component
H	Fast fading channel matrix
h_k	Channel vector for $\boldsymbol{k^{th}}$ UE
α	Path loss exponent
R_k	Steering Vector at different angle AoA
$\theta_{k,i}, \phi_{k,i}$	Elevation and Azimuthal AoA
λ	Carrier wavelength
d	Distance between adjacent antennas

A uplink channel uses K number UEs with single antenna and base station equipped with N antennas, where NLoS channel formulize as [24]

$$G = H \times D^{1/2}, \quad (1)$$

where D is the propagation matrix given as [24].

$$D = diag\{\beta_1, \beta_2, \dots, \beta_k\}, \quad (2)$$

where $\beta_k = \phi d_k^{-\alpha} \xi_k$, Φ is gain constant, d_k is the distance between k^{th} UE to base station, α is the path loss, and ξ_k is fading loss component with $10\log 10\xi_k \sim N(0, \sigma_{sh}^2)$ and $H \in \zeta^{N \times K}$ is fading channel matrix.

The elements of the channel matrix $H = [h_1, h_2 \dots, h_k]$. Here, antenna correlation is determined by clustering and scattering environments [25]. Where the channel vector h_k for the k^{th} UE is correlation matrix multiplied by standard complex Gaussian vector as

$$h_k = R_k v_k;\ k = 1,2, \dots k \quad (3)$$

Steering matrix is $R_k = \zeta^{N \times D_k}$, where D_k steering vectors and complex gaussian vector is $v_k \sim \xi N(0, I_{D_k})$. Steering matrix R_k for k^{th} user defined as [24]

$$R_k = \frac{1}{D_k}[a(\theta_{k,1},), a(\theta_{k,2}), \dots, a(\theta_{k,D_k})] \quad (4)$$

Steering vector $a(\theta_{k,i}) \in \zeta^{N \times 1}$ at Angle of Arrival $\theta_{k,i}$ for k^{th} user is given as [24].

$$a(\theta_{k,i}) = [1, e^{(\frac{j2\pi d}{\lambda})\sin\theta_{k,i}}, \dots, e^{(\frac{j2\pi(N-1)d}{\lambda})\sin\theta_{k,i}}]^T, \quad (5)$$

where d and λ are distance between antenna element and carrier wavelength respectively. Here, carrier wavelength corresponds to the free space wavelength for 28 GHz. For the rectangle antenna phased array, the steering vector steering matrix defined by [25]

$$a(\theta_{k,i}, \phi_{k,i}) = \\ = vec\{[1, e^{(\frac{j2\pi d}{\lambda})\sin\theta_{k,i}}, \dots, e^{(\frac{j2\pi(N-1)d}{\lambda})\sin\theta_{k,i}}] \otimes \\ \otimes [1, e^{(\frac{j2\pi d}{\lambda})\sin\phi_{k,i}}, \dots \dots, e^{(\frac{j2\pi(N-1)d}{\lambda})\sin\phi_{k,i}}]\}, \quad (6)$$

where $\theta_{k,i}$ and $\phi_{k,i}$ represent elevation AoA and azimuthal AoA, respectively. $vec\{.\}$ represents vectorization of matrix. In massive MIMO system, as the antenna array increase in limited space, mutual coupling also considerable in determination of channel vector redefined as [24]

$$h_k = ZR_k v_k, where\ k = 1,2,3, \dots \dots, K, \quad (7)$$

where steering matrix defined as $R_k \in \zeta^{N \times D_k}$, D_k steering vectors and $v_k \sim \zeta N(0, I_{D_k})$. According to [14, 15], the $Z \in \zeta^{N \times N}$ mutual coupling matrix expressed as

$$Z = (Z_A + Z_L)(\Gamma + Z_L I)^{-1} K, \quad (8)$$

with

$$\Gamma = \begin{bmatrix} Z_A & Z_M & 0 & \dots & 0 \\ Z_A & Z_A & Z_M & \dots & 0 \\ 0 & Z_M & Z_A & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & Z_M & Z_A, \end{bmatrix}, \quad (9)$$

where Z_A, Z_L and Z_M are antenna mutual impedance between adjacent antenna. Transmit correlation matrix defined as [26]

$$\Sigma_N = E[HH^H] = KZRR^H Z^H, \quad (10)$$

where $R = [R_1, R_2, \dots \dots, R_k]$ is correlation matrix, then mutual coupling channel model $H \sim \zeta N(0, \Sigma_N \otimes I_k)$ helpful in designing of massive MIMO antenna array and effect of antenna spacing on the performance of MIMO system.

2.2. Antenna Array

In this article, rectangular array pattern antenna array with $N \times N$ elements are considered where

N = 2, 4, 8, 16 antenna elements. For simplicity, isotropic antenna elements are considered.

An arrangement of 16×16 array is shown in Fig. 2 with aperture size of 85.71 mm and the spacing between adjacent element is 5.357 mm. An array is sub-divided into $N \times 1$ sub-arrays. These sub-arrays are in phased array configuration [4]. Each sub-array forms an element of the MIMO antenna configuration, that is, for a $N \times N$ antenna array composition, there will be N MIMO channels and each channel is comprised of $N \times 1$ phased array antennas [27].

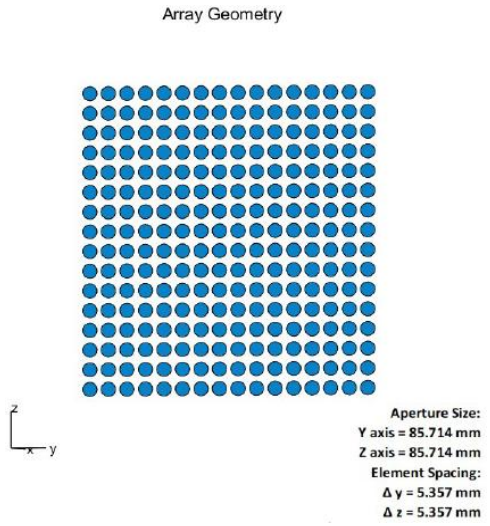


Fig. 2. Geometric Pattern rectangular antenna array $N \times M(16 \times 16)$.

A comparison of the radiation characteristics for various array sizes is presented in Table 2. It shows that as the array size increases, the directivity increases significantly. Due to this directive beam pattern, the half power beam width (HPBW) at both the azimuthal and elevation planes decreases. The same is reflected in the radiation patterns illustrated in Fig. 3. The figure further highlights that with increasing array size, the sidelobe levels decrease thereby confining maximum power within a narrow beamwidth.

Table 2. Comparative analysis of rectangular array.

Parameter	Array 2×2	Array 4×4	Array 8×8	Array 16×16
Element Spacing (mm)	3.357	5.357	5.357	5.357
Directivity (dBi)	8.63	10.09	16.51	18.90
HPBW(Az) (deg-angle)	59.92	26.30	12.00	6.36
HPBW(EI) (deg-angle)	60.00	28.00	14.00	8.00

3. Hybrid Beamforming

Hybrid beamforming technique is combination of digital beamforming and analog beamforming at its baseband and RF precoding section respectively. Analog beamforming includes a single RF chain but each antenna element is associated to a phase shifter, filter, and weights. Hence, although due to a single RF chain, power consumption is less, in massive MIMO scenario, the requirement of RF components like phase shifters becomes huge thereby increasing the hardware complexity [28].

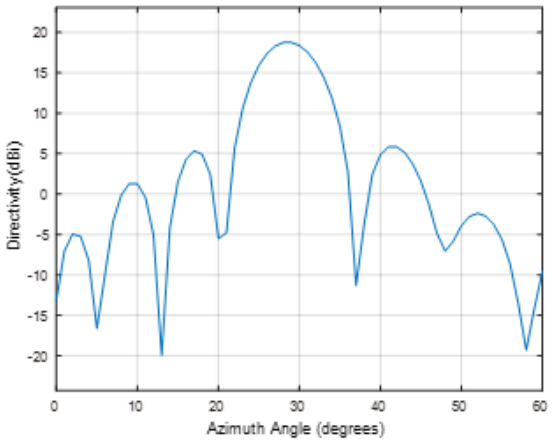


Fig. 3. Directivity 16 × 16 antenna array.

The schematic diagrams of the fully-digital beamformer, clustering-based fully connected hybrid beamformer and clustering-based partially connected hybrid beamformer are presented in Fig. 4. The article has already discussed the two stage hybrid precoding technique. In the baseband precoder, signal is adjusted by the precoding weight matrix and in the RF precoder, phase is controlled by the RF weight matrix.

A comparative study of these beamforming techniques is graphically illustrated in Fig. 5 which highlights that Spectral efficiency of optimal digital beamformer and cluster based fully connected approach is higher than the partially connected approach. Fully connected hybrid beamforming structure is outperform as compare to other structure. Furthermore, beamforming flexibility in terms of beam directions and beamwidths is limited. Thus, to reduce the hardware complexity and enhance beamforming flexibility, digital beamforming is introduced where an algorithm-based precoding system is incorporated at the baseband stage. In this technique, each antenna element is associated with a separate RF chain thus increasing the number of RF chains manifold in massive MIMO systems making the beamforming system extremely power hungry. Hybrid beamforming provides a balance between hardware complexity of analog beamforming and power consumption of digital beamforming.

A hybrid beamforming architecture is depicted in Fig. 6. In this technique, two major modifications are incorporated. Firstly, the MIMO array is divided into sub-arrays formed by a cluster of antenna elements. Each sub-array is connected to a phase shifter and associated RF hardware. The phase shift values are controlled by an analog RF precoder. Secondly, each cluster is accompanied by an RF chain connected to a

baseband precoder where digital beamforming is applied by assigning weight vectors to the signals fed to each cluster of antennas [18]. Hence, it can be comprehended that lesser number of RF chains requirement resolves the issue of power consumption faced by digital beamforming. Since the architecture deals with cluster of antenna elements as a unit, the number of RF phase shifters, filters required is also less than the number of antenna elements that simplifies the hardware complexity.

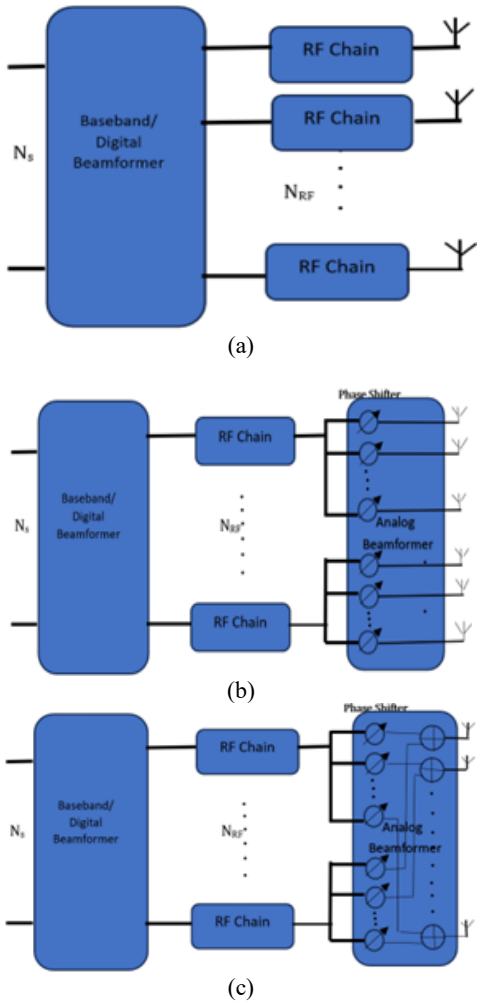


Fig. 4. Beamforming Structure: a) Digital beamformer, b) cluster based partially hybrid beamformer, c) Cluster based fully hybrid beamformer.

Multiple data streams sends through the MIMO channel, In order to derive the precoding and combining weight in hybrid beamformer from its channel matrix. Weight are applied in digital baseband significantly enhance the SINR and data rate in MIMO wireless communication [20]. N_{TRF} and N_{RRF} are always less than N_t and N_r. Encoded signal y_k transmitted by antenna array that is received at k^{th} user is given by [6, 29].

$$y_k = s_k F_{bb} F_{rf} H_k + \Sigma_{i \neq k}^{K} s_i F_{bb} F_{rf} Hk + n_k, \quad (11)$$

where n_k is i.i.d AWGN noise vector of size $N_r \times 1$ and $n \sim CN(0, \sigma^2)$. H_k is channel matrix for the k^{th} user that can be determined from (3) – (9). $\tilde{y}_k$ is post processing received signal computed at combiner computed as [21].

$$\tilde{y}_k = (s_k F_{bb} F_{rf} H_k + \Sigma_{i \neq k}^{K} s_i F_{bb} F_{rf} H_k + + n_k) W_{rf} W_{bb}, \quad (12)$$

W_{bb} and W_{rf} are the weight matrix at the hybrid beamformer at baseband and RF precoding at combiner section. Performance of the k_{th} user to estimate data stream $\tilde{y}_k$ in term signal to interference noise ratio SINR given by [30]

$$SINR_{j,k} = \frac{|F_{j,k} H_{j,j,k}^{H} W_{j,k}|^2}{\Sigma_{n,l \neq j,k} |F_{n,l} H_{n,j,k}^{H} W_{n,l}|^2 + \sigma^2} \quad (13)$$

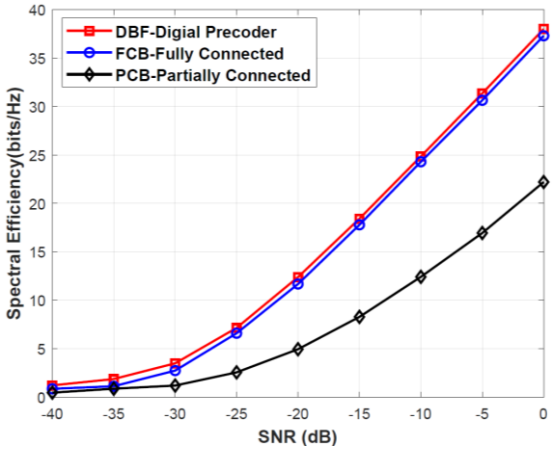


Fig. 5. Spectral Efficiency comparison digital beamformer.

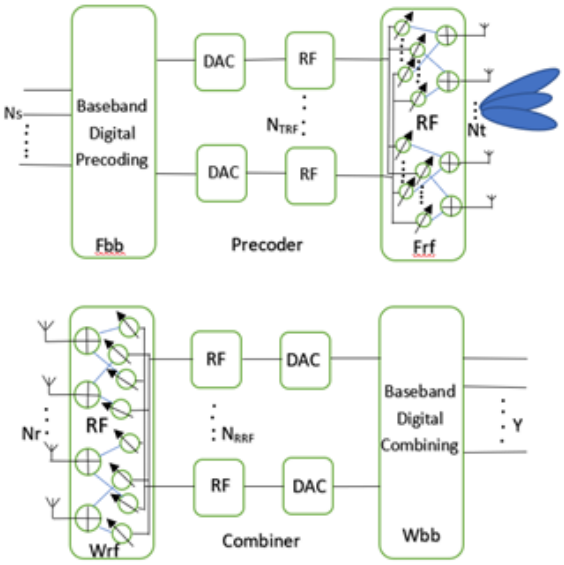


Fig. 6. Hybrid beamforming architecture.

Steering beam to the k^{th} user minimized SINR by adjusting the weight vector through efficient

optimization algorithm [29]. In digital beamformer optimal weight can be determined by diagonalization of channel matrix H [29]

$$F \times H \times W \ = \ \Lambda^{-1} H \Lambda, \tag{14}$$

where Λ, eigen vector matrix. F and W are precoding and combining weight matrix of dimension $N_s \times N_t$ and $N_r \times N_s$ respectively.

In this article hybrid beamforming algorithm method Orthogonal Matching Pursuit (OMP), Manifold Optimization and Generalized Eigen Value Decomposition (GEVD) method are implemented for precoder and combiner design. The MMSE based hybrid beamforming formulation and decomposition into subproblem are use from literature [15] and evaluate the performance under different antenna configuration and channel condition. Unlike the previous work, this study foucs on comperative performance analysis in term of spectral efficiency spectral efficiency (SE), mean square Error (MSE), bit error rate (BER), energy efficiency and computational complexity under identical framework.

BER is the important parameter to analyze the efficiency of hybrid beamforming algorithms. In BER performance analysis, the processed received data at k^{th} user equipment from (11) can be re-written as [30]

$$\tilde{y}_k \ = \ s_k F H_k W + \Sigma_{j \neq k}^{K} s_i F H_k W + n_k W \tag{15}$$

Assuming the occurance of symbol error ($s_k \neq \tilde{s}_k$) to the estimated signal, BER can be expressed in terms of SINR as [31], $BER \ = \ Q\sqrt{SINR}$

$$BER_k \ = \ Q\left(\sqrt{\frac{|F_{j,k} H_{j,j,k}^{H} W_{j,k}|^2}{\Sigma_{n,l \neq j,k} |F_{n,l} H_{n,j,k}^{H} W_{n,l}|^2 + \sigma^2}}\right) \tag{16}$$

The formulation of the other two parameters (SE and MSE) is detailed during the discussions on HBF algorithms.

3.1. Orthogonal Matching Pursuit

The orthogonal Matching Pursuit (OMP), dictionary based sparse precoding algorithm in mmWave MIMO system [30] is implemented as benchmark scheme for comparative analysis. The presented work, evaluate the performance under identical system assumption and antenna array configuration.

Orthogonal matching pursuit (OMP) algorithm are used as hybrid beamforming algorithm to calculate the precoder solution given as, precoding weight $F = [F_{BB}, F_{RF}]$ and combining weight $W = [W_{BB}, W_{RF}]$ are four set of optimal weight encode the signal and reduce the interference to produce the more steered beam. More the steered beam higher the channel capacity and spectral efficiency [32].

$$R \ = \ log_2 |I - SINR_k| \tag{17}$$

Algorithm 1. Precoding via Orthogonal Matching Pursuit [30]

Input: A_t, N_t^{RF}, N_{cl}, N_{ray}
Require: F_{opt}
$F_{RF} = EmptyMatrix$
$F_{res} = F_{opt}$
For i≤ N_t^{RF} **do**
$\Psi = A_t{}^* F_{res}$ $k = argmax_{l=1 \ldots N_{cl} N_{ray}} (\Psi \Psi^*)$
$F_{RF} = [F_{RF} | A_t^{(k)}]$
$F_{BB} = (F_{RF}^* F_{RF})^{-1} F_{RF}^* F_{RF}$
$F_{res} = \frac{F_{opt} - F_{RF} F_{BB}}{\| F_{opt} - F_{RF}^* F_{RF} \|_F}$
end for
$F_{BB} = \sqrt{N_s \frac{F_{BB}}{\| F_{RF} F_{BB} \|_F}}$
Return $F_{RF} F_{BB}$

Considering the number of data streams transmitted (Ns), SE is calculated as the maximum of R given by

$$R \ = \ log_2 (|I_{N_s} + \\ + \frac{\rho}{N_s} R_n^{-1} W_{BB}^* W_{RF}^* F_{RF} F_{BB} H F_{BB}^* F_{RF}^* W_{RF} W_{BB}|), \tag{18}$$

where $R_n = \sigma_n^2 W_{BB}^* W_{RF}^* W_{RF} W_{BB}$ is the noise covariance matrix. MSE can be estimated for HBF design as [14, 33]

$$MSE \triangleq E \parallel \beta^{-1} y - s \parallel^2, \tag{19}$$

where s is the data stream and β is the scaling factor optimize the hybrid beamformer weight, precoder and combiner. The processed data is denoted by $y = W_{BB}^H W_{RF}^H H F_{RF} F_{BB} s + W_{RF}^H W_{BB}^H u$. At the HBF receiver, W_{RF} is the analog combiner weight and W_{BB} is the baseband combiner weight. Received data with additive noise vector u is complex Gaussian distribution at N_r receiver antenna with zero mean and covariance $\sigma^2 I_{N_r}$, i.e. $u\sigma^2 I_{N_r}$, MSE (19) can be expressed as:

$$MSE = E \parallel \beta^{-1} (W_H F s + W_H u - s \parallel^2 = \\ = tr(\beta^{-1} W^H H F F^H H^H W - \beta^{-1} W^H H F \ W + \\ + I_{N_s}) \ - \beta^{-1} F^H H^H W + \sigma^2 \beta^{-2} W^H, \tag{20}$$

where F and W are baseband and combiner weight matrices over all beamformer. F_{RF} and W_{RF} are analog beamformer weights to adjust the phase. Analog beamformer should satisfy the constant modulus constraints of the phase shifter, defined as $|[F_{RF}]_{i,j}| = 1$, for $i = 1, \ldots, N_t$ and $j = 1, \ldots, N_{RF}$, and $|[W_{RF}]_{i,j}| = 1$, for $i = 1, \ldots, N_t$ and $j = 1, \ldots, N_{RF}$. MSE expression formulating the HBF optimization problem as

$$\min_{F_{RF},F_B,W_{RF},W_B} \quad MSE, \tag{21}$$

$$\text{subject to:} \quad \| F \|_B^2 \leq 1, \forall_{i,j}, \tag{22}$$

$$|[W_{RF}]_{m,l}|^2 = 1, \forall_{m,l} \tag{23}$$

Algorithm 2. Combining via Orthogonal Matching Pursuit [30]

Input: A_r, N_t^{RF}, N_{cl}, N_{ray}
Require: W_{opt}
$W_{RF} = EmptyMatrix$
$W_{res} = W_{opt}$
For i$\leq N_r^{RF}$ do
$\Psi = A_r^* E[yy^*] W_{res}$
$k = argmax_{l=1\ldots N_{cl}N_{ray}} (\Psi\Psi^*)_{l,l}$
$W_{RF} = [W_{RF} | A_r^{(k)}]$
$W_{BB} = (W_{RF}^* E[yy_*] W_{RF})^{-1} W_{RF}^* E[yy_*] W_{opt}$
$W_{res} = \frac{W_{opt} - W_{RF} W_{BB}}{\|W_{opt} - W_{RF}^* W_{opt}\|_F}$
End for
$W_{BB} = \sqrt{N_s} \frac{W_{BB}}{\|W_{RF} W_{BB}\|_F}$
Return $W_{RF} W_{BB}$

3.2. Manifold Optimization Algorithm

The Manifold optimization technique adopted in this work to solve the constrained hybrid beamforming problem directly under MMSE criteria [15]. The presented article uses these hybrid beamforming formulation for consider the antenna array and RF chain configuration.

Manifold hybrid beamforming algorithm optimize the precoder and combiner divide original problem into subproblem. MSE optimization given in equation (21) to calculate precoding matrix $F_{BB}$and scaling factor β [34].

$$\beta = (tr(F_{RF} F_{BB} F_{BB}^H F_{RF}^H))^{-\frac{1}{2}}, \tag{24}$$

and

$$F_{BB} = (F_{RF}^H H_1 H_1^H F_{RF} + \sigma^2 \omega F_{RF^H F_{RF}})^{-1} F_{RF} H_1 \tag{25}$$

Conjugate gradient function equation (26) optimized RF precoder by minimized the objective to find optimal solution to maximize the transit power level [29].

$$\nabla J(F_{RF}) = \frac{1}{\sigma_2 \omega} (F_{RF} \left(F_{RF}^H F_{RF}\right)^{-1} - I_{N_s}) \times \times H_1 P^{-2} H_1^H F_{RF} (F_{RF}^H F_{RF})^{-1} \tag{26}$$

Algorithm 3. Precoding Via MO- HBF

Input: H_1, σ^2, ω, α
Require: F_{RF}, F_{BB}, β
Set i=0, initialize $F_{RF,i}$;
For i$\leq N_{max}$ do
$\nabla J(F_{RF}) = \frac{1}{\sigma_2 \omega} (F_{RF} (F_{RF}^H F_{RF})^{-1} - I_{N_s}) \times$
$\times H_1 P^{-2} H_1^H F_{RF} (F_{RF}^H F_{RF})^{-1}$
$F_{RF,i+1} = F_{RF,i} - \alpha \nabla J(F_{RF})$
$i \leftarrow i + 1$
End for
$\beta = (tr(F_{RF} F_{BB} F_{BB}^H F_{RF}^H))^{\frac{-1}{2}}$
$F_{BB} = (F_{RF}^H H_1 H_1^H F_{RF} + \sigma^2 \omega F_{RF^H F_{RF}})^{-1} F_{RF} H_1$
Return $F_{RF} F_{BB}$

Hybrid receiver combiner is designed after updating the precoding weight matrices according to scaling factor β. The optimization problem can rewrite at hybrid receiver combiner as [34, 35]

$$\min_{W_{RF},W_B} tr(W_H H_2 H_2^H W - W^H H_2 - H_2^H W + \sigma^2 \beta - 2W^H W + I_{N_s}), \tag{27}$$

$$\text{subject to:} \ |[W_{RF}]_{i,j}| = 1, \forall_{i,j}, \tag{28}$$

where $H_2 \triangleq H F_{RF} F_{BB}$ and $W \triangleq W_{RF} W_{BB}$. Objective function differentiating W_B and setting the result to zero, optimal W_B is obtained as

$$W_B = (W_{RF}^H H_2 H_2^H W_{RF} + \sigma^2 \beta^{-2} W_{RF^H W_{RF}})^{-1} W_{RF} H_2 \tag{29}$$

RF combiner optimization problem can be reduced to the following optimization of W_{RF} as

$$\min_{W_{RF}} J(W_{RF}) \triangleq tr(I_{N_s} + \sigma^2 \beta^{-2} H_2^2 W_{RF} \times \times (W_{RF}^H W_{RF})^{-1} W_{RF}^H H_2)^{-1}, \tag{30}$$

$$\text{subject to:} \ |[W_{RF}]_{i,j}| = 1, \forall_{i,j} \tag{31}$$

MO-HBF is directly applied and implemented in the precoder matrix.

3.3. Generalized Eigen Value Decomposition

The generalized eigenvalue decomposition (GEVD) based hybrid beamforming framework proposed in Lin et al. [15] implemented to evaluate the performance under unified framework. The hybrid beamforming algorthim structure follows the MMSE reformulation and columnn wise optimization strategy match with antenna array configuration and simulation enviornment.

The gradient based algorithm optimizes the analog precoding weight matrix F_{RF}. The complexity of GEVD algorithm dependent on number of iterations, which is trade-off with suprior performance of spectral efficiency. According to [8, 36] for MIMO system, $F_{RF}^H F_{RF} \approx N_t I_{N_{RF}}$ different data streams are orthogonal to each for optimized analog beamforming, with this approximation, is MSE defined as:

$$J(F_{RF}) \approx tr(I_{N_s} + \frac{1}{\sigma^2 \omega Nt} H_1^H F_{RF} F_{RF}^H H_1)^{-1}) \quad (32)$$

Precoding matrix is optimized column by column, V_m is sub matrix of F_{RF} after the optimization of the m_{th} column. Furthermore, $A_m \triangleq I_{N_s} + \frac{1}{\sigma^2 \omega Nt} H_1^H V_m V_m^H H_1$. Using the identity $(A+B)^{-1} = A^{-1} - \frac{A^{-1}BA^{-1}}{1+tr(A^{-1}B)}$, the MSE can be expressed as [37]:

$$J(F_{RF}) \approx tr(A_m)^{-1} - \\ - \frac{tr\left(\frac{1}{\sigma^2 \omega Nt} A_m^{-1} H_1^H v_m v_m^H H_1 A_m^{-1}\right)}{1+tr\left(\frac{1}{\sigma^2 \omega Nt} A_m^{-1} H_1^H v_m v_m^H H_1\right)} = tr(A_m^{-1}) - \\ - \frac{v_m^H U_m v_m}{v_m^H W_m v_m}, \quad (33)$$

where $U_m \triangleq \frac{1}{\sigma^2 \omega Nt} H_1 A_m^{-2} H_1^H$ and $W_m \triangleq \frac{1}{N_t} + \frac{1}{\sigma^2 \omega Nt} H_1 A_m^{-2} H_1^H$ are both Hermitian matrices. In the equation (33), MSE is divided into two terms, V_m and v_m. By keeping the V_m fixed, $J(F_{RF})$ is the function of v_m in second term in (33). The optimal v_m maximizes the second term and minimizes the whole term in equation (33). Minimizing the MSE in equation (33) is accomplished by maximizing the generalized eigen value between the U_m and W_m, obtained by GEVD operation. Phase extraction operations, extract the phase of antenna element in each column. The steps for the whole V_{RF} matrix are repeat until the stop condition is satisfied.

Algorithm 4. GEVD- Based Hybrid Beamforming Alogorithm [15]

Input: H_1, σ^2, ω,
Input: H_1, σ^2, ω,
Require: F_{RF}, F_{BB}, β
Set $i = 0$, Initialize $F_{RF,i}$;
For $1 \leq m \leq N_{RF}$ **do**
Evaluate A_m, U_m, W_m
$A_m \triangleq I_{N_s} + \frac{1}{\sigma^2 \omega Nt} H_1^H V_m V_m^H H_1$.
$U_m \triangleq \frac{1}{\sigma^2 \omega Nt} H_1 A_m^{-2} H_1^H$
$W_m \triangleq \frac{1}{N_t} + \frac{1}{\sigma^2 \omega Nt} H_1 A_m^{-2} H_1^H$
Evaluate the maximum generalized eigenvector z of U_m and W_m
Set $v_{m,i}$ =exp(J∠ (z), where $\angle(z)$ extract the phase of each element of z
$i \leftarrow i + 1$
End for
Compute β and F_{BB}
$\beta = (tr(F_{RF} F_{BB} F_{BB}^H F_{RF}^H))^{\frac{-1}{2}}$
$F_{BB} = (F_{RF}^H H_1 H_1^H F_{RF} + \sigma^2 \omega F_{RF^H F_{RF}})^{-1} F_{RF} H_1$
Return $F_{RF} F_{BB}$

The overall GEVD HBF algorithm optimizes the precoding analog matrix. By fix the precoding matrices F_{RF}, F_B and scaling factor β, the optimization problem rewrite for hybrid receiver combiner [14, 32]

$$\min_{W_{RF} W_B} tr(W_{RF}^H W_B^H H F_{RF} F_B F_B^H F_{RF}^H H^H W_{RF} W_B - \\ - W_{RF}^H W_B^H H F_{RF} F_B - H_2^H W_{RF} W_B + \\ + \sigma^2 \beta^{-2} W_{RF}^H W_B^H W_{RF} W_B + I_{N_s}), \quad (34)$$

$$\text{subject to } |[W_{RF}]_{i,j}| = 1, \forall_{i,j} \quad (35)$$

Objective function differentiating with respect to W_B in in equation (34) and setting to zero, to find the resultant optimal combiner in equation defined as [14]

$$W_B = (W_{RF}^H H F_{RF} F_B F_B^H H^H W_{RF} + \\ + \sigma^2 \beta^{-2} W_{RF}^H W_{RF})^{-1} W_{RF}^H H F_{RF} F_B F_B^H F_{RF}^H H^H \quad (36)$$

RF combiner matrix W_{RF}, after optimization reduce to the following optimization problem

$$\min_{W_{RF}} J(W_{RF}) \triangleq tr((I_{N_s} + \sigma^2 \beta^{-2} F_B^H F_{RF}^H H_H W_{RF}) \times \\ \times (W_{RF}^H W_{RF})^{-1} W_{RF}^H H F_{RF} F_B)^{-1} \quad (37)$$

$$\text{subject to } |[W_{RF}]_{i,j}| = 1, \forall_{i,j} \quad (38)$$

GEVD -HBF algorithm is implemented directly to optimize the hybrid combiner.

3.4. Broadband Scenario

The hybrid beamforming algorithm can be extended to the broadband mmWave system in which the OFDM technology is to be implemented to avoid the channel frequency selectivity [38, 39]. The work follows the wideband OFDM MMSE framework [39] 40]. This section include the theoretical extension and future research extension.

Precoder weight matrix optimization is required for all the subcarriers. MSE for the k^{th} subcarrier is defined by [35, 40]

$$MSE_k = E(\| s_k - \beta^{-1} y_k \|^2), \quad (39)$$

where, β_k is scaling factor the k^{th} subcarrier, S_k is transmitted symbol and y_k is received power at the k^{th} subcarrier. Optimization objective for broadband scenario can be defined as [36]

$$\min_{F_{B,k}, F_{RF}, W_{B,k}, W_{RF}} \Sigma_{k=0}^{N-1} MSE_k, \quad (40)$$

$$\text{subject to } \| F_k \|^2 \leq 1, \quad (41)$$

$$|[F_{RF}]_{i,j}|^2 = 1, \forall_{i,j}, \quad (42)$$

$$|[W_{RF}]_{m,l}|^2 = 1, \forall_{m,l}, \quad (43)$$

Maximization of the spectral efficiency for the broadband WMMSE HBF optimization objective formulated as [36]

$$\min_{F_k, W_k, \beta_k, \Delta_k} \Sigma_{k=0}^{N-1} (tr(\Delta_k T_k) - log|\Delta_k|, \quad (44)$$

$$\text{subject to } \| F_k \|^2 \leq 1, \quad (45)$$

$$|[F_{RF}]_{i,j}|^2 = 1, \forall_{i,j}, \quad (46)$$

$$|[W_{RF}]_{m,l}|^2 = 1, \forall_{m,l}, \quad (47)$$

where, T_k and Δ_k represent MSE and weight for the k^{th} sub carrier respectively. When the objective function narrowband scenario is taken into consideration, it is seen that only constant weight matrix Δ_k is replaced by I_{N_s}.

3.5. Computational Complexity Analysis

In this section we discuss the computation complexity of the hybrid beamforming algorithm, focusing on complex multiplication. Both the transmit precoding and receiving combining subproblems follow unified procedure. Hybrid beamforming iteratively selects the best transmit antennas based on channel characteristics. The complexity of this step depends on the number of RF chains per sub-array (N_{RF}) and the number of data streams (N_s) [31]. In orthogonal Matching Pursuit (OMP), complexity primarily depends on the antenna (A_t) and optimal beamforming matrices (F_{opt}). With a loop iteration of (N_{RF}), the overall complexity is $O(N_{RF}(N_t \times N_{cl} \times N_{ray} + N_{RF}^3))$ [41]. In Manifold Optimization (MO) precoder maximize the signal power and complexities lies in the conjugate gradient function, contributing $O(N_{RF}^3 + N_t^3 + N_t^2 \times N_{RF} + N_{RF}^2 \times N_t)$ per iteration, resulting in an overall complexity of $O(N_{max}(N_{RF}^3 + N_t^3 + N_t^2 \times N_{RF} + N_{RF}^2 \times N_t)$. GEVD performs a decomposition of the channel matrix into its eigen vectors and eigen values. The computational complexity of this step depends on the size of the channel matrix. Generalized Eigenvalue decomposition (GEVD) involve computing U_m and W_m, contributing $O(N_s^3 + N_t \times N_s^2 + N_t^2 \times N_s)$ per iteration. The overall complexity is $O(N_{RF}(N_s^2 + N_t \times N_s^2 + N_t^2 \times N_s) + N_{RF}^3 + N_{RF}^2 \times N_t))$. The comparison of the computational complexity of the all-beamforming algorithms is given below Table 3.

Table 3. Comparative Complexity Analysis of HBF Algorithms.

HBF Algorithm	Computational Complexity
OMP	$O(N_{RF}(N_t \times N_{cl} \times N_{ray} + N_{RF}^3))$
MO	$O(N_{max}(N_{RF}^3 + N_t^3 + N_t^2 \times \times N_{RF} + N_{RF}^2 \times N_t)$
GEVD	$O(N_{RF}(N_s^2 + N_t \times N_s^2 + N_t^2 \times \times N_s) + N_{RF}^3 + N_{RF}^2 \times N_t))$

For $N_{RF} = 4, N_s = 4$ and $N_t = 64$ computational complexity $O(256N_{cl}N_{ray} + 256)$ for precoding via OMP HBF, higher complexity depend on N_{cl} and N_{rayl}. OMP HBF algorithm is relatively simple to implement, making it suitable for real-time applications and systems with hardware constraints. Complexity calculate for given parameter in manifold hybrid beamforming is $O(N_{max}(279616))$. It shows complexity dependency on N_{max}, which is significantly higher complexity. MO HBF algorithm can achieve improved beamforming accuracy by jointly optimizing both analog and digital precoders and providing potential benefits for interference management. GEVD provides the optimal solution for maximizing the received signal-to-noise ratio. While GEVD offers higher spectral efficiency and complexity for GEVD is $O(70976)$.

4. Simulation Results

In this section, simulation evaluate the performance of the described HBF algorithms through Monte Carlo simulation in a mmWave massive MIMO communication enviornment operating at carrier frequency of 28 GHz. The narrowband channel model is designed with cluster size $N_C = 6$, number of the rays $N_{Ray} = 8$ per cluster and scattered channel is $N_{Ray} \times N_C$. The transmitter and reciever equiped with uniform rectuagular array (URAs) with antenna element sapcing $\frac{\lambda}{2}$ = 5.357 mm. It is assumed that angle of arrival and departure distributed with angular spread 10 degree with in mean cluster angle in $[0,2\pi]$. Uniform rectangular array (URA) subdivided in uniform in Uniform linear array. The hybrid beamforming architecture employs number of RF chain $N_{RF} = 4$ and the number of data stream N_s varies from 1 to 4, subject to constraint $N_s \leq N_{RF}$. The constatnt modulus phase shifter assumed analog beamforming network and it satisfy unit magnitude constraint. The simulation results obtained with 400 Motnte Carlo realization. Simulation provided to implement the HBF optimization algorithm at precoder and combiner stages with different size array as $2 \times 2, 4 \times 4, 8 \times 8, 16 \times 16$ and length of data stream in QPSK modulated at baseband precoder. The simulation section evaluate the spectral effciency, MSE, BER and energy efficiency by implementation of persented HBF algorithms in common framework and parameter in given parameter Table 4.

Table 4. List of Parameter.

Parameter	Value
Carrier Frequency	28 GHz
Number of Cluster (N_c)	6
Rays per Clusters (N_{ray})	8
Modulation Scheme	QPSK
Data Stream (N_s)	1-4
Antenna Arrays	2×2, 4×4, 8×8, 16×16
Antenna Spacing	$\frac{\lambda}{2}$ = 5.357 mm
SNR Range	-40 dB - 0 dB
Phase Shifter	Constant Modulus

4.1. Spectral Efficiency Performance

Fig. 7 and Fig. 8 graphically depict the spectral efficiency achieved by the beamforming algorithms.

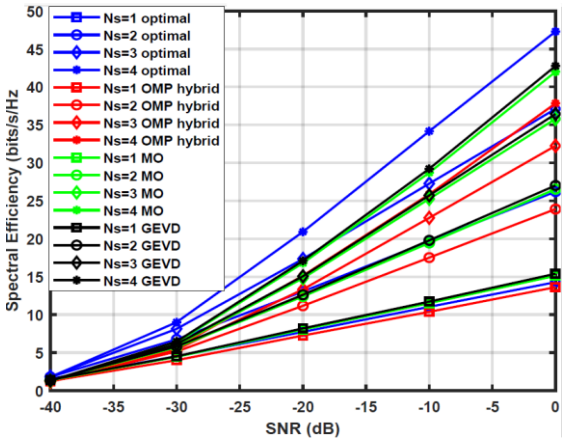


Fig. 7. Spectral Efficiency at different length DataStream, Array size 16 × 16.

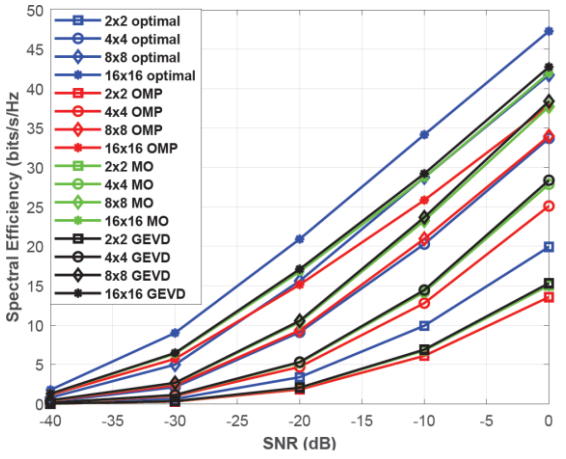


Fig. 8. Spectral Efficiency at different Array Size, DataStream length $Ns = 4$.

From Fig. 7, it is observed that when the data stream length (N_s) is less GEVD and MO algorithms give slightly better performance than optimal algorithm with GEVD being marginally improved than MO. As N_S increases to 4 and beyond, performance of optimal algorithm exceeds the others. Fig. 8 highlights that with different antenna array size the optimal algorithm gives overwhelmingly better result than any HBF algorithm. For HBF algorithms, GEVD and MO have almost similar performances.

It is also observed that as the array size is increasing, the difference between full digital beamforming (optimal) and the HBF algorithms (GEVD and MO) is decreasing steadily thus indicating that for extremely large massive MIMO cases, GEVD and MO will give almost similar spectral efficiency as optimal with GEVD being more advantageous due to its cost effectiveness. In all the cases, OMP algorithm lags behind thereby concluding that OMP is not a suitable HBF technique where spectral efficiency is the primary objective.

Table 5 tabulates the spectral efficiency values at 0 dB, to comprehend the performance of the HBF algorithms numerically. When array size is less, GEVD HBF is comparable to optimal beamforming. As array size increases, with increasing number of data streams, the performance of GEVD and MO become comparable to optimal beamforming. MO surpasses GEVD when data stream is less.

Table 5. Spectral Efficiency (bits/Hz) Comparison (SNR = 0 dB).

Beamforming	N_S	Array 2×2	Array 4×4	Array 8×8	Array 16×16
Optimal (Full Digital)	1	10	11	13	26
	2	16	20	23	26
	3	17	26	32	37
	4	17	30	40	47
OMP HBF	1	10	11	13	14
	2	14	18	22	24
	3	14	23	28	31
	4	15	25	23	37
MO HBF	1	10	11	13	15
	2	12	20	24	36
	3	13	25	32	36
	4	15	28	37	42
GEVD HBF	1	11	12	14	17
	2	15	20	27	27
	3	15	26	32	37
	4	16	29	38	43

4.2. Mean Square Error (MSE) Performance

MSE is another performance parameter that evaluates the 5G test bench performance after implementation of HBF algorithm. For transmitted RF signal through scattered channel H, while taking into account noisy channel and interference from adjacent user, MSE is formulated for various HBF algorithms as discussed in Section 3 which is mainly dependent on the channel condition, array dimension and size. MSE is evaluated for the 20 iterations and SNR from -40 dB to 0 dB.

Fig. 9 exhibits the MSE performance for the HBF algorithms for various MIMO array sizes. With increase in array size, the directional gain increases due to larger number of subarray selection. Keeping channel condition fixed, increase in directional gain lowers the interference from adjacent users that reduces the MSE. As a result, for all the algorithms, MSE decreases with increase in array size although the differences become small with increased SNR. Among the HBF algorithms, GEVD has minimum MSE due its control ability at the RF precoder level in transmitter section. It is observed that MSE for OMP algorithm tends to catch up MO algorithm as array size increases.

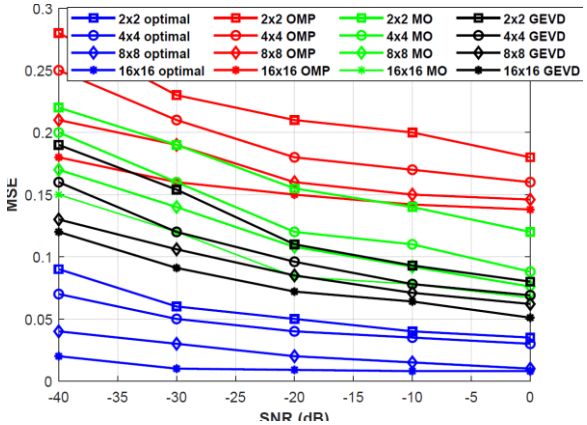


Fig. 9. MSE performance HBF algorithms at different array size.

MSE performance is evaluated with different length of data stream transmitted through the precoder circuit and presented in Fig. 10. It is seen that for higher data stream, as SNR increases, the MSE flattens and approaches constant value for GEVD and optimal beamforming techniques. GEVD is still the better candidate among the HBF algorithms. For higher data streams, with increasing SNR, the MSE decreases drastically for OMP algorithm and almost exceeds MO performance at near 0 dB SNR value.

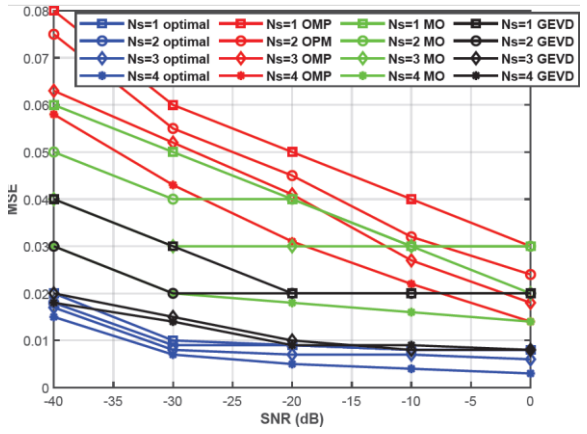


Fig. 10. MSE performance with different Data Stream.

4.3. Bit Error Rate (BER) Performance

BER performance of the beamforming algorithms has been evaluated based on (16). Fig. 11 shows the BER performance based on the length of data stream. The figure indicate, in noisy condition GEVD and MO HBF perform slightly improved BER in certain simulation realization such as finite Monte Carlo simulation and MMSE based optimization framework. As the SNR increase, fully digital beamforming remain upperbound benchmark of the perormance.

Furthermore, with an increase in the length of data stream, the BER decreases at low SNR values for all the beamforming techniques. As SNR increases, the BER decreases significantly and the length of data stream does not play any role to affect the BER. Overall, OMP algorithm has the worst performance.

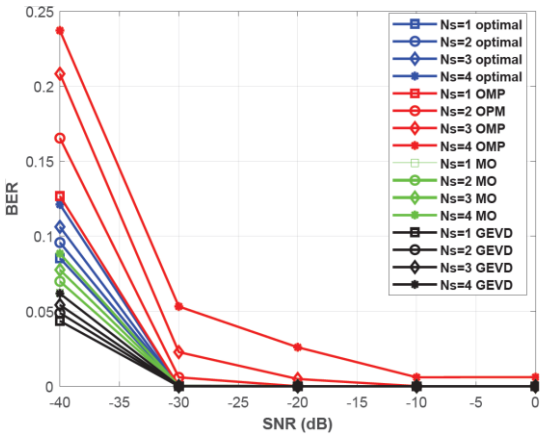


Fig. 11. BER at different length of Data Stream, Array Size 16×16.

Fig. 12 exhibits the effect of array size on the BER performance of the beamforming algorithms. For any particular array size OMP HBF exbhits highest BER, where MO HBF and GEVD HBF achieve BER peroformance close to Optimal digital beamforming. Among all the hybrid beamforming hybrid beamforming, the GEVD HBF consistently perform better BER with increase the SNR. Energy efficiency gains in sub-array precoding depend on various factors, including system design, specific algorithms used, and environmental conditions. Therefore, proper optimization and careful consideration of the system parameters are crucial to achieving maximum energy efficiency benefits.

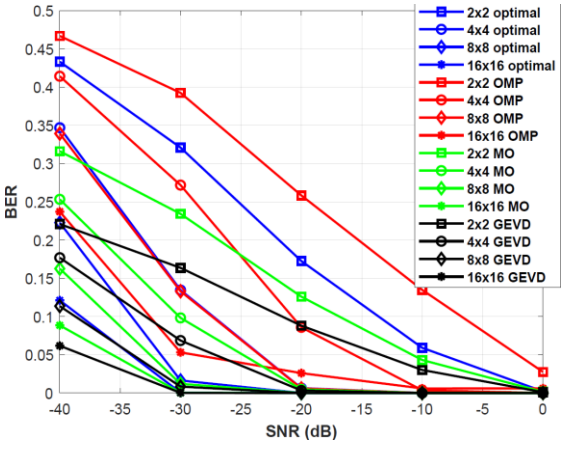


Fig. 12. BER at different Array Size, $Ns = 44.4$ Energy Efficiency Performance.

In the present work, the transmitter power $PT = 1\,W$ is considered. The energy efficiency is defined as $\eta = R/PT$, where R is achievable rate

[37]. The number of RF chain N_{RF} and number of receiver antenna Nr are 4 and 64 respectively. Fig. 13 shows the energy efficiency performance with the number of transmitter antennas (Nt). It indicates that the energy efficiency increases with increase in N_t with a tendency to flatten out as N_t increases without bound. GEVD and MO HBF algorithms have almost identical performance while OMP HBF has the poorest performance. Fig. 14 exhibits energy efficiency performance with respect to SNR for various array configuration which reveals that for lower configuration energy efficicny GEVD HBF and MO HBF outperform optimal digital precoder due to reduced number RF chain and assoiciated hardware power consumtion.

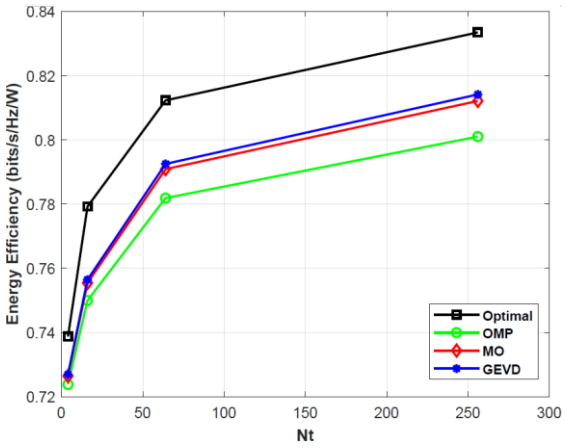


Fig. 13. Energy Efficiency Vs N_t.

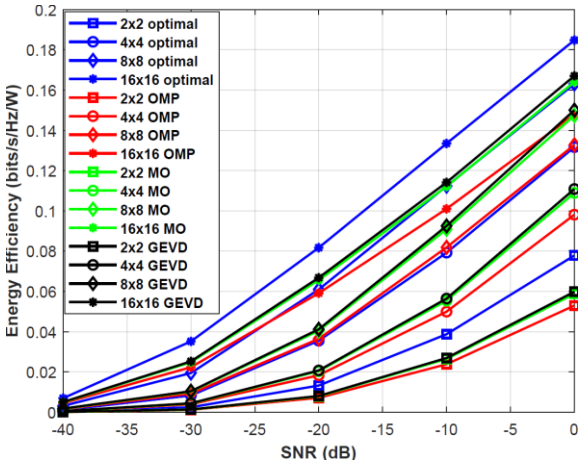


Fig. 14. Energy Efficiency Vs SNR (Ns = 1,2,3,4).

The performance of the presented HBF algorithms in the manuscript are compared with other HBF algorithms found in literature in terms of spectral efficiency (SE), mean square error (MSE) and bit error rate (BER). The comparison table is given in Table 6. From the table, it is revealed that with comparatively lesser number of antenna array elements and RF chains (N_{RF}), better performance is achieved by the presented HBF algorithms relative to the other HBF algorithms found in literature.

Table 6. Comparison of presented HBF algorithms with other reported HBF algorithms.

Cite	Methodology	Parameter	SE	MSE	BER
[42]	Zero forcing beamforming algorithm	$N_{RF} = 8$, Ns = 4	32	0.24	10^{-1}
[43]	MMSE based Rate Balancing Hybrid precoding	$N_{RF} = 3$, Ns = 2	20	...	...
[44]	Constrained-SVD based hybrid beamforming	$N_{RF} = 4$, Ns = 4	28	...	10^{-1}
[45]	GMD based hybrid beamforming	$N_{RF} = 3$, Ns = 3	14	0.16	10^{-2}
[46]	Gram-Schmidt based greedy algorithm	$N_{RF} = 3$, Ns = 2	10	0.21	10^{-1}
[24]	Eigen-Beamforming	$N_{RF} = 10$, Ns = 4	35	0.12	10^{-2}
Present work	Orthogonal Matching Pursuit	$N_{RF} = 4$, Ns = 4	37	0.14	0.0276
Present work	Manifold Optimization	$N_{RF} = 4$, Ns = 4	42	0.067	0.0015
Present work	GEVD HBF Algorithm	$N_{RF} = 4$, Ns = 4	43	0.051	0.0011

5. Conclusion

This article proposed HBF technique in mm-wave massive MIMO antenna configuration. 4, 16, 64 and 256 isotropic antenna elements have been considered in rectangular array pattern to configure the massive MIMO arrangement. $N \times 1$ elements have been clustered together to form a phased array antenna. N rows of phased array antennas form the N-ports of the MIMO antenna. It has been demonstrated that as the number of antenna elements increases in the phased array, directivity gain increases while beamwidth decreases and number of side lobes also increases. In order to minimize inter-element interference and increase channel capacity, HBF technique has been employed with optimizing the precoding weights at the baseband and combining weights at the RF sections. The various HBF algorithms used are MO, OMP and GEVD which are compared with full digital beamforming algorithm in terms of SE, MSE and BER. It has been demonstrated that with increased data streams, all the algorithms perform very well even under low SINR condition. Although digital beamforming repersents the theoretical performance upperbound, the simulation resuts with GEVD HBF and MO HBF algorithm better BER performce close to digital beamforming under considered MMSE based framework and simulation assumption. The overall, GEVD provide suprior BER, MSE performance and maintaining lower complexity requirement than manifold optimization technique. It

is most favourable trade-off between the performance and complexity cost. The investigations indicate that for large number of antenna elements, SE for MO and GEVD become comparable to full digital beamforming. The analysis and performance evaluation provided in this work will help as guidance to judge the best suited HBF optimization technique to be used for any particular mm-wave massive MIMO scenario.

References

[1]. I. Ahmed, H. Khammari, A. Shahid, A. Musa, et al., A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives, *IEEE Communications Surveys & Tutorials*, Vol. 20, Issue 4, 2018, pp. 3060-3097.

[2]. E. Ali, M. Ismail, R. Nordin, N. F. Abdulah, Beamforming techniques for massive MIMO systems in 5G: Overview, classification, and trends for future research, *Frontiers of Information Technology & Electronic Engineering*, Vol. 18, Issue 6, 2017, pp. 753-772.

[3]. R. N. Revankar, A. N. Kashyap, S. Shravan, M. C. Nithya Shree, et al., Hybrid beamforming in 5G massive MIMO, in *Proceedings of the International Conference on the Confluence of Advancements in Robotics, Vision and Interdisciplinary Technology Management (IC-RVITM)*, 2023, pp. 1-6.

[4]. S. I. Farghaly, M. M. Fouda, M. M. Emara, Beamforming of transmit antennas using grey wolf optimization and L2-norm for performance enhancement of beyond 5G communications, *IEEE Open Journal of Antennas and Propagation*, Vol. 5, 2024, pp. 1-12.

[5]. A. Mahmud, A. Ali, R. S. Dhar, S. G. Razavipour, et al., Scanning voltage microscopy for emerging electronic and photonic devices: Integrating nanotips with a single atom end for SVM, *IEEE Nanotechnology Magazine*, Vol. 11, Issue 1, 2017, pp. 4-11.

[6]. A. Garcia-Rodriguez, V. Venkateswaran, P. Rulikowski, C. Masouros, Hybrid analog-digital precoding revisited under realistic RF modeling, *IEEE Wireless Communications Letters*, Vol. 5, Issue 5, 2016, pp. 528-531.

[7]. I. Chatterjee, J. Singh, S. Srivastava, A. K. Jagannatham, Frequency selective hybrid beamforming and optimal power loading for multiuser millimeter-wave cognitive radio networks, *IEEE Access*, Vol. 11, 2023, pp. 116310-116326.

[8]. L. Jiang, H. Jafarkhani, Multi-user analog beamforming in millimeter-wave MIMO systems based on path angle information, *IEEE Transactions on Wireless Communications*, Vol. 18, Issue 1, 2019, pp. 608-619.

[9]. R. S. Dhar, L. Li, H. Ye, X. Wang, et al., Nanoscopically resolved dynamic charge-carrier distribution in operating interband cascade lasers, *Laser & Photonics Reviews*, Vol. 9, Issue 2, 2015, pp. 224-230.

[10]. L. Dai, X. Gao, Z. Wang, Energy-efficient hybrid precoding based on successive interference cancellation for millimeter-wave massive MIMO systems, in *Proceedings of the Radio and Antenna Days of the Indian Ocean (RADIO)*, 2015, pp. 1-2.

[11]. F. Sohrabi, W. Yu, Hybrid digital and analog beamforming design for large-scale antenna arrays, *IEEE Journal of Selected Topics in Signal Processing*, Vol. 10, Issue 3, 2016, pp. 501-513.

[12]. K. Zheng, S. Ou, X. Yin, et al., Massive MIMO channel models: A survey, *International Journal of Antennas and Propagation*, Vol. 2014, 2014, 848071.

[13]. H.-H. Tseng, Y.-F. Chen, S.-M. Tseng, Hybrid beamforming and resource allocation designs for mmWave multi-user massive MIMO-OFDM systems on uplink, *IEEE Access*, Vol. 11, 2023, pp. 133070-133085.

[14]. E. G. Larsson, O. Edfors, F. Tufvesson, T. L. Marzetta, Massive MIMO for next generation wireless systems, *IEEE Communications Magazine*, Vol. 52, Issue 2, 2014, pp. 186-195.

[15]. T. Lin, J. Cong, Y. Zhu, J. Zhang, et al., Hybrid beamforming for millimeter-wave systems using the MMSE criterion, *IEEE Transactions on Communications*, Vol. 67, Issue 5, 2019, pp. 3693-3708.

[16]. X. Gao, O. Edfors, F. Rusek, F. Tufvesson, Linear pre-coding performance in measured very large MIMO channels, in *Proceedings of the IEEE Vehicular Technology Conference (VTC Fall)*, 2011, pp. 1-5.

[17]. R. S. Dhar, S. G. Razavipour, E. Dupont, Z. Wasilewski, et al., Scanning voltage microscopy study of lasing and non-lasing terahertz quantum cascade lasers, in *Proceedings of the Conference on Lasers and Electro-Optics (CLEO)*, 2015, STu1G.4.

[18]. C. Masouros, M. Sellathurai, T. Ratnarajah, Large-scale MIMO transmitters in fixed physical spaces: The effect of transmit correlation and mutual coupling, *IEEE Transactions on Communications*, Vol. 61, Issue 7, 2013, pp. 2794-2804.

[19]. R. S. Dhar, D. Ban, Nanoscopic voltage distribution of operating cascade laser devices in cryogenic temperature, *Journal of Microscopy*, Vol. 262, Issue 3, 2016, pp. 226-231.

[20]. C. Masouros, J. Chen, K. Tong, M. Sellathurai, et al., Towards massive MIMO transmitters: On the effects of deploying increasing antennas in fixed physical space, in *Proceedings of the Future Network & Mobile Summit*, 2013, pp. 1-10.

[21]. G. S. Sahoo, A. Ghosh, Antenna array design for higher spectral efficiency for hybrid beamforming 5G MIMO wireless communication system, in *Proceedings of the IEEE Wireless Antenna and Microwave Symposium (WAMS)*, 2022, pp. 1-5.

[22]. R. Singh, P. Chawla, Performance analysis of OMP hybrid beamforming algorithm for OFDM millimeter-wave MIMO systems, in *Proceedings of the International Conference on Computing, Communication and Power Technology (IC3P)*, 2022, pp. 71-74.

[23]. R. S. Dhar, C. Xu, D. Ban, L. Li, et al., Direct observation of non-uniform electric field in the active regions of an interband cascade laser, in *Proceedings of the Conference on Lasers and Electro-Optics (CLEO)*, 2014, STh3G.3.

[24]. D. Ying, F. W. Vook, T. A. Thomas, D. J. Love, Hybrid structure in massive MIMO: Achieving large sum rate with fewer RF chains, in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2015, pp. 2344-2349.

[25]. F. Sohrabi, W. Yu, Hybrid analog and digital beamforming for mmWave OFDM large-scale antenna arrays, *IEEE Journal on Selected Areas in*

Communications, Vol. 35, Issue 7, 2017, pp. 1432-1443.

[26]. L. Dai, X. Gao, J. Quan, S. Han, et al., Near-optimal hybrid analog and digital precoding for downlink mmWave massive MIMO systems, in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2015, pp. 1334-1339.

[27]. Z. C. Phyo, A. Taparugssanagorn, Hybrid analog-digital downlink beamforming for massive MIMO system with uniform and non-uniform linear arrays, in *Proceedings of the 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, 2016, pp. 1-6.

[28]. L. Leyva, D. Castanheira, A. Silva, A. Gameiro, Hybrid beamforming design for communication-centric ISAC, *IEEE Sensors Journal*, Vol. 24, Issue 15, 2024, pp. 24567-24579.

[29]. D. H. N. Nguyen, L. B. Le, T. Le-Ngoc, R. W. Heath, Hybrid MMSE precoding and combining designs for mmWave multiuser systems, *IEEE Access*, Vol. 5, 2017, pp. 19167-19181.

[30]. O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, et al., Spatially sparse precoding in millimeter-wave MIMO systems, *IEEE Transactions on Wireless Communications*, Vol. 13, Issue 3, 2014, pp. 1499-1513.

[31]. Y. Chen, D. Chen, T. Jiang, L. Hanzo, Millimeter-wave massive MIMO systems relying on generalized sub-array-connected hybrid precoding, *IEEE Transactions on Vehicular Technology*, Vol. 68, Issue 9, 2019, pp. 8940-8950.

[32]. P.-A. Absil, R. Mahony, R. Sepulchre, Optimization Algorithms on Matrix Manifolds, *Princeton University Press*, Princeton, 2008.

[33]. D. H. N. Nguyen, L. B. Le, T. Le-Ngoc, Hybrid MMSE precoding for mmWave multiuser MIMO systems, in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2016, pp. 1-6.

[34]. X. Yu, J.-C. Shen, J. Zhang, K. B. Letaief, Alternating minimization algorithms for hybrid precoding in millimeter-wave MIMO systems, *IEEE Journal of Selected Topics in Signal Processing*, Vol. 10, Issue 3, 2016, pp. 485-500.

[35]. J. Cong, T. Lin, Y. Zhu, Hybrid MMSE beamforming for multiuser millimeter-wave communication systems, *IEEE Communications Letters*, Vol. 22, Issue 11, 2018, pp. 2390-2393.

[36]. J. Lee, Y. H. Lee, AF relaying for millimeter-wave communication systems with hybrid RF/baseband MIMO processing, in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2014, pp. 5838-5842.

[37]. V. Stankovic, M. Haardt, Generalized design of multi-user MIMO precoding matrices, *IEEE Transactions on Wireless Communications*, Vol. 7, Issue 3, 2008, pp. 953-961.

[38]. P. R. Gautam, L. Zhang, Hybrid precoding for millimeter-wave MIMO: Trace optimization approach, *IEEE Access*, Vol. 10, 2022, pp. 66874-66885.

[39]. Y. Chen, D. Chen, T. Jiang, L. Hanzo, Channel-covariance and angle-of-departure aided hybrid precoding for wideband multiuser millimeter-wave MIMO systems, *IEEE Transactions on Communications*, Vol. 67, Issue 12, 2019, pp. 8315-8328.

[40]. Y. Chen, Y. Xiong, D. Chen, T. Jiang, et al., Hybrid precoding for wideband millimeter-wave MIMO systems in the face of beam squint, *IEEE Transactions on Wireless Communications*, Vol. 20, Issue 3, 2021, pp. 1847-1860.

[41]. Y. Chen, D. Chen, Y. Tian, T. Jiang, Spatial lobes division-based low-complexity hybrid precoding and diversity combining for mmWave IoT systems, *IEEE Internet of Things Journal*, Vol. 6, Issue 2, 2019, pp. 3228-3239.

[42]. T. Kebede, Y. Wondie, J. Steinbrunn, H. B. Kassa, et al., Precoding and beamforming techniques in mmWave-massive MIMO: Performance assessment, *IEEE Access*, Vol. 10, 2022, pp. 16365-16387.

[43]. W. Park, J. Choi, Hybrid precoding and combining strategy for MMSE-based rate balancing in mmWave multiuser MIMO systems, *IEEE Access*, Vol. 10, 2022, pp. 88043-88057.

[44]. G. M. Zilli, W.-P. Zhu, Constrained-SVD based hybrid beamforming design for millimeter-wave communications, in *Proceedings of the 92nd IEEE Vehicular Technology Conference (VTC Fall)*, 2020, pp. 1-5.

[45]. K. Ying, Z. Gao, S. Lyu, Y. Wu, et al., GMD-based hybrid beamforming for large reconfigurable intelligent surface assisted millimeter-wave massive MIMO, *IEEE Access*, Vol. 8, 2020, pp. 19530-19539.

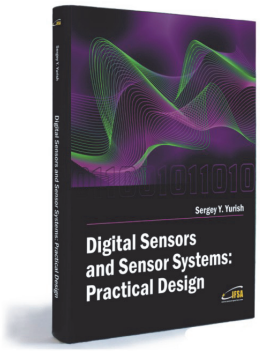
[46]. A. Alkhateeb, R. W. Heath, Gram-Schmidt based greedy hybrid precoding for frequency-selective millimeter-wave MIMO systems, in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 3396-3400.

Published by International Frequency Sensor Association (IFSA) Publishing, S. L., 2026 (http://www.sensorsportal.com).

International Frequency Sensor Association (IFSA) Publishing

Digital Sensors and Sensor Systems: Practical Design

Sergey Y. Yurish



The goal of this book is to help the practicians achieve the best metrological and technical performances of digital sensors and sensor systems at low cost, and significantly to reduce time-to-market. It should be also useful for students, lectures and professors to provide a solid background of the novel concepts and design approach.

Book features include:

- Each of chapter can be used independently and contains its own detailed list of references
- Easy-to-repeat experiments
- Practical orientation
- Dozens examples of various complete sensors and sensor systems for physical and chemical, electrical and non-electrical values
- Detailed description of technology driven and coming alternative to the ADC a frequency (time)-to-digital conversion

Formats: printable pdf (Acrobat) and print (hardcover), 419 pages

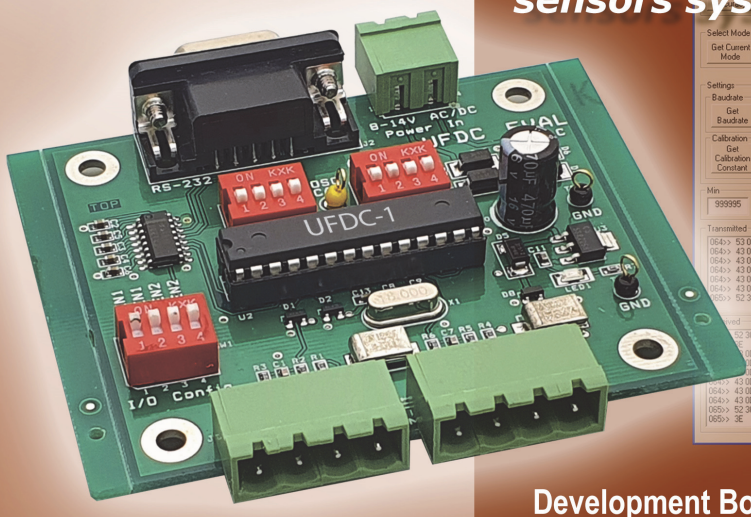
ISBN: 978-84-616-0652-8,
e-ISBN: 978-84-615-6957-1

Digital Sensors and Sensor Systems: Practical Design will greatly benefit undergraduate and at PhD students, engineers, scientists and researchers in both industry and academia. It is especially suited as a reference guide for practicians, working for Original Equipment Manufacturers (OEM) electronics market (electronics/hardware), sensor industry, and using commercial-off-the-shelf components

http://sensorsportal.com/HTML/BOOKSTORE/Digital_Sensors.htm

IFSA

Easy and quick
sensors systems development





Development Board UFDC-1/UFDC-1M-16

- 16 measuring modes, 2 channels for frequency measurements
- Frequency range from 0.05 Hz up to 7.5 MHz (120 MHz)
- Programmable accuracy from 1 % up to 0.001 %
- RS232 (USB optional)

sales@sensorsportal.com
http://www.sensorsportal.com/HTML/E-SHOP/PRODUCTS_4/Evaluation_board.htm

Call for Chapters

IFSA Publishing

Your Chapter may be in the next Volume of the popular Book Series

ADVANCES IN SENSORS 9

Open Access Book Series

Sinse 2012



- Physical, chemical and biosensors
- Sensor uncertainty, accuracy, calibration
- Smart sensors and systems
- Intelligent sensors and systems
- Artificial Intelligence enabled sensors
- Virtual and soft sensors
- Sensor instrumentation
- Sensor buses and networks
- Signal processing, sensor fusion
- Nanosensors
- Photonic sensors
- Sensor interfacing and signal conditioning
- Materials for sensors
- MEMS, BioMEMS and NEMS
- Remote Sensors and Telemetry
- Sensor packages and reliability

https://sensorsportal.com/ifsa_publishing.html

IFSA

International Frequency Sensor Association Publishing

www.sensorsportal.com

ISSN 1726-5479