

Antennas for Emerging 5G Systems

Lead Guest Editor: Khalil Sayidmarie

Guest Editors: Neil J. McEwan, Peter Excell, Raed A. Abd-Alhameed,
and Chan H. See





Antennas for Emerging 5G Systems

International Journal of Antennas and Propagation

Antennas for Emerging 5G Systems

Lead Guest Editor: Khalil H. Sayidmarie

Guest Editors: Neil J. McEwan, Peter Excell,

Raed A. Abd-Alhameed, and Chan H. See



Copyright © 2019 Hindawi. All rights reserved.

This is a special issue published in “International Journal of Antennas and Propagation.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

- Kush Agarwal, Singapore
Ana Alejos, Spain
Mohammad Ali, USA
Jaume Anguera, Spain
Rodolfo Araneo, Italy
Alexei Ashikhmin, USA
Hervé Aubert, France
Paolo Baccarelli, Italy
Xiulong Bao, Ireland
Pietro Bolli, Italy
Stefania Bonafoni, Italy
Paolo Burghignoli, Italy
Shah Nawaz Burokur, France
Giuseppe Castaldi, Italy
Luca Catarinucci, Italy
Felipe Catedra, Spain
Marta Cavagnaro, Italy
Maggie Y. Chen, USA
Shih Yuan Chen, Taiwan
Renato Cicchetti, Italy
Riccardo Colella, Italy
Laura Corchia, Italy
Lorenzo Crocco, Italy
Claudio Curcio, Italy
Francesco D'Agostino, Italy
María Elena de Cos Gómez, Spain
Tayeb A. Denidni, Canada
- Giuseppe Di Massa, Italy
Michele D'Urso, Italy
Francisco Falcone, Spain
Miguel Ferrando Bataller, Spain
Flaminio Ferrara, Italy
Vincenzo Galdi, Italy
Claudio Gennarelli, Italy
Farid Ghanem, Algeria
Sotirios K. Goudos, Greece
Rocco Guerriero, Italy
Kerim Guney, Turkey
Tamer S. Ibrahim, USA
M. Ramlee Kamarudin, Malaysia
Dmitry V. Kholodnyak, Russia
Slawomir Koziel, Iceland
Luis Landesa, Spain
Ping Li, Hong Kong
Ding-Bing Lin, Taiwan
Angelo Liseno, Italy
Pierfrancesco Lombardo, Italy
Lorenzo Luini, Italy
Atsushi Mase, Japan
Diego Masotti, Italy
Ch. F. Mecklenbräuker, Austria
A. Toaha Mobashsher, Australia
Ananda S. Mohan, Australia
J. -M. Molina-Garcia-Pardo, Spain
- Giuseppina Monti, Italy
Giorgio Montisci, Italy
Andrea Francesco Morabito, Italy
N. Nasimuddin, Singapore
Mourad Nedil, Canada
Symeon Nikolaou, Cyprus
Giacomo Oliveri, Italy
Mauro Parise, Italy
Ikmo Park, Republic of Korea
Josep Parrón, Spain
Matteo Pastorino, Italy
Ahmad Safaai-Jazi, USA
Safieddin Safavi-Naeini, Canada
Magdalena Salazar-Palma, Spain
Stefano Selleri, Italy
Raffaele Solimene, Italy
Gino Sorbello, Italy
Seong-Youp Suh, USA
Sheng Sun, Hong Kong
Larbi Talbi, Canada
Luciano Tarricone, Italy
Chien-Jen Wang, Taiwan
Wen-Qin Wang, China
Mustapha C E Yagoub, Canada
Shiwen Yang, China
Yuan Yao, China
Muhammad Zubair, Pakistan

Contents

Antennas for Emerging 5G Systems

Khalil H. Sayidmarie , Neil J. McEwan, Peter S. Excell, Raed A. Abd-Alhameed, and Chan H. See
Editorial (3 pages), Article ID 9290210, Volume 2019 (2019)

Downlink Multiuser Hybrid Beamforming for MmWave Massive MIMO-NOMA System with Imperfect CSI

Jing Jiang , Ming Lei , and Huanhuan Hou
Research Article (10 pages), Article ID 9764958, Volume 2019 (2019)

Energy-Efficient Hybrid Precoding Scheme Based on Antenna Selection Technology in Massive Multiple-Input Multiple-Output Systems

Jian Jun Ding and Jing Jiang 
Research Article (12 pages), Article ID 4196513, Volume 2019 (2019)

A Statistical Channel Model for Stochastic Antenna Inclination Angles

Gang Liu, Ming Zhang, and Yaming Bo 
Research Article (12 pages), Article ID 3487149, Volume 2019 (2019)

Covariance Matrix Reconstruction for Direction Finding with Nested Arrays Using Iterative Reweighted Nuclear Norm Minimization

Weijie Tan  and Xi'an Feng
Research Article (13 pages), Article ID 7657898, Volume 2019 (2019)

A Single-Layer Dual-Band Reflectarray Cell for 5G Communication Systems

Sandra Costanzo , Francesca Venneri, Antonio Borgia , and Giuseppe Di Massa
Research Article (9 pages), Article ID 9479010, Volume 2019 (2019)

Experimental Evaluation of MIMO-OFDM System with Rateless Space-Time Block Code

Ali H. Alqahtani , Khaled Humadi , Ahmed Iyanda Sulyman , and Abdulhameed Alsanie 
Research Article (8 pages), Article ID 6804582, Volume 2019 (2019)

Phase-Only Antenna Array Reconfigurability with Gaussian-Shaped Nulls for 5G Applications

Giulia Buttazzoni , Massimiliano Comisso , Federico Ruzzier, and Roberto Vescovo 
Research Article (8 pages), Article ID 9120530, Volume 2019 (2019)

Large-Scale Channel Modeling and Measurements for 10 GHz in Indoor Environments

Iury S. Batalha , Andréia V. R. Lopes, Jasmine P. L. Araújo, Fabrício J. B. Barros, Bruno L. S. Castro, Gervásio P. S. Cavalcante, and Evaldo G. Pelaes
Research Article (10 pages), Article ID 9454271, Volume 2019 (2019)

Editorial

Antennas for Emerging 5G Systems

Khalil H. Sayidmarie ¹, **Neil J. McEwan**,² **Peter S. Excell**,^{2,3}
Raed A. Abd-Alhameed,² and **Chan H. See**⁴

¹Department of Communication Engineering, Ninevah University, Mosul, Iraq

²Department of Biomedical and Electronics Engineering, University of Bradford, Bradford, UK

³Faculty of Art, Science and Technology, Wrexham Glyndwr University, Wrexham, UK

⁴School of Engineering and the Built Environment, Edinburgh Napier University, Edinburgh, UK

Correspondence should be addressed to Khalil H. Sayidmarie; kh.sayidmarie@uoninevah.edu.iq

Received 28 May 2019; Accepted 30 May 2019; Published 1 July 2019

Copyright © 2019 Khalil H. Sayidmarie et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

5G mobile communications technology will bring new experiences to industry and society, including higher data rates or greater capacity, higher bandwidth, increased security, and lower latency. These developments will create new opportunities for advancing society and businesses: 5G is going to make a significant change in how we live in this world.

The key enabling 5G technologies under development include novel multiple access strategies, massive MIMO, full digital beamforming or hybrid beamforming, ultra-dense networking, etc. The implementation of these high-level technologies will bring about new challenges for the designers of the physical infrastructure. These challenges undoubtedly include the antenna, but also the associated microwave systems and the characterisation of the radio propagation environment. In spite of very active current research on 5G antenna systems, including academic and industrial efforts, many challenges still need to be addressed, or they need better and more efficient solutions [1, 2].

5G will support significantly faster mobile broadband speeds and extensive mobile data usage as well as enabling the full potential of the Internet of Things [3]. From virtual reality and autonomous cars to the industrial Internet and smart cities, 5G will be at the heart of the future of communications. As well as conventional bands, 5G applications will have to utilize higher frequency bands in the millimetre wave region to enable wider bandwidth and higher data rates. The higher frequencies, wider bandwidths, and the need for beamforming, beam steering, and multiple beams are significant challenges to antenna designers [4]. Low-profile

efficient antennas and antenna arrays to ensure reliable and interference-free communications are needed, but requirements for increased power, larger bandwidth, higher gain, and insensitivity to the presence of the human user further complicate the antenna and propagation aspects. This indicates the need for novel ideas and very innovative solutions in antenna design [5–7].

This special issue collects eight papers concerning various aspects of 5G antennas and their system applications. Two papers deal with modelling issues, while three papers are concerned with multiple-input multiple-output (MIMO) systems which are expected to be extensively used in the future 5G systems. Another paper considers an important issue of synthesizing appropriate radiation patterns of linear and planar antenna arrays of arbitrary geometry to provide the desired coverage reconfiguration. One of the papers deals with a single-layer dual-band reflectarray cell that is proposed for future 5G systems. Another paper discusses the direction finding problem in a background of unknown nonuniform noise using nested arrays, where a novel gridless direction finding method is proposed via the low-rank covariance matrix approximation. The following paragraphs give more elaboration about these subjects.

There are two papers dealing with the issue of modelling of the 5G systems. The paper by I. S. Batalha et al. provides an investigation into large-scale modelling of the propagation channel for frequencies above 6 GHz. The modeling concentrates on indoor environments, which will be a significant part of the implemented 5G systems. The paper considers a

measurement investigation at 10 GHz with directional horn antennas in a computer room and a corridor within an academic office building. The paper presents experimental data and channel modelling with various polarization scenarios such as copolarizations H-H and V-V and cross-polarization V-H. Cases of both line-of-sight (LOS) and non-line-of-sight (NLOS) conditions were considered. The large-scale close-in reference is sustained by comprehensive analysis, studying propagation mechanisms such as reflection and diffraction. The results demonstrated that the established model had an inferior standard deviation in relation to measured data, proving itself more significant to propagation in indoor environments.

The other paper by G. Liu et al. also considers modelling and deals with the interaction between the mobile user, who is in dynamic movement, and the wave propagation mechanism. The movements of a person holding a mobile handset can be considered as a stochastic process since the users can change the way they handle the mobile handset very frequently in a short time. These random actions of the hand-held mobile will change the antenna inclination angles resulting in various multipath propagation pathways which will lead to changes in the received signal intensity. Very few studies or conventional channel models have been performed to investigate these features. In this paper, the relationships between the statistical characteristics of the electric field and antenna inclination angles are investigated. They are modeled, based on a three-dimensional fast ray-tracing method which considers both the reflections and diffraction events: then the radiation patterns of an antenna with arbitrary inclination angles are deducted and included in the method. Two different conditions of the line-of-sight (LOS) and non-line-of-sight (NLOS) pathways in an indoor environment are discussed. Furthermore, based on a statistical analysis, a semiempirical probability density function of the antenna inclination angles is presented. Finally, a novel statistical channel model for stochastic antenna inclination angles is proposed and the ergodic channel capacity is analyzed.

Three other papers deal with an important application of 5G systems, deploying the transmitting and receiving antennas in MIMO configurations. The paper by J. J. Ding and J. Jiang investigates the application of hybrid precoding in massive multiple-input multiple-output (MIMO) systems. This technique can reduce the number of radio frequency (RF) chains. However, the consumed power is still very high owing to the use of a large-scale antenna array. In this paper, the authors propose a precoding scheme based on antenna selection technology: this precoding scheme greatly increases the energy efficiency (EE) of the system. Firstly, they derive an exact closed-form expression for energy efficiency. They further study the relationship between the number of transmit antennas and energy efficiency on the basis of the exact closed-form expression of efficiency. An optimal value was found. When the number of transmit antennas equals this value, the EE of the system can reach a maximum by using an appropriate hybrid precoding scheme. An antenna selection algorithm to select from the transmit antennas was then proposed, where the number of selected antennas equals the optimal value. Subsequently, design of

the analog precoder based on a codebook to maximize the equivalent channel gain is presented, and then the EE can be improved by baseband digital precoding. The proposed precoding algorithm offers a compromise between spectral efficiency (SE) and EE in millimetre wave massive MIMO systems. Finally, simulation results are given to validate the theoretical analysis and show that a substantial EE gain can be obtained over the proposed precoding scheme without large performance loss.

Another paper deals with beamforming for millimetre wave massive MIMO-NOMA systems. The authors of the paper, J. Jiang et al., aim to provide a comprehensive scheme with limited feedback for downlink millimetre wave multiuser MIMO nonorthogonal multiple access (NOMA) systems. Based on the feedback of the best beam and the channel quality information (CQI) for this beam, the users are grouped into a cluster which has the same or coherent best beam and the maximal value of the CQI-difference. To further reduce the intercluster interference, only the candidate cluster can join the cluster set whose intercluster correlation with the existing clusters is lower than the threshold value. Based on the results of clustering, a hybrid beamforming strategy is designed. To improve the user experience, each cluster selects the best beam of the user with a higher guaranteed rate requirement as the analog beamforming vector. For digital beamforming, the weak user applies the block diagonalization algorithm based on the strong user's effective channel in order to reduce its intracluster interference. Finally, an intracluster power allocation algorithm is developed to maximize the power-difference in each cluster: this is beneficial to improve the successive interference cancellation (SIC) performance of the strong user. Simulation results are provided to show that the proposed MIMO-NOMA scheme offers a higher sum rate compared to the traditional orthogonal multiple access (OMA) schemes under practical conditions.

One of the papers considers the experimental evaluation of a MIMO-OFDM system with rateless space-time block code (STBC). A. H. Alqahtani et al. discuss the deployment of MIMO wireless technology in combination with orthogonal frequency division multiplexing (MIMO-OFDM) as an attractive technique for next-generation wireless systems. However, the performance of the wireless links is severely degraded because of the various channel impairments which cause a decoding failure and lead to packet loss at the receiver. One solution to such a problem is the rate-less space-time block code (RSTBC). The paper presents experimental results on the performance of a 2×2 MIMO-OFDM system with RSTBC as measured in a testbed implemented with a field programmable gate array (FPGA). The average bit error rate (BER) performance of the proposed scheme is evaluated experimentally, and the results agree closely with the simulation and analytical upper bound. The authors show that RSTBC can be implemented in real-world scenarios and will guarantee the reliability of loss-prone wireless channels.

The paper by G. Buttazzoni et al. considers the issue of antenna array reconfiguration by changing phase-only excitations to achieve Gaussian-shaped nulls. The proposed technique is intended for 5G applications. The paper presents

a fast iterative method for the synthesis of linear and planar antenna arrays of arbitrary geometry that offer reconfigurable patterns for 5G applications. The method enables the generation of wide null regions shaped according to a Gaussian distribution: the results correlate with recent measurements on millimetre-wave angular dispersion. An experimental approach of phase-only control considers moving from the pattern provided by a uniformly excited array and iteratively modifying the phases of the excitations. This allows simplification in the feed network of the array as attenuators are not needed to vary the excitation magnitudes, therefore reducing the cost of realization of 5G base stations and mobile terminals. The proposed algorithm is based on the method of successive projections and relies on closed-form expressions for both the projectors and the null positions. It thus allows fast computation of the excitation phases at each iteration. Numerical examples compliant with 5G millimetre wave standards are presented to check the effectiveness of the proposed solution through scenarios that involve linear and concentric ring arrays.

In another paper, S. Costanzo et al. present a design for a single-layer dual-band reflectarray cell intended for future 5G system applications. A reflectarray unit cell operating at 28/38 GHz is designed by adopting two pairs of miniaturized fractal patches, offering low losses (<0.7 dB) and full phase ranges (about 320°) at both operating frequencies. Dual-band operation achieved by utilizing fractal geometry is another application of this principle, which has been used for increasing the phase range at lower phase slope [8]. The proposed configuration offers very small interelement spacing and negligible mutual coupling effects between the two bands, thus enabling an independent phase tuning mechanism for both of the desired frequency bands. The compact cell demonstrated the reflectarray's abilities in achieving fixed scanned-beam and/or multibeam patterns, under the dual-band operation mode: such characteristics are beneficial and desirable in 5G systems. Full-wave numerical validations, performed on the synthesized reflectarray structures, confirmed the effectiveness of the designed dual-band configuration in achieving independent radiation patterns and good bandwidths, at the two designed frequencies.

Another paper considers yet another important issue in 5G systems: the reconstruction of the direction of the received waves. W. Tan and X. Feng address the direction finding problem in the background of unknown nonuniform noise with the nested array. A novel gridless direction finding method is proposed using the low-rank covariance matrix approximation. This approach is based on reweighted nuclear norm optimization. For the proposed method, the noise variance variable is eliminated by a linear transform, and then the covariance fitting criterion is utilized to determine the regularization parameter for ensuring robustness. The reconstructed low-rank covariance matrix is then obtained by iteratively reweighted nuclear norm optimization that imposes a nonconvex penalty. The search-free DOA estimation method is then utilized to perform the parameter estimation. The results of numerical simulations are given to verify the effectiveness of the proposed method. Moreover, results indicate that the proposed method gives a more

accurate DOA estimation of both the nonuniform noise and off-grid cases compared to a state-of-the-art DOA estimation algorithm.

Conflicts of Interest

The editors declare that they have no conflicts of interest regarding the publication of this special issue.

Acknowledgments

The guest editorial team would like to express their gratitude to all the authors for their interest in selecting this special issue for their publications. The editors also wish to extend their appreciation to the anonymous reviewers for their careful reading of the twelve manuscripts submitted to this special issue and their valuable comments, suggestions, and following up of the revised versions of papers.

Khalil H. Sayidmarie
Neil J. McEwan
Peter S. Excell
Raed A. Abd-Alhameed
Chan H. See

References

- [1] W. H. Chin, Z. Fan, and R. Haines, "Emerging technologies and research challenges for 5G wireless networks," *IEEE Wireless Communications Magazine*, vol. 21, no. 2, pp. 106–112, 2014.
- [2] M. Shafi, A. F. Molisch, P. J. Smith et al., "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, 2017.
- [3] C. H. See, R. A. Abd-Alhameed, A. A. Atojoko, N. J. McEwan, and P. S. Excell, "Link budget maximization for a mobile-band subsurface wireless sensor in challenging water utility environments," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 1, pp. 616–625, 2018.
- [4] W. Hong, Z. H. Jiang, C. Yu et al., "Multibeam antenna technologies for 5G wireless communications," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6231–6249, 2017.
- [5] N. O. Parchin, Y. I. Al-Yasir, A. H. Ali et al., "Eight-element dual-polarized MIMO slot antenna system for 5G smartphone applications," *IEEE Access*, vol. 7, pp. 15612–15622, 2019.
- [6] W. Hong, K.-H. Baek, and S. Ko, "Millimeter-wave 5G antennas for smartphones: overview and experimental demonstration," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6250–6261, 2017.
- [7] M. Alibakhshikenari, M. Khalily, B. S. Virdee, C. H. See, R. A. Abd-Alhameed, and E. Limiti, "Mutual coupling suppression between two closely placed microstrip patches using EM-bandgap metamaterial fractal loading," *IEEE Access*, vol. 7, pp. 23606–23614, 2019.
- [8] K. H. Sayidmarie and M. E. Bialkowski, "Fractal unit cells of increased phasing range and low slopes for single-layer microstrip reflectarrays," *IET Microwaves, Antennas & Propagation*, vol. 5, no. 11, pp. 1371–1379, 2011.

Research Article

Downlink Multiuser Hybrid Beamforming for MmWave Massive MIMO-NOMA System with Imperfect CSI

Jing Jiang , Ming Lei , and Huanhuan Hou

Shaanxi Key Laboratory of Information Communication Network and Security, Xi'an University of Posts and Telecommunications, Xi'an 710121, China

Correspondence should be addressed to Ming Lei; leiming19940812@163.com

Received 29 August 2018; Revised 1 March 2019; Accepted 28 April 2019; Published 13 May 2019

Guest Editor: Khalil Sayidmarie

Copyright © 2019 Jing Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper aims to provide a comprehensive scheme with limited feedback for downlink millimeter wave (mmWave) multiuser multiple-input multiple-output (MIMO) nonorthogonal multiple access (NOMA) system. Based on the feedback of the best beam and the channel quality information (CQI) on this beam, the users are grouped into a cluster having the same or coherent best beam and the maximal CQI-difference. To further reduce the intercluster interference, only the candidate cluster can join the cluster set whose intercluster correlation with the existing clusters is lower than threshold. Based on the results of clustering, mmWave hybrid beamforming is designed. To improve the user experience, each cluster selects the best beam of the user with the higher guaranteed rate requirement as the analog beamforming vector. For digital beamforming, the weak user applies the block diagonalization algorithm based on the strong user's effective channel to reduce its intracluster interference. Finally, an intracluster power allocation algorithm is developed to maximize the power difference in each cluster which is beneficial to improve the successive interference cancelation (SIC) performance of the strong user. Finally, simulation results show that the proposed MIMO-NOMA scheme offers a higher sum rate than the traditional orthogonal multiple access (OMA) scheme under practical conditions.

1. Introduction

In contrast to the conventional orthogonal multiple access (OMA), nonorthogonal multiple access (NOMA) allows different users to efficiently share the same resources (i.e., time, frequency, space, and code) at different power levels. Exploiting their respective channel gain differences, multiuser signals are separated at receivers by successive interference cancelation (SIC). Thus, it is an important technique for 5G systems which can significantly enhance system capacity and overall spectral efficiency [1–3].

NOMA can flexibly combine other 5G technologies, e.g., MIMO, cognitive radio, cooperative communication, and channel coding [3]. Because MIMO-NOMA systems can significantly increase the number of supportable users and in turn improve the system sum capacity, the research of MIMO-NOMA has been carried out widely [4–7]. Multiple users with distinct channel gains are grouped into MIMO-NOMA clusters. The users in the same cluster are scheduled

on NOMA basis. The intracluster interference is mitigated by SIC based on the obvious channel gain difference. Multiple clusters may utilize the multiuser MIMO principles to cancel the intercluster interference. References [4–8] prove that NOMA combined with MIMO techniques could achieve great system performance gains over OMA.

Thanks to the abundant bandwidth resources, millimeter wave (mmWave) frequency becomes a natural choice to achieve Gigabit data rates [9, 10]. The number of supportable users can be more than the number of radio frequency (RF) chains at the same resources by exploiting MIMO-NOMA in mmWave system. On the other hand, the channel of mmWave is highly directional. For mmWave MIMO-NOMA system, the users in a same cluster use the same beam and the different clusters use the distinct different beams. For the users in a same cluster, the obvious power difference in a same beam could assure good SIC performance. For different clusters, the distinct different beams are beneficial to greatly reduce the intercluster inference [11, 12]. Therefore,

mmWave MIMO-NOMA can effectively improve the spectral efficiency and cope with the demands of massive connectivity.

Considering mmWave MIMO-NOMA system, [11] analyzes the achievable sum rate of the proposed beamspace MIMO-NOMA in a typical mmWave channel mode, which shows an obvious performance gain compared with the existing beamspace MIMO. Zero-forcing beamforming (ZF-BF) is designed to reduce the interbeam interference in the beamspace. Furthermore, a dynamic power allocation is proposed to maximize the achievable sum rate which contains the intrabeam and interbeam power optimization. Reference [12] proposes a hybrid analog/digital beamforming scheme with a power allocation algorithm aiming to maximize the energy efficiency. Users having a high channel correlation and large channel gain difference are grouped as a cluster. Reference [13] investigates the maximization of the sum rate of the system with clusters of 2-user or more-user. It decomposes the original problem into power and beam gain allocation problems. Reference [14] proposes a hybrid analog/digital beamforming framework to maximize the sum rate and analyzes the effect of beam misalignment on rate performance.

Most of the existing works on mmWave MIMO-NOMA, such as [11–14], have assumed perfect knowledge of channel state information (CSI) at the transmitter, which is difficult to realize in practice. Reference [15] proposes a hybrid precoding scheme based on one-stage feedback, which makes use of all feedback overhead to enable precise beam information and takes advantage of the second-order channel statistics to mitigate multiuser interference. Reference [16] proposes the random beamforming in which base station (BS) broadcasts the random beams and the users with the enough signal strength feed their channel quality information (CQI) back to the BS.

MmWave MIMO-NOMA system with limited feedback has been studied but with limited results. Reference [15] has not yet considered user clustering and the intracluster interference cancelation with limited feedback. Reference [16] reduces the feedback overhead, but the channel measurement is time-consuming and complicated. In this paper, we put forward a comprehensive scheme with imperfect CSI for mmWave MIMO-NOMA system. To reduce the hardware complexity, we apply low RF chains structure at the BS, where the hybrid analog/digital beamforming is considered. The major contributions of this paper can be listed as follows:

(1) The user clustering is designed based on the limited feedback scheme. References [16–21] put forward some feedback designs for MIMO-NOMA, e.g., on bit feedback, the path loss information, and other second-order statistics. Considering mmWave user could acquire the big gain in the specific beam, one-bit feedback or the path loss information is not accurate enough. In this paper, the limited feedback uses the best beam and CQI value based on this beam. The users having the same or adjacent best beam are sorted based on their CQI values, and the users with the maximal CQI-difference are formed a cluster. To further reduce the intercluster interference, the user clustering algorithm applies

the intercluster correlation threshold to assure only the cluster with the distinct different beam can join the cluster set. Numerical results prove that it acquires obvious sum rate improvement by adequately exploiting the high directional feature of mmWave.

(2) We propose a novel hybrid analog/digital beamforming algorithm. In general, the users in a same cluster will use the beam of the strong user as the beamforming vector [11, 12]. The beamforming gains are assured to achieve first for the strong users. However, the weak users would require higher data rate. To improve the quality of user experience and the user fairness, for each cluster, we select the beam of the user with the higher rate requirement as the analog beamforming vector. For digital beamforming, the prior work [12] applies ZF-BF to cancel the intercluster interference. For mmWave communication, the distinct different beams cause very little interference as to the sparsity of mmWave channel. In this paper, the digital beamforming is designed for the strong user and the weak user, respectively. The strong user utilizes the singular value decomposition (SVD) of its effective channel to acquire more gains from digital beamforming. Especially, the weak user exploits the block diagonalization (BD) algorithm to mitigate the intracluster interference from the strong user in the same cluster.

(3) As the power difference between the two users in a same cluster increases, not only is the SIC performance of the strong user improved, but also the intracluster interference of the weak user is reduced. So we derive an intracluster power allocation scheme that maximizes the power difference under the users' quality of service (QoS) requirements. We derive a closed-form optimal power allocation for the two users in a cluster by deriving the exact bounds for the power allocation coefficient region.

(4) Finally, we evaluate the sum rate performance of the downlink MIMO-NOMA system using the proposed beamforming, user clustering and power allocation algorithms. The simulations are done with a wide range of beam correlation thresholds, including the interuser and intercluster thresholds in the user clustering algorithm. Both of the perfect CSI and random selection cases are considered in the simulations. Numerical results also compare the sum rate performances of MIMO-OMA and the proposed MIMO-NOMA and illustrate the significance of the proposed MIMO-NOMA scheme.

The rest of this paper is organized as follows. Section 2 presents the system and channel model of mmWave MIMO-NOMA system. Sections 3, 4, and 5 describe the designs of user clustering, hybrid beamforming and power allocation, respectively. The simulation configuration and results are presented and discussed in Section 6, and finally, Section 7 concludes this paper.

Notation. In the reminder of this paper, we use the following notations: $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and Hermitian transpose, respectively, $\|\cdot\|$ denotes the Frobenius norm, and $\mathcal{E}\mathcal{N}(a, b)$ denotes the distribution of circularly symmetric complex Gaussian random variable with mean a and covariance b .

2. System and Channel Model

Consider a downlink mmWave beamforming NOMA transmission system, in which one BS communicates with L clusters. Each cluster consists of multiple single-antenna users. The BS is equipped with N_{TX} antennas and N_{RF} ($N_{\text{TX}} \geq N_{\text{RF}}$) RF chains for reducing the hardware complexity. We consider the fully connected architecture that each RF chain is linked to all antennas by using multiple phase shifters. We assume that the number of clusters served by the BS is equal to that of RF chains ($L = N_{\text{RF}}$). Although each cluster can contain more than two users, we suppose here that there are two users in each cluster for simplicity. This is consistent with the standard implementation of NOMA in long term evolution advanced (LTE-A) as well [22].

On the downlink, the BS applies an $N_{\text{RF}} \times 2L$ baseband beamforming $\mathbf{W} = [\mathbf{w}_{1,1}, \mathbf{w}_{1,2}, \dots, \mathbf{w}_{L,1}, \mathbf{w}_{L,2}]$ followed by an $N_{\text{TX}} \times N_{\text{RF}}$ RF beamforming $\mathbf{F}_{\text{RF}} = [\mathbf{f}_1^{\text{RF}}, \mathbf{f}_2^{\text{RF}}, \dots, \mathbf{f}_{N_{\text{RF}}}^{\text{RF}}]$. The sampled transmitted signal is therefore

$$\mathbf{x} = \mathbf{F}_{\text{RF}} \mathbf{W} \mathbf{s}, \quad (1)$$

where $\mathbf{s} = [\sqrt{P_1} \alpha_1 s_{1,1}, \sqrt{P_1} (1 - \alpha_1) s_{1,2}, \dots, \sqrt{P_L} \alpha_L s_{L,1}, \sqrt{P_L} (1 - \alpha_L) s_{L,2}]^T$ is the $2L \times 1$ vector of transmitted symbols in which $s_{l,1}$ and $s_{l,2}$ denote the signals for the 1-th and 2-th in the l -th cluster, respectively. Similarly, $0 < \alpha_l < 1$ and $1 - \alpha_l$ stand for the power allocation coefficients for the two users in the l -th cluster. P_l is the power allocated to the l -th cluster.

The received signal of the i -th user in the l -th cluster can be represented as follows:

$$\begin{aligned} y_{l,i} &= \mathbf{h}_{l,i} \mathbf{x} + n_{l,i} \\ &= \mathbf{h}_{l,i} \mathbf{F}_{\text{RF}} \left(\mathbf{w}_{l,1} \sqrt{P_l} \alpha_l s_{l,1} + \mathbf{w}_{l,2} \sqrt{P_l} (1 - \alpha_l) s_{l,2} \right) \\ &\quad + \mathbf{h}_{l,i} \mathbf{F}_{\text{RF}} \sum_{j \neq l} \left(\mathbf{w}_{j,1} \sqrt{P_j} \alpha_j s_{j,1} + \mathbf{w}_{j,2} \sqrt{P_j} (1 - \alpha_j) s_{j,2} \right) \\ &\quad + n_{l,i}, \end{aligned} \quad (2)$$

where $\mathbf{h}_{l,i}$ is the $1 \times N_{\text{TX}}$ vector that represents the mmWave channel of the i -th user in the l -th cluster. $n_{l,i} \sim \mathcal{CN}(0, \sigma^2)$ denotes additive white complex Gaussian noise. Moreover, the second term is the intercluster interference from other clusters. Without loss of generality, for each cluster, we assume that $\|\mathbf{h}_{l,1} \mathbf{F}_{\text{RF}} \mathbf{w}_{l,1}\| \geq \|\mathbf{h}_{l,2} \mathbf{F}_{\text{RF}} \mathbf{w}_{l,2}\|$ where $1 \leq l \leq L$. Following this, the 1-th and 2-th are defined as the strong user and weak user in each cluster, respectively.

Due to the limited scattering in mmWave channel, we adopt a well-established geometric channel model with P scatterers [23, 24]. We assume that each scatterer contributes a single propagation path from the BS to user. Then, the channel $\mathbf{h}_{l,i}$ ($i \in \{1, 2\}$, $l \in \{1, \dots, L\}$) can be expressed as

$$\mathbf{h}_{l,i} = \sqrt{N_{\text{TX}}} \frac{\beta_{l,i}^0 \mathbf{a}(\theta_{l,i}^0)^{\text{H}}}{\sqrt{1 + d_{l,i}^{\eta_{\text{LOS}}}}} + \sqrt{N_{\text{TX}}} \sum_{f=1}^P \frac{\beta_{l,i}^f \mathbf{a}(\theta_{l,i}^f)^{\text{H}}}{\sqrt{1 + d_{l,i}^{\eta_{\text{NLOS}}}}}, \quad (3)$$

where the first and second terms denote the line-of-sight (LoS) component and non-line-of-sight (NLoS) components of the i -th user in the l -th cluster, respectively. $d_{l,i}$ denotes the distance from the BS to the i -th user in the l -th cluster. η_{LOS} and η_{NLOS} are the path loss exponents corresponding to the LoS and NLoS paths, respectively. $\beta_{l,i}^f \in \mathcal{CN}(0, 1)$ where $0 \leq f \leq P$ denotes the complex gain of the f -th path between the BS and the i -th user in the l -th cluster. $\theta_{l,i}^f \in [-1, 1]$ is the f -th path's normalized direction, whereas $\mathbf{a}(\theta_{l,i}^f)$ is the antenna array steering vector with respect to $\theta_{l,i}^f$. If we assume a uniform linear array (ULA) is used at the BS, $\mathbf{a}(\theta_{l,i}^f)$ can be written as [25]

$$\mathbf{a}(\theta_{l,i}^f) = \frac{1}{\sqrt{N_{\text{TX}}}} \left[1, e^{-j\pi \theta_{l,i}^f}, \dots, e^{-j\pi (N_{\text{TX}}-1) \theta_{l,i}^f} \right]^T. \quad (4)$$

where $\theta_{l,i}^f \in [-1, 1]$ is related to the angle of departure $\vartheta \in [-\pi/2, \pi/2]$ as $\theta_{l,i}^f = (2D/\lambda) \sin(\vartheta)$ [26]. D denotes the distance between antenna elements and λ denotes the signal wavelength satisfying $D = 2/\lambda$ at mmWave frequencies.

3. User Clustering

Appropriately selecting two users which are served in a cluster can help improve the performance of NOMA multiuser beamforming. On the one hand, the big difference of the channel gains can improve SIC performance of the users with high channel gains and reduce the intracluster interference of the users with low channel gains. On the other hand, the beamforming vector is shared by all users in the same cluster. If the channels of those users in a same cluster are highly correlated, the beamforming vectors can acquire the definite array gains and properly cancel the intercluster interference from other clusters. Reference [4] proposed a clustering algorithm selecting user-pair having a high channel correlation and large channel gain difference as set in the same cluster, which requires full CSI. However, the full CSI feedback to the BS is not feasible in practice due to prohibitively high feedback overhead. So we propose a user clustering algorithm with limited feedback information, which only requires the indexes of the best beams and the channel quality information (CQI) values.

The index of the best beam and the user's CQI are measured and sent back as follows. In the first stage, the BS broadcasts pilot signals on the every vector in the beamforming codebook. In the second stage, each user chooses the best beam and measures the CQI with this beamforming vector. At last, the user sends the index of the best beam and CQI to the BS. The best beam of the i -th user $\hat{\mathbf{b}}_i$ can be selected according to the following criterion as [27]

$$\hat{\mathbf{b}}_i = \arg \max_{\mathbf{b}_i \in \Phi} |\mathbf{h}_i \mathbf{b}_i|^2, \quad (5)$$

where Φ represents the beamforming codebook with size $|\Phi| = 2^{B_{\text{RF}}}$ (i.e., B_{RF} feedback bits) and consists of the steering vectors $\mathbf{a}(-1 + 2k_Q/2^{B_{\text{RF}}})$, $k_Q = 0, 1, \dots, 2^{B_{\text{RF}}} - 1$. They have

the same form of the array response vector in (4). \mathbf{b}_i is the beamforming vector selected from Φ .

As discussed in [24, 28], the effect of LoS component is dominant in mmWave channels, compared to those of NLoS components. Therefore, the above mmWave channel model in (3) can be simplified as

$$\mathbf{h}_{l,i} = \sqrt{N_{\text{TX}}} \frac{\beta_{l,i} \mathbf{a}(\theta_{l,i})^{\text{H}}}{\sqrt{1 + d_{l,i}^n}}, \quad (6)$$

where the subscript 0 of the LoS component is omitted for the simplification of notation. Then, the effective channel gain of the user i on the beam \mathbf{b}_i in (5) can be expressed as

$$\begin{aligned} |\mathbf{h}_i \mathbf{b}_i|^2 &= \frac{N_{\text{TX}} |\beta_i|^2 |\mathbf{a}(\theta_i)^{\text{H}} \mathbf{a}(\varphi_i)|^2}{1 + d_{l,i}^n} \\ &= \frac{|\beta_i|^2 \left| \sum_{l=0}^{N_{\text{TX}}-1} e^{-j\pi l(\theta_i - \varphi_i)} \right|^2}{N_{\text{TX}} (1 + d_{l,i}^n)} \\ &= \frac{|\beta_i|^2 \sin^2(\pi N_{\text{TX}} (\theta_i - \varphi_i) / 2)}{N_{\text{TX}} (1 + d_{l,i}^n) \sin^2(\pi (\theta_i - \varphi_i) / 2)} \\ &= \frac{|\beta_i|^2}{1 + d_{l,i}^n} F_{N_{\text{TX}}}(\pi [\theta_i - \varphi_i]), \end{aligned} \quad (7)$$

where φ_i is the predefined normalized direction of \mathbf{b}_i . $F_{N_{\text{TX}}}(x)$ is the Fejér kernel, which goes to zero quickly when x increases. Therefore, based on the criterion in (5), the normalized direction $\hat{\varphi}_i$ of the selected beam $\hat{\mathbf{b}}_i$ has the least difference from θ_i in the codebook. This means that the normalized direction of each user can be approximately the normalized direction of its selected beam. The correlation of the selected beam between user i and user j can be rewritten as

$$\begin{aligned} \rho_{(i,j)} &= \frac{|\hat{\mathbf{b}}_i^{\text{H}} \hat{\mathbf{b}}_j|}{\|\hat{\mathbf{b}}_i\| \|\hat{\mathbf{b}}_j\|} = \frac{|\mathbf{a}(\hat{\varphi}_i)^{\text{H}} \mathbf{a}(\hat{\varphi}_j)|}{\|\mathbf{a}(\hat{\varphi}_i)\| \|\mathbf{a}(\hat{\varphi}_j)\|} \\ &= \frac{\left| \sum_{l=0}^{N_{\text{TX}}-1} e^{-j\pi l(\hat{\varphi}_i - \hat{\varphi}_j)} \right|}{N_{\text{TX}}} = \sqrt{\frac{F_{N_{\text{TX}}}(\pi [\hat{\varphi}_i - \hat{\varphi}_j])}{N_{\text{TX}}}}. \end{aligned} \quad (8)$$

Based on the above beam correlation expression in (8) and channel expression in (6), the channel correlation between users i and user j can be simplified as

$$\begin{aligned} \text{Corrh}_{(i,j)} &= \frac{|\mathbf{h}_i \mathbf{h}_j^{\text{H}}|}{\|\mathbf{h}_i\| \|\mathbf{h}_j\|} = \frac{|\mathbf{a}(\theta_i)^{\text{H}} \mathbf{a}(\theta_j)|}{N_{\text{TX}}} \\ &= \frac{\left| \sum_{l=0}^{N_{\text{TX}}-1} e^{-j\pi l(\theta_i - \theta_j)} \right|}{N_{\text{TX}}} \end{aligned}$$

$$\begin{aligned} &= \sqrt{\frac{F_{N_{\text{TX}}}(\pi [\theta_i - \theta_j])}{N_{\text{TX}}}} \\ &\stackrel{(a)}{\approx} \sqrt{\frac{F_{N_{\text{TX}}}(\pi [\hat{\varphi}_i - \hat{\varphi}_j])}{N_{\text{TX}}}} = \frac{|\hat{\mathbf{b}}_i^{\text{H}} \hat{\mathbf{b}}_j|}{\|\hat{\mathbf{b}}_i\| \|\hat{\mathbf{b}}_j\|} = \rho_{(i,j)}, \end{aligned} \quad (9)$$

where $\text{Corrh}_{(i,j)}$ denotes the channel correlation between user i and user j . (a) follows from that θ_i and θ_j are approximately $\hat{\varphi}_i$ and $\hat{\varphi}_j$ based on the above analysis, respectively. So the correlation between the users' channels $\text{Corrh}_{(i,j)}$ can be replaced with the correlation between their selected beams $\rho_{(i,j)}$.

CQI is an indicator carrying information about the channel condition. In LTE, considering a single BS system, the received SNR for each user i can be written as [29]

$$\kappa_i = \frac{P_t \|\mathbf{h}_i \hat{\mathbf{b}}_i\|^2}{\sigma^2}, \quad (10)$$

where P_t is the transmit power and σ^2 is the power of received noise and interference. The measured SNR κ_i is then mapped to a discrete CQI value g_i using a chipset vendor specific mapping table [30]. Because the positive correlation between channel gain and CQI value, we will just calculate the difference of CQI values between users instead of their channel gain difference.

It is proved in [8] that the more users are admitted to a cluster, the lower is the achieved sum rate, which illustrates the tradeoff between the sum rate and maximum number of admitted users. Considering the tradeoff, we assume that each cluster admits 2 users. Based on the above analysis, a limited feedback downlink MIMO-NOMA user clustering algorithm is described in Algorithm 1, in which the number of clusters $L = N_{\text{RF}}$ and each cluster contains 2 users.

Algorithm 1 can be summarized as follows. In the first step, the BS calculates the beam correlation of all possible user-pairs and puts the user-pair having a beam correlation higher than the threshold μ into set T . In the second step, during each loop, the BS searches for a new user-pair (i^*, j^*) having the maximum CQI value difference from T . Then the BS calculates the beam correlations between the user i^* and the selected strong users in set A . If all the correlations are lower than the threshold ζ , then the user-pair (i^*, j^*) is added to a new cluster and the user i^* is put into A . Otherwise, no users are added. The loop is repeated consecutively until L user clusters are selected.

4. Two-Stage Hybrid Beamforming

4.1. The Optimization of Analog Beamforming. For analog beamforming, one analog beamforming vector is designed for one user in traditional design [31], while two users in a same cluster should share one analog beamforming vector in MIMO-NOMA. Furthermore, considering the fairness of QoS, the rate requirement of the weak user may be greater than that of the strong user in a cluster. So we select the

Input and Initialization:
Input: the number of clusters L , feedback beam $\hat{\mathbf{b}}_i$ and CQI value g_i , for $\forall i$.
Initialize: $T = \emptyset$, $A = \emptyset$, $l = 0$, $SC_l = \emptyset$, $Z = \emptyset$.
where T is the candidate set of user-pair that can be selected as a cluster. l is the number of selected clusters. A is the set of the strong users of the selected clusters. SC_l is the set of the selected users in the l -cluster. Z is the set of weak users that have been selected.
Output: the user set selected for each cluster SC_l , for $l = 1, 2, \dots, L$.
Step (1) Select all user-pairs having a beam correlation higher than the threshold:
The BS sorts users according to the ascending CQI value: $g_1 \geq g_2 \geq \dots \geq g_K$.
for $i = 1 : K-1$
 for $j = K : -1 : i + 1$
 if $(\rho_{(i,j)} \geq \mu)$ AND $(j \notin Z)$
 $T = T \cup (i, j)$
 $Z = Z \cup \{j\}$
 break
 end if
 end for
end for
where μ is the predefined beam correlation threshold ($0 \leq \mu \leq 1$).
Step (2) Pair selection for clustering:
while $l < L$
(a) Select the user-pair that has the maximum CQI value difference in T :
 $(i^*, j^*) = \arg \max_{(i,j) \in T} d_{(i,j)} = \arg \max_{(i,j) \in T} g_i - g_j$
(b) Put the selected user-pair having correlations with the users in A below the threshold into a new cluster:
if $(\rho_{(i^*,k)} \leq \zeta, \forall k \in A)$
 $SC_l = SC_l \cup \{(i^*, j^*)\}$
 remove the pairs in T that include user i^* or j^*
 $A = A \cup \{i^*\}$
 $l \leftarrow l + 1$
else
 $T = T \setminus \{(i^*, j^*)\}$
end if
end while
where ζ is the predefined inter-cluster beam correlation threshold ($0 \leq \zeta \leq 1$).

ALGORITHM 1: User clustering with beam and CQI.

feedback beam of the user with the maximum guaranteed rate requirement as the analog beamforming vector in each cluster, i.e., $\hat{\mathbf{t}}_l = \arg \max_{i \in SC_l} \tilde{R}_{l,i}$, where $\tilde{R}_{l,i}$ is the guaranteed rate requirement of the i -user in the l -cluster.

Based on the above analysis, we can design the analog beamforming vector for the l -th cluster according to the following criterion:

$$\mathbf{f}_l^{\text{RF}} = \hat{\mathbf{b}}_{\hat{\mathbf{t}}_l}. \quad (11)$$

Then, the analog beamforming matrix can be obtained as $\mathbf{F}_{\text{RF}} = [\mathbf{f}_1^{\text{RF}}, \mathbf{f}_2^{\text{RF}}, \dots, \mathbf{f}_{N_{\text{RF}}}^{\text{RF}}]$.

4.2. The Design of Digital Beamforming. In this subsection we design the digital beamforming. After the analog beamforming, the BS trains the effective channels with users as

$$\tilde{\mathbf{h}}_{l,i} = \mathbf{h}_{l,i} \mathbf{F}_{\text{RF}}, \quad \text{for } l = 1, \dots, L \text{ and } i = 1, 2. \quad (12)$$

The dimension of the effective channel $\tilde{\mathbf{h}}_{l,i}$ is $1 \times N_{\text{RF}}$ which is much less than that of its original channel $\mathbf{h}_{l,i}$.

This is different from the algorithms proposed in [32, 33] in which the effective channel has a larger $1 \times N_{\text{TX}}$ dimension. Then, each user quantizes its effective channel utilizing a random vector quantization (RVQ) codebook η of size $2^{B_{\text{EC}}}$ and feeds the index of the vector $\hat{\mathbf{h}}_{l,i}$ back to the BS with B_{EC} bits, where $\hat{\mathbf{h}}_{l,i} = \arg \max_{\hat{\mathbf{h}}_{l,i} \in \eta} \|\hat{\mathbf{h}}_{l,i} \hat{\mathbf{h}}_{l,i}^{\text{H}}\|$ [34]. Finally, the BS designs its digital beamforming based on the quantized channels.

The user clustering and analog beamforming mitigate the intercluster interference. Then we can design the digital beamforming vector for the weak user to reduce its intracluster interference. We take the design of the digital beamforming vector $\mathbf{w}_{l,2}$ for the weak user in the l -th cluster as an example. First, we construct a complementary matrix $\bar{\mathbf{H}}_{l,2}$ that contains the other user in the same cluster, i.e., $\bar{\mathbf{H}}_{l,2} = \hat{\mathbf{h}}_{l,1}$. Note that $\bar{\mathbf{H}}_{l,2} \in \mathbb{C}^{1 \times N_{\text{RF}}}$ and $N_{\text{RF}} \geq 2$. It is obviously $\text{rank}(\bar{\mathbf{H}}_{l,2}) = 1$. Then, we define the SVD of $\bar{\mathbf{H}}_{l,2}$ as

$$\bar{\mathbf{H}}_{l,2} = \bar{\mathbf{U}}_{l,2} \bar{\Sigma}_{l,2} [\bar{\mathbf{V}}_{l,2}^{(1)} \quad \bar{\mathbf{V}}_{l,2}^{(0)}]^{\text{H}}, \quad (13)$$

where $\bar{\mathbf{V}}_{l,2}^{(0)}$ holds the last $(N_{\text{RF}} - 1)$ right singular vectors and forms an orthogonal basis for the null space of $\bar{\mathbf{H}}_{l,2}$. We perform SVD on $\hat{\mathbf{h}}_{l,2} \bar{\mathbf{V}}_{l,2}^{(0)}$ as

$$\hat{\mathbf{h}}_{l,2} \bar{\mathbf{V}}_{l,2}^{(0)} = \bar{\mathbf{U}}_{l,2} \bar{\boldsymbol{\Sigma}}_{l,2} \bar{\mathbf{V}}_{l,2}. \quad (14)$$

Finally, we can design the digital beamforming vector $\mathbf{w}_{l,2}$ as

$$\mathbf{w}_{l,2} = \frac{\bar{\mathbf{V}}_{l,2}^{(0)} \bar{\mathbf{V}}_{l,2}^{(1)}}{\|\mathbf{F}_{\text{RF}} \bar{\mathbf{V}}_{l,2}^{(0)} \bar{\mathbf{V}}_{l,2}^{(1)}\|}, \quad \text{for } l = 1, \dots, L, \quad (15)$$

where $\bar{\mathbf{V}}_{l,2}^{(1)}$ denotes the first column of $\bar{\mathbf{V}}_{l,2}$. For the strong user, we can design a digital beamforming vector to increase its capacity and in turn reduce its interuser interference. More exactly, we define the SVD of $\hat{\mathbf{h}}_{l,1}$ as

$$\hat{\mathbf{h}}_{l,1} = \mathbf{U}_{l,1} \boldsymbol{\Sigma}_{l,1} \mathbf{V}_{l,1}^H. \quad (16)$$

Then we can design the digital beamforming vector $\mathbf{w}_{l,1}$ for the strong user as

$$\mathbf{w}_{l,1} = \frac{\mathbf{V}_{l,1}^{(1)}}{\|\mathbf{F}_{\text{RF}} \mathbf{V}_{l,1}^{(1)}\|}, \quad \text{for } l = 1, \dots, L. \quad (17)$$

5. Power Allocation

In the proposed system, all the clusters contain two users (same cluster size), then the total transmit powers are equally allocated to the clusters, i.e., $P_1 = \dots = P_L = P_t/L$. This approach can be said as nearly optimal since each cluster contains users with nearly similar channel gain distinctness [6].

With the increasing of the power difference between two users in a same cluster, the SIC performance of the strong can be better, and the intracluster interference of the weak user can be also reduced because the power allocated to the signal of the strong user is decreased. So we propose a power allocation scheme for intracluster power allocation which maximizes the power difference to help perform SIC and decrease the intracluster interference for the weak user, while keeping the data rates of the two users are greater than or equal to their own minimum rate requirements. The formulation is as follows:

$$\alpha_l^* = \arg \max_{\alpha_l} P_{\text{dif}} = \arg \max_{\alpha_l} [(1 - \alpha_l) - \alpha_l] \quad (18)$$

$$\text{subject to } \log_2 \left(1 + \frac{\|\mathbf{h}_{l,1} \mathbf{F}_{\text{RF}} \mathbf{w}_{l,1}\|^2 P_l \alpha_l}{\sigma^2} \right) \geq \bar{R}_{l,1}, \quad (19)$$

$$\log_2 \left(1 + \frac{\|\mathbf{h}_{l,2} \mathbf{F}_{\text{RF}} \mathbf{w}_{l,2}\|^2 P_l (1 - \alpha_l)}{\|\mathbf{h}_{l,2} \mathbf{F}_{\text{RF}} \mathbf{w}_{l,1}\|^2 P_l \alpha_l + \sigma^2} \right) \geq \bar{R}_{l,2}, \quad (20)$$

where P_{dif} is the power coefficient difference between the two users. Here, $\bar{R}_{l,1}$ and $\bar{R}_{l,2}$ are defined as the strong and the weak users' rates in the l -cluster if the two users would

be supported by conventional TDMA hybrid beamforming system, respectively, i.e., $\bar{R}_{l,i} = 1/2 R_{l,i, \text{TDMA-BF}}$. In conventional hybrid beamforming system with L RF chains, $2L$ users need two time slots to be supported, while it only takes one time slot to support $2L$ users in MIMO-NOMA system. So $R_{l,i, \text{TDMA-BF}}$ is multiplied by $1/2$. $R_{l,i, \text{TDMA-BF}}$ is defined as

$$R_{l,i, \text{TDMA-BF}} = \log_2 \left(1 + \frac{\|\mathbf{h}_{l,i} \mathbf{B}_i \mathbf{f}_{l,i}^{\text{BB}}\|^2 P_l}{\sigma^2} \right). \quad (21)$$

where \mathbf{B}_i is the analog beamforming matrix containing all the feedback beams of the i -user in each cluster. $\mathbf{f}_{l,i}^{\text{BB}}$ is designed for the i -user in the l -th cluster by using ZF-BF based on the effective channels of the i -user in each cluster presented in [34].

For the strong user, the power allocation coefficient α_l that ensures its rate is found by solving (19). Solving (19) for α_l gives

$$\implies \alpha_l \geq \frac{\sqrt{\gamma_{l,1} + 1} - 1}{\varepsilon_{l,1}} = \alpha_{\text{inf}}, \quad (22)$$

where $\gamma_{l,i} = \|\mathbf{h}_{l,i} \mathbf{B}_i \mathbf{f}_{l,i}^{\text{BB}}\|^2 P_l / \sigma^2$ and $\varepsilon_{l,1} = \|\mathbf{h}_{l,1} \mathbf{F}_{\text{RF}} \mathbf{w}_{l,1}\|^2 P_l / \sigma^2$. The least value of power allocation coefficient α_l to ensure that NOMA is fair to the strong user is given by the right side of (22), and any α_l satisfying (22) will lead to (19).

Similarly, if the rate of the weak user is to be at least as good as conventional hybrid BF, then (20) leads to

$$\alpha_l \leq \frac{1 + \varepsilon_{l,2} - \sqrt{1 + \gamma_{l,2}}}{\varepsilon_{l,2} + \beta_{l,2} (\sqrt{1 + \gamma_{l,2}} - 1)} = \alpha_{\text{sub}}, \quad (23)$$

where $\beta_{l,2} = \|\mathbf{h}_{l,2} \mathbf{F}_{\text{RF}} \mathbf{w}_{l,1}\|^2 P_l / \sigma^2$ and $\varepsilon_{l,2} = \|\mathbf{h}_{l,2} \mathbf{F}_{\text{RF}} \mathbf{w}_{l,2}\|^2 P_l / \sigma^2$. Therefore, the greatest value of the power allocation coefficient α_l to ensure that NOMA is fair to the weak user is given by the right side of (23), and any α_l satisfying (23) will lead to (20).

If $\alpha_{\text{inf}} \leq \alpha_{\text{sub}}$ is satisfied, the NOMA power allocation region is therefore defined as $\Lambda_{\text{FN}} = [\alpha_{\text{inf}}, \alpha_{\text{sub}}]$. Since the objective function is monotonically decreasing function of α_l , the optimal power allocation coefficient for the strong user in the l -cluster, α_l^* , can be given as

$$\alpha_l^* = \arg \max_{\alpha_l \in \Lambda_{\text{FN}}} (1 - 2\alpha_l) = \alpha_{\text{inf}}. \quad (24)$$

In the case where $\alpha_{\text{inf}} > \alpha_{\text{sub}}$, (19) and (20) cannot be satisfied simultaneously. We will set $\alpha_l^* = \alpha_{\text{sub}}$ to allocate more power to the strong user to increase the system sum rate while keeping the rate of the weak user is equal to its minimum rate requirement, because the strong user has a larger channel gain and receives no intracluster interference by SIC.

6. Simulation Results

This section will present the computer simulation results to evaluate the performance of the proposed MIMO-NOMA scheme. We consider a multiuser downlink mmWave MIMO

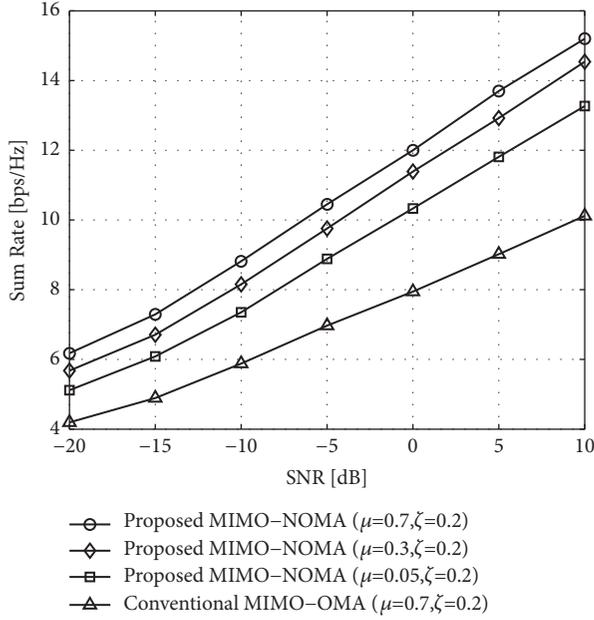


FIGURE 1: Sum rate performances of the proposed MIMO-NOMA and conventional hybrid ZF-BF systems against SNR ($\mu=0.7, 0.3$, and 0.05 ; $\zeta=0.2$).

system with a BS employing a ULA of $N_{TX} = 100$ transmitting antennas and $N_{RF} = 3$ RF chains, i.e., 3 clusters in the MIMO-NOMA system. The mmWave bandwidth of the system is assumed to be 2 GHz, and its carrier frequency is assumed to be 60 GHz. The number of users $K = 50$. And the all K users are randomly distributed in a cell with a radius of 20m. We assume that there is a LoS path and 2 NLoS paths ($P = 2$) for all users' channels. We set the channel parameters of each user as follows: (1) $\eta_{LOS} = 2$, $\eta_{NLOS} = 3$; (2) the normalized direction $\theta_{i,i}^f$ for $f = 0, 1, 2$ follows the uniform distribution over $[-1, 1]$. The beamforming codebook and RVQ codebook are quantized with $B_{RF} = 5$ bits and $B_{EC} = 13$ bits, respectively.

Figure 1 shows the sum rate performances of the proposed MIMO-NOMA system and the conventional TDMA hybrid beamforming system mentioned in the Section 5. The key observation from Figure 1 is that the proposed MIMO-NOMA system outperforms conventional MIMO-OMA system when the beam correlation and intercluster correlation thresholds are set as $\mu = 0.7$ and $\zeta = 0.2$, respectively. Moreover, by comparing the three cases, i.e., $\mu = 0.7, 0.3$, and 0.05 , Figure 1 shows an improvement of sum rate of the MIMO-NOMA system as μ increases. This is because, with the increase of the beam correlation threshold μ , more highly correlated users are clustered through the proposed clustering algorithm. The MIMO-NOMA system with a higher channel correlation between the users can help reduce the intercluster interference for the users by user clustering and analog beamforming.

Figure 2 shows the sum rate performances of the proposed MIMO-NOMA system, where $\mu = 0.3$. Observe all the curves; we can see that the sum rate of the proposed

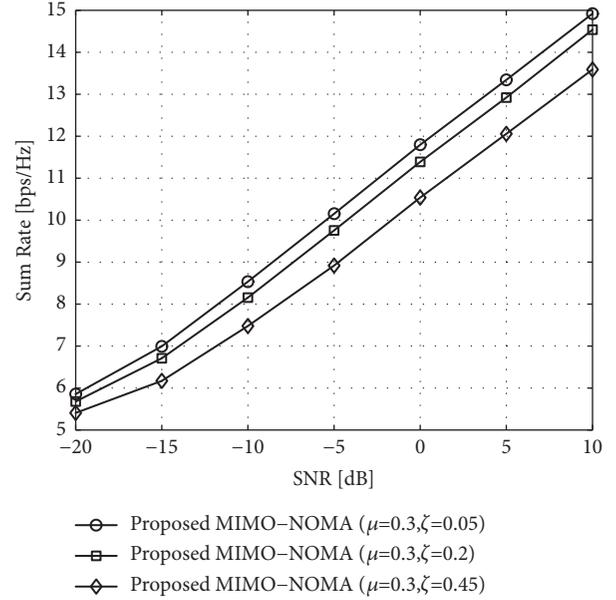


FIGURE 2: Sum rate performances of the proposed MIMO-NOMA system against SNR ($\mu=0.3$; $\zeta=0.05, 0.2$, and 0.45).

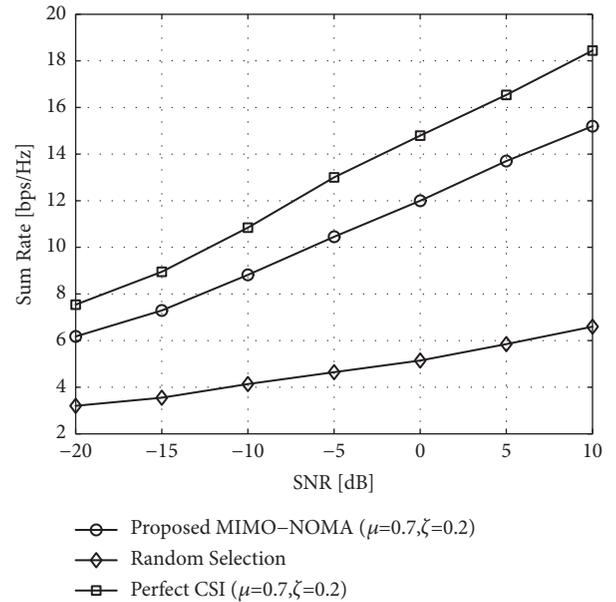


FIGURE 3: Comparison of sum rate performance between different schemes for MIMO-NOMA system against SNR ($\mu=0.7$; $\zeta=0.2$).

system increases as the intercluster correlation threshold ζ decreases. These results are expected because when ζ is reduced, clusters having lower intercluster correlations can be selected per the proposed clustering algorithm. A lower correlation between the selected clusters can help reduce the intercluster interference for the users.

Figure 3 shows the sum rate performances in three cases: randomly selecting, perfect CSI, and the proposed scheme. Randomly selecting means selecting two users from all the K users as a cluster randomly. Perfect CSI means that perfect

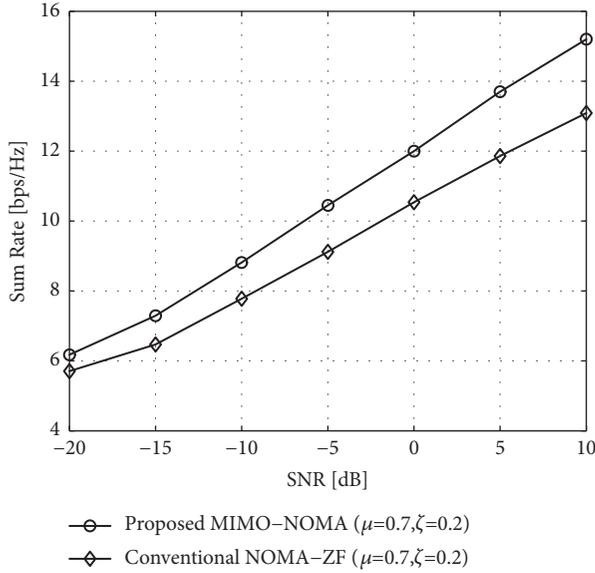


FIGURE 4: Sum rate performances of the proposed hybrid beamforming and conventional NOMA-ZF against SNR ($\mu=0.7; \zeta=0.2$).

CSI is assumed to be available at BS to generate beamforming matrices and cluster users with their channel correlation and channel gains instead of beam correlation and CQI values by Algorithm 1. Randomly selecting is the simplest of the three cases, but results in a degraded sum rate performance because this scheme randomly clusters users ignoring their intercluster interferences. On the other hand, the results show that the sum rate of the proposed scheme is lower than that of perfect CSI but requires less feedback overhead.

Next, we examine the advantage of the proposed hybrid beamforming for the MIMO-NOMA system with the conventional NOMA-ZF, which is performed by utilizing ZF-BF based on the effective channel information of the strong user of each cluster [12], cf. Figure 4. At very low SNR, the two beamforming schemes asymptotically approach the same sum rate. As SNR increases, the proposed beamforming scheme achieves a larger performance gain. This is because the proposed beamforming scheme reduces the interference of the weak users and the conventional NOMA-ZF does not.

Finally, we compare the proposed mmWave massive MIMO-NOMA scheme with limited feedback with the random beamforming scheme with one-bit feedback in [16]. The central angle of the sector is set as $2\Delta = 0.2$, and the power allocation coefficients of the two users in each cluster are set as $1/4$ and $3/4$. The number of the sectors in the random beamforming scheme is set as 3, which is the same as the number of RF chains in the proposed scheme. As can be observed from Figure 5, the performance of the proposed scheme degrades with decreasing the number of quantization bit B_{EC} . The figure also shows that the feedback overhead of the proposed scheme is larger than that of the random beamforming scheme with one-bit feedback, but it achieves better performance.

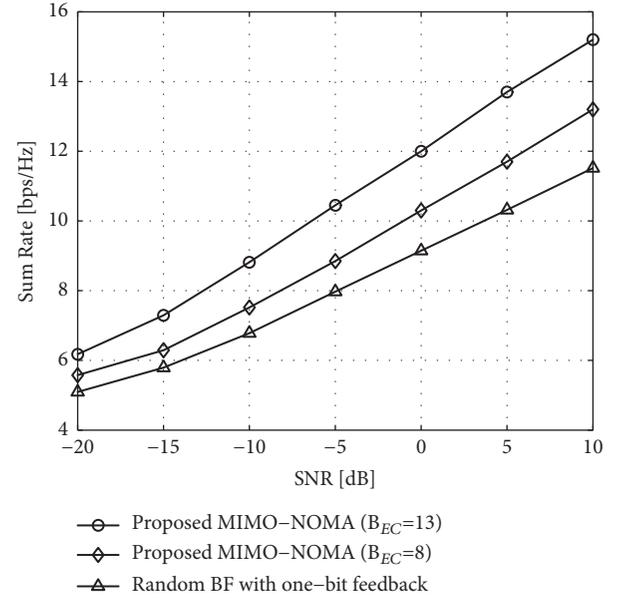


FIGURE 5: Sum rate performances of the proposed schemes with different quantization bits ($\mu=0.7; \zeta=0.2$) and the random beamforming scheme with one-bit feedback ($\Delta = 0.1$) against SNR.

7. Conclusions

In this paper, we propose a novel mmWave MIMO-NOMA system with limited feedback. A low-complexity user clustering algorithm and hybrid beamforming scheme have been designed with the limited feedback CSI. On this basis, we propose a power allocation solution to maximize the power difference between the two users in each cluster under the rate constraints of the two users. Simulation results show that the proposed MIMO-NOMA scheme achieves a better sum rate performance than conventional MIMO-NOMA, and the inter- and intracluster interferences and feedback overhead are reduced by applying the proposed clustering, beamforming, and effective power allocation algorithms. Therefore, the proposed scheme represents a feasible option to ensure that the mmWave MIMO-NOMA system provides an improved sum rate performance with limited feedback.

Data Availability

All data included in this study are available upon request by contact with the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This paper was supported by National Natural Science Foundation of China under Grant 61871321, National Science and Technology Major Project under Grant 2016ZX03001016, and

Innovation Team Project of Shaanxi Province under Grant 2017KCT-30-02.

References

- [1] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple Access Techniques for 5G Wireless Networks and Beyond*, Springer International Publishing, 2019.
- [2] Z. Ding, X. Lei, G. K. Karagiannidis, R. Schober, J. Yuan, and V. K. Bhargava, "A survey on non-orthogonal multiple access for 5g networks: research challenges and future trends," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2181–2195, 2017.
- [3] S. M. R. Islam, N. Avazov, O. A. Dobre, and K.-S. Kwak, "Power-domain non-orthogonal multiple access (NOMA) in 5G systems: potentials and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 721–742, 2017.
- [4] B. Kimy, S. Lim, H. Kim et al., "Non-orthogonal multiple access in a downlink multiuser beamforming system," in *Proceedings of the IEEE Military Communications Conference, MILCOM '13*, pp. 1278–1283, San Diego, Calif, USA, 2013.
- [5] X. Chen, A. Benjebbour, Y. Lan, A. Li, and H. Jiang, "Impact of rank optimization on downlink non-orthogonal multiple access (NOMA) with SU-MIMO," in *Proceedings of the IEEE International Conference on Communication Systems, IEEE ICCS '14*, pp. 233–237, Macau, China, 2014.
- [6] M. S. Ali, H. Tabassum, and E. Hossain, "dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems," *IEEE Access*, vol. 4, pp. 6325–6343, 2016.
- [7] M. S. Ali, E. Hossain, and D. I. Kim, "Non-orthogonal multiple access (NOMA) for downlink multiuser MIMO systems: user clustering, beamforming, and power allocation," *IEEE Access*, vol. 5, pp. 565–577, 2017.
- [8] M. Zeng, A. Yadav, O. A. Dobre, G. I. Tsiropoulos, and H. V. Poor, "Capacity comparison between MIMO-NOMA and MIMO-OMA with multiple users in a cluster," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2413–2424, 2017.
- [9] T. Rappaport, S. Sun, R. Mayzus et al., "Millimeter wave mobile communications for 5G cellular: it will work!," *IEEE Access*, vol. 1, pp. 335–349, 2013.
- [10] T. S. Rappaport, R. W. Heath Jr., R. C. Daniels, and J. N. Murdock, *Millimeter Wave Wireless Communications*, Pearson Education, 2014.
- [11] B. Wang, L. Dai, Z. Wang, N. Ge, and S. Zhou, "Spectrum and energy-efficient beamspace MIMO-NOMA for millimeter-wave communications using lens antenna array," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2370–2382, 2017.
- [12] W. Hao, M. Zeng, Z. Chu, and S. Yang, "Energy-efficient power allocation in millimeter wave massive MIMO with non-orthogonal multiple access," *IEEE Wireless Communications Letters*, vol. 6, no. 6, pp. 782–785, 2017.
- [13] Z. Xiao, L. Zhu, J. Choi, P. Xia, and X.-G. Xia, "Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter-wave communications," *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 2961–2974, 2018.
- [14] M. A. Almasi, M. Vaezi, and H. Mehrpouyan, "Impact of beam misalignment on hybrid beamforming NOMA for mmWave communications," *IEEE Transactions on Communications*, 2018, <https://arxiv.org/pdf/1808.00591.pdf>.
- [15] M. Dai and B. Clerckx, "Multiuser millimeter wave beamforming strategies with quantized and statistical CSIT," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7025–7038, 2017.
- [16] Z. Ding, P. Fan, and H. V. Poor, "Random beamforming in millimeter-wave noma networks," *IEEE Access*, vol. 5, pp. 7667–7681, 2017.
- [17] Z. Ding and H. V. Poor, "Design of massive-MIMO-NOMA with limited feedback," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 629–633, 2016.
- [18] Z. Yang, Z. Ding, P. Fan, and G. K. Karagiannidis, "On the performance of non-orthogonal multiple access systems with partial channel information," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 654–667, 2016.
- [19] P. Xu and K. Cumanan, "Optimal power allocation scheme for non-orthogonal multiple access with α -fairness," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 10, pp. 2357–2369, 2017.
- [20] P. Xu, Y. Yuan, Z. Ding, X. Dai, and R. Schober, "On the outage performance of non-orthogonal multiple access with 1-bit feedback," *IEEE Transactions on Wireless Communications*, vol. 15, no. 10, pp. 6716–6730, 2016.
- [21] Q. Yang, H.-M. Wang, D. W. K. Ng, and M. H. Lee, "NOMA in downlink SDMA with limited feedback: Performance analysis and optimization," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 10, pp. 6716–6730, 2016.
- [22] "Study on Downlink Multiuser Superposition Transmission for LTE, 3rd Generation Partnership Project, 3GPP, Document RP250456, 2015".
- [23] M. N. Kulkarni, A. Ghosh, and J. G. Andrews, "A comparison of MIMO techniques in downlink millimeter wave cellular networks with hybrid beamforming," *IEEE Transactions on Communications*, vol. 64, no. 5, pp. 1952–1967, 2016.
- [24] G. Lee, Y. Sung, and J. Seo, "Randomly-directional beamforming in millimeter-wave multiuser miso downlink," *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1086–1100, 2016.
- [25] X. Gao, L. Dai, S. Han, I. Chih-Lin, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmwave MIMO systems with large antenna arrays," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 998–1009, 2016.
- [26] X. Gao, L. Dai, Z. Chen, Z. Wang, and Z. Zhang, "Near-optimal beam selection for beamspace mmwave massive mimo systems," *IEEE Communications Letters*, vol. 20, no. 5, pp. 1054–1057, 2016.
- [27] J. Choi, "Beam selection in mm-wave multiuser MIMO systems using compressive sensing," *IEEE Transactions on Communications*, vol. 63, no. 8, pp. 2936–2947, 2015.
- [28] T. S. Rappaport, E. Ben-Dor, J. N. Murdock, and Y. Qiao, "38 GHz and 60 GHz angle-dependent propagation for cellular & peer-to-peer wireless communications," in *Proceedings of the IEEE International Conference on Communications (ICC '12)*, pp. 4568–4573, Ottawa, ON, Canada, 2012.
- [29] M. Cordina and C. J. Debono, "A support vector machine based sub-band CQI feedback compression scheme for 3GPP LTE systems," in *Proceedings of the International Symposium on Wireless Communication Systems, ISWCS '17*, pp. 325–330, Bologna, Italy, 2017.

- [30] A. Alorainy and M. J. Hossain, "Cross-layer performance of downlink dynamic cell selection with random packet scheduling and partial CQI feedback in wireless networks with cell sleeping," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5353–5369, 2017.
- [31] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, 2014.
- [32] A. Kammoun, A. Müller, E. Björnson, and M. Debbah, "Linear precoding based on polynomial expansion: large-scale multi-cell mimo systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 861–875, 2014.
- [33] L. D. Nguyen, T. Q. Duong, H. Q. Ngo, and K. Tourki, "Energy efficiency in cell-free massive MIMO with zero-forcing precoding design," *IEEE Communications Letters*, vol. 21, no. 8, pp. 1871–1874, 2017.
- [34] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 11, pp. 6481–6494, 2015.

Research Article

Energy-Efficient Hybrid Precoding Scheme Based on Antenna Selection Technology in Massive Multiple-Input Multiple-Output Systems

Jian Jun Ding and Jing Jiang 

College of Communication and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an, Shanxi, China

Correspondence should be addressed to Jing Jiang; jiangjing@xupt.edu.cn

Received 31 August 2018; Accepted 4 March 2019; Published 2 May 2019

Guest Editor: Raed A. Abd-Alhameed

Copyright © 2019 Jian Jun Ding and Jing Jiang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Hybrid precoding is a promising technology for massive multiple-input multiple-output (MIMO) systems. It can reduce the number of radio frequency (RF) chains. However, the power consumption is still very high owing to the large-scale antenna array. In this paper, we propose an energy-efficient precoding scheme based on antenna selection technology. The precoding scheme greatly increases the energy efficiency (EE) of the system. In the first step, we derive an exact closed-form expression of EE. Meanwhile, we further study the relationship between the number of transmit antennas and EE on the basis of the exact closed-form expression of EE. We prove that there exists an optimal value. When the number of transmit antennas equals to the value, the EE of the system can reach the maximum by a proper hybrid precoding scheme. Then, we propose an antenna selection algorithm to select antennas from the transmit antennas. And the number of selected antennas equals to the optimal value. Subsequently, we design the analog precoder based on a codebook to maximize the equivalent channel gain. At last, we further improve the EE by baseband digital precoding. The precoding algorithm we proposed offers a compromise between spectral efficiency (SE) and EE in millimeter wave (mmWave) massive MIMO systems. Finally, simulation results validate our theoretical analysis and show that a substantial EE gain can be obtained over the precoding scheme we proposed without large performance loss.

1. Introduction

Millimeter wave band communication recently acquires more and more attention owing to its great advantages [1–4]. A beneficial feature of mmWave is that many antenna arrays can be packed into a small dimension owing to the small wavelength [5–11]. However, the problem is that the availability of multiantennas causes high interference between different users and high hardware complexity. Fortunately, precoding can eliminate interference between different users and reduce hardware complexity. Precoding is therefore more favored over mmWave MIMO communication systems. Simple linear precoding schemes, for example, zero forcing (ZF) and minimum mean square error precoding (MMSE), are virtually optimal. However, the digital processing in the MIMO system requires a dedicated

base-band and radio frequency (RF) chain for every antenna element. Owing to the large amount of antennas deployed in mmWave massive MIMO systems, the costs are very huge and it is impossible to popularize in practice [12]. Researchers have widely studied on the reduction of the hardware cost. A hybrid precoding scheme comprising both digital and analog processing is capable of reducing the number of RF chains greatly. Therefore, the hybrid precoding scheme is widely used in mmWave MIMO systems. In consideration of a single-user scenario, a fully connected architecture-based hybrid precoding scheme is proposed in [13, 14], where each RF chain is connected to all antennas by analog phase shifters (APSs) and RF adders. The number of RF chains is greatly reduced in the fully connected architecture. However, in the fully connected architecture, the number of APSs is equal to the product of the number

of RF chains and the number of antennas. The cost of hardware is still excessive owing to the large amount of antenna elements. Different from the fully connected architecture, a partially connected architecture-based hybrid precoding scheme is proposed in [15–19], where each RF chain only connected to an antenna subarray and the number of APSs equals to the number of antennas. Compared with the fully connected architecture, the number of phase converters and energy consumption is greatly reduced. In the partially connected precoding scheme, the mapping relationship between the antennas and RF chains is predetermined, but the channel condition is time-varying. Thus, the partially connected architecture cannot guarantee that the mapping relationship is optimal under different channel conditions. A dynamic subarray architecture-based hybrid precoding scheme is proposed in [20], where an adaptive antenna selection network is added between RF chains and antenna elements to enhance spectral efficiency (SE). All the previous work is aimed at improving the SE of the system. In the fully connected architecture-based and partially connected architecture-based architecture hybrid precoding schemes, all the transmitted antennas are activated. In this paper, we propose an energy-efficient precoding scheme with considerably reduced energy consumption and assume that not all transmitting antennas are activated. We jointly optimize the number of activated antennas, analog precoding matrix, and digital precoding matrix to maximize the EE of the system. Firstly, instead of designing the digital and analog precoder directly, we perform an antenna selection procedure before digital and analog precoding. Then, we further optimize the performance in terms of EE by digital and analog precoding.

The contributions existing in this paper are summarized as follows:

- (1) An exact closed-form expression of EE was derived in this paper, and we derived an optimum number of antennas for maximizing the EE according to the exact closed-form expression of EE
- (2) In previous precoding schemes, all the antennas are activated. In this paper, not all the transmit antennas are activated (an antenna is activated, which means the antenna is used to transmit message). Based on channel state information and the exact closed-form expression of EE, we select antennas from the transmitting antennas to activate and the remaining antennas will be temporarily closed
- (3) In the prior hybrid precoding scheme, we design the analog precoder and the digital precoder according to the channel state information \mathbf{H} which was obtained by channel estimation to improve the sum rate and reduce the interference between different users. The hybrid precoding scheme we proposed differs from the prior hybrid precoding scheme. The transmitter first performs an antenna selection process according to the channel state information and then optimizes system performance in terms of EE by analog and digital precoding. Through the scheme, the number

of antennas and power consumption is reduced, and the EE is improved

- (4) An energy-efficient hybrid precoding scheme for a single user in mmWave systems is developed in this paper. First, we calculate the optimal number of transmit antennas and then the antenna selection algorithm is used to select the subset of transmit antennas. Then, we use the analog precoding scheme to maximize the gain of the equivalent channel between BS and objective users. Furthermore, we use the digital precoding scheme to maximize the EE of the system

The rest of this paper is organized as follows: Section 2, Section 3, and Section 4 introduce the system model, channel model, and power consumption model of the system, respectively. In Section 5, we introduce the EE optimization problem and propose an antenna selection algorithm. In Section 6, we proposed a hybrid beamforming algorithm. Computer simulation results are shown in Section 7. Finally, conclusions are drawn in Section 8.

Notations: we use the following notation throughout the paper. \mathbf{A} denotes a matrix; \mathbf{a} is a vector; α denotes a scalar; \mathbf{A}^T denotes the transpose of \mathbf{A} . \mathbf{A}^H denotes the conjugate of \mathbf{A} . \mathbf{A}^{-1} denotes the inversion of \mathbf{A} . $\mathbf{A}(i, j)$ denotes the (i, j) -th element of \mathbf{A} . We express $\|\mathbf{A}\|_F^2$ as $\text{Tr}(\mathbf{A}\mathbf{A}^H)$. \mathbf{I} is an identity matrix. The acronym s.t. denotes “subject to,” and i.i.d. denotes “independent and identically distributed.” $\mathbb{E}(\mathbf{A})$ denotes the mathematical expectation of \mathbf{A} , and \mathcal{CN} denotes complex Gaussian distribution.

2. System Model

In this paper, we consider a downlink SU-MIMO mmWave system where a base station (BS) is equipped with N_t antennas and N_{RF}^t RF chains to serve a single user which has N_{RF}^r RF chains and N_r antennas. The number of RF chains satisfies $N_s \leq N_{\text{RF}}^t \leq N_t$ to guarantee multistream transmission.

First, N_s transmit data streams at the BS are processed by a digital precoder \mathbf{F}_{BB} in the baseband and then processed by an analog precoder \mathbf{F}_{RF} (RF precoder using analog circuitry). \mathbf{F}_{RF} of size $N_t \times N_{\text{RF}}^t$ denotes the transmitting analog beam former. \mathbf{F}_{BB} denotes the base band digital precoder satisfying $\|\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F^2 = \rho$, and ρ denotes the total transmit power. Notably, the RF precoder can realize only phase changes because of phase-only control. For hybrid precoding systems as shown in Figure 1, the received signal vector of the objective user can be expressed by

$$\mathbf{y} = \mathbf{H}\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\mathbf{s} + \mathbf{n}, \quad (1)$$

where $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ denotes the received signal vector. $\mathbf{H} = [h_1, h_2, \dots, h_{N_r}]^T$ denotes the $N_r \times N_t$ channel matrix with h_k denoting the channel vector between the BS and the k -th receiving antenna. \mathbf{s} denotes the transmitted signal vector, satisfying $\mathbb{E}(\mathbf{s}\mathbf{s}^H) = \mathbf{I}_{N_s}$. And N_s represents the number of data streams. \mathbf{n} is the vector of additive complex

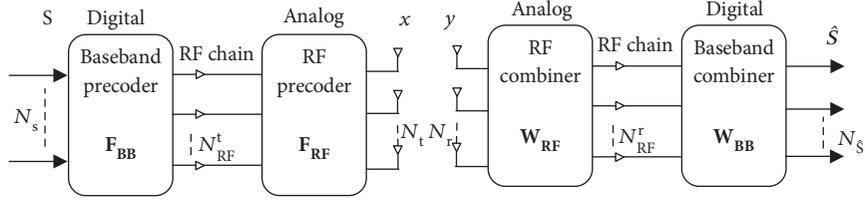


FIGURE 1: A mmWave single-user MIMO system.

Gaussian noise with zero mean and variance σ^2 . σ^2 denotes the power of noise.

3. Channel Model

Since mmWave channels have high free-space path loss, the mmWave propagation can be perfectly characterized by a clustered channel model [21] in this paper. The Saleh-Valenzuela model was used in this paper.

$$\mathbf{H} = \sqrt{\frac{N_r N_t}{N_{\text{cl}} N_{\text{ray}}}} \sum_{i=1}^{N_{\text{cl}}} \sum_{l=1}^{N_{\text{ray}}} a_{il} a_r(\phi_{il}^r, \theta_{il}^r) \alpha_t(\phi_{il}^t, \theta_{il}^t)^H, \quad (2)$$

$$\alpha(\varphi_{il}, \theta_{il}) = \frac{1}{\sqrt{M}} \left[1, \dots, e^{j(2\pi/\lambda)d(p \sin \varphi_{il} \theta_{il} + q \cos \theta_{il})}, \dots, e^{j(2\pi/\lambda)((\sqrt{M}-1)d \sin \varphi_{il} \sin \theta_{il} + ((\sqrt{M}-1) \cos \theta_{il})} \right]^{1/2}. \quad (3)$$

d denotes the distance between antennas, and λ denotes the wavelength of the signal. p satisfies $0 \leq p \leq \sqrt{M}$, the same with q . p and q are the antenna indices in the 2D plane.

4. Power Consumption Model

This section discusses the power consumption model of the downlink single-user MIMO mmWave system. The power consumption model of the system [22–24] can be denoted by

$$P = \rho P_t + P_c, \quad (4)$$

where ρ denotes the inefficiency factor of the power amplifier (PA). P_t denotes the power consumption of the transmitter, and $P_t = \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2$. P_c consists of the dynamic power consumption P_{dyn} and static power consumption P_{stac} . P_{dyn} can be denoted by

$$P_{\text{dyn}} = N_{\text{RF}}^t P_{\text{RF}} + N_{\text{RF}}^t N_t P_s, \quad (5)$$

where $N_{\text{RF}}^t P_{\text{RF}}$ denotes the power consumption of all RF chains. The number of APSs is $N_{\text{RF}}^t N_t$; thus, the power consumption of phase shifters is $N_{\text{RF}}^t N_t P_s$, where P_s

where the number of cluster and the number of rays can be expressed by N_{cl} and N_{ray} , and the gain of the i -th ray in the j -th cluster can be expressed as α_{ij} . We assumed that α_{ij} satisfies the complex Gaussian distribution $\mathcal{CN}(0, \sigma_{\alpha,i}^2)$ and $\sigma_{\alpha,i}^2$ satisfies $\sum_{j=1}^{N_{\text{cl}}} \sigma_{\alpha,i}^2 = \hat{\beta}$, and $\hat{\beta}$ is a normalization factor, where $\alpha(\phi_{il}^r, \theta_{il}^r)$ and $\alpha(\phi_{il}^t, \theta_{il}^t)$ denotes the receive and transmit array response vectors, respectively. $\phi_{il}^r(\theta_{il}^r)$ and $\theta_{il}^r(\phi_{il}^r)$ denote the azimuth and elevation angles of arrival and departure, respectively.

In this paper, we adopt a uniform square planar array with $\sqrt{M} \times \sqrt{M}$ antenna elements.

denotes the energy consumption of APS and P_{RF} denotes the power consumption of the RF chain.

5. Problem Formulation and Antenna Selection Algorithm

5.1. Problem Formulation. First, we consider the relationship between system capacity and signal-to-noise ratio. According to Shannon's theorem, the relationship between system capacity and SNR can be expressed by the following formula:

$$C = \log_2 \det(\mathbf{I} + \text{SNR}) \\ = \log_2 \det \left(\mathbf{I} + \frac{(\mathbf{H}_{\text{RF}} \mathbf{F}_{\text{BB}})^* (\mathbf{H}_{\text{RF}} \mathbf{F}_{\text{BB}})}{\sigma^2} \right). \quad (6)$$

Once the antenna set is determined, the corresponding channel matrix is determined. We can define the channel matrix as $\mathbf{H}_{\text{sel}}^t$. Equation (6) can be denoted as

$$C(S_{\text{sel}}^t) = \log_2 \det \left(\mathbf{I} + \frac{(\mathbf{H}_{\text{sel}}^t \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}})^* (\mathbf{H}_{\text{sel}}^t \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}})}{\sigma^2} \right). \quad (7)$$

The objective of this paper is to select the transmit antenna subset according to channel state information and design the hybrid precoders to maximize the EE of the system. The problem can be formulated as

$$\begin{aligned} \{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}, S_{\text{sel}}^t\} = \arg \max_{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}, S_{\text{sel}}^t} & \left(\frac{\log \left| \mathbf{I} + (1/\sigma_n^2) \|\mathbf{H}(S_{\text{sel}}^t) \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2 \right|}{P_c + \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2} \right) \\ \text{s.t. } & \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2 \leq P_{\text{tot}}, \end{aligned} \quad (8)$$

where S_{sel}^t denotes the selected antenna subset. The elements of S_{sel}^t are the index of selected antennas (e.g., $S_{\text{sel}}^t = \{1, 3, 4\}$ denotes that the selected antennas are the first, second, and fourth antennas of the BS). It is worth pointing out that the constraint in equation (8) is nonconvex, because of the coupling between the analog precoding matrix and digital precoding matrix. This makes equation (8) to be a difficult problem to solve. Fortunately, we can decouple the optimization problem into two convex optimization problems. First, we focus on the antenna selection process. We select the optimal subset of the antennas from the transmit antennas according to channel matrix \mathbf{H} . Then, with the channel matrix between the selected antenna and objective user, the hybrid precoding scheme can be adopted to improve the EE of the system. Specially, we adopt the analog precoding to maximize the gain of the equivalent channel and digital precoding to maximize the EE.

At first, we would briefly introduce the EE of the system without consideration of hybrid precoding. The EE is defined as the throughput divided by the total energy consumed, which can be denoted by

$$\text{EE} = \frac{\mathbb{E}(\mathbf{I})}{P_{\text{tot}}} = \frac{\mathbb{E}(\mathbf{I})}{P_c + P_t}, \quad (9)$$

where $\mathbb{E}(\mathbf{A})$ denotes the mathematical expectation of \mathbf{A} , and \mathbf{I} denotes the mutual information. If the power consumption of the receiver is taken into account, the circuit power consumption model can be expressed:

$$\begin{aligned} P_c \approx & N_t(P_{\text{dac}} + P_{\text{mixer}} + P_{\text{filt}}) + 2P_{\text{syn}} \\ & + N_r(P_{\text{ln } a} + P_{\text{mixer}} + P_{\text{ifa}} + P_{\text{adc}}), \end{aligned} \quad (10)$$

where N_t and N_r denote the number of transmitters and receivers, respectively, and P_{dac} , P_{mixer} , P_{filt} , P_{syn} , P_{lna} , P_{ifa} , and P_{adc} are the power consumption of the digital analog

converter, the mixer, the filter at the transmit side, the synthesizer, the low-noise amplifier, the intermediate-frequency amplifier, the filters at the receiver side, and the analog digital converter, respectively. We set

$$P_1 = 2P_{\text{syn}} + N_r(P_{\text{lna}} + P_{\text{mixer}} + P_{\text{ifa}} + P_{\text{adc}}) + P_t, \quad (11)$$

$$P_2 = (P_{\text{ln } a} + P_{\text{mixer}} + P_{\text{ifa}} + P_{\text{adc}}), \quad (12)$$

$$\text{EE} = \frac{\mathbb{E}(\mathbf{I})}{P_1 + N_t P_2}, \quad (13)$$

where P_2 is a very small value much less than one.

Lemma 1. Suppose there exists a communication system. N_t transmit antennas were equipped by the BS, and N_{sel} antennas are selected from these transmitter antennas; the distribution of mutual information is approximately [25] given by

$$\begin{aligned} I_{\text{sel}} \sim \mathcal{FN} \left(\log_2 \left(1 + \left(1 + \ln \frac{N_t}{N_{\text{sel}}} \right) \rho N_{\text{sel}} \right), \right. \\ \left. \frac{(\log_2 e)^2 \rho^2 N_{\text{sel}} (2 - (N_{\text{sel}}/N_t))}{(1 + (1 + \ln (N_t/N_{\text{sel}})) \rho N_{\text{sel}})} \right). \end{aligned} \quad (14)$$

Thus, $\mathbb{E}(I_{\text{sel}}) = \log_2(1 + (1 + \ln(N_t/N_{\text{sel}}))\rho N)$. So, the closed-form expression of the EE becomes

$$\text{EE} = \frac{\mathbb{E}(I_{\text{sel}})}{P_{\text{tot}}} \approx \frac{\log_2[1 + (1 + \ln(N_t/N_{\text{sel}}))\rho N_{\text{sel}}]}{\rho + P_1 + N_{\text{sel}} P_2}. \quad (15)$$

Next, we will analyze the effects of the number of transmitted antennas on the EE.

Lemma 2. If the power consumption of the mmWave MIMO system can be modeled as the addition of the transmit power P_t and the circuit power consumption P_c , EE increases at first and then decreases as the number of selected transmit antennas N_{sel}^t increases.

Proof 1. EE can be expressed as a function of N_{sel} :

$$\eta(N_{\text{sel}}) = \frac{\log_2[1 + (1 + \ln(N_t/N_{\text{sel}}))\rho N_{\text{sel}}]}{\rho + P_1 + N_{\text{sel}} P_2}. \quad (16)$$

We can take the first derivative of $\eta(N_{\text{sel}})$

$$\begin{aligned} \frac{\partial(\eta(N_{\text{sel}}))}{\partial N_{\text{sel}}} &= \frac{(\log_2(1 + (1 + \ln(N_t/N_{\text{sel}}))\rho N_{\text{sel}}))'(\rho + P_1 + N_{\text{sel}} P_2)}{(\rho + P_1 + N_{\text{sel}} P_2)^2} + \frac{(\log_2(1 + (1 + \ln(N_t/N_{\text{sel}}))\rho N_{\text{sel}}))' P_2}{(\rho + P_1 + N_{\text{sel}} P_2)^2} \\ &= \frac{(\rho + P_1 + N_{\text{sel}} P_2) \rho \ln(N_t/N_{\text{sel}})}{q(N_{\text{sel}})} - \frac{P_2 [1 + (1 + \ln(N_t/N_{\text{sel}}))\rho N_{\text{sel}}] \ln[1 + (1 + \ln(N_t/N_{\text{sel}}))\rho N_{\text{sel}}]}{q(N_{\text{sel}})}, \end{aligned} \quad (17)$$

setting $q(N_{\text{sel}}) = (\rho + P_1 + N_{\text{sel}}P_2)^2$. Because the denominator of (17) is positive, the denominator of (17) has no effect on the monotonicity of function (17). Setting the numerator of (17) as $p(N_{\text{sel}})$,

$$p(N_{\text{sel}}) = (\rho + P_1 + N_{\text{sel}}P_2)\rho \ln \frac{N_t}{N_{\text{sel}}} - P_2 \left[1 + \left(1 + \ln \frac{N_t}{N_{\text{sel}}} \right) \rho N_{\text{sel}} \right] * \ln \left(1 + \left(1 + \ln \frac{N_t}{N_{\text{sel}}} \right) \rho N_{\text{sel}} \right). \quad (18)$$

Then, we take the first derivative of $p(N_{\text{sel}})$ again:

$$p'(N_{\text{sel}}) = -\frac{(\rho + P_1 + N_{\text{sel}}P_2)}{N_{\text{sel}}} - P_2\rho \ln \left[1 + \left(\ln \frac{N_t}{N_{\text{sel}}} + 1 \right) \rho N_{\text{sel}} \right]. \quad (19)$$

It is obvious that $p'(N_{\text{sel}})$ is a negative value. So, $p(N_{\text{sel}})$ is a decreasing function of (N_{sel}) ; thus, $p(N_t) \leq p(N_{\text{sel}}) \leq p(1)$:

$$p(N_t) = -P_2(1 + N_t\rho) < 0, \quad (20)$$

and

$$p(1) = (\rho + P_1 + P_2)\rho \ln N_t - P_2 \left[1 + (1 + \ln(N_t))\rho \right] \cdot \ln \left[1 + (1 + \ln(N_t))\rho \right]. \quad (21)$$

We do not know whether $p(1)$ is a positive or negative value. However, equation (21) can also be viewed as a function of ρ . We set $P(1) = f(\rho)$. Then, we take the first derivative of ρ .

$$f'(\rho) = (2\rho + P_1 + P_2) \ln(N_t) - P_2 \left[(1 + \ln(N_t)) \right] \cdot \ln \left[(1 + \ln(N_t))\rho \right] - \left(P_2(1 + \ln(N_t))\rho \right) \cdot \left(\frac{1 + \ln(N_t)}{1 + \ln(N_t)} \right). \quad (22)$$

We take the first derivative of (22) again in respective of ρ , we can conclude that

$$f''(\rho) \geq 2 \ln(N_t) - P_2(1 + \ln(N_t)). \quad (23)$$

As mentioned earlier in this paper, P_2 is a very small number much less than one, so $f''(\rho) > 0$. Then, $f'(\rho) > f'(0) > 0$, so $f(\rho)$ is an increasing function of ρ . Besides,

$$\lim_{\rho \rightarrow 0} f(\rho) = 0, \quad (24)$$

so $P(1) = f(\rho) > 0$. So, η increases first and then decreases with the increase in N_{sel} . Assume that N_0 satisfies $p(N_0) = 0$. With mathematical knowledge, we can conclude that

- (1) if $N_{\text{sel}} \in [1, N_0]$, the **EE** of the system increases with the increase in N_{sel}
- (2) if $N_{\text{sel}} \in [N_0, +\infty)$, the **EE** of the system decreases with the increase in N_{sel}

That is to say, there exists an optimal number of antenna $N_{\text{sel}}^{\text{opt}}$ that can maximize the **EE** of the system.

5.1.1. Discussion. Based on the system model and power consumption model that we proposed, **EE** increases at first and then decreases with the increase in the number of antennas. This is because with the increase in the number of selected antennas, the power consumption increases linearly. However, this is not true for the sum rate of the system. When the number of antennas increases to a certain value, the increase in the sum rate is not enough to compensate for the increase in the energy consumption of the system, so the **EE** of the system will reduce.

Denote the selected transmit antenna subset as S_{sel}^t , the number of active transmits L_t is $|S_{\text{sel}}^t|$ where $|A|$ denotes the cardinality of A . **EE** could also be also expressed as

$$\text{EE} = \frac{\log_2 \det \left(I + \left(\| \mathbf{H}(S_{\text{sel}}^t) \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \|^2 / \sigma_n^2 \right) \right)}{P_c + \rho P_t} \text{ (bit/Hz/joule)}, \quad (25)$$

where $\mathbf{H}(S_{\text{sel}}^t)$ is the channel matrix of size $N_r \times N_{\text{sel}}$ between the L_t transmit antenna and the N_r receive antennas. The objective of antenna selection can be expressed as

$$\{ S_{\text{sel}}^t, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}} \} = \arg \max_{S_{\text{sel}}^t, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}} \text{EE}(S_{\text{sel}}^t, \mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}) \quad \text{s.t.} \quad \text{SNR} > \gamma, \quad (26)$$

where (26) is a nonconvex optimization problem [26–29]. It is often intractable to directly solve, because it is a joint optimization problem over three variables. We propose a suboptimal algorithm to decouple the nonconvex optimization problem into a convex optimization problem that can be solved directly. First, just as it is proved above, there exists an optimal value N_{opt} . When the number of transmitter antennas equals to the value, the **EE** of the system can achieve an optimal value. Thus, the proposed algorithm will be divided into three parts.

At first, as we know, there exists an optimal value N_{opt} that maximizes the **EE**. Thus, we can select N_{opt} antennas from N_t transmit antennas to optimize the **EE** of the system before hybrid precoding.

Then, an analog precoder is designed to maximize the gains of the equivalent channel, where the equivalent channel \mathbf{H}_{eff} is defined as $\mathbf{H}_{\text{sel}} \mathbf{F}_{\text{RF}}$. At this step, it is assumed that the digital precoding matrix \mathbf{F}_{bb} is fixed.

The design of digital precoder will become an easily solved optimization problem with a single variable after the analog precoding. The digital precoding matrix will be designed based on the maximized EE criterion.

The proposed antenna select algorithm and hybrid precoding scheme will be discussed in the next section.

5.2. Antenna Selection Algorithm. In this section, a low-complexity antenna select algorithm is proposed which enables the BS to reduce the number of transmit antennas and without large performance loss. The antenna selection algorithm is described as below.

As discussed above, EE increases first and then decreases with the increase in transmit antenna in mmWave massive MIMO systems. So, we can select N_{opt} antennas from N_t transmit antennas to optimize the EE of the system before hybrid precoding at first.

The pseudocode of the proposed antenna selection algorithm is summarized in Algorithm 1, which can be explained as follows. At the beginning, we formulate the EE of the system as the sum rate R divided by energy consumption P_c and set the EE denoted in (13). During the antenna selection, in step 1, step 2, and step 3, we calculate the EE of the mmWave MIMO system with different numbers of antennas. In step 4, we find the optimal number of antennas N_{opt} by comparing EE under different conditions. Then, in step 6, step 7, and step 8, we choose N_{opt} optimal antennas from N_t transmitted antennas by comparing the Frobenius norm of the channel vector of each antenna. Finally, the N_{opt} antennas are used to transmit information. Note that the antenna selection algorithm enables the system to adaptively choose the transmit antenna under different channel conditions to maximize the EE of the system.

6. The Design of the Hybrid Precoder

In this section, we will discuss the design of the hybrid precoder for the mmWave massive MIMO system under the condition that the transmit antenna subset S_{sel} has been obtained.

As the antenna selection process effectively reduces the number of transmit antennas, the power consumption of the system is reduced. After the antenna subset has been obtained, analog precoding and base band digital precoding will be utilized to improve the EE of the system. The optimization problem (26) can be written as

$$\{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}\} = \arg \max_{\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}} \left(\frac{R(\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}})}{P(\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}})} \right) \quad \text{s.t.} \quad \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2 \leq P_{\text{tot}}, \quad (27)$$

where $R(\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}})$ denotes the sum rate of the proposed system and $P(\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}})$ denotes the energy consumption of the whole system.

$$\begin{aligned} R(\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}) &= \log \left| \mathbf{I}_{N_r} + \mathbf{H}_{\text{sel}}^t \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{F}_{\text{BB}}^H \mathbf{F}_{\text{RF}}^H (\mathbf{H}_{\text{sel}}^t)^H \right|, \\ P(\mathbf{F}_{\text{RF}}, \mathbf{F}_{\text{BB}}) &= P_c + \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2. \end{aligned} \quad (28)$$

At first, an analog precoding algorithm is proposed to maximize the power of the received signal for the objective user [30–32]. Further, a digital precoder will be designed to maximize the EE of the system. However, it is difficult to solve problem (27) because of the coupling between two matrix variables: \mathbf{F}_{RF} and \mathbf{F}_{BB} . We solve this problem by decoupling the joint optimization problem into two convex optimization problems.

Firstly, an analog precoder is designed that is aimed at maximizing the equivalent channel gain and assuming that the base-band precoding matrix \mathbf{F}_{BB} is fixed at this stage.

After the analog precoding matrix is obtained, the original optimization problem (27) becomes an easily solved convex optimization problem with one matrix variable \mathbf{F}_{BB} , which can be expressed as

$$\{\mathbf{F}_{\text{BB}}\} = \arg \max_{\mathbf{F}_{\text{BB}}} \text{EE}(\mathbf{F}_{\text{BB}}) \quad \text{s.t.} \quad \|\mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}\|_F^2 \leq P_{\text{tot}}. \quad (29)$$

Then, the optimal baseband precoding matrix \mathbf{F}_{BB} can be obtained by solving (29).

6.1. The Design of the Analog Precoder. In this section, we will discuss the optimal analog precoding matrix \mathbf{F}_{RF} , where $\mathbf{F}_{\text{RF}} = (f_1, f_2, \dots, f_{N_{\text{rf}}})$; the equivalent channel \mathbf{H}_{eff} is defined as $\mathbf{H} \mathbf{F}_{\text{RF}}$. We rewrite the receive signal of the objective user

$$\mathbf{y} = \mathbf{H}_{\text{eff}} \mathbf{F}_{\text{BB}} \mathbf{s} + \mathbf{n}. \quad (30)$$

The gain of the equivalent channel can be expressed as

$$G(\mathbf{H}_{\text{eff}}) = \|\mathbf{H} \mathbf{F}_{\text{RF}}\|_F^2. \quad (31)$$

We aim to find an analog precoding matrix \mathbf{F}_{RF} to maximize the equivalent channel gain based on codebook \mathcal{F} . \mathcal{F} is a predefined RF beamforming codebook. \mathcal{F} can be specified as a quantized matrix $\alpha_T (2\pi i_\varphi / M_\varphi, 2\pi j_\theta / M_\theta)$. For i_φ , it takes the value $0, 1, 2, 3, \dots, M_\varphi - 1$, and for j_θ , it takes the value $0, 1, 2, 3, \dots, M_\theta - 1$. M_φ and M_θ denote the quantized precision of azimuth and elevation angles, respectively. The most optimal precoding matrix \mathbf{F}_{RF} is designed based on the following metric.

$$\mathbf{f}_i = \arg \max_{\mathbf{f}_i \in \mathcal{F}} \|\mathbf{H} \mathbf{f}_i\|_F^2, \quad i = 1, 2, 3, \dots, N_{\text{rf}}. \quad (32)$$

The pseudocode of the proposed analog precoding algorithm is summarized in Algorithm 2.

At the beginning, we set $i = 1$, and we search in the codebook \mathcal{F} to find out an optimal vector $\mathbf{f}_i \in \mathcal{F}$ which has the

Input: the number of transmit antenna N_t , the power consumption of digital analog, converter, filter, synthesizer, low-noise amplifier, the intermediate-frequency amplifier, and the analog digital amplifier: P_{dac} , P_{filt} , P_{syn} , P_{lna} , P_{ifa} , P_{dac} , the transmit power p , channel matrix \mathbf{H} .

Initialization: $S_{sel}^t = \phi$, $EE = \phi$

- 1: **for** $1 \leq l \leq N_t$ **do**
- 2: Calculate: $f(l) = \log_2[1 + (1 + \ln(N_t/l))\rho l]/\rho + P_1 + lP_2$
- 3: Update $EE(l) = f(l)$;
- 4: $l = l + 1$;
- 5: **end for**
- 6: Sort $\{EE(i)\}_{i=1}^{N_t}$ in a descend order as:
 $EE(i_1) \geq EE(i_2) \geq \dots \geq EE(i_{N_t})$
- 7: Setting $N_{opt} = i_1$
- 8: Setting $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3, \dots, \mathbf{h}_{N_t}]$
- 9: Sort $\{\|\mathbf{h}_i\|_F^2\}_{i=1}^{N_t}$ in a descend order as
 $\|\mathbf{h}_{i_1}\|_F^2 \geq \|\mathbf{h}_{i_2}\|_F^2 \geq \dots \geq \|\mathbf{h}_{i_{opt}}\|_F^2 \geq \dots \geq \|\mathbf{h}_{i_{N_t}}\|_F^2$
- 10: $S_{sel}^t = \{i_1, i_2, i_3, \dots, i_{i_{opt}}\}$
- 11: **Output:** S_{sel}^t

ALGORITHM 1: The proposed antenna selection scheme.

Input: Channel matrix \mathbf{H} , analog beamforming codebook \mathcal{F} , the number of RF chains N_{RF}^t

Initialization: $\mathbf{F}_{RF} = \phi$

- 1: **for** $1 \leq i \leq N_{RF}^t$ **do**
- 2: **while** $\{r_i = \arg \max_{r_i \in \{1, 2, \dots\}} \|\mathbf{H}\mathcal{F}_{r_i}\|_F^2\}$ **do**
- 3: User feed back the index of codebook vector r_i to BS. BS finds corresponding vector \mathcal{F}_{r_i} from codebook \mathcal{F} and setting \mathcal{F}_{r_i} as the i th column of \mathbf{F}_{RF} , that is $f_i = \mathcal{F}_{r_i}$
- 4: $\mathcal{F} = \mathcal{F} \setminus \mathcal{F}_{r_i}$
- 5: **end while**
- 6: $i = i + 1$
- 7: **end for**

ALGORITHM 2: Analog precoding scheme.

largest inner product with channel matrix \mathbf{H} and make it as the first column of the analog precoding matrix \mathbf{F}_{RF} .

Then, i is updated by $i + 1$, and the procedure above is repeated to find optimal vector f_{i+1} and f_{i+1} is set as the $(i + 1)$ -th column of analog precoding matrix \mathbf{F}_{RF} .

The procedure is repeated until $i = N_{RF}^t$. Finally, \mathbf{F}_{RF} could be obtained, i.e., $\mathbf{F}_{RF} = [f_1, f_2, f_3, \dots, f_{N_{RF}^t}]$.

6.2. Digital Precoding Design to Improve EE. In this section, we discuss the design of digital beamformer to further improve the EE of the system. After the analog precoding matrix is obtained, the original nonconvex problem is conveyed to a univariate optimization [33] problem which is given by

$$\mathbf{F}_{BB} = \arg \max_{\mathbf{F}_{BB}} (EE(\mathbf{F}_{BB})) \quad \text{s.t.} \quad \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 \leq P_{\text{tot}}, \quad (33)$$

$$EE(\mathbf{F}_{BB}) = \frac{\log_2 \left[\mathbf{I}_{N_t} + \mathbf{H}_{\text{sel}}^t \mathbf{F}_{RF} \mathbf{F}_{BB} \mathbf{F}_{BB}^* \mathbf{F}_{RF}^* (\mathbf{H}_{\text{sel}}^t)^* \right]}{P_c + \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2}. \quad (34)$$

There exist many approaches to solve this classical fractional programming problem [34]. At first, we transform it into an equivalent parametric programming problem, $E(\alpha) = 0$, where $E(\alpha)$ can be expressed as

$$E(\alpha) = \max_{\mathbf{F}_{BB}} (R(\mathbf{F}_{BB}) - \alpha P(\mathbf{F}_{BB})), \quad (35)$$

$D = \{\mathbf{F}_{BB} \mid \|\mathbf{F}_{RF}\mathbf{F}_{BB}\|_F^2 \leq P_{\text{tot}}\}$, where α satisfies $0 \leq \alpha \leq (R_{\text{max}}/P_c)$ and $R_{\text{max}} = \log(1 + (P_{\text{tot}}/\sigma_n^2) \|\mathbf{H}_{\text{sel}}^t\|_F^2)$ denotes the max sum rate of the system. The optimal digital precoding matrix $\mathbf{F}_{BB}^{\text{opt}}$ which is the solution of (35) is also the optimal solution of (33). The general method to solve (35) is the half-interval search method. It is summarized as in Algorithm 3.

In Algorithm 3, the key step is solving (35). In this section, we will discuss how to solve (35). Problem (35) can be expressed as

Initialization: $\alpha^{\min} = 0$, $\alpha^{\max} = S_{\max}/P_c$ and $\xi \geq 0$

- 1: **while** $|\alpha^{\max} - \alpha^{\min}| \leq \xi$ **do**
- 2: $\alpha = (\alpha^{\max} + \alpha^{\min})/2$
- 3: Solve the problem (35) for a given α ,
- 4: update $E(\alpha)$ and $\mathbf{F}_\alpha^{\text{opt}}$.
- 5: If $G(\alpha) \leq 0$ let $\alpha^{\max} = \alpha$. Otherwise, $\alpha^{\min} = \alpha$.
- 6: **end while**
- 7: $\mathbf{F}_{\text{BB}}^{\text{opt}} = \mathbf{F}_\alpha^{\text{opt}}$

ALGORITHM 3: Half-interval search method.

$$\max_{\mathbf{F}} L(\mathbf{F}), \quad \text{s.t.} \quad \|\mathbf{F}\|_F^2 \leq P_{\text{tot}},$$

$$\begin{aligned} L(\mathbf{F}) &= \log \det \left(\mathbf{I}_{N_{\text{rx}}} + \frac{1}{\sigma_n^2} \mathbf{H} \mathbf{F} \mathbf{F}^H \mathbf{H}^H \right) - \eta \xi \|\mathbf{F}\|_F^2 \\ &= \log \det \left(\mathbf{I}_{N_{\text{tx}}} + \frac{1}{\sigma_n^2} \mathbf{F} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \right) - \eta \xi \|\mathbf{F}\|_F^2. \end{aligned} \quad (36)$$

According to the knowledge of the determinant, it is obvious that $|\mathbf{I} + \mathbf{M}\mathbf{N}| = |\mathbf{I} + \mathbf{N}\mathbf{M}|$; (35) could be conveyed into

$$\begin{aligned} T(\mathbf{F}) &= \log \left(\mathbf{I}_{N_H} + \sigma_n^{-2} \sum_1^{1/2} \mathbf{U}_1 \mathbf{F} \mathbf{F}^H \mathbf{U}_1^H \sum_1^{1/2} \right) - \alpha \xi \|\mathbf{F}\|^2 \\ &= \log \left(\mathbf{I}_{N_H} + \sigma_n^{-2} \sum_1^{1/2} (\mathbf{U}_1 (\mathbf{F} \mathbf{F}^H) \mathbf{U}_1^H) \sum_1^{1/2} \right) - \alpha \xi \|\mathbf{F}\|^2 \\ &= \log \left(\mathbf{I}_{N_H} + \sigma_n^{-2} \sum_1^{1/2} \mathbf{U}_1 \mathbf{P} \mathbf{U}_1^H \sum_1^{1/2} \right) - \alpha \xi \|\mathbf{F}\|^2 \\ &= \log \left(\mathbf{I}_{N_H} + \sigma_n^{-2} \sum_1^{1/2} \tilde{\mathbf{P}} \sum_1^{1/2} \right) - \alpha \xi \|\mathbf{F}\|^2, \end{aligned} \quad (37)$$

where $\mathbf{P} = \mathbf{F} \mathbf{F}^H$ and $\tilde{\mathbf{P}} = \mathbf{U}_1 \mathbf{P} \mathbf{U}_1^H$; it is obvious that for any matrix \mathbf{B} which is nonnegative, $|\mathbf{B}| \leq \prod_i \mathbf{B}_{i,i}$. Thus, only if \mathbf{P} is a diagonal matrix will (37) achieve its maximum value. Then, (38) will further expressed as

$$\begin{aligned} \max & \left(\log \left[\prod_{j=1}^{N_H} (1 + p_j v_j \sigma_n^{-2}) \right] - \eta \xi \sum_{j=1}^{N_H} P_j \right) \\ &= \max \sum_{j=1}^{N_H} \log(1 + p_j v_j \sigma_n^{-2}) - \eta \xi P_j \quad \text{s.t.} \quad \sum_{j=1}^{N_H} P_j \leq P_{\text{tot}}. \end{aligned} \quad (38)$$

The problem in (38) was a convex optimization problem, and its general solution method is to convert it into an unconstrained optimization problem with the Lagrange multiplier method [35, 36]. Equation (38) could be converted into (39) by the Lagrange multiplier method:

$$\begin{aligned} P(\tilde{P}, \lambda) &= - \sum_{i=1}^{N_H} \left[\log \left(1 + \frac{p_i v_i}{\sigma_n^2} \right) - \eta \xi q_i \right] + \lambda \left(\sum_{i=1}^{N_H} p_i - P_{\text{tot}} \right) \\ &= - \sum_{i=1}^{N_H} \left[\log \left(1 + \frac{v_i}{\sigma_n^2} p_i \right) - (\eta + \lambda) p_i \right] - \lambda P_{\text{tot}}. \end{aligned} \quad (39)$$

By solving the KKT conditions of (39), the optimal diagonal entries of P^{opt} are obtained for a given λ via the water-filling algorithm [37, 38].

$$p_j^{\text{opt}}(\lambda) = \left[\frac{1}{\alpha \xi + \lambda} - \frac{\sigma_n^2}{v_j} \right]^+. \quad (40)$$

p_j^{opt} is a decreasing function of λ ; thus, λ^{opt} could be found by a one-dimensional search algorithm. We proposed a binary water-filling algorithm in this paper. The binary water-filling algorithm is summarized in Algorithm 4.

7. Simulation Results

In this section, we present simulation results to evaluate the performance of the proposed algorithm in terms of EE and the sum rate. Specially, the simulation parameters are summarized in Table 1. In this section, we first compare the performance of different algorithms in terms of EE with different antenna array sizes. Then, we compare the performance of the proposed algorithm with a conventional algorithm under different power constraints in terms of EE. At last, we compare the performance of different precoding schemes in terms of the sum rate.

Figure 2 compares the performance between the energy-efficient algorithm and spectral efficient algorithm under different SNRs in terms of SE. As Figure 2 shows, with the increase in SNR, the system performance of different algorithms increases, but the performance of the pure digital precoding system is optimal. The performance of the spectral efficient hybrid precoding scheme is close to that of the digital precoding scheme, and the performance of the energy-efficient algorithm is lower than that of the spectral efficient algorithm in terms of SE. This is because the energy-efficient algorithm sacrifices part of the spectrum efficiency to improve the energy efficiency of the system. And, for the same algorithm (energy-efficient algorithm/spectral efficient algorithm), using a different number of RFs will lead to a different performance in terms of spectral efficiency. As can be seen from the figure, under the same SNR condition, the higher the number of radio frequency chains, the better the system performance.

Figure 3 compares the EE performance between the energy-efficient algorithm and spectral efficient algorithm under different transmit power constraints. As Figure 3 shows, at low power constraint, the energy-efficient performance of the proposed algorithm is exactly close to that of the spectral efficient algorithm. As the transmit power increases, the EE of the proposed algorithm remains a constant but shows a decline in the spectral efficient algorithm.

Initialization: $\rho_{\min} = 0, \rho_{\max} = \max (v_i/\sigma_n^2 - \alpha\epsilon) i = 1 \cdots N_H$

- 1: **for** $|\rho_{\max} - \rho_{\min}| < \epsilon$ **do**
- 2: $\rho = (\rho_{\max} + \rho_{\min})/2$
- 3: calculate $P(\rho)$ based on (39), (40)
- 4: update ρ_{\max} and ρ_{\min}
- 5: if $P(\rho) < P_{tot}$, $\rho_{\max} = \rho$
- 6: otherwise
- 7: $\rho_{\min} = \rho$
- 8: **end for**
- 9: $\rho_{opt} = \rho$
- 10: Calculate P_{opt} according to (40)
- 11: $F_{\alpha}^{opt} = U_1^H \sqrt{(P_{opt})}$

ALGORITHM 4: Binary iterative water-filling algorithm.

TABLE 1: Simulation parameters.

Parameters	Value
Number of antennas	0~180
Channel model	IEEE 802.11ad living room
Number of users	1
Carrier frequency (GHz)	60
Bandwidth	2.56
Power amplifier efficiency ρ (%)	26
Modulation coding scheme	16PSK, 1/2 turbo codes
Simulation frames	20000
Number of RF chains	2~20

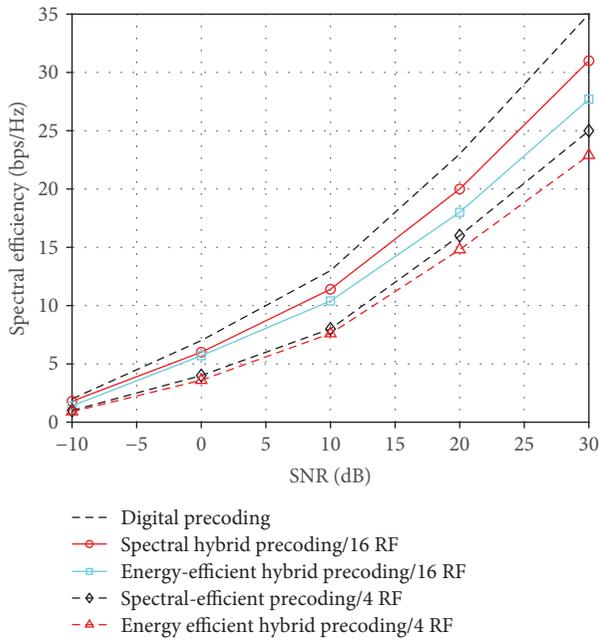


FIGURE 2: SE comparison of different precoding algorithms under different SNRs.

The reason is that in the spectral efficient algorithm, the increase in the sum rate cannot keep path with the increase in power consumption.

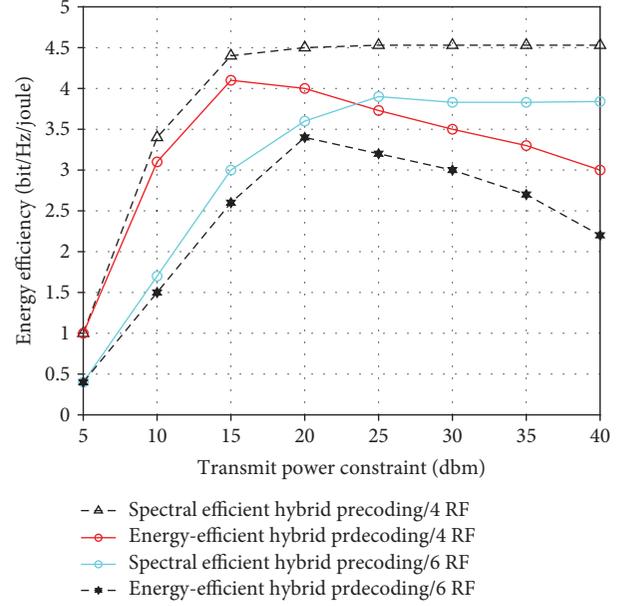


FIGURE 3: EE comparison of different precoding algorithms under different power constraints.

Figure 4 compares the performance between the energy-efficient algorithm and spectral efficient algorithm under different antenna array sizes in terms of EE. As Figure 4 shows, there exists an optimal antenna array size to maximize the EE of the system.

The EE increases at first and then decreases as the number of transmit antennas increases. The more power constraint there is, the larger the antenna array size is. It is reasonable, because once the number of transmit antennas N_t increases over a certain value, the increase in EE cannot compensate for the negative influence of power consumption increase P_c .

Figure 5 compares the energy efficiency achieved under a given optimal number of antennas N_{opt} and transmits power but a different number of RF chains. As Figure 5 shows, under the condition that the optimal number of transmitting antennas N_{opt} is 80, with the increase in the number of RF chains, the EE of the system increases first and then decreases. It can be seen from the figure that the optimum RF number is 8. Meanwhile, it can be seen from Figure 5 that when the number of transmitting antennas is the same, the optimal number of the RF chain is the same. At this time, the larger the transmitting power of the system, the higher the EE that the system can achieve.

Figure 6 compares the SE achieved under a given optimal number of antennas N_{opt} and transmit power but a different number of RF chains. As Figure 5 shows, when the optimal number of transmitting antennas is 80, the EE of the system increases at first and then is kept constant with the increase in the number of radio frequency chains. As can be seen from the figure, when the number of RF chains increases from 1 to 10 under the condition that $N_t = 80$ and $p = 30$ dbm, the RF of the system increases. When the

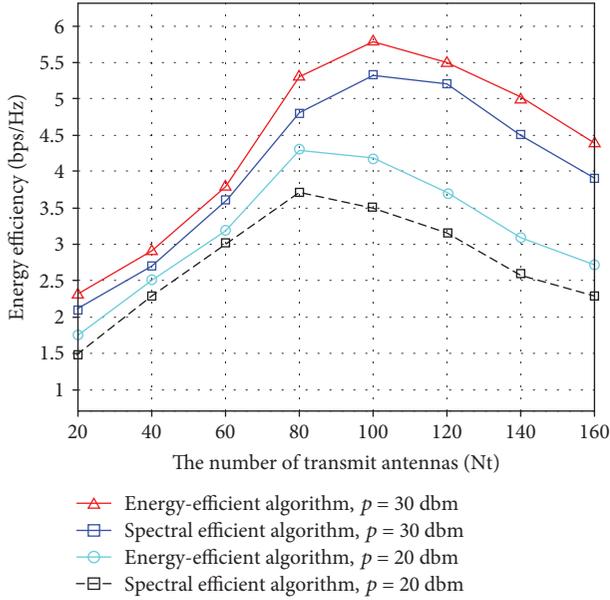


FIGURE 4: EE comparison of different precoding algorithms under different transmit antennas.

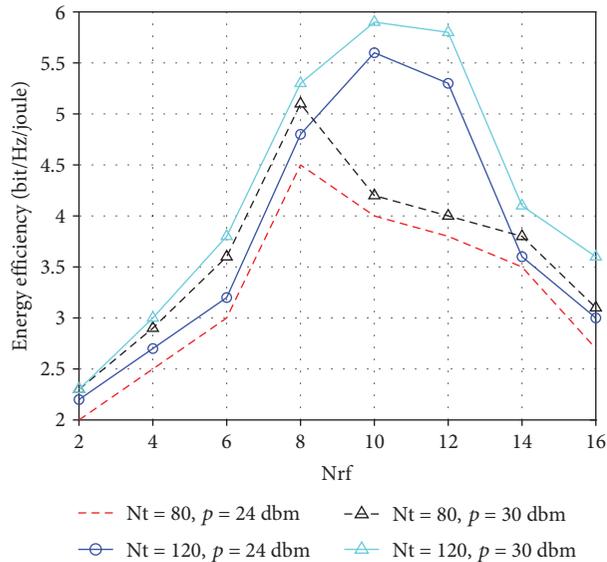


FIGURE 5: EE comparison of the energy-efficient precoding scheme under different RF chains.

number of RF chains is greater than 10, the SE of the system will be constant. However, with the increase in the number of RF, the power consumption of the system will increase, which will lead to the decrease in the EE of the system. This is consistent with the result of Figure 6. In Figure 6, we can conclude that the EE of the system reaches its maximum when the number of RF chains is 8 under the condition that $p = 30$ dbm and $N_t = 80$. Meanwhile, the larger the number of transmitting antennas N_t , the greater the maximum SE the system can achieve.

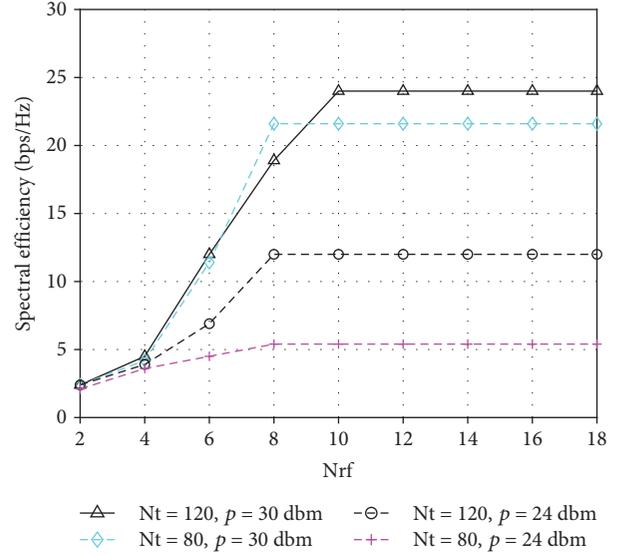


FIGURE 6: SE comparison of the energy-efficient precoding scheme under different radio frequency chains.

8. Conclusions

In this paper, an energy-efficient hybrid precoding scheme for single-user mmWave massive MIMO systems was considered. First, we calculate the optimal number of transmit antennas according to the channel state information and then design an antenna selection algorithm to select the antenna to maximize the channel gain. Finally, the EE of the system is optimized by digital precoding. Numerical results show that the near-optimal hybrid precoding algorithm achieved a near-optimal EE performance. Simulation results show that the algorithm we proposed can achieve higher EE than the traditional precoding scheme can. In future work, we will study on the joint optimization of transmit antennas, receive antennas, and precoder to improve the EE of the system.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper is sponsored by the National Natural Science Foundation of China under Grant 61871321, National Science and Technology Major Project under Grant 2016ZX03001016, and Innovation Team Project of Shaanxi Province under Grant 2017KCT-30-02.

References

- [1] D. W. Bliss, A. M. Chan, and N. B. Chang, "Mimo wireless communication channel phenomenology," *IEEE Transactions on Antennas and Propagation*, vol. 52, no. 8, pp. 2073–2082, 2004.
- [2] K.-K. Wong, R. D. Murch, and K. B. Letaief, "Performance enhancement of multiuser mimo wireless communication systems," *IEEE Transactions on Communications*, vol. 50, no. 12, pp. 1960–1970, 2002.
- [3] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wide-band millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3029–3056, 2015.
- [4] C. Rusu, R. Mendez-Rial, N. Gonzalez-Prelcic, and R. W. Heath, "Low complexity hybrid precoding strategies for millimeter wave communication systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8380–8393, 2016.
- [5] T. E. Bogale and L. B. le, "Massive mimo and mmwave for 5g wireless hetnet: potential benefits and challenges," *IEEE Vehicular Technology Magazine*, vol. 11, no. 1, pp. 64–75, 2016.
- [6] T. Halinen, A. A. Dowhuszko, and J. Hamalainen, "Performance of distributed beamforming for dense relay deployments in the presence of limited feedback information," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, 2013.
- [7] J. Zhang, X. Huang, V. Dyadyuk, and Y. Guo, "Massive hybrid antenna array for millimeter-wave cellular communications," *IEEE Wireless Communications*, vol. 22, no. 1, pp. 79–87, 2015.
- [8] J. Singh and S. Ramakrishna, "On the feasibility of beamforming in millimeter wave communication systems with multiple antenna arrays," *IEEE Transactions on Wireless Communications*, vol. 14, no. 5, pp. 3802–3808, 2014.
- [9] X. Gao, L. Dai, S. Han, C. L. I, and R. W. Heath, "Energy-efficient hybrid analog and digital precoding for mmwave mimo systems with large antenna arrays," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 998–1009, 2016.
- [10] D. W. Browne, M. Manteghi, M. P. Fitz, and Y. Rahmat-Samii, "Experiments with compact antenna arrays for mimo radio communications," *IEEE Transactions on Antennas and Propagation*, vol. 54, no. 11, pp. 3239–3250, 2006.
- [11] A. M. Sayeed and V. Raghavan, "Maximizing mimo capacity in sparse multipath with reconfigurable antenna arrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 156–166, 2007.
- [12] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broad-band systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101–107, 2011.
- [13] X. Zhu, Z. Wang, L. Dai, and Q. Wang, "Adaptive hybrid precoding for multi-user massive mimo," *IEEE Communications Letters*, vol. 20, no. 4, pp. 776–779, 2016.
- [14] P. Zhao and Z. Wang, "Hybrid precoding for millimeter wave communications with fully connected subarrays," *IEEE Communications Letters*, vol. 22, no. 10, pp. 2160–2163, 2018.
- [15] C. C. Hu and J. H. Zhang, "Hybrid precoding design for adaptive subconnected structures in millimeter-wave mimo systems," *IEEE Systems Journal*, vol. 13, no. 1, pp. 137–146, 2019.
- [16] T. Huang, M. Linhua, H. Xing, H. Shaocheng, S. Kangning, and L. Shipping, "Practical hybrid precoding algorithm for millimeter wave massive MIMO," *Journal of Electronics & Information Technology*, vol. 39, no. 8, pp. 1788–1795, 2017.
- [17] M. Majidzadeh, A. Moilanen, N. Tervo, H. Pennanen, A. Tolli, and M. Latva-Aho, "Hybrid beamforming for single-user mimo with partially connected RF architecture," in *2017 European Conference on Networks and Communications (EuCNC)*, Oulu, Finland, 2017.
- [18] M. Hanif, H. C. Yang, G. Boudreau, E. Sich, and H. Seyedmehdi, "Practical hybrid precoding for multi-user massive mimo systems," in *2016 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, Vancouver, BC, Canada, 2016.
- [19] N. Li, Z. Wei, H. Yang et al., "Hybrid Precoding for mmWave Massive MIMO Systems With Partially Connected Structure," *IEEE Access*, vol. 5, no. 1, pp. 15142–15151, 2017.
- [20] S. Park, A. Alkhateeb, and R. W. Heath, "Dynamic subarrays for hybrid precoding in wideband mmWave MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2907–2920, 2017.
- [21] Y. Kai and B. E. Ottersten, "Models for MIMO propagation channels: a review," *Wireless Communications and Mobile Computing*, vol. 2, no. 7, 666 pages, 2002.
- [22] F. Heliot, M. A. Imran, and R. Tafazolli, "On the energy efficiency gain of mimo communication under various power consumption models," in *Future Network & Mobile Summit*, Warsaw, Poland, 2011.
- [23] K. Wang, T. Nirmalathas, C. Lim, and S. Skafidas, "Power-consumption comparison of high-speed indoor millimeter-wave 60GHz communication system and optical wireless systems," in *Asia Pacific Microwave Photonics Conference in Japan*, Kyoto, 2012.
- [24] C. Ma, C. Pan, Y. Pan, X. Han, and C. Ming, "Energy-efficient hybrid precoding for millimeter wave MIMO systems," in *International Conference on Communications & Networking in China*, Shanghai, China, 2016.
- [25] H. Li, L. Song, and M. Debbah, "Energy efficiency of large-scale multiple antenna systems with transmit antenna selection," *IEEE Transactions on Communications*, vol. 62, no. 2, pp. 638–647, 2014.
- [26] A. Moudafi, "A proximal iterative approach to a non-convex optimization problem," *Nonlinear Analysis: Theory, Methods & Applications*, vol. 72, no. 2, pp. 704–709, 2010.
- [27] X. Y. Wang, X. W. Jiang, and J. X. Dai, "Homotopy method for nonconvex optimization problem," *Journal of Jilin University*, vol. 52, no. 2, pp. 273–276, 2014.
- [28] Y. Bian, B. Mirzasoleiman, J. M. Buhmann, and A. Krause, "Guaranteed non-convex optimization: submodular maximization over continuous domains," 2016, arXiv: 1606.05615.
- [29] P. Jain and P. Kar, "Non-convex optimization for machine learning," *Foundations and Trends® in Machine Learning*, vol. 10, no. 3-4, pp. 142–336, 2017.
- [30] J. Jing, C. Xiaoxue, and X. Yongbin, "Energy-efficiency based downlink multi-user hybrid beamforming for millimeter wave massive mimo system," *The Journal of China Universities of Posts and Telecommunications*, vol. 23, no. 4, pp. 53–62, 2016.

- [31] J. Jing, K. Deting, L. Guangyue, W. Junxuan, and Y. Wujun, "User scheduling algorithm for mmwave FDMA massive MU-MIMO system with hybrid beamforming," in *Proceedings of the 8th International Conference on Signal Processing Systems - ICSPS 2016*, pp. 194–198, Auckland, New Zealand, 2016.
- [32] T. E. Bogale and L. B. Le, "Pilot optimization and channel estimation for multiuser massive mimo systems," in *2014 48th Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, USA, 2014.
- [33] A. Zilinskas and G. Gimbutiene, "On one-step worst-case optimal trisection in univariate bi-objective Lipschitz optimization," *Communications in Nonlinear Science and Numerical Simulation*, vol. 35, pp. 123–136, 2016.
- [34] J. P. Crouzeix and J. A. Ferland, "Algorithms for generalized fractional programming," *Mathematical Programming*, vol. 52, no. 1-3, pp. 191–207, 1991.
- [35] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, 1982.
- [36] G. Haeser and M. L. Schuverdt, "On approximate KKT condition and its extension to continuous variational inequalities," *Journal of Optimization Theory and Applications*, vol. 149, no. 3, pp. 528–539, 2011.
- [37] M. Kobayashi and G. Caire, "An iterative water-filling algorithm for maximum weighted sum-rate of Gaussian MIMO-BC," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1640–1646, 2006.
- [38] R. S. Prabhu and B. Daneshrad, "An energy-efficient water-filling algorithm for OFDM systems," in *2010 IEEE International Conference on Communications*, Cape Town, South Africa, 2010.

Research Article

A Statistical Channel Model for Stochastic Antenna Inclination Angles

Gang Liu, Ming Zhang, and Yaming Bo 

College of Electronic and Optical Engineering & College of Microelectronics, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Correspondence should be addressed to Yaming Bo; [ymbo@njupt.edu.cn](mailto:yambo@njupt.edu.cn)

Received 28 August 2018; Accepted 3 February 2019; Published 28 March 2019

Guest Editor: Raed A. Abd-Alhameed

Copyright © 2019 Gang Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The actions of a person holding a mobile device are not a static state but can be considered as a stochastic process since users can change the way they hold the device very frequently in a short time. The change in antenna inclination angles with the random actions will result in varied received signal intensity. However, very few studies and conventional channel models have been performed to capture the features. In this paper, the relationships between the statistical characteristics of the electric field and the antenna inclination angles are investigated and modeled based on a three-dimensional (3D) fast ray-tracing method considering both the diffraction and reflections, and the radiation patterns of an antenna with arbitrary inclination angles are deducted and included in the method. Two different conditions of the line-of-sight (LOS) and non-line-of-sight (NLOS) in the indoor environment are discussed. Furthermore, based on the statistical analysis, a semiempirical probability density function of antenna inclination angles is presented. Finally, a novel statistical channel model for stochastic antenna inclination angles is proposed, and the ergodic channel capacity is analyzed.

1. Introduction

Wireless communication technology has been widely used in communication systems for its mobility, convenience, flexibility, and lower cost compared with wired transmission. However, the signals are significantly affected by the surrounding environment and undergo fading and time variation before arriving at the receiver. In order to achieve a higher rate and more reliable communication, the acquisition of accurate channel state information (CSI) and channel modelling is fundamental and crucial in designing a wireless communication system and has been attracting researchers' attention.

A basic framework of the geometry-based stochastic channel modelling approach (GSCMA) is developed in [1] for three different scenarios with the corresponding channel parameters such as delay spread, angle spread, shadow fading, angle of departure, angle of arrival, and delay power spectrum extracted from a large number of measurements. Additionally, a polarized channel model is also proposed based on crosspolarization discrimination (XPD) when

considering the depolarization effect of channels on electromagnetic waves. The WINNER II channel model [2] extends the number of scenarios to more than a dozen but follows the same channel modelling approach. Furthermore, the WINNER II channel model allows propagation between line-of-sight (LOS) and non-line-of-sight (NLOS) conditions for a same scenario. Analogously, a number of channel models are established using GSCMA but assuming that the scatterers are distributed on regular geometries in two or three dimensions such as the one-ring model [3], twin-cluster model [4], and elliptical model [5] considering only the azimuth angle and the double-cylinder model [6], two-sphere model [7], and multiconfocal ellipsoid model [8] considering the influence of the elevation angle. However, these conventional channel models are assumed to be generally stationary, but this is not sufficiently applicable for the channels of the massive multiple-input multiple-output (MIMO) recognized as one of the most important candidate technologies for the fifth-generation (5G) mobile communication systems due to the potential and additional advantages compared with conventional MIMO technologies [9–11].

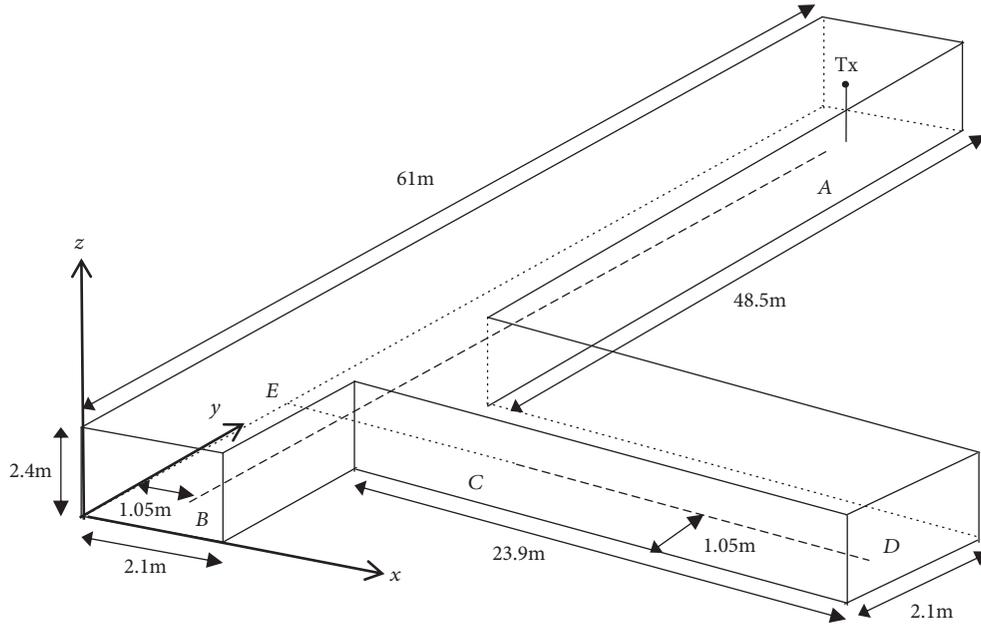


FIGURE 1: The three-dimensional layout of the T-shaped corridor.

Consequently, several novel models [12–14] are developed to capture the new features observed from the measurements such as the spherical wave front, nonstationary effect on the antenna array axis and the time axis.

The above channel models are independent of the antenna configurations and element radiation patterns. Instead, the correlative channel models such as the Kronecker model [15] and the Weichselberger model [16] use the correlation matrices at the mobile station (MS) and base station (BS) without knowing the distribution of scatterers or clusters resulting in a lower complexity. However, few studies of the channel modelling for stochastic antenna inclination angles have been done. It is known that mobile devices are not fixed on walls or people’s desks as routers or computers. Instead, people communicate using a mobile device whenever and wherever possible; for instance, they are lying down, standing, and walking. The way people hold a mobile device is not a static state but a stochastic process since users can change the way they hold the mobile device very frequently even in a few seconds. Consequently, the antenna inclination angles will change with the rotation of the mobile devices leading to the variation of received signals due to the polarization mismatch between the signals and antennas. In this paper, a statistical channel model for stochastic antenna inclination angles in the indoor environment is developed based on a modified three-dimensional (3D) fast ray-tracing method. Two different conditions of LOS and NLOS for a common scenario of the T-shaped corridor for an indoor environment are investigated. Furthermore, in order to capture the stochastic characteristics of people holding a mobile device, a semiempirical probability density function (PDF) of antenna inclination angles is proposed, and closed expressions for the radiation patterns of a half-wave antenna for arbitrary inclination angles are deduced based on the principle of coordinate transformation. Finally, the

ergodic capacities under two different conditions are analyzed based on the proposed channel model.

This paper is organized as follows. A modified 3D fast ray-tracing method is introduced, and the validity and accuracy of the method in predicting the electromagnetic fields are verified in Section II. In Section III, the statistical channel model for stochastic antenna inclination angles is presented in detail. The numerical results are analyzed in Section IV, and conclusions are drawn in Section V.

2. Simulation Environment, Method, and Validation

In this section, a modified 3D fast ray-tracing method based on space subdivision is introduced and used to predict the electromagnetic fields in a T-shaped corridor. The layout and corresponding sizes of the corridor are shown in Figure 1. It is composed of brick walls, concrete floor, and ceiling with the corresponding parameters: the relative permittivity and conductivity are $\epsilon_r = 4.0$ and $\sigma_2 = 0.343$ s/m for the walls and $\epsilon_r = 6.14$ and $\sigma_2 = 1.005$ s/m for the floor and ceiling [17]. If all the angles in the corridor are assumed to be right angles, the whole space of the corridor can be divided into many hexahedrons and each hexahedron can be further split into five tetrahedrons. Each hexahedron and tetrahedron must be seamless and nonoverlapping. It is worth noting that there are two different types of faces or lines in the model. One is called real face (RF) or real line (RL) since the face or line exists in the realistic scene such as the surface or edge of a wall. The other one is invisible in fact and only introduced for subsequent analyses and computations so that it is called the virtual face (VF) or virtual line (VL). The space meshing should make each face of a tetrahedron have only one property, the real or the virtual.

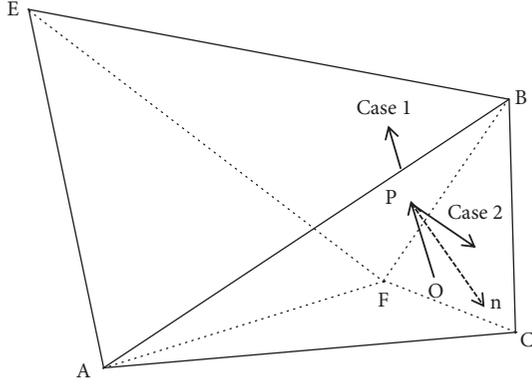


FIGURE 2: The demonstration of determining the next tetrahedron.

Since the size of an antenna is relatively small compared with the realistic propagation environment, a single antenna can be approximated as a transmitting point. If the coordinates of the transmitting point are given, the initial tetrahedron where the transmitting point is located in can be determined. The rays depart from the transmitting point and arrive at a receiving point through a number of tetrahedrons, and the path of each ray can be traced using the method proposed in [18]. However, the calculation is so cumbersome that we modify the expression and give a more compact form as

$$\alpha_i = \frac{\langle \overrightarrow{OB}, \vec{n}_i \rangle}{\langle \vec{r}, \vec{n}_i \rangle}, \quad i = 1, 2, 3, 4, \quad (1)$$

where \overrightarrow{OB} denotes the vector from the transmitting point O to an arbitrary vertex of the i -th face of the initial tetrahedron as shown in Figure 2 and \vec{n}_i and a_i represent the normal vector and extension coefficient of the i -th face, respectively. \vec{r} is the unit vector of propagation, and $\langle \cdot, \cdot \rangle$ represents the inner product of two vectors. Consequently, the face hit by the ray corresponds to the minimum and positive extension coefficient.

Note that there are two cases of propagation for different type of faces. If the face is a VF, the incident rays will pass through the face and reach the adjacent tetrahedron as in the case 1 shown in Figure 2. Otherwise, the face is a RF, and the incident rays will be reflected in the current tetrahedron as in the case 2. If the roughness of all surfaces in the environment is assumed to be neglected, the reflected field can be determined according to Fresnel's laws of reflection. The reflection coefficients for perpendicular polarization and parallel polarization are given as

$$\Gamma_{\perp} = \frac{\cos \theta_i - \sqrt{\epsilon_r - \sin^2 \theta_i}}{\cos \theta_i + \sqrt{\epsilon_r - \sin^2 \theta_i}}, \quad (2)$$

$$\Gamma_{\parallel} = \frac{\epsilon_r \cdot \cos \theta_i - \sqrt{\epsilon_r - \sin^2 \theta_i}}{\epsilon_r \cdot \cos \theta_i + \sqrt{\epsilon_r - \sin^2 \theta_i}}, \quad (3)$$

where ϵ_r is the relative permittivity and θ_i is the incident angle.

Furthermore, if an obstacle with the size much larger than the wavelength of the incident wave is present in the propagation path, the diffraction should be taken into consideration. In order to determine the diffraction field, Holm's heuristic diffraction coefficients are selected to calculate the diffraction field due to the simple expressions and the good consistency with the rigorous solution for finite conductivity as shown in [19]. The diffraction coefficients for perpendicular polarization and parallel polarization are expressed in a more compact form as

$$D_{\parallel}^{\pm}(L, n, \varphi, \varphi') = \frac{-e^{-j(\pi/4)}}{2n\sqrt{2\pi k}} \cdot \left\{ \cot \left[\frac{\pi - (\varphi - \varphi')}{2n} \right] \cdot F \left[2kLn^2 \sin^2 \left(\frac{\pi - (\varphi - \varphi')}{2n} \right) \right] + R_{0_{\parallel}^{\pm}} R_{n_{\parallel}^{\pm}} \cot \left[\frac{\pi + (\varphi - \varphi')}{2n} \right] \cdot F \left[2kLn^2 \sin^2 \left(\frac{\pi + (\varphi - \varphi')}{2n} \right) \right] + R_{0_{\parallel}^{\pm}} \cot \left[\frac{\pi - (\varphi + \varphi')}{2n} \right] \cdot F \left[2kLn^2 \sin^2 \left(\frac{\pi - (\varphi + \varphi')}{2n} \right) \right] + R_{n_{\parallel}^{\pm}} \cot \left[\frac{\pi + (\varphi + \varphi')}{2n} \right] \cdot F \left[2kLn^2 \sin^2 \left(\frac{\pi + (\varphi + \varphi')}{2n} \right) \right] \right\},$$

$$F(x) = 2j\sqrt{x}e^{jx} \int_{\sqrt{x}}^{\infty} e^{-j\tau^2} d\tau,$$

$$L = \frac{ss'}{s + s'}, \quad (4)$$

where $R_{0_{\parallel}^{\pm}}$ and $R_{n_{\parallel}^{\pm}}$ are the reflection coefficients for the perpendicular polarization and parallel polarization referred to in the formulas expressed in (2) and (3) for the 0-face and n -face [19]. φ' and φ represent the incident angle and diffraction angle, respectively. k is the wave number, $(2 - n)\pi$ is the inner angle of the wedge and here $n = 1.5$ due to the previous assumption of right angles in the corridor, s is the distance between the diffraction point and the diffraction observation

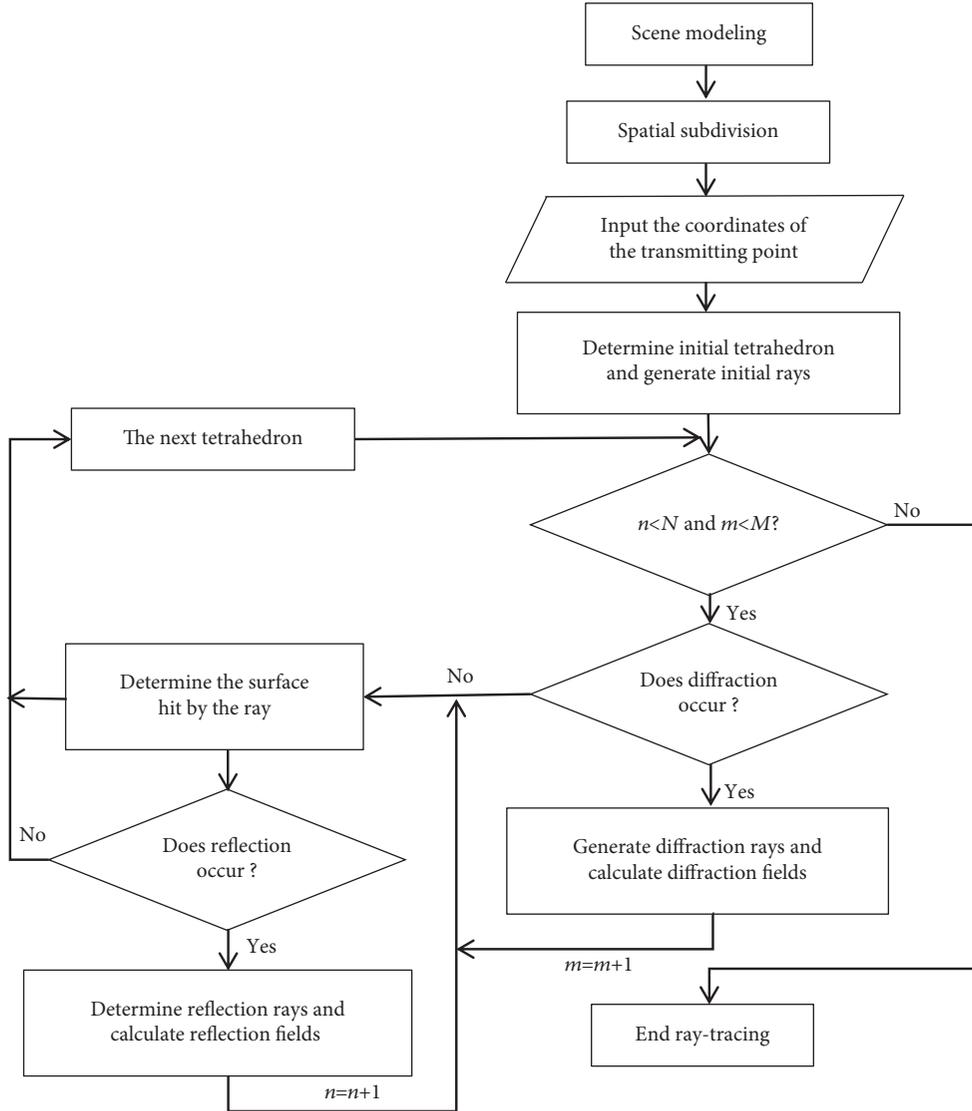


FIGURE 3: Flow chart of the 3D fast ray-tracing method.

point, and s' is the distance between the source point of the incident ray and the diffraction point.

The process of the 3D fast ray-tracing is shown in Figure 3, where N and M denote the number of reflections and the number of reflections after diffraction, respectively.

In the process of ray-tracing, it is assumed that each ray is independent of each other and the field around a ray should be represented only by the ray. In addition, it is necessary to determine whether a ray contributes to the field at a receiving point. One of the effective methods is using the reception sphere [20]. The radius of the reception sphere for each ray is expressed as

$$r = \frac{\alpha d}{\sqrt{3}}, \quad (5)$$

where α is the angle between two adjacent rays and d is the path length from the transmitting point to the receiving point.

Note that if the receiving point locates in the overlapping area of the rays, the double counting error [21] will be generated. A method of reducing the double count error is presented as follows.

Firstly, in order to determine whether a ray is received, all its adjacent rays need to be tested simultaneously. Secondly, if an adjacent ray is received, it must be determined whether the number of reflections and the surfaces in the whole paths of the two rays are equal. If it is true, we compare the distances from the receiving point to the two rays and discard the further one as a repeated ray. Otherwise, the two rays will be retained.

It is known that when people are communicating with mobile devices, the rapidly changing actions of holding their mobile devices result in the antenna inclination angles varying randomly. Assuming that half-wave dipole antennas are used on both the transmitting and receiving sides, and if the coordinate system of the antenna is taken as the local coordinate system and the coordinate system of the corridor

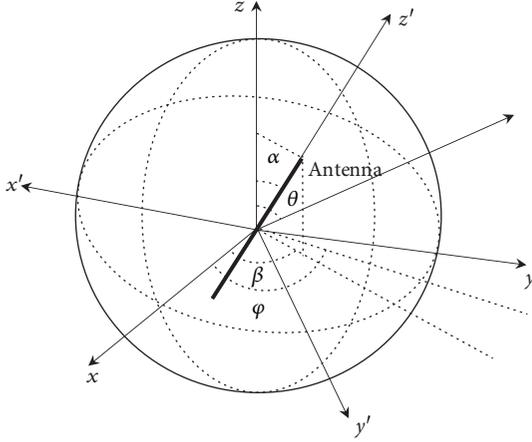


FIGURE 4: The demonstration of two coordinate systems. x - y - z is the global coordinate system, and x' - y' - z' is the local coordinate system.

environment is considered as the global coordinate system as shown in Figures 1 and 4, the radiation patterns of the antenna also change with the antenna inclination angles in the global coordinate system. According to the coordinate transformation, radiation patterns of a half-wave dipole antenna with arbitrary inclination angles are expressed as

$$\begin{aligned}
 F_{\theta}(\alpha, \beta, \theta, \varphi) &= (\cos \alpha \sin \theta - \sin \alpha \cos \beta \cos \theta \cos \varphi \\
 &\quad - \sin \alpha \sin \beta \cdot \cos \theta \sin \varphi) \cdot \frac{\cos((\pi/2)\chi)}{1 - \chi^2}, \\
 F_{\varphi}(\alpha, \beta, \theta, \varphi) &= (\sin \alpha \sin \beta \cos \varphi - \sin \alpha \cos \beta \sin \varphi) \\
 &\quad \cdot \frac{\cos((\pi/2)\chi)}{1 - \chi^2}, \\
 \chi &= \cos \alpha \cos \theta + \sin \alpha \cos \beta \sin \theta \cos \varphi \\
 &\quad + \sin \alpha \sin \beta \sin \theta \sin \varphi,
 \end{aligned} \tag{6}$$

where F_{θ} and F_{φ} represent the radiation pattern of the θ direction and φ direction in the global coordinate system, respectively. As shown in Figure 4, α and β represent the zenith angle and azimuth angle of the tilted antenna in the global coordinate system, respectively. θ and φ are the zenith angle and azimuth angle in the global coordinate system, respectively. For $\alpha = 0^\circ$, the electric field patterns are simplified to

$$\begin{aligned}
 F_{\theta} &= \frac{\cos((\pi/2) \cos \theta)}{\sin \theta}, \\
 F_{\varphi} &= 0.
 \end{aligned} \tag{7}$$

These are the electric field patterns of a vertical polarized half-wave dipole antenna.

Assuming that the transmitting antenna is fixed in the vertical polarization whereas the angle of the receiving antenna varies randomly, the intensity of the electric field received by the antenna will be different at the same receiving point since only the electric field components parallel to the

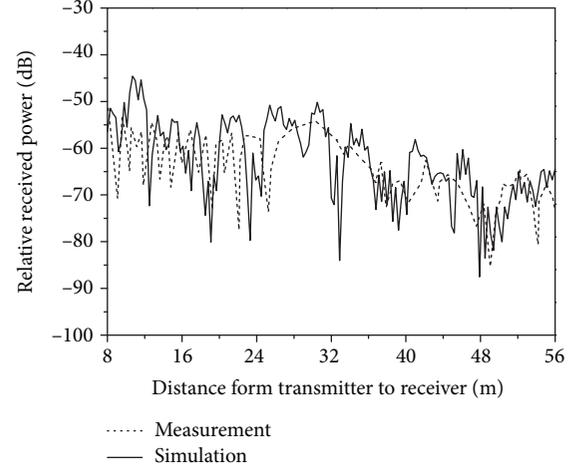


FIGURE 5: The relative received power from A to B in the LOS corridor ($N = 10$, $M = 8$, $G_T = 1$ dBi, and $G_R = 1$ dBi).

antenna act on the received signals. Consequently, after all the effective rays are determined, the total complex E-field can be obtained as

$$\begin{aligned}
 \vec{E}_{\text{total}} &= \left\langle \sum_{i=1}^K \left[\vec{E}_i(\theta_i, \varphi_i) \cdot \vec{e}_{\theta} \cdot F_{\theta}(\alpha, \beta, \theta_i, \varphi_i) + \vec{E}_i(\theta_i, \varphi_i) \right. \right. \\
 &\quad \left. \left. \cdot \vec{e}_{\varphi} \cdot F_{\varphi}(\alpha, \beta, \theta_i, \varphi_i) \right], \vec{S}_{\text{Ant}}(\alpha, \beta) \right\rangle \cdot \vec{S}_{\text{Ant}}(\alpha, \beta),
 \end{aligned} \tag{8}$$

where K is the total number of the effective rays. i denotes the i -th effective subpath. \vec{e}_{θ} and \vec{e}_{φ} are the unit direction vectors in the θ and φ directions, respectively. \vec{S}_{Ant} is the antenna axial directional vector. $\langle \cdot, \cdot \rangle$ represents the inner product of two vectors. Consequently, the relative received power (RRP) (relative to transmitting power) can be obtained as [17]

$$\text{RRP}(\text{dB}) = G_T + G_R + 20 \log_{10} \left| \frac{\vec{E}_{\text{total}}}{E_{\text{total}}} \right| - \text{PL}_0, \tag{9}$$

where G_T and G_R represent the transmitting antenna gain and receiving antenna gain, respectively. PL_0 is the free space path loss from the transmitter to receiver.

In order to verify the accuracy and effectiveness of the modified 3D fast ray-tracing method used for predicting the propagation characteristics of the electromagnetic waves in the T-shaped corridor, the simulation results of RRP obtained from the method are compared with the measurement results in [17] as shown in Figures 5 and 6. The simulation using the ray-tracing method is performed at 5.3 GHz according to the configuration and parameter setting of the measurement system in [17], but the antennas are assumed to be half-wave dipole antennas at the transmitting and receiving sides. Both transmitting and receiving

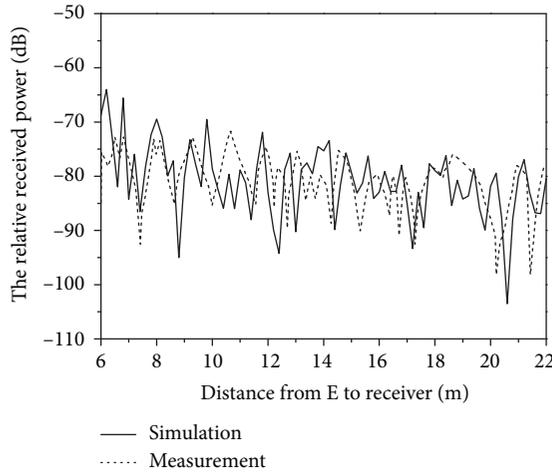


FIGURE 6: The relative received power from C to D in the NLOS corridor ($N = 10$, $M = 8$, $G_T = 13$ dBi, and $G_R = 1$ dBi).

antennas are vertically polarized, i.e., $\alpha = 0^\circ$. The transmitting antenna is fixed at one end of the corridor, and the receiving antenna moves along the path from points A to B in the LOS corridor and points C to D in the NLOS corridor as shown in Figure 1. The heights of the transmitting antenna and the receiving antenna are set to 1.8 m and 1.6 m, respectively. Furthermore, the transmitting power is set to be 29 dBm in the simulation, and omnidirectional vertically polarized antennas with different gains but with the same transmitting power are assumed at the transmitting side for different conditions of the LOS and NLOS corridors. As shown in Figure 5, the transmitting antenna gain and receiving antenna gain are equal to 1 dBi, i.e., $G_T = G_R = 1$ dBi, the relative received power (relative to the transmitting power) ranges from -85 dB to -45 dB and decreases along the path from points A to B in the LOS corridor as shown in Figure 1. It is known that the direct path plays a dominant role in the receiving power, and signal attenuation is mainly due to the energy diffusion and reflections on the walls, floors, and ceilings when scattering and transmission are assumed to be neglected. In another case, i.e., $G_T = 13$ dBi and $G_R = 1$ dBi, the relative received power ranges from -100 dB to -60 dB along the path from points C to D in the NLOS corridor as shown in Figure 6, and it is less than the LOS condition even for the larger transmitting antenna gain. This is because there is no direct path in the NLOS corridor, and multiple reflections and diffraction result in the heavy attenuation. The simulation results show good consistency with the measurements for both LOS and NLOS conditions even though there are some differences due to the complicated realistic corridor environment. Consequently, the method is verified to be accurate and effective in the prediction of electric field for the T-shaped corridor environment.

The intensity of the electric field received by an antenna can change with different inclination angles at a receiving point due to the polarization mismatch and the changing electric field patterns and can be analyzed using the modified 3D fast ray-tracing method. For instance, assuming the azimuth angle is equal to zero degrees, the relationships between the relative received power and the zenith angles along the

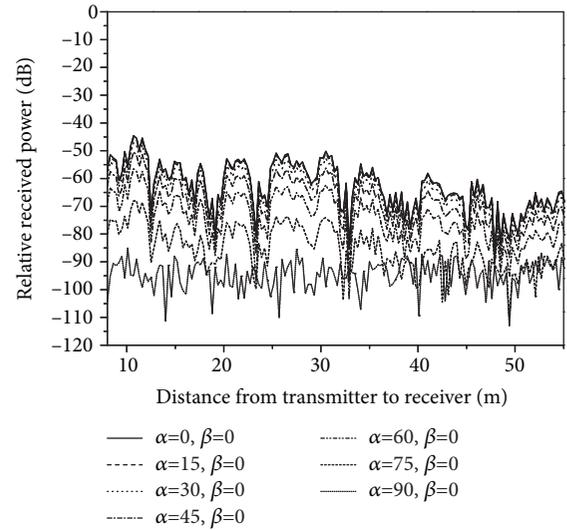


FIGURE 7: The relative received power for different inclination angles for the LOS corridor ($\beta = 0^\circ$, $N = 10$, $M = 8$, $G_T = 1$ dBi, and $G_R = 1$ dBi).

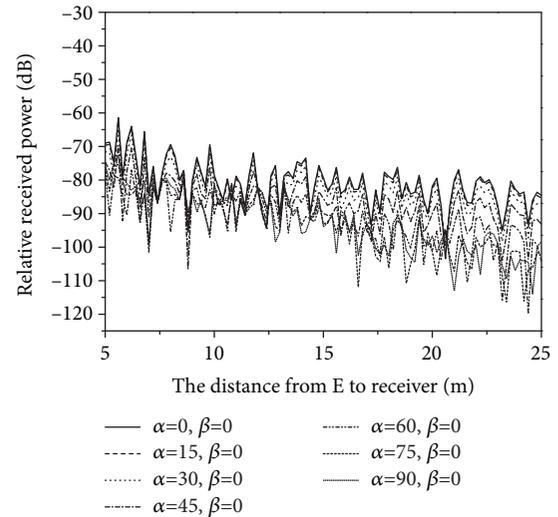


FIGURE 8: The relative received power for different inclination angles for the NLOS corridor ($\beta = 0^\circ$, $N = 10$, $M = 8$, $G_T = 13$ dBi, and $G_R = 1$ dBi).

paths from point A to point B in the LOS corridor and from point C to point D in the NLOS corridor are shown in Figures 7 and 8, respectively. In the LOS corridor, the relative receiving power decreases with the increase of the zenith angle, and the offsets relative to the vertically polarized antenna, i.e., $\alpha = 0^\circ$, increase but retain almost the same trend from 0 to 75 degrees. However, for the horizontally polarized antenna, i.e., $\alpha = 90^\circ$, the relative received power is far less than the vertical polarization and shows a different trend compared with others. This is because the transmitting antenna is vertically polarized and the main propagation mechanisms are the direction and reflections. The direct path plays a dominant role in the LOS corridor resulting in the relatively weaker depolarization effect.

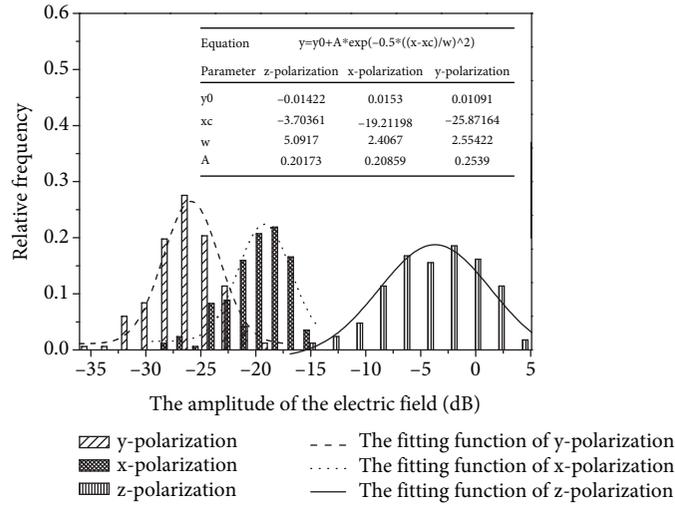


FIGURE 9: The probability density function (PDF) of the amplitude of the electric field in the LOS corridor.

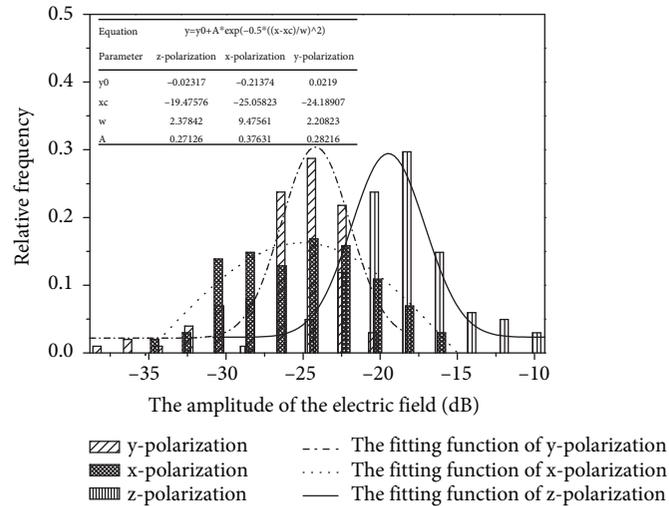


FIGURE 10: The probability density function (PDF) of the amplitude of the electric field in the NLOS corridor.

In the NLOS corridor, the relative received power for the vertically polarized antenna is still the largest whereas the offsets relative to the vertically polarized antenna for other angles are less than those in the LOS corridor especially for the horizontally polarized receiving antenna. This is because there is no direct path in the NLOS corridor and the main propagation mechanisms are the diffraction and reflections leading to the serious channel depolarization effect.

3. A Statistical Channel Model for Stochastic Antenna Inclination Angles

In order to obtain the statistical CSI of the T-shaped corridor, 169 different user positions from A to B for the LOS corridor and 101 different user positions from C to D for the NLOS corridor are selected for simulation and investigation. The half-wave dipole antennas are assumed to be used at both ends of the transceiver with the transmitting power of 30 dBm, and the transmitting antenna is fixed in the vertical

polarization whereas the inclination angles of the receiving antenna are variable.

The simulation results of the electric field at different receiving points can be obtained according to (8) using the ray-tracing method. According to the statistical analyses, the statistical characteristics of the amplitude and the phase of the electric field can also be obtained. The results show that the amplitude of the electric field when expressed in decibels is subject to Gauss distribution for both LOS and NLOS corridors. For instance, the statistical distribution of the amplitude of the received electric field for the vertical polarization or z-polarization, i.e., $\alpha = 0^\circ$, and two horizontal polarizations or x-polarization and y-polarization, i.e., $\alpha = 90^\circ$, $\beta = 0^\circ$, and $\beta = 90^\circ$, are shown in Figures 9 and 10. For the LOS condition, the mean value is -3.7 dB for the z-polarization and much larger than -19.2 dB for the x-polarization and -25.9 dB for the y-polarization as shown in Figure 9 since the vertical polarization and horizontal polarization are copolarized and crosspolarized relative to the transmitting antenna, respectively, and the existence

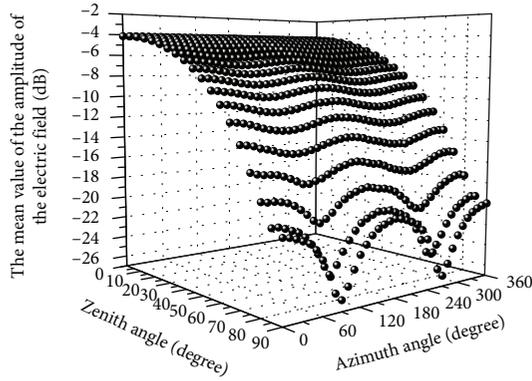


FIGURE 11: The mean value of the amplitude of the electric field for different inclination angles in the LOS corridor (scatter diagram).

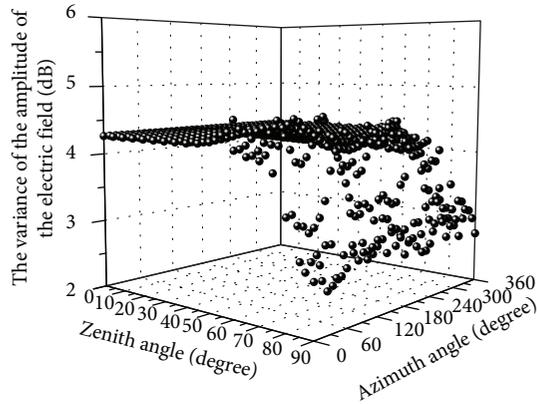


FIGURE 12: The variance of the amplitude of the electric field for different inclination angles in the LOS corridor (scatter diagram).

of the direct path leads to relatively weaker depolarization as mentioned above. Furthermore, the value of the x-polarization is larger than that of the y-polarization since the reflection surfaces in the LOS corridor are in the y direction resulting in the heavier depolarization for the component of the y-polarization than that of the x-polarization and z-polarization. For the LOS corridor, as shown in Figure 10, the mean value is -19.5 dB for the z-polarization and approximately equal to -25 dB for the x-polarization and -24.2 dB for the y-polarization due to the strong depolarization after the diffraction and multiple reflections. Furthermore, the reflection surfaces in the NLOS corridor are in the x direction leading to nearly equal mean values of the two horizontal polarizations.

In order to obtain the statistical characteristics of received signals, the mean value and variance of the amplitude of the electric field expressed in decibels for different zenith angles and azimuth angles in the LOS corridor are depicted in Figures 11 and 12, respectively. The results show that the mean values range from -26 dB to -2 dB and decrease with the increase of the zenith angle due to the increased polarization mismatch. However, the effect of the azimuth angle is not obvious for the lower zenith angles but becomes

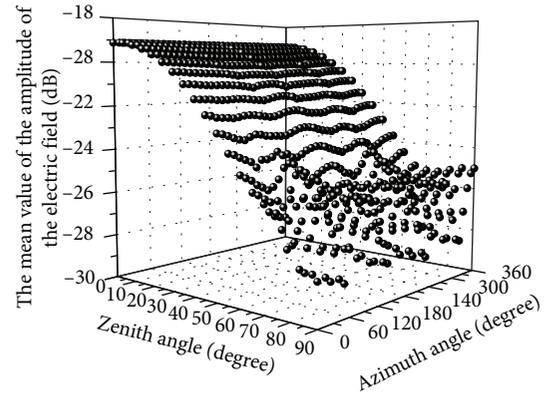


FIGURE 13: The mean value of the amplitude of the electric field for different inclination angles in the NLOS corridor.

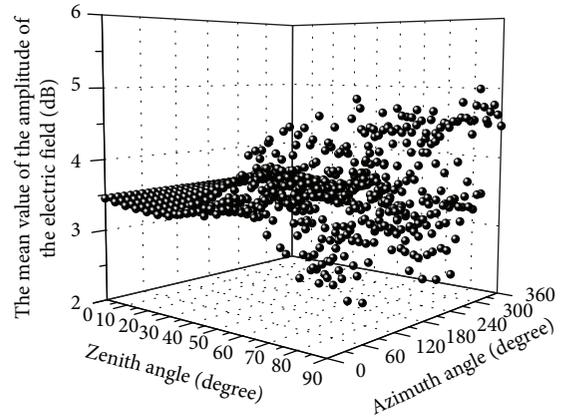


FIGURE 14: The variance of the amplitude of the electric field for different inclination angles in the NLOS corridor.

significant for the larger zenith angles, especially when the zenith angle is equal to zero degrees; it is known that the mean value is independent of azimuth angles. This is because there are three basic polarizations as mentioned above, and the polarization mismatch also exists between two horizontal polarizations. Furthermore, the main reflection surfaces in the LOS corridor are in the y direction, which results in the heavier depolarization for the y-polarization than that of the x-polarization. However, the variance approximates to a constant of 4.25 dB for the lower zenith angles, but fluctuations occur for the larger angles as shown in Figure 12.

In the NLOS corridor, the mean values and variances of the amplitude of the electric field changing with antenna inclination angles are shown in Figures 13 and 14, respectively. The mean values range from -30 dB to -19 dB and are smaller than those of the LOS condition due to the heavy attenuation after the diffraction and multiple reflections. The depolarization transforms more power from the copolarization components into the crosspolarization ones. In addition, since the main reflection surfaces in the NLOS corridor are in the x direction, the gaps among three polarizations are reduced. Similar to the LOS condition, the variance also

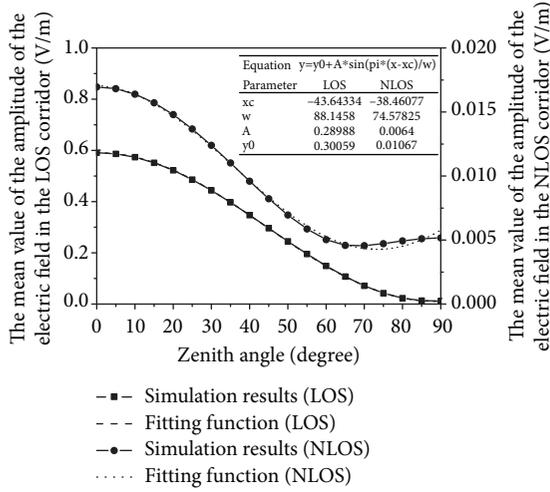


FIGURE 15: The mean value of the amplitude of the electric field for different inclination angles in the LOS and NLOS corridors ($\beta = 0^\circ$).

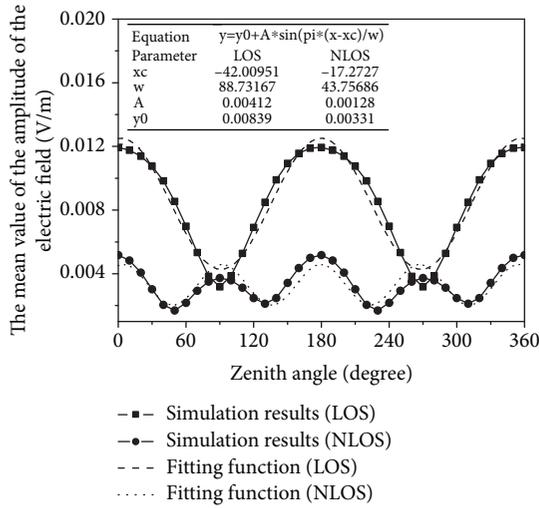


FIGURE 16: The mean value of the amplitude of the electric field for different inclination angles in the LOS and NLOS corridors ($\alpha = 90^\circ$).

tends to be a constant for the lower zenith angles with the value of 3.5 dB but shows approximately uniform distribution between 2.5 dB and 5 dB for the larger zenith angles.

According to the fitting results, the mean value of the amplitude of the electric field shows sinusoidal variation with the zenith angle or azimuth angle for both LOS and NLOS corridors as shown in Figures 15 and 16, and the fitting function can be expressed as

$$y = y_0 + A \sin\left(\frac{\pi(x - x_c)}{w}\right), \quad (10)$$

where y_0 , A , x_c , and w are the coefficients of the fitting function. Consequently, the functional relationship between the mean value of the amplitude of the electric field and antenna inclination angles can be deduced and expressed in decibels as follows.

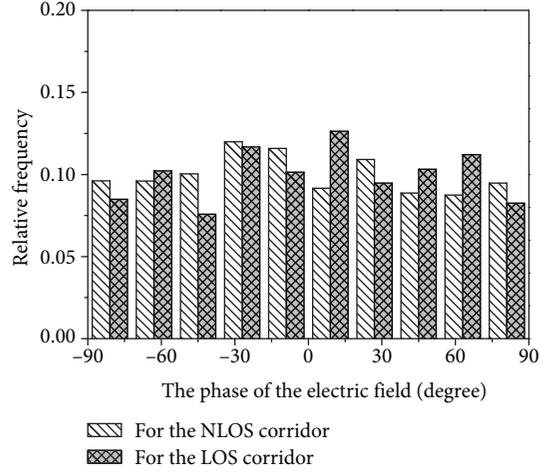


FIGURE 17: The statistical distribution of the phase of the electric field for the LOS and NLOS corridors.

For the LOS condition:

$$A_{\text{mean-LOS}}(\alpha, \beta) = 10 \log_{10} \left\{ A_{\text{LOS}} \sin \left[\frac{\pi(\alpha - x_{c,\text{LOS}})}{w_{\text{LOS}}} \right] + 0.1976 \right\}, \quad 0^\circ \leq \alpha \leq 90^\circ, 0^\circ \leq \beta \leq 360^\circ,$$

$$A_{\text{LOS}} = 0.0025 \sin(0.0357\beta - 1.7263) + 0.1902,$$

$$w_{\text{LOS}} = 0.9540 \sin(0.0356\beta - 2.009) + 88.3722,$$

$$x_{c,\text{LOS}} = 2.2633 \sin(0.0351\beta + 1.4202) - 45.474. \quad (11)$$

For the NLOS condition:

$$A_{\text{mean-NLOS}}(\alpha, \beta) = 10 \log_{10} \left\{ A_{\text{NLOS}} \sin \left[\frac{\pi(\alpha - x_{c,\text{NLOS}})}{w_{\text{NLOS}}} \right] + y_{\text{NLOS}} \right\}, \quad 0^\circ \leq \alpha \leq 90^\circ, 0^\circ \leq \beta \leq 360^\circ,$$

$$A_{\text{NLOS}} = 0.0003 \sin(0.0702\beta - 1.589) + 0.0052,$$

$$x_{c,\text{NLOS}} = -42.5295 + 2.6605 \sin(0.0384\beta + 0.7388),$$

$$w_{\text{NLOS}} = 81.103 + 3.7265 \sin(0.037\beta - 1.85),$$

$$y_{\text{NLOS}} = 0.0072 + 0.0004 \sin(0.0696\beta - 4.7047). \quad (12)$$

where α and β represent the zenith angle and azimuth angle of the tilted antenna as shown in Figure 4.

According to the statistical analysis, as shown in Figure 17, the phase of the electric field is subject to uniform distribution between -90° and 90° for both LOS and NLOS corridors.

A statistical channel model for stochastic antenna inclination angles can be established as

$$H_{M_R \times N_T} = A(\alpha, \beta) \exp(j\Phi(\alpha, \beta)), \quad (13)$$

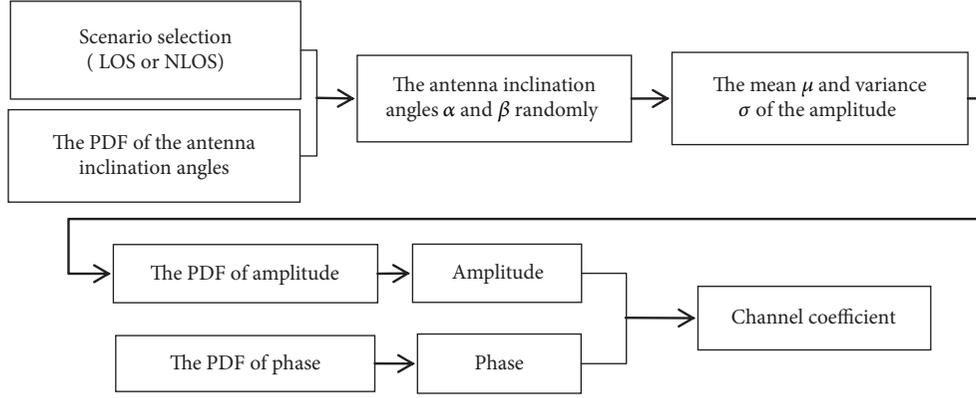


FIGURE 18: The flow chart of the process of generating channel coefficients.

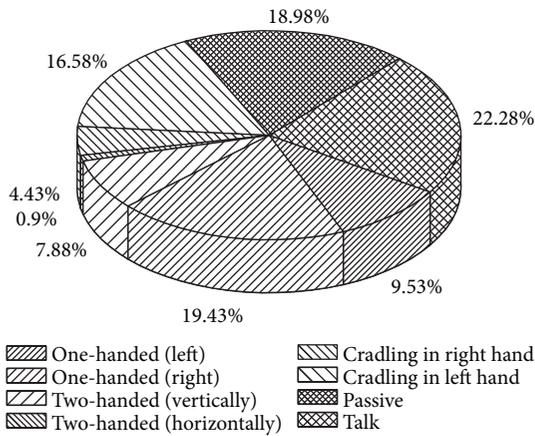


FIGURE 19: The pie chart of the survey results for people holding their mobile devices.

TABLE 1: The relationship between the way of holding the mobile devices and the antenna inclination angles.

The ways of holding the mobile devices	The antenna inclination angles (zenith angle)
Talk	0°-30°
One-handed use, two-handed use (vertically), cradling	30°-60°
Two-handed use (horizontally), passive	60°-90°

where $H_{M_R \times N_T}$ is the $M_R \times N_T$ channel matrix, M_R and N_T are the number of antennas at the transmitting and receiving sides, respectively. A and Φ represent the amplitude and phase of the channel impulse response, respectively. α and β represent the zenith angle and azimuth angle as shown in Figure 4, respectively. j is the square root of -1.

The process for generating channel coefficients is described as follows. Firstly, select the simulation scenario including the LOS or NLOS conditions and randomly generate the inclination angle of the antenna. Based on the obtained antenna inclination angles, the mean value of the amplitude of the channel coefficient for these angles can be obtained from (11) and (12). For simplicity, the variance is

considered to be a constant 4.25 dB for the LOS condition or 3.5 dB for the NLOS condition as shown in Figures 12 and 14. Consequently, the amplitude of the channel coefficient is randomly generated according to the statistical distribution. Furthermore, the phase factor can also be generated in a similar way. Finally, the channel coefficient can be obtained, and the flow chart of the process for generating channel coefficients is shown in Figure 18.

4. Numerical Analysis

As mentioned above, the changing actions for people holding their mobile devices result in random variation for the antenna inclination angles. In fact, it is not a static state but can be considered as a stochastic process since it may change frequently even in a very short time. However, very few studies have been performed on the statistical characteristics of the way people hold their mobile devices. Additionally, the effect of the stochastic antenna inclination angles on received signals has not been taken into account in conventional channel models.

In [22], an interesting study of the way that people naturally hold and interact with their mobile devices is performed and 1333 observations of people using mobile devices in different situations at different places in seven cities are made. The pie chart of the survey result is shown in Figure 19. It shows that talking in the way of voice calls occupies 22.28% of the users, while 18.98% is passive activities such as listening to audio or watching a video. Furthermore, the most common way of holding the mobile devices is the one-handed use accounting for 28.96% including 67% of right-handed use and 33% of left-handed use. The 21.01% of users is engaged in cradling their mobile devices in the left hands or right hands with 79% and 21%, respectively. However, only 8.78% is two-handed use with two different situations. The first one accounting for 90% is to hold the mobile devices vertically, i.e., in the portrait mode, and the other one is to hold the mobile devices horizontally, i.e., in the landscape mode.

Although the data in [22] were used to evaluate how people actually use their mobile phones, we can empirically associate these data with the antenna inclination angles as

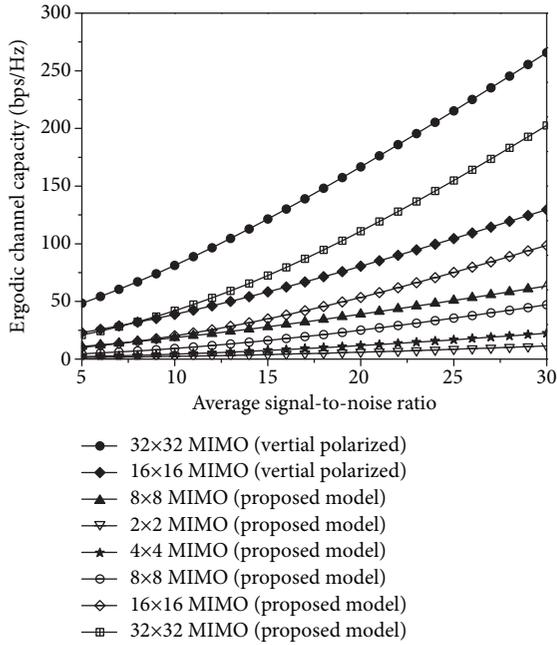


FIGURE 20: The ergodic channel capacity for the LOS condition.

shown in Table 1. Consequently, a semiempirical PDF is proposed based on the survey results. According to the statistical results, the antenna inclination angle obeys Gauss distribution with the corresponding mean and standard deviation of 45.5° and 9.44° , respectively.

After obtaining the statistical distribution of antenna inclination angles, the channel coefficients can be generated according to the method introduced in Section III. In order to evaluate the channel model, taking the MIMO system of $M_T = M_R = 4$ as an example, the ergodic capacity for MIMO channels can be obtained according to the formula [23]

$$C = E \left\{ \log_2 \det \left[I_{M_R} + \frac{\rho}{M_T} \mathbf{H}\mathbf{H}^H \right] \right\} \text{bps/Hz}, \quad (14)$$

where I_{M_R} is the $M_R \times M_R$ identity matrix; ρ is the system signal-to-noise ratio (SNR); $(\cdot)^H$ is the Hermitian transpose; M_R and M_T denote the number of receivers and transmitters, respectively; and \mathbf{H} is a matrix whose entries are the $M_R \times M_T$ channel gains.

The ergodic channel capacity of two different conditions of LOS and NLOS is depicted in Figures 20 and 21. The results show that the ergodic capacity increases with the increase of average SNR and it is larger in the LOS condition compared with the NLOS channel under the same SNR. Furthermore, the ergodic channel capacity increases with the increase of the number of antennas at the transmitting and the receiving sides, and for the same number of antennas, the channel capacity of the channel model with stochastic antenna inclination angles is considered to be smaller than that of the antenna fixed as a vertical polarization due to the effect of the polarization mismatch.

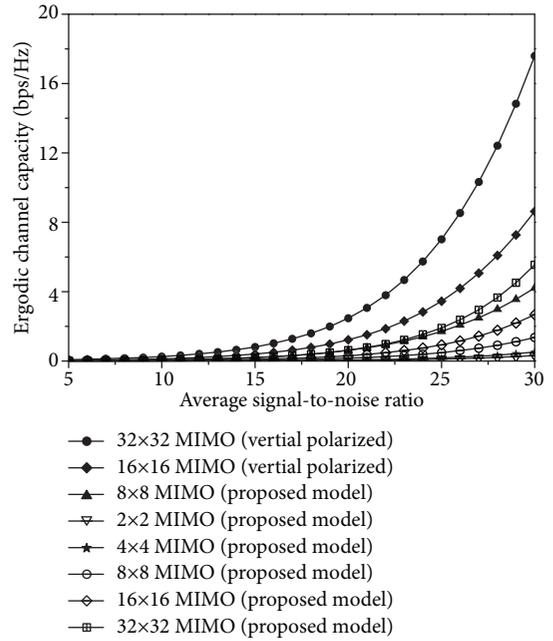


FIGURE 21: The ergodic channel capacity for the NLOS condition.

5. Conclusion

A 3D statistical channel model for stochastic antenna inclination angles in a T-shaped corridor environment based on a modified 3D fast ray-tracing method has been proposed in this paper. The radiation patterns for arbitrary inclination angles of a half-wave antenna have been deduced and considered in the method. Based on the statistical analyses, the relationships between the statistical characteristics of the electric field and antenna inclination angles are analyzed for both of the LOS and NLOS corridors. Furthermore, a semiempirical probability density function of antenna inclination angles has been proposed and used to depict the stochastic process of people holding their mobile devices. For future work, the channel model can serve as a preliminary attempt; the parameters used in the channel model can be verified and extracted from the channel measurements and extended to more scenarios. Combined with the existing channel models, a more general, accurate, and low-complexity channel model may be developed and applied to the next-generation wireless communication systems.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (NSFC) grant no. 61427801 and the Postgraduate Research & Practice Innovation Program of Jiangsu Province grant no. KYCX18_0861.

References

- [1] 3GPP, *3rd Generation Partnership Project; Technical Specification Group Radio Access Networks; Spatial Channel Model for MIMO Simulations (Release 6)*, Tech. Rep. 3GPP TR 25.996 V6.1.0, 2003.
- [2] IST-4-027756 WINNER II D111, *WINNER II Interim Channel Models, v1.0*, WINNER, 2006.
- [3] M. Pätzold and B. O. Hogstad, "A space-time channel simulator for MIMO channels based on the geometrical one-ring scattering model," *Wireless Communications and Mobile Computing*, vol. 4, no. 7, 737 pages, 2004.
- [4] H. Hofstetter, A. F. Molisch, and N. Czink, "A twin-cluster MIMO channel model," in *2006 First European Conference on Antennas and Propagation*, Nice, France, 2006.
- [5] M. Pätzold and B. O. Hogstad, "A wideband MIMO channel model derived from the geometric elliptical scattering model," in *2006 3rd International Symposium on Wireless Communication Systems*, Valencia, Spain, 2006.
- [6] R. He, G. L. Stüber, B. Ai, and Z. Zhong, "A two-cylinder based polarized MIMO channel model," in *2017 XXXII Ind General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, Montreal, QC, Canada, 2017.
- [7] Y. Yuan, X. Cheng, C. X. Wang, D. I. Laurenson, X. H. Ge, and F. Zhao, "Space-time correlation properties of a 3D two-sphere model for non-isotropic MIMO mobile-to-mobile channels," in *2010 IEEE Global Telecommunications Conference GLOBECOM*, Miami, FL, USA, 2010.
- [8] L. Bai, C. X. Wang, S. B. Wu, J. Sun, and W. S. Zhang, "A 3-D wideband multi-confocal ellipsoid model for wireless massive MIMO communication channels with uniform planar antenna array," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, Sydney, NSW, Australia, 2017.
- [9] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.
- [10] F. Rusek, D. Persson, Buon Kiong Lau, E. G. Larsson, T. L. Marzetta, and F. Tufvesson, "Scaling up MIMO: opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, 2013.
- [11] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, 2014.
- [12] S. Wu, C.-X. Wang, H. Haas, e.-H. M. Aggoune, M. M. Alwakeel, and B. Ai, "A non-stationary wideband channel model for massive MIMO communication systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1434–1446, 2015.
- [13] S. Wu, C.-X. Wang, e.-H. M. Aggoune, M. M. Alwakeel, and Y. He, "A non-stationary 3-D wideband twin-cluster model for 5G massive MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1207–1218, 2014.
- [14] K. Zheng, S. Ou, and X. Yin, "Massive MIMO channel models: a survey," *International Journal of Antennas and Propagation*, vol. 2014, Article ID 848071, 10 pages, 2014.
- [15] J. P. Kermaol, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, 2002.
- [16] W. Weichselberger, M. Herdin, H. Özcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Transactions on Wireless Communications*, vol. 5, no. 1, pp. 90–100, 2006.
- [17] X. Zhao, S. Geng, L. Vuokko, J. Kivinen, and P. Vainikainen, "Polarization behaviours at 2, 5 and 60 GHz for indoor mobile communications," *Wireless Personal Communications*, vol. 27, no. 2, pp. 99–115, 2003.
- [18] Z. Zhang, Z. Yun, and M. F. Iskander, "3D tetrahedron ray tracing algorithm," *Electronics Letters*, vol. 37, no. 6, pp. 334–335, 2001.
- [19] P. D. Holm, "A new heuristic UTD diffraction coefficient for nonperfectly conducting wedges," *IEEE Transactions on Antennas and Propagation*, vol. 48, no. 8, pp. 1211–1219, 2000.
- [20] S. Y. Seidel and T. S. Rappaport, "Site-specific propagation prediction for wireless in-building personal communication system design," *IEEE Transactions on Vehicular Technology*, vol. 43, no. 4, pp. 879–891, 1994.
- [21] G. Durgin, N. Patwari, and T. S. Rappaport, "An advanced 3D ray launching method for wireless propagation prediction," in *1997 IEEE 47th Vehicular Technology Conference. Technology in Motion*, pp. 785–789, Phoenix, AZ, USA, 2002.
- [22] S. Hooper, *How do people really hold and touch their mobile devices*, Online Forums, 2013.
- [23] G. J. Foschini and M. J. Gans, "On the limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

Research Article

Covariance Matrix Reconstruction for Direction Finding with Nested Arrays Using Iterative Reweighted Nuclear Norm Minimization

Weijie Tan  and Xi'an Feng

School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China

Correspondence should be addressed to Weijie Tan; tanweijie@hotmail.com

Received 30 July 2018; Accepted 10 December 2018; Published 18 March 2019

Guest Editor: Raed A. Abd-Alhameed

Copyright © 2019 Weijie Tan and Xi'an Feng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we address the direction finding problem in the background of unknown nonuniform noise with nested array. A novel gridless direction finding method is proposed via the low-rank covariance matrix approximation, which is based on a reweighted nuclear norm optimization. In the proposed method, we first eliminate the noise variance variable by linear transform and utilize the covariance fitting criteria to determine the regularization parameter for insuring robustness. And then we reconstruct the low-rank covariance matrix by iteratively reweighted nuclear norm optimization that imposes the nonconvex penalty. Finally, we exploit the search-free DoA estimation method to perform the parameter estimation. Numerical simulations are carried out to verify the effectiveness of the proposed method. Moreover, results indicate that the proposed method has more accurate DoA estimation in the nonuniform noise and off-grid cases compared with the state-of-the-art DoA estimation algorithm.

1. Introduction

Source localization is always a significant research direction in the past decades and today which is widely used in various fields including radar, sonar, wireless communication, and acoustics [1–4]. However, achieving superresolution is a great challenge for more sources with the least elements. Compared with the uniform linear arrays (ULAs), the sparse linear arrays (SLAs) can provide much more degree of freedoms (DOFs) and resolve more sources with the same number of physical elements. Meanwhile, the SLAs can efficiently reduce the cost and power consumption. Therefore, the study of SLAs has attracted more and more researchers' great attention [5, 6].

So far, there are some works about SLA configurations. The minimum redundancy arrays (MRAs) are class optimum lag SLAs [5], which provide a complete set of spatial lags between pairs of elements with minimum redundancies. However, it is not easy to attain the optimum design of

MRAs, especially for a large array; it is time-consuming to search the optimum structure. In the past decades, two potential array configurations, i.e., coprime arrays (CPA) and nested arrays (NA), have drawn the researchers' attentions since they have exactly closed-form expressions for sensor locations and it is easy to predict the attainable DOFs [7, 8]. Coprime arrays can resolve up to $o(MN)$ sources with only $M + N - 1$ elements; however, the CPAs do not generate a filled coarray since CPA has holes in the difference coarray [7]. So it cannot directly apply the augmentation techniques. By contrast, nested arrays can generate a filled difference coarray except the mentioned advantages [8]. Although the NAs and CPAs are not optimum lag arrays like MRAs, they are the most attractive array configuration because they are easy to build in the last decade.

The processing of SLA mainly has array interpolation (AI) [9–11] and the Toeplitz completion method [12–14]. The classical AI technique [10] can effectively obtain the element data of virtual ULAs, which imposes a linear

interpolation process on the element data of a real SLA and selects interpolator coefficients by minimizing the interpolation error for a source coming from a certain angular sector. The drawback is that this technique needs to know the angular sector. In AI techniques, Wiener array interpolation (WAI) [15] is a practically attractive method, which exploits a maximum likelihood method to estimate the power of signal and noise and use the calibration angles to recover the array steering matrix; hence, it can approximate the MSE optimum solution. However, WAI requires the initial DoA estimation. In the classical Toeplitz completion method, the direct augmentation approach (DAA) is a widely used method for improving the covariance matrix of the SLA [12]. DAA constructs a Toeplitz matrix using the sample covariance matrix; since there is one-to-one correspondence between covariance lags and spatial lags, the diagonal elements could be obtained by redundant averaging. Unfortunately, Toeplitz completion method does not guarantee positive semidefinite augmented covariance matrix. In order to construct a positive definite augmented matrix, the authors in [13, 14] proposed an iterative DAA algorithm; however, the complicated iteration procedure cannot guarantee the global convergence. Besides, the coarray MUSIC algorithm is used to estimated angle parameters [16] which utilizes the Toeplitz properties of matrices.

Sparse signal representation (SSR) framework has attracted a great interest in direction finding [17–23], which exploits the spatial sparsity of the signal arriving angles. Despite these aforementioned SSR-based methods have some attractive features, there are still two problems need to be considered. The first problem is nonuniform noise [23, 24]; many of SSR-based methods assume that the element noises are spatially uniform, and the reason for this is the uniform noise assumption benefits to choose an appropriate regularization parameter in a sparse representation model. However, in some practical applications, due to the nonidealistic of the practical arrays, such as the nonideality of the receiving channel, the nonuniformity of the element response, and the mutual coupling between elements, the uniform white noise assumption among all elements may not be satisfied. Therefore, the diagonal elements in noise covariance matrix should be considered as arbitrary values which represent the noise levels. Once the assumption misfits the true noise levels, the performance of conventional SSR-based DoA estimation approaches may thereby degrade severely. The second problem is the off-grid problem (also called basis mismatch) [17, 25]; in fact, a great deal of SSR-based DoA estimation methods assume that the true target directions exactly lie on the prespecified angular grid. In practical application, the angle parameters are continuous variables, the on-grid assumption rarely holds, and the basis mismatch always exists, which leads to the accuracy degradation of DoA estimation.

In this paper, we propose a gridless direction finding method under the unknown nonuniform noise over with nested array, which is based on low-rank covariance matrix approximation by an iteratively reweighted nuclear norm minimization. The advantages of the proposed method are in three aspects: Firstly, we mitigate the noise variance variable by linear transform to reduce the effect of

unknown nonuniform noise. Secondly, we utilize the covariance fitting criteria to determine the regularization parameter for the accuracy error fitting. Last but not least, we develop a novel reweighted nonconvex penalty objective function for exactly approximating the low-rank covariance matrix. After attaining the covariance matrix, we exploit the search-free DoA method to estimate parameters. Numerical simulation is carried out to verify the effectiveness of the proposed method.

The rest of this paper is organized as follows: In Section 2, we present the signal model with nested array. In Section 3, the DoA estimation with iteratively covariance matrix reconstruction (ICMR) is introduced to the nested array, which is based on an iteratively reweighted nuclear norm minimization algorithm. Section 4 validates the proposed method by numerical simulations and comparing the algorithm with state-of-the-art DoA estimation methods based on nested array. Section 5 concludes this paper.

Notation. We define a vector \mathbf{x} with a boldface lowercase and a matrix \mathbf{X} with a boldface uppercase. The symbol $(\cdot)^*$ denotes the complex conjugate of vector \mathbf{x} or matrix \mathbf{X} , and the transpose and the Hermitian transpose of vector or matrix are expressed by $(\cdot)^T$ and $(\cdot)^H$, respectively. $\mathbb{E}\{\cdot\}$ denote the statistical expectation, $\text{vec}(\cdot)$ denotes stacking element-wise in a column, $\text{rank}(\cdot)$ and $\|\cdot\|_*$ denote the rank and nuclear norm of the matrix, respectively, \otimes denotes the Kronecker product, \odot denotes the element-wise multiplication between two matrices with the same dimensional, $\text{diag}(\cdot)$ denotes the diagonal operator that takes a vector to a matrix with vector on the diagonal, sign is signum function, $\sigma(\mathbf{X})$ denote the eigenvalue of matrix \mathbf{X} , $\sigma_i(\mathbf{X})$ denote the i -th largest eigenvalue value, \mathbf{I}_k denotes a $K \times K$ dimensional identity matrix, $\|\cdot\|_2$ denotes the l_2 norm, and $\text{Pr}(\cdot)$ denotes probability density function.

2. Array Model for Nested Array

We consider a two-level nested array consisting of two ULA subarrays, where the first ULA subarray has N_1 elements with the interelement spacing $d_1 = d$, while the second ULA subarray has N_2 elements with the interelement spacing $(N_1 + 1)d_1$; the total number of element is $M = N_1 + N_2$, as shown in Figure 1.

Let the first element of the array as the phase reference point, the position set of element in ascending is denoted as

$$\mathcal{P} = \{p_1, \dots, p_M\}. \quad (1)$$

Assume that K uncorrelated narrow-band far-field sources from $\boldsymbol{\theta} = [\theta_1, \dots, \theta_K]^T$ impinge on the nested array. The array output at time t can be represented as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t), \quad (2)$$

where $\mathbf{x}(t) \in \mathbb{C}^{M \times 1}$ is the output of nested array. $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_K(t)]^T$ is the $K \times 1$ receive signal source vector, and $\mathbf{n}(t) \in \mathbb{C}^{M \times 1}$ denotes the additive white Gaussian

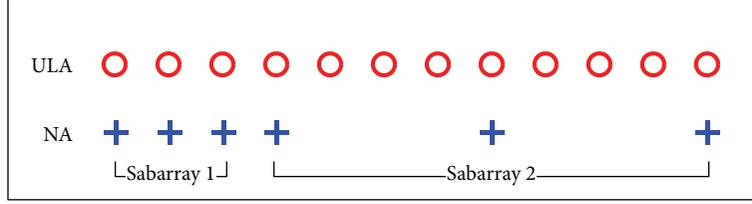


FIGURE 1: The nested array structure.

noise vector. The $M \times K$ array manifold matrix in array model (2) is defined by

$$\mathbf{A} = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_K)], \quad (3)$$

where $\mathbf{a}(\theta_k) = [e^{j\pi p_1 \sin(\theta_k)}, e^{j\pi p_2 \sin(\theta_k)}, \dots, e^{j\pi p_M \sin(\theta_k)}]$ is the array steering vector and the element position is defined in (1). We note that each position p_i is measured with half wavelength of indent signal source. Let the \tilde{M} denote the maximum interelement distance (i.e., the array aperture). Then \tilde{M} is an integral multiple of the half signal wavelength. Nested array possesses the same set of interelement distances as the ULA which composed of \tilde{M} elements separated by half signal wavelength.

From the above discussion, we can derive that nested array can be regarded as spatial sparse sampling on the ULA, i.e., we can take the position set \mathcal{P} of nested array from the position subset $\tilde{\mathcal{P}} = [0, 1, \dots, \tilde{M} - 1]$ of the elements of the ULA, as shown in Figure 1. Thus, nested array can be represented by its element index set $\mathcal{P} \subset \tilde{\mathcal{P}}$. For the k -th signal source, the steering vector $\mathbf{a}_{\tilde{\mathcal{P}}}(\theta_k)$ of the ULA can be given as $\mathbf{a}_{\tilde{\mathcal{P}}}(\theta_k) = [e^{j2\pi\tilde{\mathcal{P}}_1(d/\lambda) \sin(\theta_k)}, \dots, e^{j2\pi\tilde{\mathcal{P}}_{\tilde{M}}(d/\lambda) \sin(\theta_k)}]$.

According to the index set \mathcal{P} of the NA, we design a selection matrix $\Gamma \in \{0, 1\}^{M \times \tilde{M}}$ such that the i -th row element contains $\mathbf{0}_s$ except a single 1 at the $\tilde{\mathcal{P}}_i$ -th location. From the spatial sparse sampling viewpoint, we have $\mathbf{a}(\theta_k) = \Gamma \mathbf{a}_{\tilde{\mathcal{P}}}(\theta_k)$. Then, the nested array model can be expressed as

$$\mathbf{x}(t) = \mathbf{A}(\theta_k)\mathbf{s}(t) + \mathbf{n}(t) = \Gamma \mathbf{A}_{\tilde{\mathcal{P}}}\mathbf{s}(t) + \Gamma \mathbf{n}_{\tilde{\mathcal{P}}}(t). \quad (4)$$

Under ideal condition, the array covariance matrix of the output $\mathbf{x}(t)$ of nested array can be calculated as

$$\begin{aligned} \mathbf{R} &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{x}(t_n) \mathbf{x}^H(t_n) = \mathbf{A} \mathbf{R}_s \mathbf{A}^H + \text{diag}(\boldsymbol{\sigma}^2) \\ &= \Gamma \mathbf{A}_{\tilde{\mathcal{P}}} \mathbf{R}_s \mathbf{A}_{\tilde{\mathcal{P}}}^H \Gamma^T + \Gamma \text{diag}(\boldsymbol{\sigma}_{n, \tilde{\mathcal{P}}}^2) \Gamma^T \\ &= \Gamma \text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \Gamma^T + \Gamma \text{diag}(\boldsymbol{\sigma}_{n, \tilde{\mathcal{P}}}^2) \Gamma^T, \end{aligned} \quad (5)$$

where N is the number of collected snapshot, $\mathbf{R}_s = \text{diag}([\sigma_{s,1}^2, \dots, \sigma_{s,K}^2])$, and $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) = \mathbf{A}_{\tilde{\mathcal{P}}} \mathbf{R}_s \mathbf{A}_{\tilde{\mathcal{P}}}^H$ for some $\mathbf{u}_{\tilde{\mathcal{P}}} \in \mathbb{C}^{\tilde{M} \times 1}$. Furthermore, the noise-free covariance matrix $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$ is a Hermitian Toeplitz matrix, which is

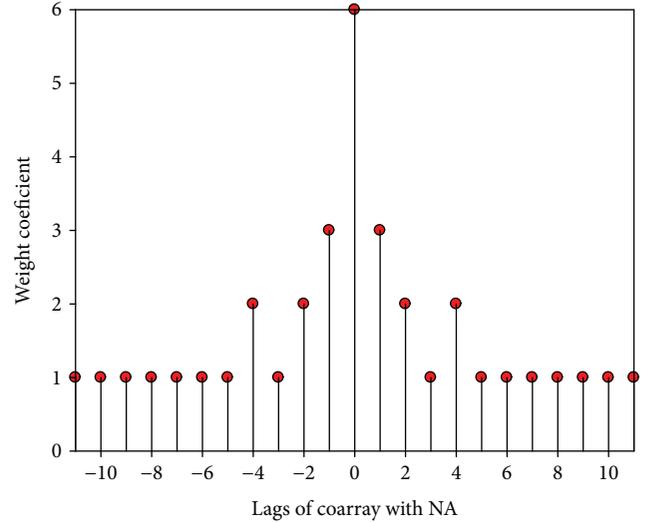


FIGURE 2: Weight coefficient of NA coarray with six elements.

determined by M complex numbers.

$$\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) = \begin{bmatrix} u_1 & u_2 & \cdots & u_{\tilde{M}} \\ u_2^* & u_1 & \cdots & u_{\tilde{M}-1} \\ \vdots & \vdots & \ddots & \vdots \\ u_{\tilde{M}}^* & u_{\tilde{M}-1}^* & \cdots & u_1 \end{bmatrix}. \quad (6)$$

It is evident that $\text{rank}(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})) = K \leq \tilde{M} - 1$, and $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$ is positive definite, i.e., $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \geq 0$. For nested array, $\text{Toep}(\mathbf{u})$ holds that

$$\begin{aligned} \text{Toep}(\mathbf{u})_{m,n} &= \sum_{k=1}^K \sigma_{s,k}^2 a_m(\theta_k) a_n^*(\theta_k) \\ &= \sum_{k=1}^K \sigma_{s,k}^2 e^{j2\pi d/\lambda (m-n) \sin(\theta_k)}. \end{aligned} \quad (7)$$

Let $\mathcal{D} = \{m - n : m, n \in \mathcal{P}, m \geq n\} \subset \tilde{\mathcal{P}}$. For the NA, \mathcal{D} is a redundancy array which can be regarded as a coarray defines on the NA. The matrix $\text{Toep}(\mathbf{u})$ contains all elements $\mathbf{u}_{\tilde{\mathcal{P}}}$ of explicitly by (6). It implies that the relation $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \rightarrow \text{Toep}(\mathbf{u})$ is one-to-one.

A coarray of nested array with six elements is shown in Figure 2. When the positions of nested array are $\mathcal{P} = \{0, 1, 2, 3, 6, 10\}$, the coarray of nested array composes of $\tilde{M} + 1 = 11$

elements with the position $\tilde{\mathcal{P}} = \{0, 1, \dots, 10\}$. From Figure 2, we can observe that the coarray is a redundancy array, which guaranteed the reconstruction of the covariance matrix [17].

One of the aims of this paper is to recover the noise-free covariance matrix $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$ of the nested array output under nonuniform noise environment. In the next section, we introduce the reweighted nuclear norm minimization to approximate the low-rank noise-free covariance matrix $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$.

3. Iterative Covariance Matrix Reconstruction for Direction Finding

3.1. Nuclear Norm Minimization Optimization. As mentioned above, the key problem is recovery of the noise-free covariance matrix $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$. On one hand, we note that the covariance matrix $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$ has three key features: Toeplitz-structured, low rank, and positive definite (PSD); on the other hand, in application, the sample covariance matrix of NA is calculated by $\hat{\mathbf{R}} = \sum_{n=1}^N \mathbf{x}(t_n) \mathbf{x}^H(t_n) / N$. Due to the finite samples, there exists the perturbation error term between the $\hat{\mathbf{R}}$ and \mathbf{R} , whose distribution benefits to choose the regularized parameter [17]. Making full use of the two constraints, we formulate the recovery problem of the covariance matrix as the following optimization problem:

$$\begin{aligned} \min_{\mathbf{u}, \sigma} \quad & \text{rank}(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})) \\ \text{s.t.} \quad & \left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \right\|_2^2 \leq \eta \\ & \text{Toep}(\mathbf{u}) \geq 0, \end{aligned} \quad (8)$$

where $\hat{\mathbf{C}} = (1/N) \mathbf{R}^T \otimes \mathbf{R}$ represents the prewhitening weighted matrix, $\Delta \mathbf{R} = \hat{\mathbf{R}} - \mathbf{R}$ is the perturbation error term, and η is an upper bound of $\left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \right\|_2^2$, which determines the quality of fitting term. According to the covariance fitting criteria [17, 18, 26], the vectorization form of the perturbation error term $\Delta \mathbf{R}$ satisfies the asymptotic normal (AsN) distribution as follows:

$$\text{vec}(\Delta \mathbf{R}) \sim \text{AsN} \left(\mathbf{0}_{M^2, 1}, \frac{1}{N} \mathbf{R}^T \otimes \mathbf{R} \right). \quad (9)$$

After prewhitening processing, we obtain

$$\hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \sim \text{AsN}(0, \mathbf{I}_{M^2}). \quad (10)$$

We can further derive the following formulation:

$$\left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \right\|_2^2 \sim \text{As}\chi^2(M^2), \quad (11)$$

where $\text{As}\chi^2(M^2)$ represents the asymptotic chi-square distribution with M^2 degrees of freedom. According to (11), it is quite obvious how to choose the parameter η . We choose the parameter η large enough such that the inequality $\left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \right\|_2^2 \leq \eta$ is satisfied with a high probability $1 - p$,

e.g., 0.999, $(\Pr(\left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \right\|_2^2 \leq \eta) = 1 - p)$, which is calculated by the matlab function $\text{chi2inv}(1 - p, M^2)$, so we have $\Pr\{\chi_{M^2}^2 \leq \eta\} = p$ and $\eta = \chi_p^2\{M^2\}$.

Unfortunately, this problem in (8) is nonconvex optimization problem and it is also a NP-hard problem. To sidestep the nonconvexity, we exploit convex relaxation technique, i.e., we relax the rank norm of $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$ to the nuclear norm of $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$ for approximating the low-rank matrix. Therefore, the convex relaxation form for (8) can be represented as

$$\begin{aligned} \min_{\mathbf{u}, \sigma} \quad & \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})\|_* \\ \text{s.t.} \quad & \left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \right\|_2^2 \leq \eta \\ & \text{Toep}(\mathbf{u}) \geq 0. \end{aligned} \quad (12)$$

Combining (5) and (12), the optimization problem is rewritten as

$$\begin{aligned} \min_{\mathbf{u}, \sigma} \quad & \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})\|_* \\ \text{s.t.} \quad & \left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\hat{\mathbf{R}} - \Gamma(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) + \text{diag}(\sigma_n^2))\Gamma^T) \right\|_2^2 \leq \eta \\ & \text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \geq 0. \end{aligned} \quad (13)$$

Remark 1. It is worth noting that there is noise variance variable that needs to be optimized in model (13). In fact, we can eliminate this variable by estimating the noise variance in advance, but the estimation of the nonuniform noise variance is not easy to attain directly, especially in the underdetermined DoA estimation circumstances while the sample covariance matrix no longer contains a low-rank component spanning the noise subspace. We can use two methods to handle this case, the first method is that we treat noise variance as a variable to be optimized in the optimal problem such as (13). The second method, we get rid of noise variables by linear transform firstly and then solve the optimization problem; in this paper, we focused on the latter.

It is interesting to observe that the noise term has M nonzero elements which contain in the diagonal element of the covariance matrix and correspond to the unknown noise variance $\{\sigma_n^2\}$. Based on this observation, we carry out the linear transform to remove the element of the noise variance terms. Mathematically, this operation can be formulated as

$$\begin{aligned} \text{Jvec}(\Delta \mathbf{R}) &= \text{Jvec}(\hat{\mathbf{R}} - (\Gamma \text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \Gamma^H + \Gamma \text{diag}(\sigma_n^2) \Gamma^H)) \\ &= \text{Jvec}(\hat{\mathbf{R}}) - \text{Jvec}(\Gamma \text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \Gamma^H), \end{aligned} \quad (14)$$

where \mathbf{J} is an $M(M-1) \times M^2$ linear transform selecting matrix and can be represented as

$$\mathbf{J} = [\mathbf{J}_1, \mathbf{J}_2, \dots, \mathbf{J}_{M-1}]^T, \quad (15)$$

where

$$\mathbf{J}_m = \left[\mathbf{e}_{(m-1)(M-1)+2}, \mathbf{e}_{(m-1)(M-1)+3}, \dots, \mathbf{e}_{m(M-1)} \right], \quad (16)$$

$$m = 1, 2, \dots, M-1,$$

with $\mathbf{e}_i (i = (m-1)(M-1) + 2, \dots, m(M-1))$ being an $M^2 \times 1$ column vector with one at the i th position and zeros elsewhere. $\mathbf{J}_m \in \mathbb{R}^{M^2 \times M}$. By applying the above operation, the unknown noise variance is eliminated effectively and easily. This elimination operation avoids the estimation of noise variance, therefore facilitating a noise-free low-rank covariance matrix reconstruction.

Through the above derivation, (13) can further reformatted as

$$\begin{aligned} \min_{\mathbf{u}} \quad & \|\text{Toep}_{\tilde{\varphi}}(\mathbf{u})\|_* \\ \text{s.t.} \quad & \left\| \tilde{\mathbf{C}}^{-1/2} (\text{Jvec}(\hat{\mathbf{R}}) - \text{Jvec}(\mathbf{\Gamma} \text{Toep}_{\tilde{\varphi}}(\mathbf{u}) \mathbf{\Gamma}^H)) \right\|_2^2 \leq \hat{\eta} \\ & \text{Toep}_{\tilde{\varphi}}(\mathbf{u}) \geq 0. \end{aligned} \quad (17)$$

From (17), we observe that there is only one vector \mathbf{u} that needs to be optimized by the above processing. At the same time, the prewhitening matrix and regularization parameter should be changed. Combining the distribution of the perturbation error term with (14) yields

$$\text{Jvec}(\Delta \mathbf{R}) \sim \text{AsN} \left(\mathbf{0}_{M(M-1),1}, \frac{1}{N} \mathbf{J}(\mathbf{R}^T \otimes \mathbf{R}) \mathbf{J}^T \right). \quad (18)$$

Applying the prewhitening processing [19], we have

$$\tilde{\mathbf{C}}^{-1/2} \text{Jvec}(\Delta \mathbf{R}) \sim \text{AsN} \left(\mathbf{0}_{M(M-1),1}, \mathbf{I}_{M(M-1)} \right), \quad (19)$$

where the prewhitening matrix $\tilde{\mathbf{C}}^{-1/2}$ is Hermitian square root of $\tilde{\mathbf{C}}^{-1}$, $\tilde{\mathbf{C}} = \mathbf{J}(\mathbf{R}^T \otimes \mathbf{R}) \mathbf{J}^T / N$, and $\mathbf{I}_{M(M-1)}$ is an $M(M-1) \times M(M-1)$ identity matrix. Like the choosing parameter $\hat{\eta}$, the parameter $\tilde{\eta}$ can be determined by $\tilde{\eta} = \chi_p^2 \{M(M-1)\}$.

3.2. Iteratively Reweighted Nuclear Norm Minimization. We utilize the nuclear norm $\|\text{Toep}_{\tilde{\varphi}}(\mathbf{u})\|_* = \sum_i \sigma_i(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})) = \sum_i g_1(\sigma_i(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})))$ in (12) instead of the rank norm $\text{rank}(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})) = \|\sigma(\text{Toep}_{\tilde{\varphi}}(\mathbf{u}))\|_0 = \sum_i g_0(\sigma_i(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})))$ in (8), where $g_0(x) = \text{sign}(x)$, $x \geq 0$, $g_1(x) = x$, and $x \geq 0$. The advantage is effective computable and the best convex approximation, but the drawback is the worst fitting for rank approximation. We can observe that there is a large gap between the rank and nuclear norm for robust and exact reconstruction of low-rank covariance matrix. Therefore, it is

essential to find a surrogate function of rank function which not only have the better approximation but also needs a low computational complexity. Some methods utilize a class of nonconvex surrogate functions to approximate the rank function, such as the works in [27, 28].

Among the nonconvex approximation functions, we find $f(x) = 1 - e^{-x/\epsilon}$ closely matches $g_0(x)$, so we use this function $h^\epsilon(x) = f(x) = 1 - e^{-x/\epsilon}$ as the approximation function of $g_0(x)$, i.e., rank norm. Combine the model (17), the following general nonconvex optimization model is proposed:

$$\begin{aligned} \min_{\mathbf{u}} \quad & h^\epsilon(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})) \\ \text{s.t.} \quad & \left\| \tilde{\mathbf{C}}^{-1/2} (\text{Jvec}(\hat{\mathbf{R}}) - \text{Jvec}(\mathbf{\Gamma} \text{Toep}_{\tilde{\varphi}}(\mathbf{u}) \mathbf{\Gamma}^H)) \right\|_2^2 \leq \tilde{\eta} \\ & \text{Toep}_{\tilde{\varphi}}(\mathbf{u}) \geq 0, \end{aligned} \quad (20)$$

where $h^\epsilon(\sigma(\text{Toep}_{\tilde{\varphi}}(\mathbf{u}))) = \sum_i f(\sigma_i(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})))$ denotes the rank approximating function. We expect that the proposed nonconvex optimization model can bridge the gap between the rank norm model (8) and nuclear norm (12).

Since the model in (20) is nonconvex optimization problem, how to guarantee for convergence to a global minimum is a key problem. Now, we propose an iteratively reweighted nuclear norm minimization to obtain a suboptimal solution, which is based on the framework of the majorization-maximization (MM) method. The MM method is an iterative approach, it iteratively solves a sequence of weighted convex surrogate optimization problems whose weights are attained by the previous optimal value. Thus, by calculating the sequence of convex problems given in (22), we can minimize the optimization function in (20). For the k -th iteration, the weights \mathbf{u} are calculated by the previous optimal variable \mathbf{u}_{k-1} . Since $h^\epsilon(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})) = \sum_i f(\sigma_i(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})))$ is concave when $f(x) = g_1(x)$, we have

$$\begin{aligned} h^\epsilon(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})) & \leq h^\epsilon(\text{Toep}_{\tilde{\varphi}}(\mathbf{u}_k)) \\ & \quad + \|\nabla h^\epsilon[\text{Toep}_{\tilde{\varphi}}(\mathbf{u})] \text{Toep}_{\tilde{\varphi}}(\mathbf{u} - \mathbf{u}_k)\|_*. \end{aligned} \quad (21)$$

Therefore, the optimization problem for the k -th iteration becomes

$$\begin{aligned} \min_{\mathbf{u}} \quad & \|\mathbf{W}_k^\epsilon \text{Toep}_{\tilde{\varphi}}(\mathbf{u})\|_* \\ \text{s.t.} \quad & \left\| \tilde{\mathbf{C}}^{-1/2} (\text{Jvec}(\hat{\mathbf{R}}) - \text{Jvec}(\mathbf{\Gamma} \text{Toep}_{\tilde{\varphi}}(\mathbf{u}) \mathbf{\Gamma}^H)) \right\|_2^2 \leq \tilde{\eta} \\ & \text{Toep}_{\tilde{\varphi}}(\mathbf{u}) \geq 0, \end{aligned} \quad (22)$$

where $\mathbf{W}_k^\epsilon \triangleq \nabla h^\epsilon(\text{Toep}_{\tilde{\varphi}}(\mathbf{u}_k))$. In order to calculate \mathbf{W}_k^ϵ , we start by introducing a key useful proposition as follows.

Proposition 1. Suppose that $h^\epsilon(\sigma(\mathbf{Z}))$ is expressed as $h^\epsilon(\sigma(\mathbf{Z})) = \sum_i f(\sigma_i(\mathbf{Z}))$, where \mathbf{Z} is positive semidefinite

```

1: Input:  $\mathbf{X} \in \mathbb{C}^{M \times K}$ .
2: Initialize:  $k = 0$ ,  $\mathbf{u}_0 = \mathbf{0}_{M \times 1}$ ,  $\epsilon$ ,  $\gamma$ ,  $\epsilon = 10^{-4}$ ,  $\text{Iter}_{\max}$ 
3: Calculate the sample covariance matrix  $\hat{\mathbf{R}}_S$  of NA using
   the observe snapshots  $\mathbf{x}_\Omega(t)$ ;
4: while  $k \leq \text{Iter}_{\max}$  do
5:   Update the weight matrix  $\mathbf{W}_k^\epsilon$  by the equation
6:   Update  $\mathbf{u}_{k+1}$  by the optimization problem (22), (24);
7:   if the convergence condition  $|\mathbf{u}_{k+1} - \mathbf{u}_k| \leq \epsilon$  is satisfied
8:     then
9:       break
10:    else
11:       $\epsilon = \epsilon/\gamma$ ,  $k = k + 1$ ;
12:    end if
13: end while
14: Calculate the covariance matrix  $\hat{\mathbf{R}} = \text{Toep}(\hat{\mathbf{u}})$ ;
15: Using search-free DoA estimation method to estimate the
   angle parameters.

```

ALGORITHM 1: Gridless DoA estimation using iteratively reweighted covariance matrix reconstruction.

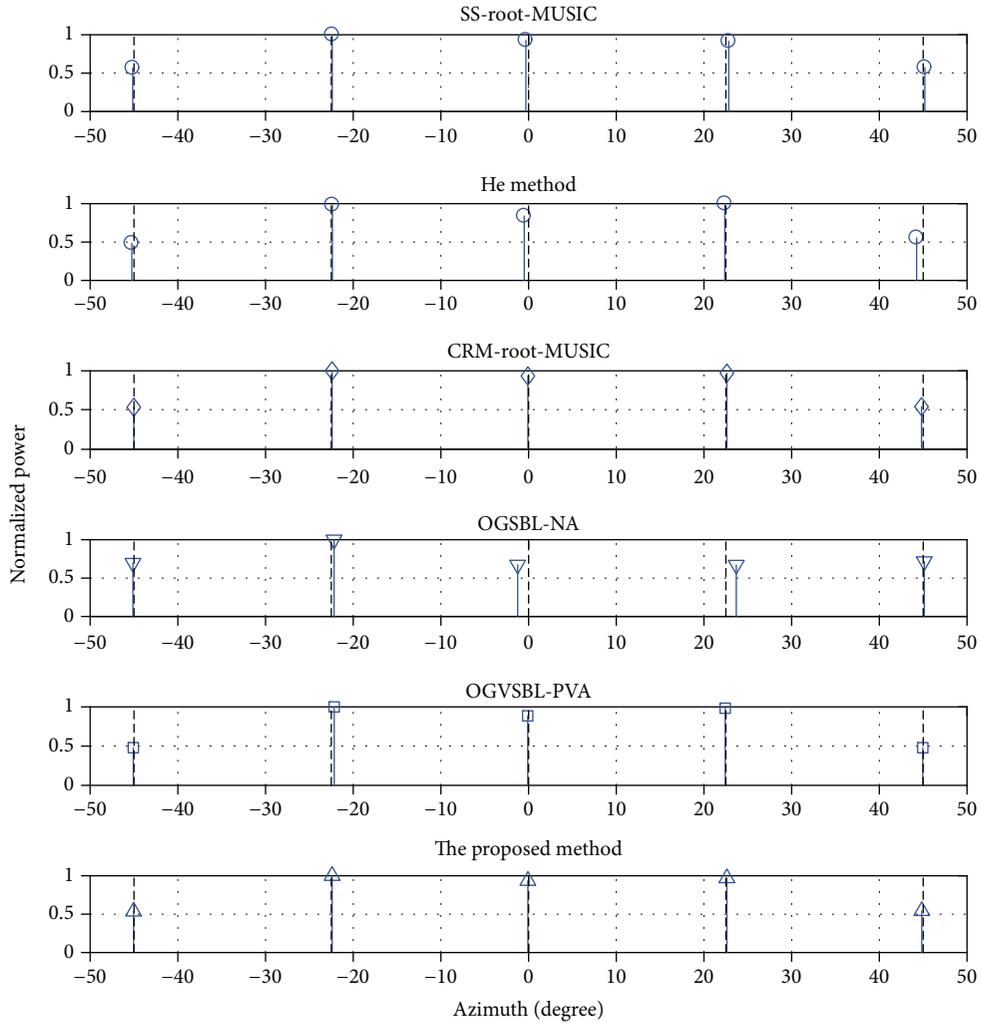


FIGURE 3: Normalized spatial spectra in nonuniform noise case for $M = 6$, $K = 5$, $N = 500$, and $\text{SNR} = 0$ dB.

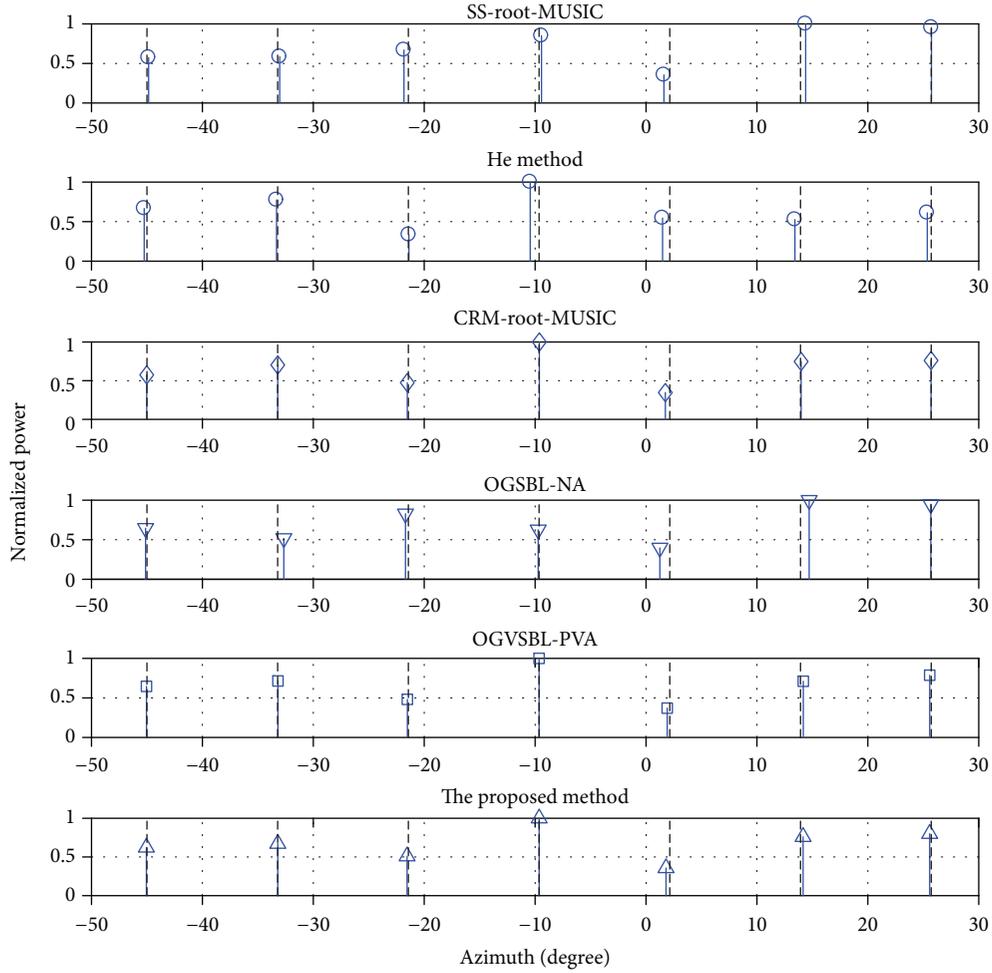


FIGURE 4: Normalized spatial power in nonuniform noise case for $M = 6$, $K = 7$, $N = 500$, and $\text{SNR} = 0$ dB.

matrices with the eigenvalue decomposition (EVD) $\mathbf{Z} = \mathbf{U} \text{diag}(\boldsymbol{\sigma}(\mathbf{Z})) \mathbf{U}^H$, $\mathbf{Z} \geq 0$, and h^ϵ and f are concave and differentiable, then the gradient of $f(\mathbf{Z})$ at \mathbf{Z} is calculated as

$$\nabla h^\epsilon(\mathbf{Z}) = \mathbf{U} \text{diag}(\boldsymbol{\eta}(\mathbf{Z})) \mathbf{U}^H, \quad (23)$$

where $\boldsymbol{\eta}(\mathbf{Z}) = \nabla h^\epsilon(\boldsymbol{\sigma}(\mathbf{Z}))$ denotes the gradient of h^ϵ at $\boldsymbol{\sigma}(\mathbf{Z})$.

In (22), $\mathbf{Z} = \text{Toep}_{\hat{\varphi}}(\mathbf{u}_k)$, for the sake of simplicity, we denote $\boldsymbol{\sigma}(\text{Toep}_{\hat{\varphi}}(\mathbf{u}_k))$ as $\boldsymbol{\sigma}_k$ and apply the result in [27]; we obtain

$$\mathbf{W}_k^\epsilon = \mathbf{U}_k \text{diag}(\nabla h^\epsilon(\boldsymbol{\sigma}_k)) \mathbf{U}_k^H, \quad (24)$$

where $\mathbf{U}_k \text{diag}(\boldsymbol{\sigma}_k) \mathbf{U}_k^H$ is the EVD of $\text{Toep}_{\hat{\varphi}}(\mathbf{u}_k)$.

We observe that ϵ is a key parameter, which controls the closeness between $f(\boldsymbol{\sigma})$ and $g_0(\boldsymbol{\sigma})$ or $g_1(\boldsymbol{\sigma})$. In particular, $f(\boldsymbol{\sigma})$ is close to $g_0(\boldsymbol{\sigma})$ when ϵ is small, but it is subjected to many local minima. On the other hand, a large value of ϵ causes $f(\boldsymbol{\sigma})$ toward the convex $g_1(\boldsymbol{\sigma})$, whereas with poor approximation quality to $g_0(\boldsymbol{\sigma})$. In order to avoid getting

local minima, the algorithm starts with a large value of ϵ , and then the value ϵ decreases gradually during the iteration.

It is worth mentioning that this strategy is adopted in model (22). We define the weight matrix \mathbf{W}_k^ϵ in each iteration, which uses the solution of the previous iteration for avoiding local minima. The convergence analysis of the proposed method for (20) is given as the following theorem, which is proven in the appendix.

Theorem 1. Denote \mathbf{u}_k as the k -th iteration solution in (22), after that, the optimal variable \mathbf{u}_k satisfies the following properties:

- (i) $h^\epsilon[\text{Toep}_{\hat{\varphi}}(\mathbf{u}_k)]$ is monotonically decreasing and bounded as k tends to ∞
- (ii) The optimal variable \mathbf{u}_k converges to a local minimum of (20)

3.3. Steps of the Proposed Method. To sum up, the steps of the proposed method are presented as follows. Once the solution \mathbf{u} of model (20) is obtained, we can determine the number of

sources $\hat{\mathbf{K}}$ by the rank of $\text{Toep}_{\hat{\mathcal{P}}}(\mathbf{u})$. Given $\text{Toep}_{\hat{\mathcal{P}}}(\hat{\mathbf{u}})$, the remaining task is to estimate the directions of the signal $\hat{\boldsymbol{\theta}}$. According to classical Vandermonde decomposition lemma [28], we can estimate the DoAs using positive semidefinite Toeplitz matrices $\text{Toep}_{\hat{\mathcal{P}}}(\mathbf{u})$. So we can apply the grid-free DoA estimation method, such as Prony's method to efficiently estimate $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\sigma}}_s^2$. Firstly, by solving a linear system involving only $\hat{\mathbf{u}}$, we can obtain $\hat{\boldsymbol{\theta}}$ from zeros of a polynomial. Then, $\hat{\boldsymbol{\sigma}}_s^2$ is solved by a polynomial rooting-based estimation method.

In the proposed method, firstly, it recovers the complete noiseless matrix $\text{Toep}_{\hat{\mathcal{P}}}(\mathbf{u})$ in advance under the nonuniform noise background, which is attained by solving an iteratively weighted nuclear norm minimization optimization. Compare with the ULAs which have the same number of physical elements, the proposed method forms a larger aperture ULA and gets more degree of freedoms, so it has the potential ability to estimate the DoA in the underdetermined case. Moreover, exploiting the grid-free DoA estimation method to estimate the DoAs, which do not need to discretize the angle grid, avoids the grid mismatch in the grid-based sparse method.

4. Simulation Results

In this section, we evaluate the DoA estimation performance of the proposed method and compare the proposed method with the others: the state-of-the-art method, spatial smoothing root MUSIC (SS-root-MUSIC) [7], covariance matrix reconstruction approach with root MUSIC (CMR-root-MUSIC) [22], He et al.'s method [23], OGSBL-NA [25], and OGVSBL-PVA [24]. To provide a benchmark for evaluating the effectiveness of these methods, the Cramer-Rao lower bound (CRLB) for nonuniform noise given in [23] is also included though it is a lower bound for unbiased estimators.

In the following simulations, we consider a two-level nested array with optimal configuration [8], whose total element number is six. The first ULA subarray contains three elements with interelement spacing $d_1 = d$; the second ULA subarray contains three elements with spacing $d_2 = 4d_1$, where d is equal to half wavelength. We assume that all the signals are far-field equal power uncorrelated signals; the additive noise is spatially nonuniform which is modeled as an independent complex Gaussian random vector with zero mean, and its covariance matrix is $\mathbf{R}_n = \text{diag}\{\boldsymbol{\sigma}_n^2\}$, since the input signal-to-noise ratio (SNR) can be defined as

$$\text{SNR} = 10 \log_{10} \left[\frac{1}{M} \sum_{m=1}^M \frac{\sigma_s^2}{\sigma_{n,m}^2} \right], \quad (25)$$

where $M = N_1 + N_2$ is the number of NA elements, σ_s^2 denotes the signal source power, and $\sigma_{n,m}^2$ denotes the noise power in difference elements. For the performance metric of DoA estimation, the root-mean-square error (RMSE) is

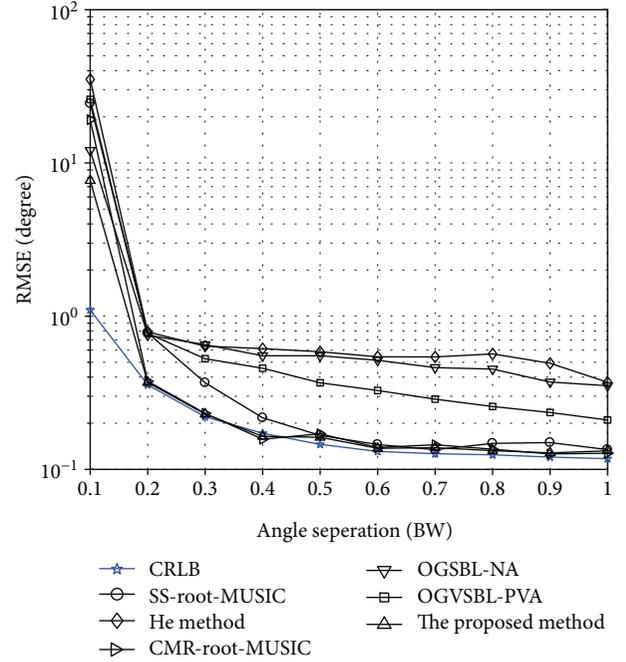


FIGURE 5: RMSEs of DoA estimation versus angle separation for $M = 6$, $N = 500$, and $\text{SNR} = 0$ dB.

selected for evaluating the estimation accuracy of the DoAs, which is defined as

$$\text{RMSE} = \left[\frac{1}{LK} \sum_{k=1}^K \sum_{l=1}^L \left(\theta_k^r - \hat{\theta}_{k,l} \right)^2 \right]^{1/2}, \quad (26)$$

where $\hat{\theta}_{k,l}$ denote the estimated θ_k^r of the k -th signal source from the l -th Monte Carlo trial.

In the first experiment, we investigate the resolution ability under the overdetermined case with the nested array geometries, respectively. We consider $K = 5$ spatial uncorrelated sources which are uniformly distributed within -45° to 45° . The SNR is set to 0 dB, and 500 snapshots are collected. The nonuniform noise is $\boldsymbol{\sigma}_n^2 = [10, 5.6, 8.5, 4.2, 9.5, 0.5]$. In He et al.'s method, we set the grid size is 0.5° . Figure 3 shows the normalized power of the five targets for the nested array, respectively. We use the dashed lines to represent the true DoAs. It is observed that the proposed method can exactly resolve all targets among all methods, and it has the most superior estimation performance since the DoA estimations are closest to the true target directions. The reason is that the proposed method makes the best of the features of covariance matrix, i.e., low rank, Toeplitz structure, and PSD, and approach the true rank by the iterative weighted idea. The OGVSBL-PVA and the CMR-root-MUSIC are a little worse than the proposed method, so we confirmed that the proposed method achieved higher DoA estimation accuracy.

In the second experiment, we investigate the detectable capability of all methods for the underdetermined case. We consider $K = 7$ narrowband uncorrelated sources whose number is greater than the number of physical elements,

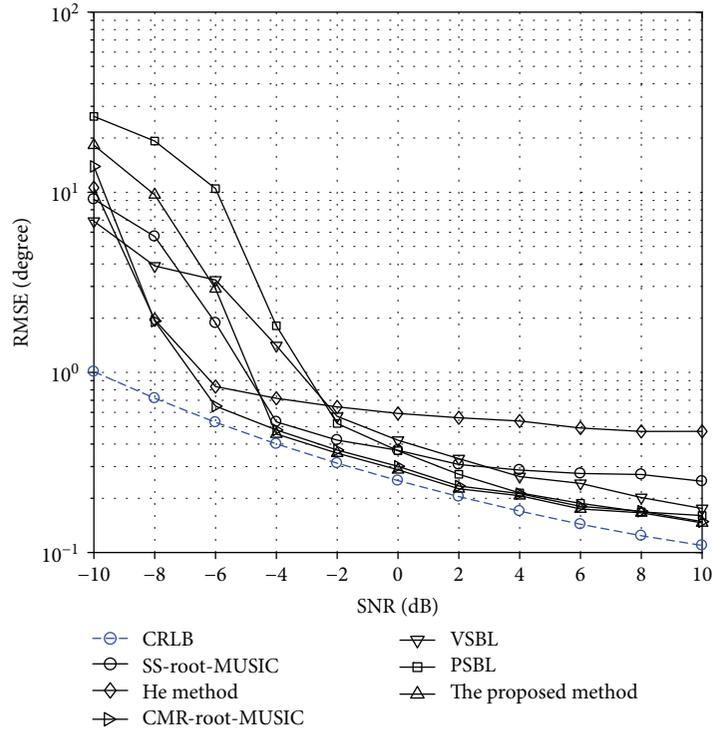


FIGURE 6: RMSEs of DoA estimation versus SNR for $M = 6$, $K = 7$, and $N = 500$.

the signal source directions are uniformly distributed within the range from $-\pi/4$ to $-\pi/7$, and the nonuniform noise power is assumed that $\sigma_n^2 = [10, 5.6, 8.5, 4.2, 9.5, 0.5]$; the SNR is set to 0 dB, and the number of snapshots is 500. Figure 4 displays the normalized power of all methods, respectively. It is observed that the OGSBL-PVA method and the proposed method can detect all the signal sources exactly. He et al.'s method has a certain deviation; this method uses the denoising operation and covariance matrix fitting to estimate the DoAs, which is based on the on-grid model. From Figure 4, it is easy to observe that the DoA estimation of the proposed method is closest to the true DoAs among all the methods; the reason is that the NA array provides more DOFs than the number of physical elements, and the low-rank covariance matrix approximation is more exact than the covariance matrix reconstruction method by exploiting the iteratively reweighted nuclear norm minimization; moreover, the proposed method is not limited by the grid size like the traditional grid-based sparsity method. Thus, the proposed methods have the superior estimation performance in all methods.

In the third experiment, in order to further illustrate the superior performance of the proposed method, the RMSE of DoA estimation performance versus the angular separation between the source is examined for all methods. In the experiment, the DoA of the first source is fixed and equal to -5.3° , whereas the DoA of the second source is varied and the varied step size is equal to 0.1 times beam width. When the angle separation is larger than 0.2 times beam width, the performance of the proposed method and

CMR-root-MUSIC is achieving the CRLB; since the covariance matrix reconstruction method is based on the low-rank property, it circumvents the traditional grid search, which is the gridless method. The OGSBL-NA method and OGVSBL-PVA method also have the better DoA performance due to the linear approximation manifold vector. As shown in Figure 5, compared with other methods, the performance of the proposed method demonstrates the better performance even when the target is very close.

In the following experiments, unless otherwise specified, the element number in nested array is set to six; seven uncorrelated equal power signals are considered which are uniformly distributed between $-\pi/4$ and $-\pi/7$. In this case, when grid size is set to 0.5° , for He et al.'s method, the DoAs do not fall into the predefined angular grid, i.e., grid mismatch. The nonuniform noise power is assumed that $\sigma_n^2 = [10, 5.6, 8.5, 4.2, 9.5, 0.5]$. All simulation results were obtained by averaging over 500 independent Monte Carlo trials.

In the fourth experiment, we examine the RMSE of DoA estimation performance with regard to the SNR for different methods. Figure 6 shows the RMSE of DoA estimation performance as a function of SNR when the number of snapshots is 500. Figure 7 shows that the RMSE versus the number of snapshots with the SNR is fixed to 0 dB. As we can see in Figures 6 and 7, the DoA estimation performance of the proposed method is improved with the SNR and increasing snapshot number, since the covariance matrix approximation is more accurate with larger SNR and more snapshots. The proposed method and the CMR-root-MUSIC outperform OGSBL-NA,

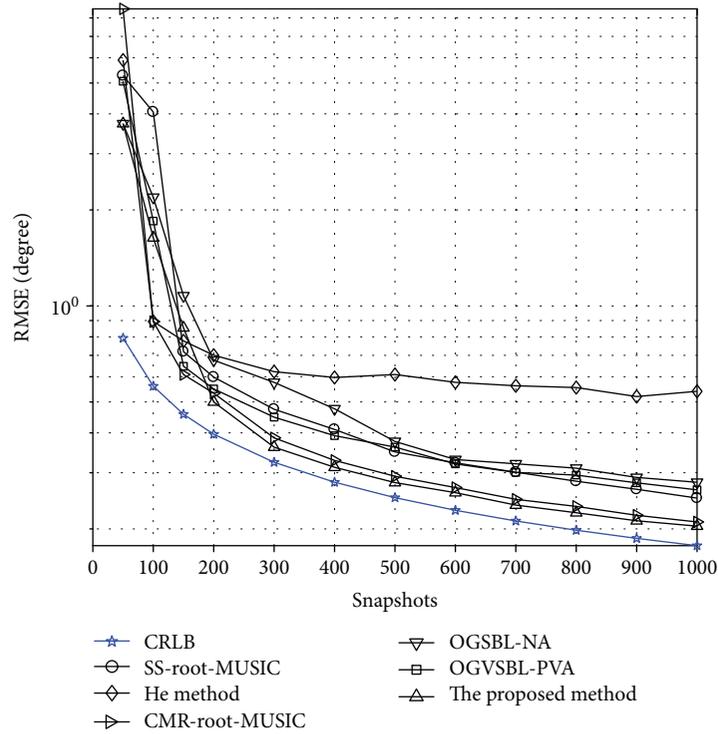


FIGURE 7: RMSEs of DoA estimation versus snapshot number for $M = 6$, $K = 7$, and $SNR = 0$ dB.

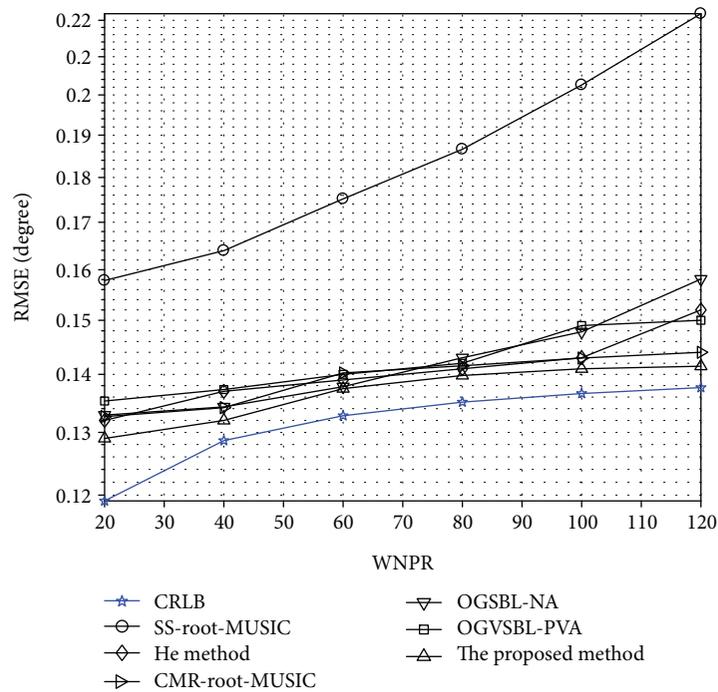


FIGURE 8: RMSEs of DoA estimation versus WNPR for $M = 6$, $N = 500$, and $SNR = 0$ dB.

OGVSBL-PVA, and the other methods. CMR-root-MUSIC method and the proposed method provide almost similar estimation performance, while the proposed method is slightly better than CMR-root-MUSIC. Since the proposed method not only exploits the full noiseless covariance matrix but also

utilizes the iteratively reweighted strategy to enhance the sparsity of matrix rank resulting in better covariance matrix approximation, we can see that in most of the cases He et al.'s method showcases the worst performance with the largest error; the reason is that He et al.'s method only utilizes

the part of the noiseless covariance matrix due to the denoising operation and when the DoAs do not fall into the predefined angular grid, the performance will degrade. OGSBL-NA method and OGVsBL-PVA method both improve the DoA performance by linear approximation manifold vector, but the precise depends on the order of the Taylor expansion. It needs to point that the proposed method considerably enhances the sparsity of the covariance matrix due to the iteratively weighted nuclear norm strategy; moreover, the search-free method is free of grid restriction. Thus, we can see that the proposed method has the best performance of DoA estimation accuracy in all tested methods.

In the last experiment, we measure the DoA estimation performance in the nonuniform noise environment. We define the worst noise power ratio (WNPR) $WNPR = \sigma_{n_{\max}}^2 / \sigma_{n_{\min}}^2$ [29], where $\sigma_{n_{\max}}^2$ and $\sigma_{n_{\min}}^2$ denote the maximum and the minimum noise power, respectively. The RMSE of DoA estimation performance for all methods is shown in Figure 8 as a function of the WNPR. The SNR is fixed to 0 dB. We vary the WNPR from 20 to 120 by changing the noise power of the first element from 10 to 60. It is observed that the proposed method performs better than the other methods at all WNPR levels; this can be explained by the fact that using denoising operation mitigates the influence of the nonuniform noise and enhances the robustness of the proposed method; furthermore, by iteratively reweighted idea, the proposed method can exactly approximate the rank of covariance matrix, and search-free DoA method is not subject to the grid constraint.

5. Conclusion

In this paper, we addressed the direction finding problem in the presence of unknown nonuniform noise over nested array. We proposed a gridless DoA estimation method based on the low-rank covariance matrix approximation. In order to mitigate the effect of nonuniform noise, we eliminated the nonuniform noise variables on the diagonal of the covariance matrix by linear transform. In order to recover the noise-free covariance matrix, we proposed an iterative reweighted non-convex optimization method, whose regularization parameter is determined by the covariance fitting criteria. Finally, we exploited the search-free DoA estimation method to estimate the exact DoA parameters. Numerical simulations were carried out to verify the effectiveness of the proposed method, compared with the state-of-the-art methods, which achieves more accurate DoA estimation performance for the nonuniform noise environment in the presence of grid bias.

Appendix

Proof. Assume $h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})) = \sum_i f(\sigma_i(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})))$. If function f is twice differentiable, strictly concave, for any bounded k and $0 \leq x \leq k$, $f'' \leq -m'_k < 0$, and then $h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}))$ is strictly concave. For any bounded $\mathbf{u}, \mathbf{v} (\mathbf{u} \neq \mathbf{v})$, there exists some $\lambda > 0$ such that

$$\begin{aligned} & h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})) - h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{v})) \\ & \leq \langle \text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u} - \mathbf{v}), \nabla f^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{v})) \rangle - \frac{\lambda}{2} \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u} - \mathbf{v})\|_F^2. \end{aligned} \quad (\text{A.1})$$

We first show that $h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k))$ is convergent. In terms of the concavity of $h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k))$ and classical Vandermonde decomposition lemma [28], we have

$$\begin{aligned} & h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_{k+1})) - h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k)) \\ & \leq \|\nabla h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k))\|_* \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_{k+1} - \mathbf{u}_k)\|_* \\ & \quad - \frac{\lambda}{2} \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_{k+1} - \mathbf{u}_k)\|_F^2, \end{aligned} \quad (\text{A.2})$$

which directly results in

$$\begin{aligned} & h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k)) - h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_{k+1})) \\ & \geq \|\nabla h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k))\|_* \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k - \mathbf{u}_{k+1})\|_*. \end{aligned} \quad (\text{A.3})$$

According to the model (22), we update \mathbf{u}_{k+1} and deduce that

$$\begin{aligned} & \|\nabla h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k))\|_* \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_{k+1})\|_* \\ & \leq \|\nabla h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k))\|_* \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k)\|_*, \end{aligned} \quad (\text{A.4})$$

which together with (A.3) confirms that

$$h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k)) - h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_{k+1})) \geq 0. \quad (\text{A.5})$$

With the aid of the result on the $h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})) = \sum_i h^\epsilon(\sigma_i \text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})) \geq 0$, we can draw the first property. In order to derive the second property, we prove that the optimal variable $\{\mathbf{u}_k\}$ is convergent. By applying classical Vandermonde decomposition lemma [28] on $h^\epsilon(\text{Toep}(\mathbf{u}))$, we have

$$\begin{aligned} & h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k)) - h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_{k+1})) \\ & \geq \|\nabla h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k))\|_* \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_k - \mathbf{u}_{k+1})\|_* \\ & \quad + \frac{\lambda}{2} \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_{k+1} - \mathbf{u}_k)\|_F^2 \geq \frac{\lambda}{2} \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_{k+1} - \mathbf{u}_k)\|_F^2. \end{aligned} \quad (\text{A.6})$$

For all $k \geq 0$ summing the inequality above, we can attain

$$\frac{\lambda}{2} \sum_{k=0}^{+\infty} \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_{k+1} - \mathbf{u}_k)\|_F^2 \leq h^\epsilon(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_0)), \quad (\text{A.7})$$

which implies that $\lim_{k \rightarrow \infty} \|\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}_{k+1} - \mathbf{u}_k)\|_F = 0$. Thus, $\{\mathbf{u}_k\}$ is convergent. In order to prove that $\{\mathbf{u}_k\}$ converges to a local minimum, it is noted that the models (20) and (22) have the same Karush-Kuhn-Tucker (KKT) conditions when the optimal variable $\{\mathbf{u}_k\}$ converges, i.e., $\mathbf{u}_k \rightarrow \mathbf{u}^*$ as $k \rightarrow \infty$. The reason is that the gradients of the objective functions and the constraints in the rank norm are the same as in nuclear norm models when the MM method

converges. We conclude that \mathbf{u}^* tend to be a local minimum of (20) since the objective function in (20) monotonically decreases. This completes the proof.

Data Availability

The authors claim that all data generated or analysed in this article are provided by their simulations and are included in this published article. The matlab data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 61671378). The authors would like to thank Dr. Weiqiang Tan and Professor Guo Wang for their comments that greatly improved the manuscript.

References

- [1] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, 1996.
- [2] W. Tan, M. Matthaiou, S. Jin, L. Fan, and G. K. Karagiannidis, "On the spectral efficiency of massive MIMO systems with hybrid DFT processing," *IEEE Transactions on Wireless Communications*, vol. 1, no. 1, pp. 1–13, 2018.
- [3] X. Wu, W. P. Zhu, and J. Yan, "A high-resolution DOA estimation method with a family of nonconvex penalties," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 6, pp. 4925–4938, 2018.
- [4] W. Tan, M. Matthaiou, S. Jin, and X. Li, "Spectral efficiency of DFT-based processing hybrid architectures in massive MIMO," *IEEE Wireless Communications Letters*, vol. 6, no. 5, pp. 586–589, 2017.
- [5] A. Moffet, "Minimum-redundancy linear arrays," *IEEE Transactions on Antennas and Propagation*, vol. 16, no. 2, pp. 172–175, 1968.
- [6] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 573–586, 2011.
- [7] P. Pal and P. P. Vaidyanathan, "Coprime sampling and the music algorithm," in *2011 Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE)*, pp. 289–294, Sedona, AZ, USA, 2011.
- [8] P. Pal and P. P. Vaidyanathan, "Nested arrays: a novel approach to array processing with enhanced degrees of freedom," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4167–4181, 2010.
- [9] A. J. Weiss and M. Gavish, "Direction finding using esprit with interpolated arrays," *IEEE Transactions on Signal Processing*, vol. 39, no. 6, pp. 1473–1478, 1991.
- [10] B. Friedlander and A. J. Weiss, "Direction finding using spatial smoothing with interpolated arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 2, pp. 574–587, 1992.
- [11] B. Friedlander, "The root-music algorithm for direction finding with interpolated arrays," *Signal Processing*, vol. 30, no. 1, pp. 15–29, 1993.
- [12] S. Pillai and F. Haber, "Statistical analysis of a high resolution spatial spectrum estimator utilizing an augmented covariance matrix," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 11, pp. 1517–1523, 1987.
- [13] Y. I. Abramovich, D. A. Gray, A. Y. Gorokhov, and N. K. Spencer, "Positive-definite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays. I. Fully augmentable arrays," *IEEE Transactions on Signal Processing*, vol. 46, no. 9, pp. 2458–2471, 1998.
- [14] Y. I. Abramovich, N. K. Spencer, and A. Y. Gorokhov, "Positive-definite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays. II. Partially augmentable arrays," *IEEE Transactions on Signal Processing*, vol. 47, no. 6, pp. 1502–1521, 1999.
- [15] T. E. Tuncer, T. K. Yasar, and B. Friedlander, "Direction of arrival estimation for nonuniform linear arrays by using array interpolation," *Radio Science*, vol. 42, no. 4, 2007.
- [16] C. L. Liu and P. P. Vaidyanathan, "Remarks on the spatial smoothing step in coarray MUSIC," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1438–1442, 2015.
- [17] Z. Yang and L. Xie, "On gridless sparse methods for line spectral estimation from complete and incomplete data," *IEEE Transactions on Signal Processing*, vol. 63, no. 12, pp. 3139–3153, 2015.
- [18] Z. M. Liu, Z. T. Huang, and Y. Y. Zhou, "Sparsity-inducing direction finding for narrowband and wideband signals based on array covariance vectors," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 1–12, 2013.
- [19] J. Yin and T. Chen, "Direction-of-arrival estimation using a sparse representation of array covariance vectors," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4489–4493, 2011.
- [20] Y. Tian and X. Sun, "Passive localization of mixed sources jointly using MUSIC and sparse signal reconstruction," *AEU - International Journal of Electronics and Communications*, vol. 68, no. 6, pp. 534–539, 2014.
- [21] W. Tan, X. Feng, X. Ye, and L. Jing, "Direction-of-arrival of strictly non-circular sources based on weighted mixed-norm minimization," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, 2018.
- [22] W. Tan and X. Feng, "Covariance matrix reconstruction for source localization over impaired uniform linear array," in *2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, pp. 1–5, Xiamen, China, 2017.
- [23] Z. Q. He, Z. P. Shi, and L. Huang, "Covariance sparsity-aware DOA estimation for nonuniform noise," *Digital Signal Processing*, vol. 28, pp. 75–81, 2014.
- [24] C. Wen, X. Xie, and G. Shi, "Off-grid DOA estimation under nonuniform noise via variational sparse Bayesian learning," *Signal Processing*, vol. 137, pp. 69–79, 2017.
- [25] J. Yang, G. Liao, and J. Li, "An efficient off-grid DOA estimation approach for nested array signal processing by using sparse Bayesian learning strategies," *Signal Processing*, vol. 128, pp. 110–122, 2016.

- [26] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digital Signal Processing*, vol. 8, no. 3, pp. 185–210, 1998.
- [27] M. Malek-Mohammadi, M. Babaie-Zadeh, and M. Skoglund, "Iterative concave rank approximation for recovering low-rank matrices," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5213–5226, 2014.
- [28] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, University of California Press, 2001.
- [29] M. Pesavento and A. B. Gershman, "Maximum-likelihood direction-of-arrival estimation in the presence of unknown nonuniform noise," *IEEE Transactions on Signal Processing*, vol. 49, pp. 1310–1324, 2002.

Research Article

A Single-Layer Dual-Band Reflectarray Cell for 5G Communication Systems

Sandra Costanzo , Francesca Venneri, Antonio Borgia , and Giuseppe Di Massa

Università della Calabria, Rende, Italy

Correspondence should be addressed to Sandra Costanzo; costanzo@dimes.unical.it

Received 29 August 2018; Revised 18 December 2018; Accepted 14 January 2019; Published 11 March 2019

Guest Editor: Khalil Sayidmarie

Copyright © 2019 Sandra Costanzo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A single-layer dual-band reflectarray cell is proposed in this work for future 5G systems. A reflectarray unit cell operating at 28/38 GHz is designed by adopting two pairs of miniaturized fractal patches, offering low losses (<0.7 dB) and almost full-phase ranges ($\approx 320^\circ$) at both operating frequencies. The proposed configuration allows to achieve very small interelement spacings and negligible mutual coupling effects between the two bands, thus assuring an independent phase-tuning mechanism at both desired frequency bands. The designed compact cell is successfully adopted to demonstrate reflectarrays' abilities in achieving fixed scanned-beam and/or multibeam patterns, under the dual-band operation mode. Full-wave numerical validations, performed on the synthesized reflectarray structures, confirm the effectiveness of the designed dual-band configuration in achieving independent radiation patterns and quite good bandwidths, at the two designed frequencies. Thanks to its compactness and versatility in achieving both frequency diversity and multibeam/scanned-beam radiation patterns, the proposed unit cell is appealing for future 5G applications.

1. Introduction

Nowadays, the development of new technologies for future fifth generation (5G) wireless communication networks is the main challenge in the telecommunications industry. 5G communication systems are expected to meet the growing demand for higher data rates (i.e., 1-10 theoretically gigabit per second (Gbps) [1]), as required by multimedia applications and the Internet of Things (IoT). To address this demand, 5G systems will use millimeter wave (mmw) frequencies, which represent one of the key enabling technologies in the development and implementation of 5G communication networks [1, 2]. However, the mmw frequencies are characterized by propagation limitations, such as higher path loss and shorter communication distances, mainly due to the atmospheric absorption of electromagnetic waves at higher frequencies [3]. For this reason, there is a need for designing high-gain antennas able to compensate for path losses. To this end, microstrip array antennas may represent a good candidate, providing also a narrow beamwidth and very thin profiles that are essential for 5G

operations. Some interesting 5G array antennas have been recently proposed in [2, 4, 5].

A very attractive alternative solution for designing 5G antennas is offered by microstrip reflectarrays [6, 7]. They consist of an array of microstrip elements illuminated by a feed antenna. Each element is designed to compensate for the phase delay in the path coming from the feed and to introduce a phase contribution giving a prescribed main beam direction in the antenna radiation pattern. Thanks to their higher efficiencies, due to the adopted spatial feeding approach [6], reflectarrays represent a promising solution for designing high-gain/directivity 5G antennas. Furthermore, reflectarrays can be properly designed to offer several reconfiguration capabilities, which are very appealing for 5G systems [2], such as beam-steering functions and/or frequency agility [6, 8–11], multibeam radiation patterns [6, 9], and multiband operation modes [6].

In order to demonstrate reflectarray versatility in satisfying most of the 5G requirements, a single-layer dual-band reflectarray cell is investigated in this paper to operate within the Ka band (at 28/38 GHz), which is currently under

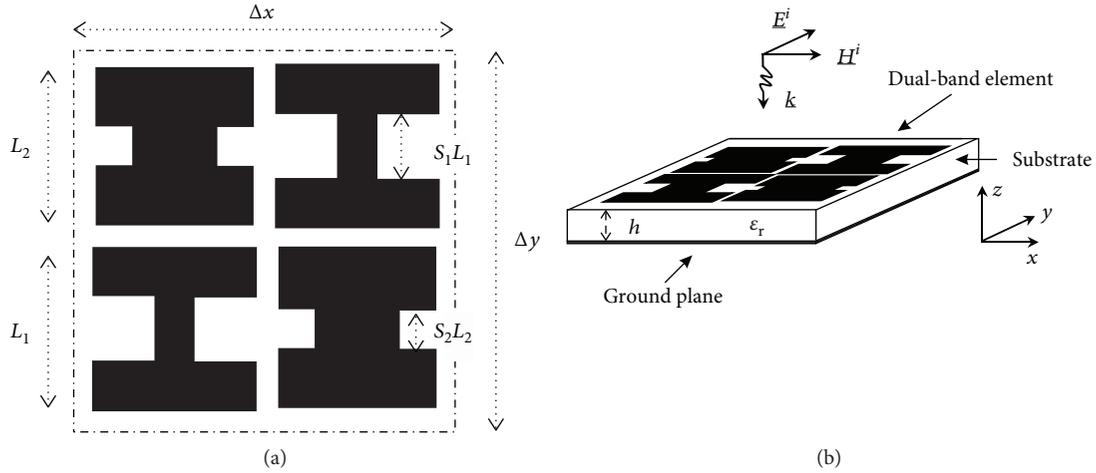


FIGURE 1: Unit cell layout: (a) top view and (b) 3D view and reference system.

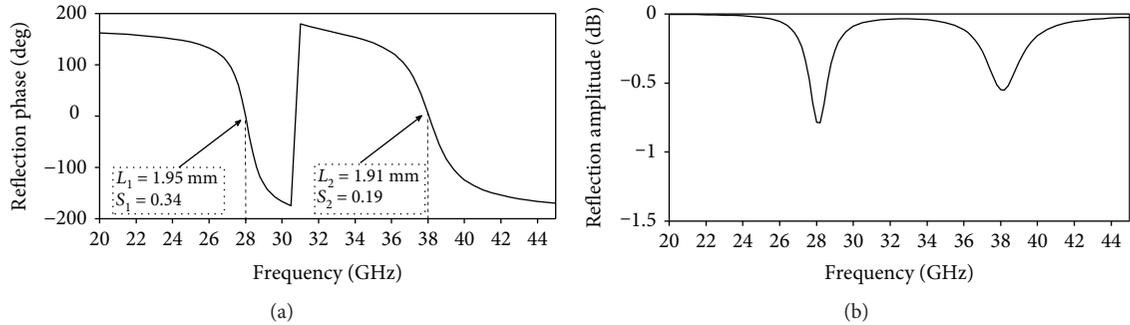


FIGURE 2: Simulated unit cell reflection coefficient vs. frequency: (a) phase and (b) amplitude.

consideration for 5G technologies [2]. The proposed cell is adopted to demonstrate reflectarrays' abilities in achieving fixed scanned-beam and/or multibeam patterns, under the dual-band operation mode.

The fractal concept has yet to be applied in the literature to reflectarray antennas, as reported in [12–16]. In this paper, a modified layout of the fractal element, originally proposed by the authors in [12, 13], is investigated to design a compact single-layer dual-band cell with single linear polarization, offering low losses (<0.7 dB) and almost full-phase ranges ($\cong 320^\circ$), at both operating frequencies. The miniaturization capabilities of the adopted fractal geometry are exploited to achieve a dual-band behavior, simply by embedding two pairs of miniaturized resonators within the same unit cell. Negligible mutual effects between the two bands are demonstrated, thus achieving an independent phase-tuning mechanism for each frequency band, by properly changing the fractal shapes. Unlike other multi-band reflectarray cells presented in the literature [17–21], the proposed dual-band fractal cell allows to achieve the following: a simpler and thinner structure with respect to the multilayer stacked configurations [17, 18]; smaller unit cell size at both operating frequencies ($\cong 0.4\lambda$ at 28 GHz, $\cong 0.54\lambda$ at 38 GHz) with respect to other single-layer configurations [19, 20], thus preserving the capability to point the main beam at large scan angles,

without occurring in grating lobe phenomena; and smaller electrical interferences between the elements operating at the different bands [21], despite the adopted small interelement spacing.

As a first proof of the concept, the proposed cell is adopted to design reflectarrays able to independently scan the main beam at the two operating frequencies. A 1 dB gain-bandwidth of about 1 GHz is simulated at both operating frequencies, potentially offering several Gbps of data rates at 28 or 38 GHz [1]. Moreover, a dual-band multibeam reflectarray design is numerically demonstrated. Finally, the dual-band behavior of the proposed unit cell configuration is experimentally validated.

2. Dual-Band Reflectarray Unit Cell

The proposed dual-band unit cell is composed of two alternately arranged pairs of fractal patches (Figure 1), each designed to operate around a specific resonant frequency. The layout of the single patch is essentially derived from the 1st iteration fixed-length Minkowski patch originally proposed by the authors in [12]. The patch geometry, reported in Figure 1(a), is characterized by a beginning of dimensions $L_n \times L_n$ ($n = 1, 2$). A smaller square of the side $S_n L_n$ is removed from the center of the two lateral sides (i.e., the resonant sides), thus obtaining a linearly polarized

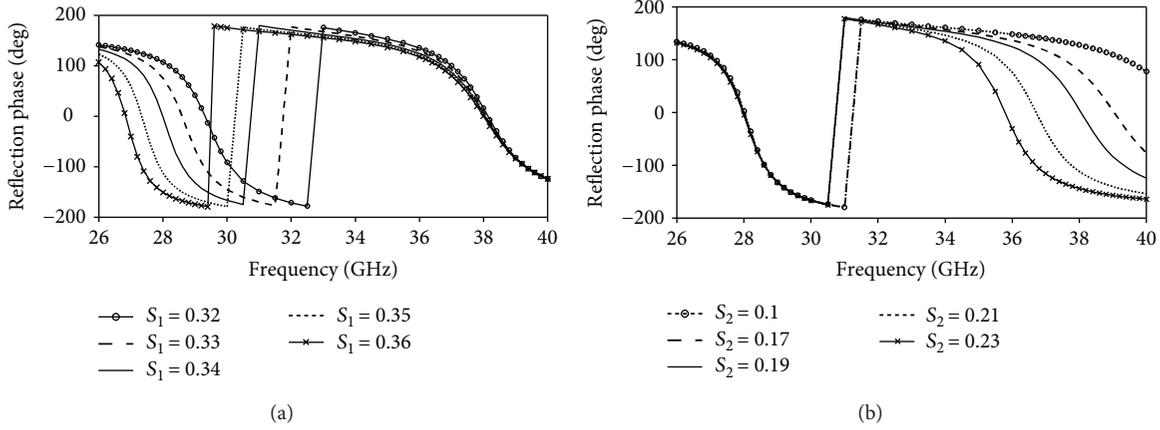


FIGURE 3: Simulated unit cell reflection coefficient vs. frequency for different scaling factors: (a) $L_1 = 1.95$ mm, $S_1 =$ variable and $L_2 = 1.91$ mm, $S_2 = 0.19$; (b) $L_1 = 1.95$ mm, $S_1 = 0.34$ and $L_2 = 1.91$ mm, $S_2 =$ variable.

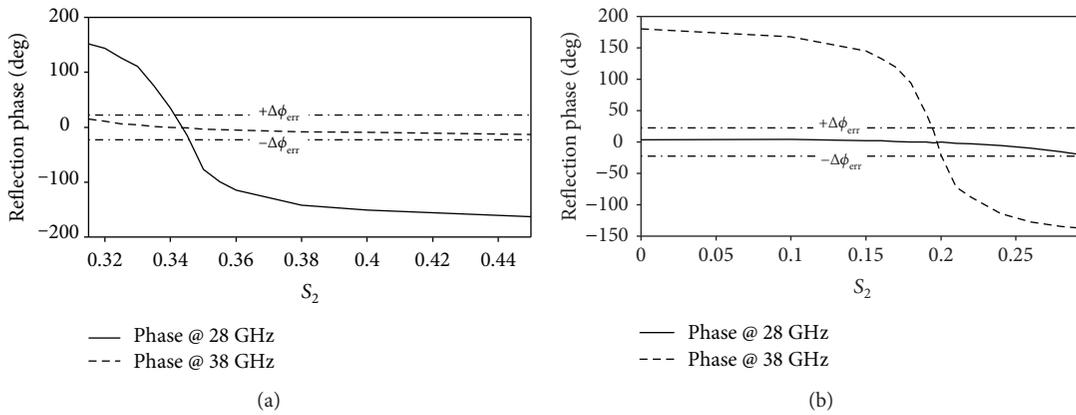


FIGURE 4: Simulated phase curve vs. the scaling factors S_1 and S_2 , at the two operating frequencies: (a) $L_1 = 1.95$ mm, $S_1 =$ variable and $L_2 = 1.91$ mm, $S_2 = 0.19$; (b) $L_1 = 1.95$ mm, $S_1 = 0.34$ and $L_2 = 1.91$ mm, $S_2 =$ variable.

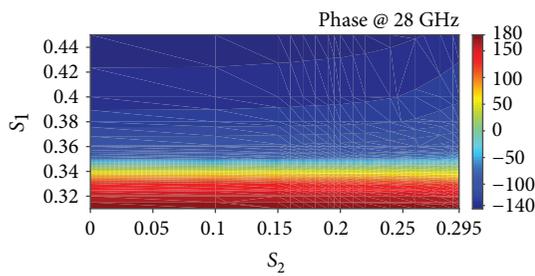


FIGURE 5: Simulated reflection phase vs. S_1 and S_2 (at 28 GHz).

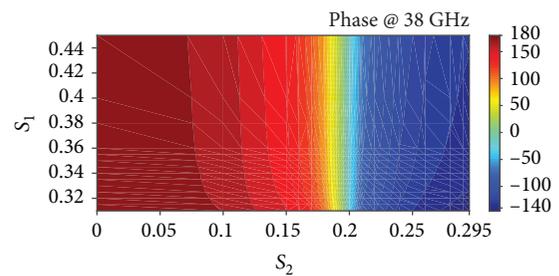


FIGURE 6: Simulated reflection phase vs. S_1 and S_2 (at 38 GHz).

element, along the y -axis (Figure 1(b)). The scaling factor S_n varies from 0 up to 0.45. The reflection phase tuning is realized by independently varying the fractal scaling factor S_n of each element, leaving unchanged the patch size $L_n \times L_n$. The above fractal shape allows to fit a longer electrical resonator into a smaller unit cell [12], thus offering very appealing miniaturization skills. Furthermore, the fixed length of the radiating sides (i.e., the upper and lower sides of the patch in Figure 1(a)) guarantees a higher independence to mutual coupling effects. As a

matter of fact, in [22], it is demonstrated how the stronger contribution to the mutual coupling between microstrip patches is that occurring along the E -plane (yz -plane in Figure 1(b)).

2.1. Unit Cell Design. The layout depicted in Figure 1 is adopted to design a dual-band unit cell operating at the following 5G frequencies: $f_1 = 28$ GHz and $f_2 = 38$ GHz. A benzocyclobutene (BCB) polymer is adopted as a substrate material. The excellent material features in terms of low loss

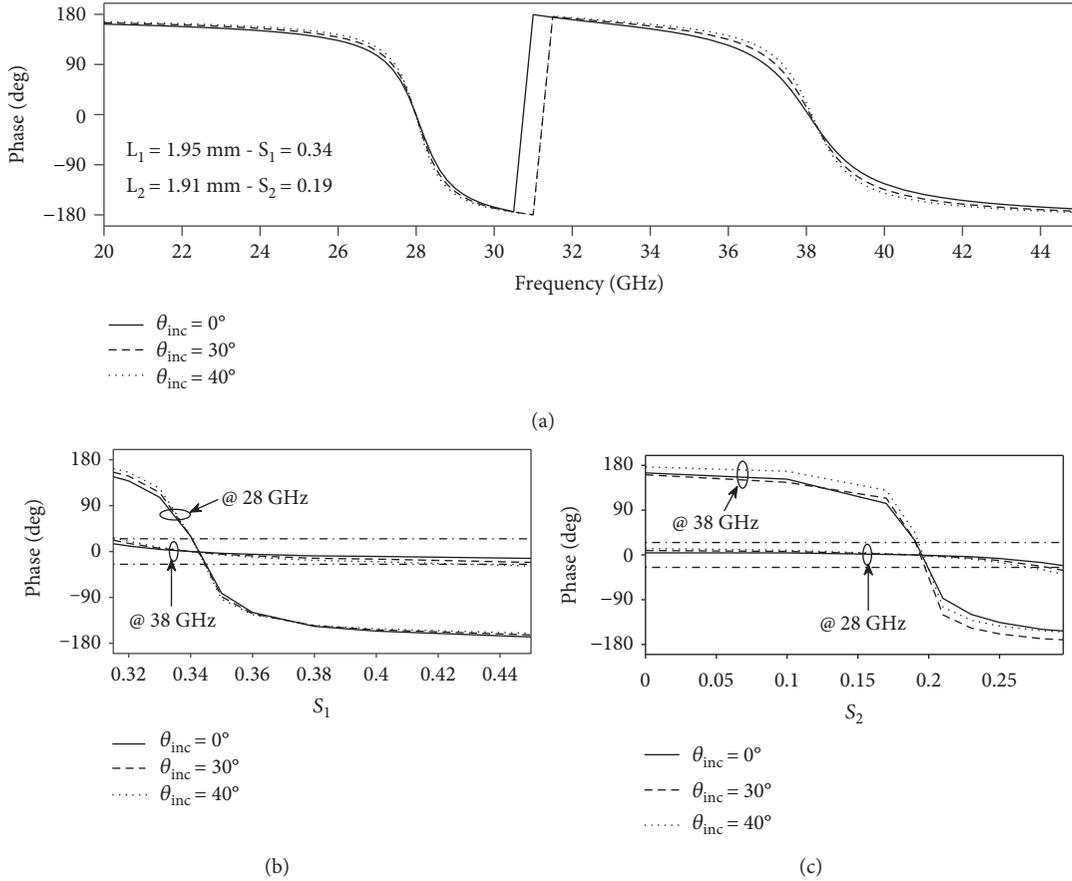


FIGURE 7: Simulated reflection phase curves for different incidence angles: (a) phase vs. frequency; (b) phase vs. the scaling factor S_1 at the two operating frequencies ($L_1 = 1.95$ mm, $S_1 =$ variable and $L_2 = 1.91$ mm, $S_2 = 0.19$); (c) phase vs. the scaling factor S_2 at the two operating frequencies ($L_1 = 1.95$ mm, $S_1 = 0.34$ and $L_2 = 1.91$ mm, $S_2 =$ variable).

($\tan \delta \leq 0.008$), low dielectric constant ($\cong 2.65$), and strong dielectric stability against frequency and temperature make the BCB polymer a good choice for designing mmw microstrip antennas [23, 24]. A substrate thickness equal to 0.26 mm is fixed [24]. A commercial full-wave code [25], based on the infinite array approach, is adopted for unit cell analysis. A normal incident plane wave is considered ($\theta_{inc} = 0^\circ$). A periodicity of 4.3 mm is fixed in both directions (i.e., $\Delta x = \Delta y = 4.3$ mm) corresponding to 0.4λ at 28 GHz and 0.54λ at 38 GHz. Following the design rules outlined in [12], the unit cell is designed to achieve the desired dual-resonant behavior. In particular, remembering that the use of smaller patch lengths L_n gives higher resonant frequencies and, conversely, the use of greater S_n values allows to move down the resonant frequencies, the size of the fractal patches is fixed to the values depicted in Figure 2. The minimum distance between two adjacent elements (i.e., $(\Delta x - L_1 - L_2)/2 = (\Delta y - L_1 - L_2)/2$, see Figure 1) is equal to 0.22 mm. Both the phase (Figure 2(a)) and the amplitude (Figure 2(b)) of the unit cell reflection coefficient are illustrated in Figure 2. Similar phase curve behavior vs. frequency can be observed at both resonances (Figure 2(a)). Furthermore, low losses smaller than 0.7 dB can be appreciated in Figure 2(b).

2.2. Parametric Analysis of the Unit Cell. The reflection phase tuning is independently realized at each resonant frequency f_n , by varying the corresponding fractal scaling factor S_n , leaving unchanged the patch size $L_n \times L_n$. Both factors, S_1 and S_2 , are properly tuned to achieve a full-phase tuning range, at the corresponding operating frequency. Figure 3 shows the effectiveness of the adopted phase-tuning mechanism that allows to independently vary the reflection phase at each desired frequency. In particular, it can be observed that, by changing the scaling factor S_1 associated to the 28 GHz resonant element and leaving unchanged the other fractal sizes, the phase response of the unit cell varies only corresponding to a neighborhood of $f_1 = 28$ GHz, whilst a negligible phase variation is obtained at $f_2 = 38$ GHz (Figure 3(a)). A similar behavior can be observed in Figure 3(b). In fact, by varying the scaling factor S_2 , a variable phase shift is introduced only corresponding to a neighborhood of $f_2 = 38$ GHz, leaving unchanged the phase response at $f_1 = 28$ GHz. The phase curves computed vs. the scaling factors S_n (Figure 4) show the admissible scaling factor variation range for both S_1 (Figure 4(a)) and S_2 (Figure 4(b)). In other words, the results depicted in Figure 4 affirm that, in order to achieve a quite full-phase tuning range ($\cong 320^\circ$) at 28 GHz, the scaling

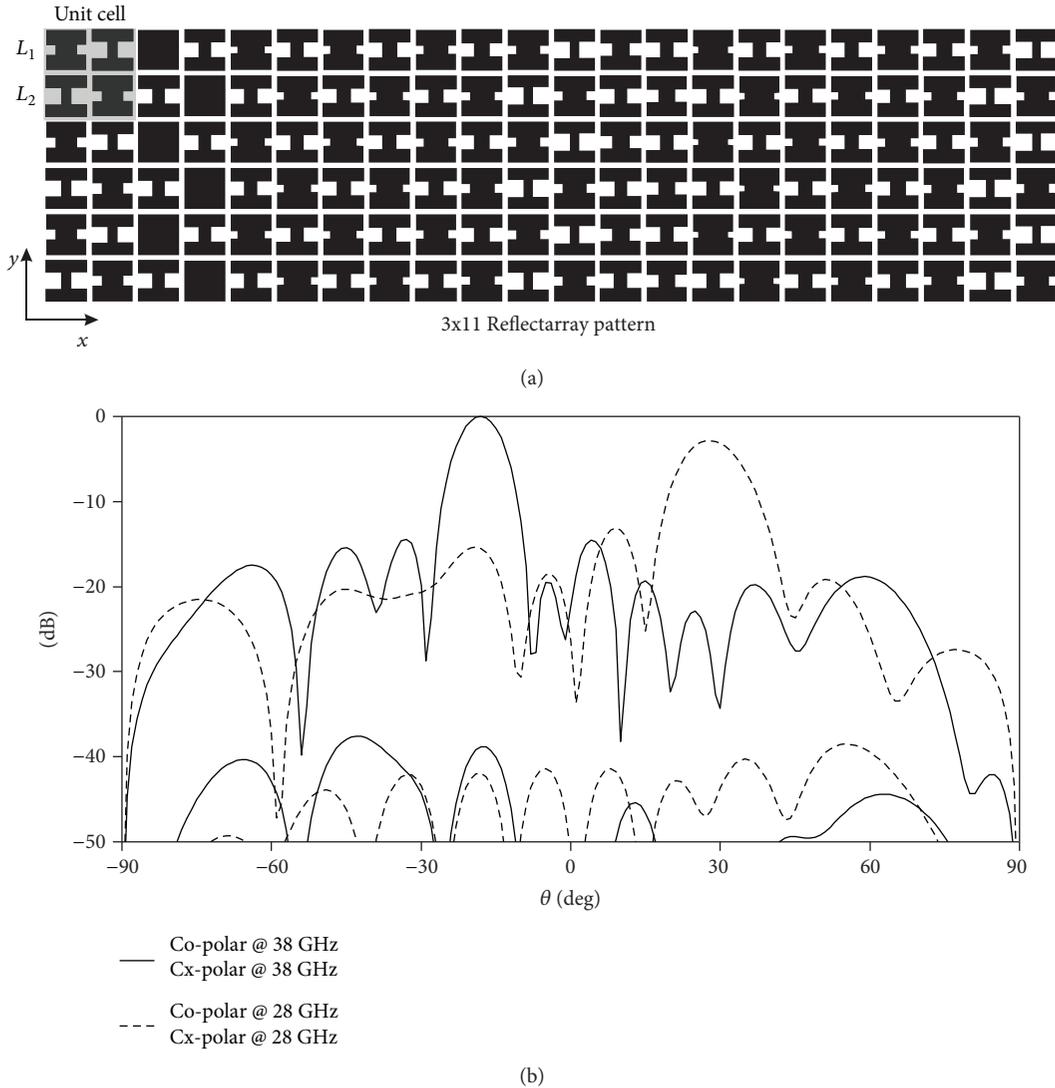


FIGURE 8: Synthesized dual-band scanned reflectarray: (a) reflectarray layout and (b) simulated radiation patterns at the designed frequencies.

factor S_1 must be varied from 0.31 up to 0.45, assuring also small phase excursions, corresponding to the frequency $f_2 = 38$ GHz, that are comprised between the phase error margins given by $\pm\Delta\phi_{\text{err}} = \pm 22.5^\circ$ (Figure 4(a)). Conversely, the scaling factor S_2 must be varied within the value range starting from 0 up to 0.29, in order to assure a quite full-phase range at 38 GHz and a confined phase excursion error corresponding to the frequency $f_1 = 28$ GHz (Figure 4(b)).

In order to give a more exhaustive description of the reflection phase response computed for the designed dual-band unit cell, a contour plot is reported under Figures 5 and 6, showing, respectively, the phase variations computed at frequency $f_1 = 28$ GHz vs. S_1 , for different S_2 values (Figure 5), and the phase computed at frequency $f_2 = 38$ GHz vs. S_2 , for different S_1 values (Figure 6). Both figures confirm how the proposed dual-band unit cell can offer an independent phase-tuning mechanism at each designed frequency band. As a matter of fact, a quite constant reflection phase can be observed at 28 GHz (Figure 5), by

changing the scaling factor S_2 for a fixed S_1 value. Similar considerations can be extrapolated from Figure 6.

The unit cell is further analyzed to investigate the behavior of its phase response, for different incidence angles of the impinging wave. Figure 7 shows acceptable phase variations under 30° and 40° oblique incidence angles, with respect to the normal case. Furthermore, as depicted in Figures 7(b) and 7(c), the independency of the phase-tuning mechanism is preserved at both frequencies also in the case of oblique incidence angles. In particular, it can be observed that, by changing the scaling factor S_1 associated to the 28 GHz resonant element, negligible phase variations are obtained at the frequency $f_2 = 38$ GHz (Figure 7(a)). A similar behavior can be observed in Figure 7(b), in the case of the 38 GHz resonant patch.

3. Dual-Band Reflectarray Designs

In order to prove the effectiveness of the proposed dual-band unit cell, two small 3×11 reflectarray designs

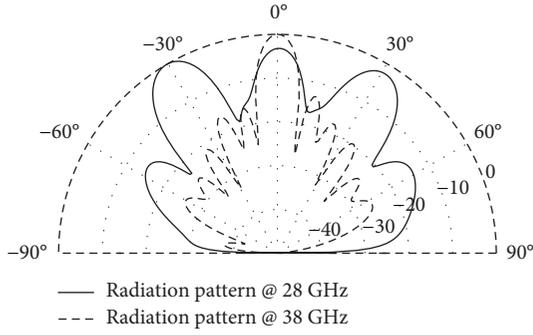


FIGURE 9: Synthesized dual-band multibeam reflectarray: simulated radiation patterns at the designed frequencies.

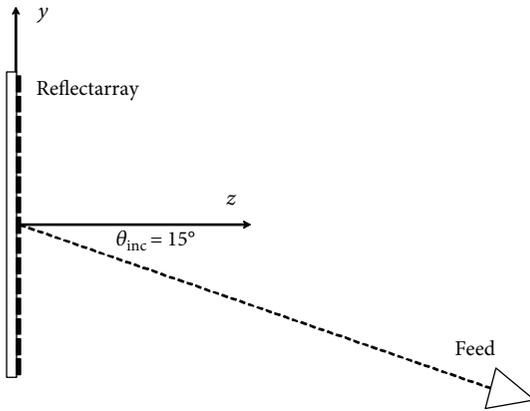


FIGURE 10: Schematic layout of the 15×15 reflectarray.

are reported in the following. The high independence between the two designed frequency bands (28/38 GHz), demonstrated in Section 2.2, allows to separately synthesize the two sets of resonant fractal patches (identified by L_1 and L_2 in Figure 1), without affecting the accuracy of the adopted design procedure. Both reflectarrays are designed by adopting a synthesis algorithm [9] that receives as input the desired radiation pattern, in terms of the main beam direction and maximum side lobe level, and automatically returns the required excitation phase on each reflectarray cell. The algorithm uses the phase design curves depicted in Figures 4(a) and 4(b) that assume fixed fractal patch sizes, except the adopted phase-tuning parameter (i.e., S_1 for $f_1 = 28$ GHz and S_2 for $f_2 = 38$ GHz). A normal incident plane wave is considered.

The first design consists of a reflectarray able to independently scan the main beam at the two different operating frequencies, whilst the second design is characterized by a multibeam pattern.

The fractal patches embedded in each cell of the first reflectarray are chosen as follows: the elements resonating within the 28 GHz band, identified by the couple (L_1, S_1) , are computed giving the main lobe steered towards the direction $\theta_{MB} = 30^\circ$, in the H -plane (xz -plane in Figure 8); whilst the elements resonating within the 38 GHz band, identified by (L_2, S_2) , are chosen by imposing a main beam direction equal to -20° , in the H -plane. Figure 8 shows the synthesized

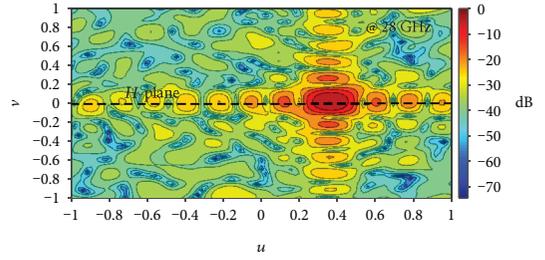


FIGURE 11: Synthesized 28 GHz pattern of the 15×15 offset-fed reflectarray.

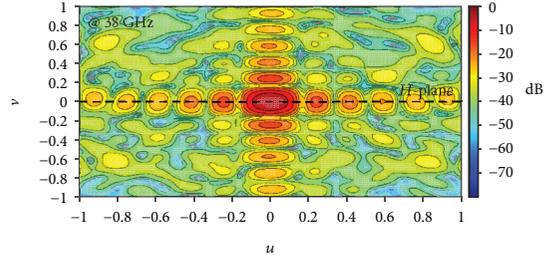


FIGURE 12: Synthesized 38 GHz pattern of the 15×15 offset-fed reflectarray.

reflectarray layout (Figure 8(a)) and the full-wave simulations of the overall structure. In particular, two independent radiation patterns are achieved at the two designed frequencies. Both patterns fully satisfy the constraints imposed during the synthesis stage, confirming the effectiveness of the proposed dual-band unit cell.

Furthermore, a dual-band 3×11 reflectarray is designed to achieve a multibeam radiation pattern corresponding to the first resonance (i.e., 28 GHz). In particular, a pattern having three beams, respectively, directed along -30° , 0° , and 30° in the H -plane, is synthesized at 28 GHz, whilst a broadside radiation pattern is imposed at 38 GHz. The full-wave simulation of the overall reflectarray structure is depicted in Figure 9. The simulated patterns match quite well the constraints imposed during the synthesis stage.

Finally, a 15×15 reflectarray is designed to independently direct the main lobes at the two different operating frequencies in the H -plane (i.e., $\theta_{MB} = 20^\circ$ at 28 GHz and $\theta_{MB} = 0^\circ$ at 38 GHz). The reflecting surface is illuminated by a Ka-band horn (characterized by a 15 dB gain and a $16.6 \text{ mm} \times 20.2 \text{ mm}$ aperture), which is placed in the E -plane (i.e., the yz -plane in Figure 10), at a distance of 16.5 cm from the reflecting surface, with an offset angle of about 15° . The two couples of elements are properly synthesized to compensate the phase delay in the paths from the feed and to introduce a proper phase contribution able to meet the synthesis constraints. Figures 11 and 12 show the contour plot of the synthesized patterns in u - v coordinates, where $u = \sin(\theta) \cos(\phi)$ and $v = \sin(\theta) \sin(\phi)$. Figure 13 illustrates the gain patterns vs. frequency, computed, respectively, along the main beam direction $(\theta_{MB}, \phi_{MB}) = (20^\circ, 0^\circ)$, in the case of the 28 GHz radiation pattern, and along the direction $(\theta_{MB}, \phi_{MB}) = (0^\circ, 0^\circ)$, in the case of

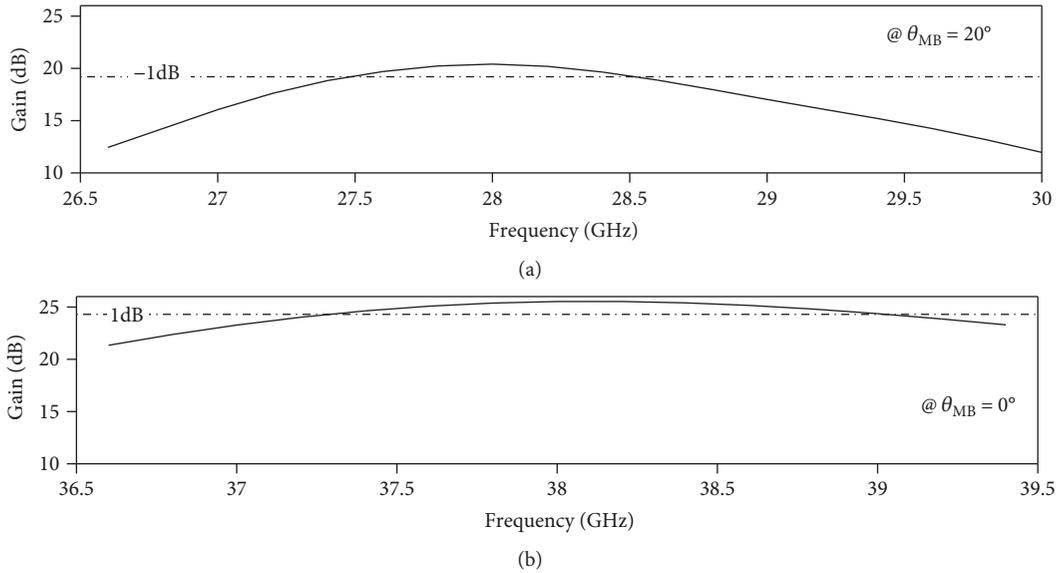


FIGURE 13: Simulated gain pattern versus frequency in the H -plane: (a) $\theta = 20^\circ$ and (b) $\theta = 0^\circ$.

the 38 GHz radiation pattern. A greater gain peak value can be observed in the case of the 38 GHz pattern (i.e., the gain difference is about 4.2 dB), mainly due to the greater electrical size and the lower spillover characterizing the antenna aperture at the higher operating frequency (38 GHz) and secondly due to lower scan losses [26]. Although the structures have not been properly optimized in terms of bandwidth, the simulated gain patterns show an acceptable 1 dB gain-bandwidth, approximately equal to 950 MHz at 28 GHz ($\approx 3.4\%$) and to 1.5 GHz at 38 GHz ($\approx 3.9\%$), that, at the considered mmw frequencies, can potentially offer high data rates [1], as required by 5G applications.

The above reflectarray designs confirm the effectiveness of the proposed dual-band unit cell in achieving an independent control of the reflectarray radiation features at each operating frequency. As future developments, the proposed configuration will be further optimized in terms of bandwidth. Furthermore, experimental validations of the configuration will be scheduled in order to deal with the problems associated with the manufacturing process, such as fabrication tolerances and illumination feed displacement, not yet considered in this work.

4. Experimental Validation of the Unit Cell

In order to give a preliminary validation of the proposed dual-band unit cell, a small array prototype, composed of 11×11 identical cells, is realized and tested in the Microwave Laboratory of the University of Calabria (Figure 14). The array is printed on a DiClad 880 substrate, having $\epsilon_r = 2.24$ and $h = 0.254$ mm [27]. A periodicity equal to 5 mm is fixed in both directions. The patches embedded in each unit cell are characterized by the following dimensions: $L_1 = 2.24$ mm, $S_1 = 0.33$ and $L_2 = 2.1$ mm, $S_2 = 0.2$, giving a minimum distance between two adjacent patches equal to about 0.33 mm. The cell operates in a dual-band mode corresponding to about 29 GHz and 39 GHz.

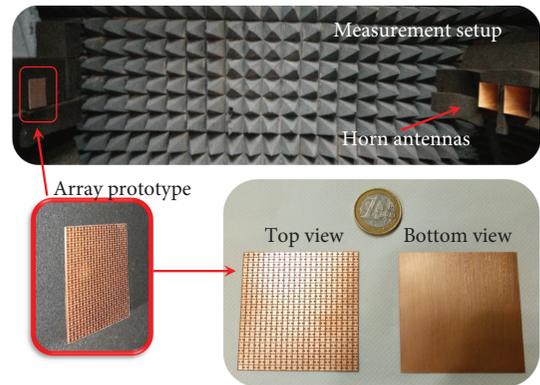


FIGURE 14: Reflection phase measurement setup and the dual-band array prototype.

A far-field measurement system is adopted (Figure 14), consisting of two identical transmitting and receiving horn antennas (operating within the $(26.5 \div 40)$ GHz frequency band), both connected to a vector network analyzer. The instrumentation detects the field reflected by the array along the broadside direction in the far-field region [9]. The reflection phase curve of the cell is measured within the frequency range $26.5 \div 40$ GHz (Figure 15). It can be observed a phase variation of about 300° around both operating frequencies. Furthermore, a good agreement between simulations and measurements can be observed in Figure 15, showing a small but appreciable frequency shift, mainly due to the fabrication error tolerance related to the adopted printed circuit board (PCB) milling process [28]. Anyway, the effects due to the above errors can be reduced through a proper reiteration of the unit cell synthesis procedure (see Section 2) or, alternatively, by adopting a more precise fabrication process, as that described in [23, 24]. In conclusion, it can be stated that the experimental results confirm the dual-band behavior of the proposed configuration.

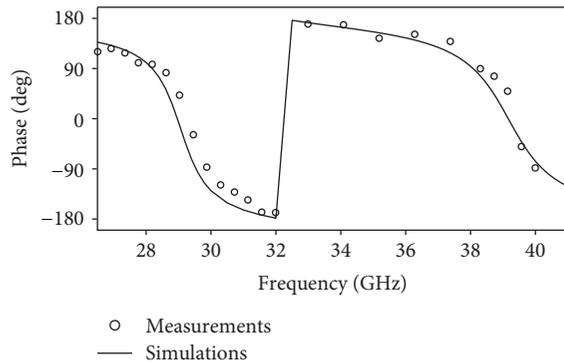


FIGURE 15: Comparison between measured and simulated phase curve vs. frequency.

5. Conclusion

A single-layer dual-band reflectarray with a single linear polarization cell has been investigated in this work for future 5G wireless communication networks. A reflectarray unit cell operating within the Ka band (at 28/38 GHz) has been designed. Two pairs of miniaturized fixed-length fractal patches have been synthesized, achieving low losses (<0.7 dB) and almost full-phase ranges ($\cong 320^\circ$), at both operating frequencies. A thorough parametric analysis of the unit cell has been performed, demonstrating negligible mutual coupling effects between the two pairs of resonant elements, so assuring the independence between the two designed frequency bands. The designed compact cell has been successfully adopted to demonstrate reflectarrays' abilities in achieving fixed scanned-beam and/or multibeam patterns, under the dual-band operation mode.

Full-wave simulations of the synthesized reflectarray structures have been performed, confirming the effectiveness of the designed dual-band configuration in achieving independent radiation patterns at the two designed frequencies. In conclusion, the designed dual-band configuration offers, at the same time, compactness, low losses, frequency diversity, high versatility in achieving fixed scanned-beam and/or multibeam radiation patterns, and quite good gain-bandwidths. The above features make the proposed dual-band reflectarray configuration appealing for the implementation of 5G antennas. A preliminary experimental validation of the dual-band behavior of the proposed unit cell has been performed.

As future developments, the proposed configuration will be further optimized in terms of bandwidth and fabrication tolerances. Further experimental validations on dual-band reflectarray prototypes will be scheduled for the future, in order to give a comprehensive proof of the proposed concept.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] T. S. Rappaport, Y. Xing, G. R. MacCartney, A. F. Molisch, E. Mellios, and J. Zhang, "Overview of millimeter wave communications for fifth-generation (5G) wireless networks—with a focus on propagation models," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6213–6230, 2017.
- [2] W. Hong, Z. H. Jiang, C. Yu et al., "Multibeam antenna technologies for 5G wireless communications," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6231–6249, 2017.
- [3] T. S. Rappaport, S. Sun, R. Mayzus et al., "Millimeter wave mobile communications for 5G cellular: it will work!," *IEEE Access*, vol. 1, pp. 335–349, 2013.
- [4] O. M. Haraz, A. Elboushi, S. A. Alshebeili, and A.-R. Sebak, "Dense dielectric patch array antenna with improved radiation characteristics using EBG ground structure and dielectric superstrate for future 5G cellular networks," *IEEE Access*, vol. 2, pp. 909–913, 2014.
- [5] W. T. Sethi, M. A. Ashraf, A. Ragheb, A. Alasaad, and S. A. Alshebeili, "Demonstration of millimeter wave 5G setup employing high-gain Vivaldi array," *International Journal of Antennas and Propagation*, vol. 2018, Article ID 3927153, 12 pages, 2018.
- [6] J. Huang and J. Encinar, *Reflectarray Antennas*, Wiley-IEEE Press, 2008.
- [7] M. H. Dahri, M. H. Jamaluddin, M. I. Abbasi, and M. R. Kamarudin, "A review of wideband reflectarray antennas for 5G communication systems," *IEEE Access*, vol. 5, pp. 17803–17815, 2017.
- [8] S. V. Hum and J. Perruisseau-Carrier, "Reconfigurable reflectarrays and array lenses for dynamic antenna beam control: a review," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 1, pp. 183–198, 2014.
- [9] F. Venneri, S. Costanzo, and G. Di Massa, "Design and validation of a reconfigurable single varactor-tuned reflectarray," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 2, pp. 635–645, 2013.
- [10] S. Costanzo, F. Venneri, A. Raffo, G. Di Massa, and P. Corsonello, "Radial-shaped single varactor-tuned phasing line for active reflectarrays," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 7, pp. 3254–3259, 2016.
- [11] S. Costanzo, F. Venneri, G. Di Massa, A. Borgia, and A. Raffo, "Bandwidth performances of reconfigurable reflectarrays: state of art and future challenges," *Radioengineering*, vol. 27, no. 1, pp. 1–9, 2018.
- [12] S. Costanzo and F. Venneri, "Miniaturized fractal reflectarray element using fixed-size patch," *IEEE Antennas and Wireless Propagation Letters*, vol. 13, pp. 1437–1440, 2014.
- [13] S. Costanzo, F. Venneri, G. Di Massa, A. Borgia, A. Costanzo, and A. Raffo, "Fractal reflectarray antennas: state of art and new opportunities," *International Journal of Antennas and Propagation*, vol. 2016, Article ID 7165143, 17 pages, 2016.
- [14] K. H. Sayidmarie and M. E. Bialkowski, "Fractal unit cells of increased phasing range and low slopes for single-layer microstrip reflectarrays," *IET Microwaves, Antennas & Propagation*, vol. 5, no. 11, p. 1371, 2011.

- [15] F. Xue, H. J. Wang, M. Yi, and G. Liu, "A broadband Ku-band microstrip reflectarray antenna using single-layer fractal elements," *Microwave and Optical Technology Letters*, vol. 58, no. 3, pp. 658–662, 2016.
- [16] D. Oloumi, S. Ebadi, A. Kordzadeh, A. Semnani, P. Mousavi, and X. Gong, "Miniaturized reflectarray unit cell using fractal-shaped patch-slot configuration," *IEEE Antennas and Wireless Propagation Letters*, vol. 11, pp. 10–13, 2012.
- [17] M. R. Chaharmir, J. Shaker, and H. Legay, "Dual-band Ka/X reflectarray with broadband loop elements," *IET Microwaves, Antennas & Propagation*, vol. 4, no. 2, pp. 225–231, 2010.
- [18] Y. Pan, Y. R. Zhang, and X. Yu, "A X/Ku dual-band reflectarray design with cosecant squared shaped beam," *Microwave and Optical Technology Letters*, vol. 56, no. 9, pp. 2028–2034, 2014.
- [19] Z. Hamzavi-Zarghani and Z. Atlasbaf, "A new broadband single-layer dual-band reflectarray antenna in X- and Ku-bands," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 602–605, 2015.
- [20] R. S. Malfajani and Z. Atlasbaf, "Design and implementation of a dual-band single layer reflectarray in X and K bands," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 8, pp. 4425–4431, 2014.
- [21] S. W. Qu, Q. Y. Chen, M. Y. Xia, and X. Y. Zhang, "Single-layer dual-band reflectarray with single linear polarization," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 1, pp. 199–205, 2014.
- [22] C. A. Balanis, *Antenna Theory Analysis and Design*, Wiley, New York, NY, USA, 1989.
- [23] S. Costanzo, I. Venneri, G. Massa, and A. Borgia, "Benzocyclobutene as substrate material for planar millimeter-wave structures: dielectric characterization and application," *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 31, p. 66, 2010.
- [24] S. Costanzo, F. Venneri, A. Borgia, I. Venneri, and G. Di Massa, "60 GHz microstrip reflectarray on a benzocyclobutene dielectric substrate," *IET Science, Measurement and Technology*, vol. 5, no. 4, pp. 134–139, 2011.
- [25] "Ansoft Designer®," <https://www.ansys.com>.
- [26] D. M. Pozar, S. D. Targonski, and H. D. Syrigos, "Design of millimeter wave microstrip reflectarrays," *IEEE Transactions on Antennas and Propagation*, vol. 45, no. 2, pp. 287–296, 1997.
- [27] <https://www.rogerscorp.com>.
- [28] "LPKF ProtoMat," <https://www.lpkf.com/>.

Research Article

Experimental Evaluation of MIMO-OFDM System with Rateless Space-Time Block Code

Ali H. Alqahtani ¹, Khaled Humadi ², Ahmed Iyanda Sulyman ²
and Abdulhameed Alsanie ²

¹College of Applied Engineering, King Saud University, Muzahimiyah Campus, P.O. Box 2454, Riyadh 11451, Saudi Arabia

²Department of Electrical Engineering, King Saud University, P.O. Box 800, Riyadh 11421, Saudi Arabia

Correspondence should be addressed to Ali H. Alqahtani; ahqahtani@ksu.edu.sa

Received 24 July 2018; Accepted 25 November 2018; Published 13 February 2019

Guest Editor: Peter Excell

Copyright © 2019 Ali H. Alqahtani et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Multiple-input multiple-output (MIMO) wireless technology in combination with orthogonal frequency-division multiplexing (MIMO-OFDM) is an attractive technique for next-generation wireless systems. However, the performance of wireless links is severely degraded due to various channel impairments which cause a decoding failure and lead to packet loss at the receiver. One technique to cope with this problem is the rateless space-time block code (RSTBC). This paper presents experimental results on the performance of a 2×2 MIMO-OFDM system with RSTBC as measured in a testbed implemented with field-programmable gate array (FPGA). The average bit error rate (BER) performance of the proposed scheme is evaluated experimentally, and the results agree closely with simulation and analytical upper bound. It has been shown that RSTBC can be implemented in real-world scenarios and guarantee the reliability of loss-prone wireless channels.

1. Introduction

OFDM techniques have been applied in various wireless communication standards, such as IEEE 802.11a [1] and IEEE 802.16a [2], since they have high spectral efficiency and robustness to multipath fading. MIMO communication systems are considered as an effective solution for high-performance wireless systems. Moreover, MIMO combined with OFDM technique, which helps to avoid the frequency selective fading and to take away the need for complex equalizers, is a good solution for reliable high spectral-efficient wireless communication systems. However, in practice, loss-prone channels lead to significant performance degradation compared to loss-free environments.

Many possible error (or erasure) control schemes are available and can be used in MIMO-OFDM systems. The classical space-time codes (STCs) have spatial diversity in a flat-fading MIMO channel, but, in general, they are not efficient in the frequency-selective fading channels [3, 4].

One approach to guarantee the reliability of lossy communication systems is the *rateless coding*. In contrast to the common fixed-rate channel codes, rateless code does not have a fixed code rate before transmission as it can generate a flexible number of encoded packets. In this way, rateless codes are able to work universally over various classes of channels and change their code rate in accordance with the specific channel state. Conceptually, rateless codes are similar to the hybrid automatic repeat request (HARQ) schemes [5], but they mainly differ in the code design and overhead redundancy, where rateless codes can be interpreted as a continuous version of the HARQ.

To our best knowledge, in the literature of rateless coding, most of the existent works is an extension of the fountain codes over the *erasure channels* [6], in which the receiver simply drops the erroneous packets. Such rateless erasure codes were first invented for computer networks. Luby transform (LT) code [7] is the first practical pattern of rateless codes. LT code was subsequently developed to the well-known Raptor code [8].

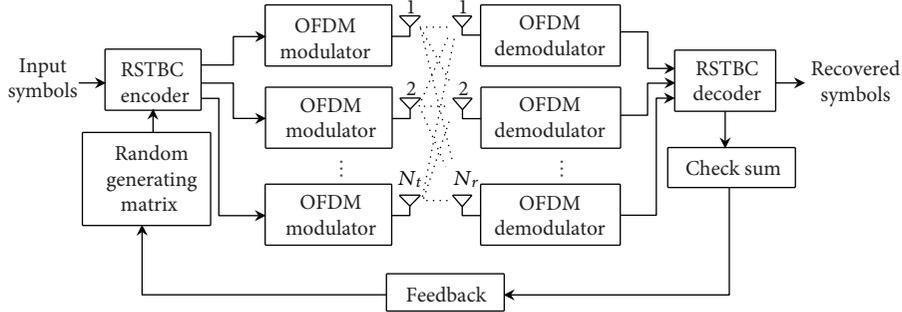


FIGURE 1: A general MIMO-OFDM system with RSTBC.

For additive white Gaussian noise (AWGN) channels, perfect rateless codes were constructed in [9]. Moreover, rateless coding was used for single-input single-output (SISO) fading channel in [10] and for MIMO fading channel in [11]. Initial work of rateless space-time coding was introduced in [12]. Nevertheless, the rateless space-time codes still require more research work, especially for massive MIMO systems. In [13], we proposed the use of RSTBC in massive MIMO systems having lossy characteristic, where we considered only simulation to test the performance of the system. Furthermore, we have applied RSTBC for antenna failure, loss-tolerant data transmission, and pilot contamination, over a massive MIMO system, in [14–16], respectively.

In this paper, however, we examine experimentally the performance of RSTBC in a MIMO-OFDM system and hence demonstrate the realization of this code to protect the system performance in a practical setup. Although we have included some sample experimental results in [16], this paper strictly provides the details of the test-bed used to get the results. Our outcomes prove that RSTBC is able to resolve high-rate losses (such as 25%) and achieve a reliable robust transmission. We consider such high-rate losses since some real traffic measurements revealed that some receivers experience very high loss rate of above 25% [17, 18].

The rest of the paper is organized as follows. System model is introduced in Section 2, then Section 3 characterizes the mechanism of RSTBC. The performance of the system is addressed in Section 4. In Section 5, the experimental setup, architecture, and operation are detailed. The results are presented and discussed in Section 6. Finally, Section 7 concludes the paper.

2. System Model

Consider a MIMO-OFDM system equipped with N_t transmit and N_r receive antennas. Figure 1 depicts the block diagram of a general MIMO-OFDM system in which RSTBC is implemented. The received matrix $\mathbf{Y} \in \mathbb{C}^{N_r \times L}$ is given by

$$\mathbf{Y} = \sqrt{\frac{E_s}{N_t N_0}} \mathbf{H} \mathbf{S} + \mathbf{W}, \quad (1)$$

where $\mathbf{S} \in \mathbb{C}^{N_t \times L}$ is the transmitted RSTBC matrix, L is the number of RSTBC blocks, $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is the channel transfer matrix, $\mathbf{W} \in \mathbb{C}^{N_r \times L}$ is the zero-mean Gaussian noise matrix

with unit-variance elements, E_s is the symbol energy, N_0 is the noise power spectral density at the receiver, and E_s/N_0 is the average signal-to-noise ratio (SNR) per receive antenna. The random generating matrix \mathbf{G} is used to construct the RSTBC matrix $\mathbf{S} = [\mathbf{s}_1 \ \mathbf{s}_2 \ \cdots \ \mathbf{s}_l \ \cdots \ \mathbf{s}_L]$ such that the column vectors of \mathbf{S} are generated as follows:

$$\mathbf{s}_l = \mathbf{x} \odot \mathbf{g}_l, \quad (2)$$

where $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_{N_t}]$ is the original vector to be transmitted by the system, $\mathbf{g}_l = [g_1 \ g_2 \ \cdots \ g_{N_t}]$ is the l th $N_t \times 1$ vector of random single-bit binary numbers ($g_n \in \{0, 1\}$) representing the instantaneous losses which occurred during transmission (namely, zeros represent losses and ones represent no losses), and \odot denotes the element-by-element multiplication (Hadamard product) [13]. In other words, the positions of zeros in \mathbf{g}_l represent dropped symbols in which $\mathbf{g}_1 \neq \mathbf{g}_2 \neq \cdots \neq \mathbf{g}_L$, and consequently, $\mathbf{s}_1 \neq \mathbf{s}_2 \neq \cdots \neq \mathbf{s}_L$.

3. Rateless Coding Mechanism

In this section, we simply exhibit the mechanism of a rateless code over a MIMO-OFDM system. Figure 2 depicts the RSTBC encoder and decoder, where a part of the encoded packets (or blocks) cannot be received due to losses [16]. In this case, the receiver keeps receiving until the minimum possible number of transmitted encoded packets are already received and enough to recover the original packets. The required number of blocks for recovery depends on the channel condition. During the transmission, for each block \mathbf{s}_l , the mutual information between the transmitter and the receiver is compared to a threshold [14]. Namely, it is desired to decode a message of total mutual information C . Assume that the required packets to deliver the message correctly are $[\mathbf{s}_1 \ \mathbf{s}_2 \ \cdots \ \mathbf{s}_l \ \cdots \ \mathbf{s}_L]$. Let c_l denote the measured mutual information after receiving the codeword block \mathbf{s}_l . If $c_l < C$, then the receiver keeps receiving further blocks, else if $c_l \geq C$, then receiver decodes the received codeword $[\mathbf{y}_1 \ \mathbf{y}_2 \ \cdots \ \mathbf{y}_l]$ to get $[\tilde{\mathbf{s}}_1 \ \tilde{\mathbf{s}}_2 \ \cdots \ \tilde{\mathbf{s}}_l]$ and a simple acknowledgement is sent to the transmitter to stop transmitting the residual encoded packets and remove them from the buffer.

It is worth noted that RSTBC is flexible, that is, the process continues until the receiver accumulates a sufficient

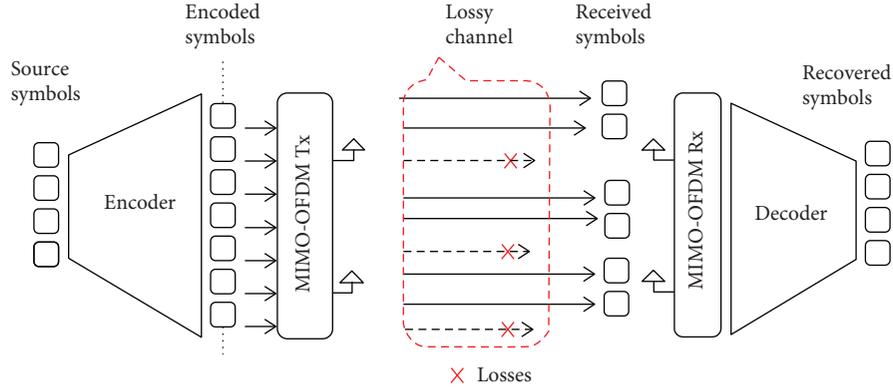


FIGURE 2: A rateless encoder and decoder in a lossy MIMO system.

number of blocks (L) to recover the message or the assigned time is over (it is limited by the channel coherence time). The decoding process is conducted sequentially; first using \tilde{s}_1 , then $[\tilde{s}_1 \ \tilde{s}_2]$, if \tilde{s}_1 is not sufficient, and so forth [13]. Once the checksum criterion is satisfied, the received blocks are linearly combined to decode the entire underlying message. We have to emphasize that the code is characterized as “rateless” because the number of required blocks (L) to recover the message is not predefined before transmission, rather it depends on the channel state. Furthermore, the code is extended ratelessly over time and space, and it belongs to the block codes. The code rate is flexible and can be calculated as the overall number of different transmitted symbols over the number of channel uses. Hence, this coding scheme is named as *rateless space-time block code* (RSTBC) [16].

4. System Performance Analysis

The received signal at the j th receive antenna can be expressed as

$$u_j = \sqrt{\frac{E_s}{LN_t N_o}} \sum_{l=1}^L \sum_{n=1}^{N_t} h_{jnl} g_{nl} x_n + w_j, \quad (3)$$

where $j = 1, 2, \dots, N_r$, $l = 1, 2, \dots, L$, $n = 1, 2, \dots, N_t$, x_n is the transmitted symbol by the n th antenna, h_{jnl} is the channel gain from the n th transmit antenna to the j th receive antenna for the l th block of RSTBC, and w_j is the noise at the j th receive antenna. Also, the loss is introduced by $g_{nl} \in \{0, 1\}$, where $g_{nl} = 0$ for loss and $g_{nl} = 1$ when there is no loss.

The maximum likelihood (ML) estimate \tilde{s}_l of the transmitted codeword s_l is expressed as

$$\tilde{s}_l = \arg \min_{s_l} \left\| \mathbf{y}_l - \sqrt{\frac{E_s}{LN_t N_o}} \sum_{l=1}^L \mathbf{H} \mathbf{s}_l \right\|_F^2, \quad (4)$$

where $\|\cdot\|_F^2$ denotes the squared Frobenius norm and $\mathbf{y}_l \in \mathbf{C}^{K \times 1}$ is the l th received vector. The conditional

probability of deciding erroneously in favor of $\tilde{\mathbf{S}}$ at the receiver while the codeword \mathbf{S} was transmitted is the pairwise error probability (PWE). Let $d_E(\mathbf{S}, \tilde{\mathbf{S}})$ denotes the Euclidean distance between \mathbf{S} and $\tilde{\mathbf{S}}$, which can be calculated by

$$d_E^2(\mathbf{S}, \tilde{\mathbf{S}}) = \sum_{l=1}^L \sum_{j=1}^{N_r} \left| \sum_{n=1}^{N_t} h_{jnl} g_{nl} (x_n - \tilde{x}_n) \right|^2. \quad (5)$$

The PWE can be expressed, in terms of Q-function, as

$$P(\mathbf{S} \rightarrow \tilde{\mathbf{S}} | \mathbf{H}) \leq Q \left(\sqrt{\frac{E_s}{2LN_t N_o}} d_E(\mathbf{S}, \tilde{\mathbf{S}}) \right). \quad (6)$$

Using the Chernoff bound, $Q(x) \leq (1/2)(e^{-x^2/2})$, the PWE in (6) is upper-bounded as

$$P(\mathbf{S} \rightarrow \tilde{\mathbf{S}} | \mathbf{H}) \leq \frac{1}{2} \exp \left(-\frac{E_s}{4LN_t N_o} d_E^2(\mathbf{S}, \tilde{\mathbf{S}}) \right). \quad (7)$$

Now, define $\mathbf{D} = (\mathbf{S} - \tilde{\mathbf{S}})(\mathbf{S} - \tilde{\mathbf{S}})^H$ as the codeword distance matrix which is a nonnegative definite Hermitian [19]. Hence, the singular value decomposition (SVD) of \mathbf{D} is calculated as

$$\mathbf{D} = \mathbf{B}^H \mathbf{\Lambda} \mathbf{B}, \quad (8)$$

where \mathbf{B} is a unitary matrix with orthogonal eigenvector columns $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_{N_t}\}$ and $\mathbf{\Lambda}$ is a diagonal matrix of nonnegative eigenvalues $\{b_1, b_2, \dots, b_{N_t}\}$. Consequently, this results in an N_t -dimensional vector space with orthonormal basis [14].

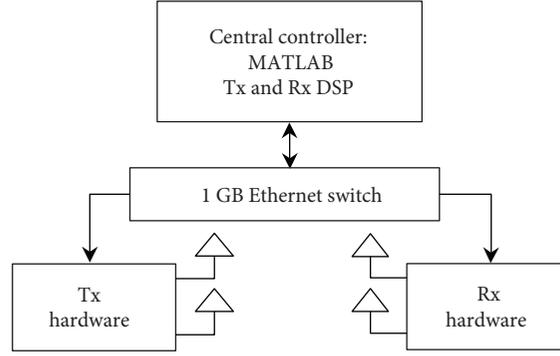


FIGURE 3: Block diagram of test-bed setup for a 2×2 MIMO system.

The squared modified Euclidean distance $d_E^2(\mathbf{S}, \tilde{\mathbf{S}})$ can be rewritten as

$$d_E^2(\mathbf{S}, \tilde{\mathbf{S}}) = \left\| \mathbf{H}(\mathbf{S} - \tilde{\mathbf{S}}) \right\|_F^2 = \sum_{l=1}^L \sum_{j=1}^{N_r} \sum_{n=1}^{N_t} g_{nl} b_n |\Omega_{jn}|^2, \quad (9)$$

where Ω_{jn} is the (j, n) th element of $\mathbf{H}\mathbf{B}$.

It is obvious from (9) that, compared to the free-loss case, and without rateless coding (i.e., $L = 1$), $d_E^2(\mathbf{S}, \tilde{\mathbf{S}})$ becomes smaller when losses occur since $g_{nl} = 0$ when there is loss and $g_{nl} = 1$ when there is no loss. Hence, as the loss rate increases, the upper bound of the PWEF, given in (7), becomes large.

This performance degradation (due to losses) is compensated by applying the RSTBC, since (9) becomes larger as L increases. Thus, the upper bound in (7) decreases, and consequently, the error performance is improved. The required number of redundant blocks of RSTBC depends on the loss rate [14].

Based on the well-known union bounding technique [20], the average bit error rate (BER) is analytically calculated and upper-bounded as

$$\text{BER} \leq \frac{1}{N_t L k^k} d_H(\mathbf{S}, \tilde{\mathbf{S}}) P(\mathbf{S} \rightarrow \tilde{\mathbf{S}} | \mathbf{H}), \quad (10)$$

where k is the number of bits per symbol and $d_H(\mathbf{S}, \tilde{\mathbf{S}})$ is the Hamming distance between the transmitted matrix \mathbf{S} and the estimated matrix $\tilde{\mathbf{S}}$.

5. Experimental Setup

The block diagram of the main components of the test-bed setup is depicted in Figure 3 which is implemented using 2 WARP v3 (Wireless Open-Access Research Project version 3) kits, 4 antennas (2 for each node) operating at 2.4 GHz, and connected through a 1 GB Ethernet switch to a personal computer (PC). Figure 4 shows the physical layout of the experimental test-bed.

WARP v3 kit, shown in Figure 5, is a software-defined radio (SDR) platform developed by Rice University and Mango Communications [21]. It is built on a Xilinx Virtex-6 LX240T FPGA with four programmable radio

frequency (RF) interfaces operating at 2.4 and 5 GHz with a 40 MHz bandwidth. Because of its accessibility and simple interface with MATLAB by which transmitted and received data are processed, the WARP v3 kit is used in many works such as [22] in which OFDM parameters are similar to what we have considered.

MATLAB is used to perform synchronization, modulation, RSTBC, channel estimation, and equalization. The power of the transmitted signal is controlled by the identifiers of both the baseband gain (BB-Gain) and RF gain (RF-Gain). In MATLAB, we can change the values of these identifiers to get a particular gain. We calibrated the transmitted power from the RF interfaces of the WARP v3 node by connecting these interfaces to a spectrum analyzer (SA) using a coaxial cable of 0.13 dB loss and transmitting a signal from the WARP v3 board to the SA. We found that a good transmission performance is obtained at transmit power values of -13.5 dBm and -15.9 dBm from the two transmit antennas, respectively.

It is obvious that the channel between the transmitter and the receiver is Rician distributed with different means and K -factors for each path. In this experiment, the Rician K -factors for all paths are estimated separately, where the values range between 27 dB and 30 dB.

The structure of the OFDM packet is based on IEEE 802.11g [23], in which each packet contains 30 OFDM symbols (each with 64 subcarriers) [22]. Twelve subcarriers are set to null, and four subcarriers are devoted for pilot signals which are utilized for frequency offset correction (FOC). The encoded data are inserted in the remaining 48 subcarriers. Each packet consists of three blocks: preamble, training symbols, and payload data. The preamble block is employed for timing synchronization and packet detection by means of two long OFDM training symbols [22, 23]. The training symbols block has 4 OFDM symbols used for channel estimation. This OFDM symbols are BPSK streams with two symbols for each transmit antenna while the remaining two OFDM symbols are null in each antenna for the purpose of beamforming.

The incoming binary data are applied to constellation mapper (QPSK or 16-QAM modulation). The modulated symbols are encoded by RSTBC encoder using a generating random matrix as in (2) to generate a number of blocks. Then, these blocks are transmitted sequentially until a feedback from the receiver terminates the transmission of the

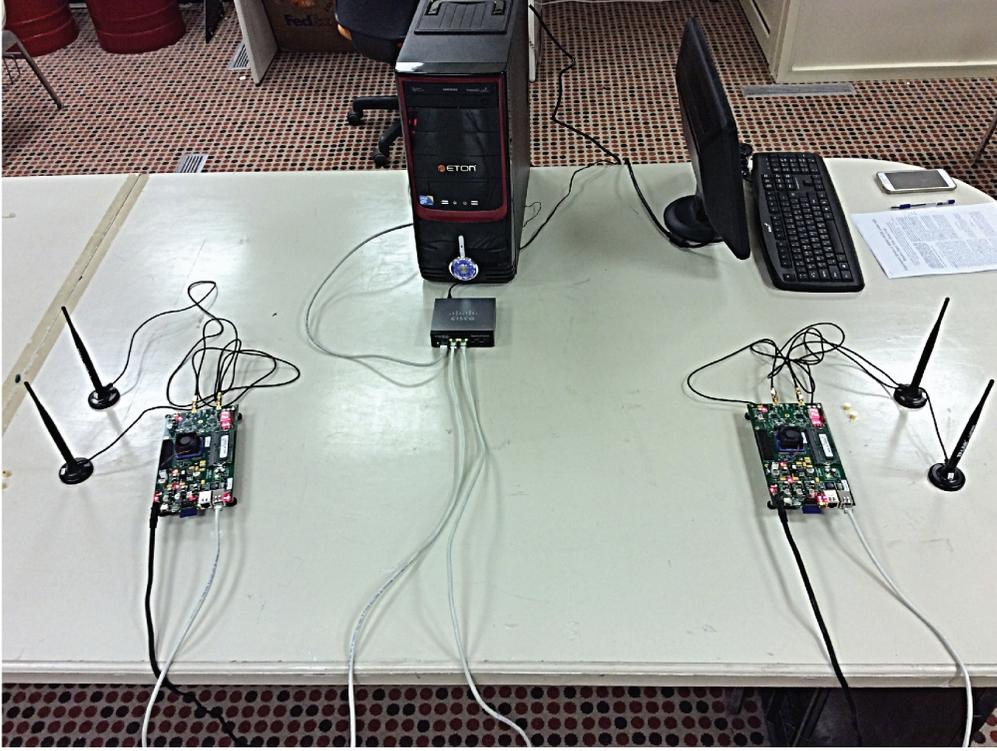


FIGURE 4: The physical layout of the experimental test-bed.



FIGURE 5: Top view of WARP v3 kit.

remaining blocks or all the blocks have been transmitted. For each transmission cycle, each block is reshaped into two streams each with 48 OFDM data subcarriers. The pilot samples are inserted and interleaved across subcarriers and OFDM symbols. At the beginning of the data OFDM words, the preamble and training symbols are embedded. The OFDM waveforms are generated using a 64-point inverse fast Fourier transform (IFFT). To avoid intersymbol interference (ISI) between OFDM symbols, a cyclic prefix (CP) is added to each OFDM symbol as a guard interval. The OFDM waveforms are converted from parallel to series. Upsampling with a factor of two is conducted to reduce the bandwidth to 20 MHz. In order to avoid attenuation near the DC, the packet is upconverted to 5 MHz (an intermediate frequency (IF)). Finally, the packet is extended to the dynamic range of the WARP digital-to-analog (D/A) and

TABLE 1: MIMO-OFDM parameters.

Parameter	Value
Number of Tx antennas, N_t	2
Number of Rx antennas, N_r	2
Number of OFDM symbols	1000
Nominal channel bandwidth	20 MHz
Number of data subcarriers, N_{data}	48
Number of pilot subcarriers, N_{pilot}	4
Number of guard subcarriers, N_{guard}	12
Ratio of guard time to useful symbol time, G	0.25
Number of used subcarriers, $N_{\text{used}} = N_{\text{data}} + N_{\text{pilot}}$	52
Number of FFT points, N_{FFT}	64
Sampling frequency, f_s	20 MHz
Subcarrier spacing, $\Delta f = \frac{f_s}{N_{\text{FFT}}}$	312.5 kHz
Used bandwidth, $BW_{\text{used}} = N_{\text{used}} \Delta f$	16.25 MHz
Used symbol time, $T_b = \frac{1}{\Delta f}$	3.2 μs
CP time, $T_g = G \cdot T_b$	0.8 μs
OFDM symbol time, $T_{\text{sym}} = T_b + T_g$	4 μs
Sampling time, $T_s = \frac{1}{f_s}$	50 ns
FFT offset	4
LTS correlation threshold	0.85

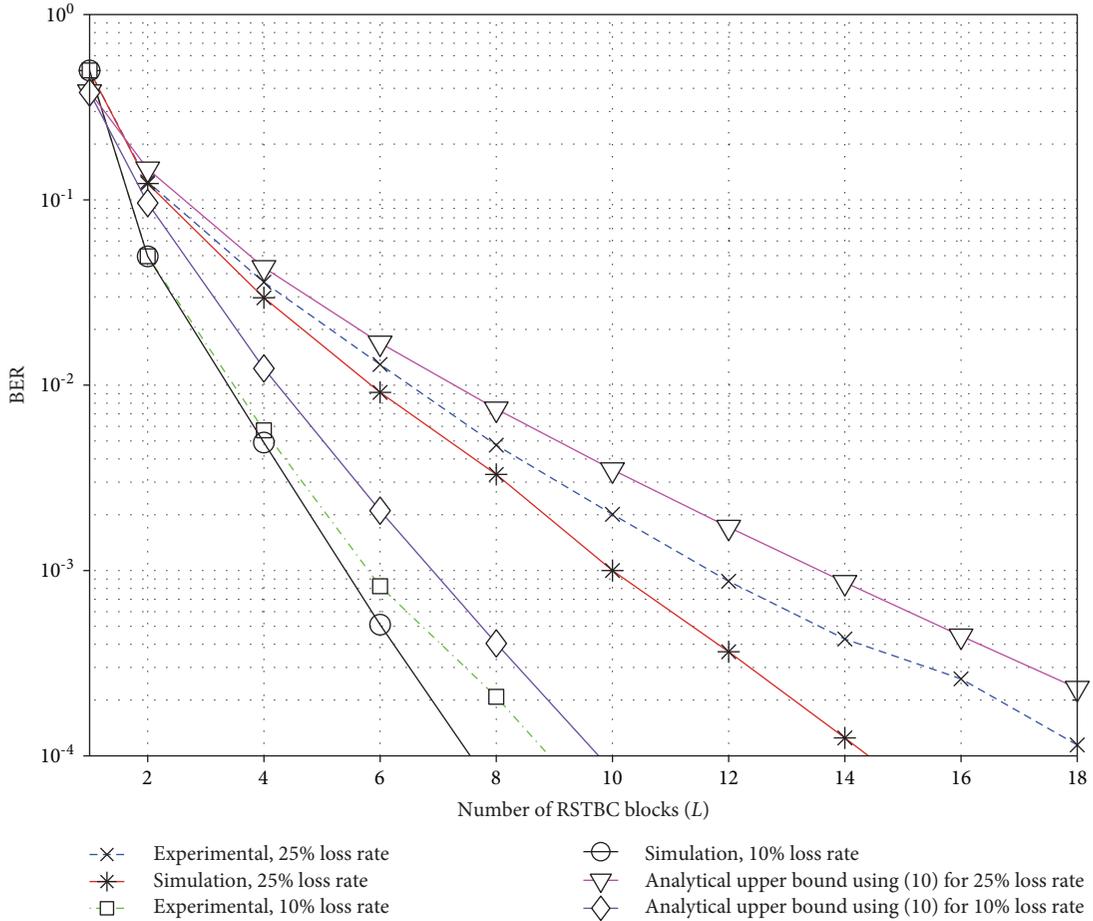


FIGURE 6: BER versus the number of blocks of RSTBC for experimental, simulation, and analytical upper bounds, in a 2×2 MIMO-OFDM system, with QPSK and loss rates of 10% and 25%. The left three curves are for 10% and the right three curves are for 25%.

analog-to-digital (A/D) converters, to enhance the quantization resolution and avoid clipping [22].

For each reception instant, the two sample streams of received data, with an additional of 300 samples for each stream, are extracted from the RF buffers of the WARP node. The added 300 samples are used for synchronization. The processes of synchronization and packet detection are obtained via cross-correlating the received data with the predetermined preamble sequence. Then, the received signal is downconverted from the RF band to baseband and downsampled by two to match the hardware bandwidth. All the received OFDM symbols are reshaped into parallel subcarriers, and the CP guard interval is removed. The OFDM symbols are then extracted from the received OFDM waveforms by applying a 64-point FFT. The MIMO channel matrix elements are estimated via the predetermined training symbols. The carrier phase offset (CFO), for each OFDM symbol, is estimated using the pilot signals and corrected by performing an inverse phase shift [22, 24]. The data payload is demodulated to extract the transmitted complex symbols. As described in Section 3, the receiver keeps receiving the transmitted blocks until the checksum criterion is achieved, then the available RSTBC blocks are linearly aggregated to recover the uncoded symbols.

6. Experimental Results and Discussion

In this section, we present the experimental results of the MIMO-OFDM setup having the parameters tabulated in Table 1. The experimental and simulation results of BER performance versus the number of RSTBC blocks in a MIMO-OFDM system, for QPSK and 16-QAM, are shown in Figures 6 and 7, respectively. For each figure, two loss rates, 10% and 25%, are considered. In Figure 6, we include also the BER upper bound calculated using (10).

It is worth noted that error performance is improved as the number of RSTBC blocks increases. That is because of the ability of RSTBC to compensate for the losses that may arise in the system. As the loss rate is high, more RSTBC redundancy is required to reach the target performance.

Specifically, for more illustration, from Figure 6, we summarized Table 2 for the 10% loss rate. It is noted that, without RSTBC, which corresponds to $L = 1$, BER is close to 0.5, while experimental BER is improved by approximately 83-fold when $L = 4$ which is, in turn, improved by 30-fold when $L = 8$. Similar discussion can be conducted for the 25% loss case but it requires more redundant RSTBC blocks to overcome the high loss rate.

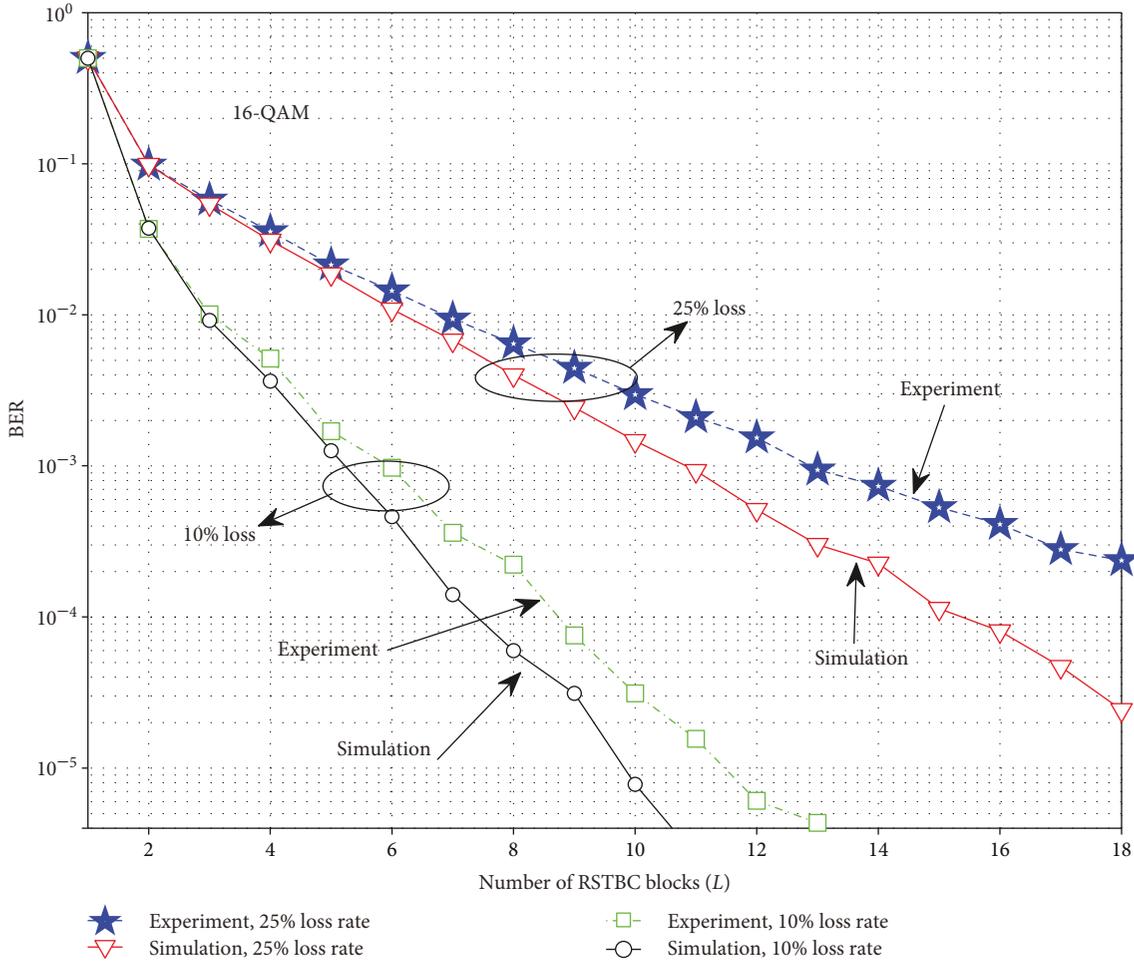


FIGURE 7: BER versus the number of blocks of RSTBC for experimental, simulation, and analytical upper bounds, in a 2×2 MIMO-OFDM system, with 16-QAM and loss rates of 10% and 25%.

TABLE 2: BER numerical results for 10% loss rate when RSTBC is applied with $L = 1, 2, 4, 6, 8$.

L	Experimental BER	Simulation BER
1	0.5	0.5
2	5×10^{-2}	5×10^{-2}
4	6×10^{-3}	5×10^{-3}
6	8×10^{-4}	5×10^{-4}
8	2×10^{-4}	0.625×10^{-4}

Comparing experimental and simulation values, it is inferred, in general, that for low number of blocks ($L \leq 4$), the experimental BER results agree approximately with that of the simulation. However, as L increases, the gap between the experimental and simulation BER increases. This difference, for high number of RSTBC blocks, results from practical errors in channel estimation or calibration. That is because of the difficulty of achieving exact channel estimation due to the noise, losses, and interference [16]. Similar discussion can be presented for Figure 7, which is for

16-QAM modulation format, taking into consideration the number of constellation symbols. Based on the previous results and discussion, this study has demonstrated the effectiveness of RSTBC for deployment in lossy wireless communication systems.

7. Conclusion

In this paper, we examined practically a rateless space-time block code (RSTBC) in a MIMO-OFDM setup and analyzed the BER performance of the scheme. The experimental results demonstrated that RSTBC can be employed to reserve the reliability of the system in a lossy environment. In a 2×2 MIMO-OFDM test-bed, with QPSK and 16-QAM, we have shown that, from a certain amount of encoded data, RSTBC is able to recover the original data even when the loss rate is 10% or 25%. The experimental results agree closely with the simulation and analytical upper bounds. The proposed model, the developed analysis, and the experimental results along with the simulation are considered as a good reference for designing, performance analyzing, and prototyping of RSTBC in a loss-prone massive MIMO systems.

Data Availability

The measurements used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors acknowledge the College of Engineering Research Centre, King Saud University, and the Deanship of Scientific Research at King Saud University in Riyadh, Saudi Arabia, for the financial support to carry out the research work reported in this paper.

References

- [1] LAN MAN Standards Committee and IEEE-SA Standards Board, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE, 2003.
- [2] C. Eklund, R. B. Marks, K. L. Stanwood, and S. Wang, "IEEE standard 802.16: a technical overview of the WirelessMAN/sup TM/ air interface for broadband wireless access," *IEEE Communications Magazine*, vol. 40, no. 6, pp. 98–107, 2002.
- [3] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 744–765, 1998.
- [4] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [5] H. El Gamal, G. Caire, and M. O. Damen, "The MIMO ARQ channel: diversity-multiplexing-delay tradeoff," *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3601–3621, 2006.
- [6] J. W. Byers, M. Luby, M. Mitzenmacher, and A. Rege, "A digital fountain approach to reliable distribution of bulk data," *ACM SIGCOMM Computer Communication Review*, vol. 28, no. 4, pp. 56–67, 1998.
- [7] M. Luby, "LT codes," in *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings*, pp. 271–280, Vancouver, BC, Canada, 2002.
- [8] A. Shokrollahi, "Raptor codes," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2551–2567, 2006.
- [9] U. Erez, M. D. Trott, and G. W. Wornell, "Rateless coding for Gaussian channels," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 530–547, 2012.
- [10] J. Castura and Y. Mao, "Rateless coding over fading channels," *IEEE Communications Letters*, vol. 10, no. 1, pp. 46–48, 2006.
- [11] Y. Fan, L. Lai, E. Erkip, and H. Poor, "Rateless coding for MIMO fading channels: performance limits and code construction," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1288–1292, 2010.
- [12] U. Erez, G. W. Wornell, and M. D. Trott, "Rateless space-time coding," in *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005*, pp. 1937–1941, Adelaide, South Australia, Australia, 2005.
- [13] A. H. Alqahtani, A. I. Sulyman, and A. Alsanie, "Rateless space time block code for massive MIMO systems," *International Journal of Antennas and Propagation*, vol. 2014, Article ID 154261, 10 pages, 2014.
- [14] A. H. Alqahtani, A. I. Sulyman, and A. Alsanie, "Rateless space time block code for antenna failure in massive MU-MIMO systems," in *2016 IEEE Wireless Communications and Networking Conference*, pp. 1–6, Doha, Qatar, April 2016.
- [15] A. H. Alqahtani, A. I. Sulyman, and A. Alsanie, "Loss-tolerant large-scale MU-MIMO system with rateless space time block code," in *2016 22nd Asia-Pacific Conference on Communications (APCC)*, pp. 342–347, Yogyakarta, Indonesia, August 2016.
- [16] A. H. Alqahtani, A. I. Sulyman, and A. Alsanie, "Rateless space-time block code for mitigating pilot contamination effects in multi-cell massive MIMO system with lossy links," *IET Communications*, vol. 10, no. 16, pp. 2252–2259, 2016.
- [17] S. K. Chin and R. Braun, "A survey of UDP packet loss characteristics," in *Conference Record of Thirty-Fifth Asilomar Conference on Signals, Systems and Computers (Cat.No.01CH37256)*, vol. 1, pp. 200–204, Pacific Grove, CA, USA, 2001.
- [18] M. Fleury and N. Qadri, *Streaming Media with Peer-to-Peer Networks: Wireless Perspectives*, IGI Global, 2012.
- [19] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications with MATLAB®*, John Wiley & Sons, 2010.
- [20] M. Di Renzo and H. Haas, "Bit error probability of SM-MIMO over generalized fading channels," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 3, pp. 1124–1144, 2012.
- [21] P. Murphy, A. Sabharwal, and B. Aazhang, "Design of WARP: a wireless open-access research platform," in *Signal processing conference, 2006 14th European. IEEE*, pp. 1–5, Florence, Italy, 2006.
- [22] R. Measel, D. J. Bucci, C. S. Lester et al., "A MATLAB platform for characterizing MIMO-OFDM communications with software-defined radios," in *2014 IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR)*, pp. 1–6, Tucson, AZ, USA, May 2014.
- [23] ICSLMS Committee and IEEE Standards Board, *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE, 1997.
- [24] M. Jankiraman, *Space-Time Codes and MIMO Systems*, Artech House, 2004.

Research Article

Phase-Only Antenna Array Reconfigurability with Gaussian-Shaped Nulls for 5G Applications

Giulia Buttazzoni , Massimiliano Comisso , Federico Ruzzier, and Roberto Vescovo 

Department of Engineering and Architecture, University of Trieste, Via A. Valerio 10, 34127 Trieste, Italy

Correspondence should be addressed to Giulia Buttazzoni; gbuttazzoni@units.it

Received 30 August 2018; Accepted 19 November 2018; Published 7 February 2019

Guest Editor: Raed A. Abd-Alhameed

Copyright © 2019 Giulia Buttazzoni et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a fast iterative method for the synthesis of linear and planar antenna arrays of arbitrary geometry that provides pattern reconfigurability for 5G applications. The method enables to generate wide null regions shaped according to a Gaussian distribution, which complies with recent measurements on millimeter-wave (mmWave) angular dispersion. A phase-only control approach is adopted by moving from the pattern provided by a uniformly excited array and iteratively modifying the sole phases of the excitations. This allows the simplification of the array feeding network, hence reducing the cost of realization of 5G base stations and mobile terminals. The proposed algorithm, which is based on the method of successive projections, relies on closed-form expressions for both the projectors and the null positions, thus allowing a fast computation of the excitation phases at each iteration. The effectiveness of the proposed solution is checked through numerical examples compliant with 5G mmWave scenarios and involving linear and concentric ring arrays.

1. Introduction

By now, the expected outbreak of the 5G technology represents one of the main challenging opportunities for antenna designers, whose efforts will have to deal with several significant designs, realizations, and prototyping issues. One of the distinctive features of next-generation cellular systems will in fact consist in the adoption of the millimeter-wave (mmWave) spectrum, between 30 and 300 GHz, to satisfy the huge capacity demand that will require the implementation of the Internet of Things (IoT) and Internet of Everything (IoE) paradigms [1]. The main reasons that have driven the attention of the 5G developers towards the extremely high-frequency (EHF) band are the considerable amount of unused bandwidth and the possibility of packaging many radiating elements in a single mobile terminal (MT) or base station (BS), with the aim of providing a high gain, so as to compensate the significant attenuation of the mmWave channel. For the first time, from the deployment of the previous 1-4G systems, the 5G network will indeed have to necessarily rely on beamforming both in transmission and reception, to make the communications feasible in

a propagation environment much more sensitive to blockages due to obstacles [2–4].

In this context, a basic approach to find the most suitable beamforming techniques for mmWave links may be that of considering the significant amount of algorithms that have been developed along the years and that gave outstanding solutions for a wide variety of synthesis problems. To make a preliminary effective selection, it may be also useful to properly identify the main antenna requirements that are expected to characterize the forthcoming 5G transceivers. Firstly, since mmWave systems will be designed to manage a dense environment, the spatial filtering capabilities of the antenna array will represent the main topic, to the point that many studies have assumed a significant change of perspective for the analytical characterization of the 5G performance, consisting in the translation from the so far assumed 1-4G interference-limited regime to a 5G noise-limited one [5]. In this scenario, the shape of the pattern and, precisely, its beamwidth and side-lobe level (SLL), as well as the depth and width of its possible nulls, will play a key role for the sustainability of the 5G links. A second relevant antenna requirement concerns the feeding network.

Since the significant mmWave attenuations will force the developers to adopt small cells, a huge number of BSs is expected to be deployed. This leads to the need of containing the fabrication costs not only of the MTs but also of the BSs and, in turn, of the corresponding antenna systems, thus orienting the choice of the beamforming algorithm towards a phase-only control approach, which leads to a simplified feeding network.

In light of these considerations, this paper proposes a fast synthesis algorithm for linear and planar antenna arrays of arbitrary geometry that allows the reconfigurability of the pattern and the shaping of the null regions according to a Gaussian power azimuth spectrum. The algorithm takes, as starting point, a uniformly excited array, which provides a narrow beamwidth, and then iteratively modifies the sole phases of the excitations to impose wide Gaussian nulls in specified directions. Furthermore, the algorithm, which relies on the method of successive projections, is developed to obtain closed-form expressions for the projectors and analytical formulas for the positions of the nulls. This approach has the advantage of leading to a fast computation of the excitation phases, since no numerical estimations are required at each iteration. Numerical examples considering linear and concentric ring arrays are presented to assess the effectiveness of the proposed solution in a 5G interfered context, discussing, beside the performance in terms of obtained patterns, also the computational time necessary for their synthesis.

The paper is organized as follows. Section 2 presents the related work. Section 3 formulates the problem. Section 4 describes the development of the algorithm. Section 5 discusses the numerical results. Section 6 summarizes the most relevant conclusions.

2. Related Work

Many phase-only null synthesis techniques for antenna arrays proposed in the recent years [6–18] may be nowadays reconsidered for addressing the expected 5G requirements. In particular, a deterministic method for linear arrays is presented in [6], where the phase-only approach and the possibility to generate multiple close nulls are outlined. Two iterative solutions for arbitrary arrays are derived [7], with the aim of exploiting a root mean square approximation to enable null synthesis by phase-only control. In [8], the problem of assigning prescribed nulls in the pattern of a linear array by phase-only control is solved through the development of a genetic algorithm. A phase-only null synthesis method for linear arrays is also discussed in [9], where the authors adopt a sequential quadratic programming approach. Scenarios specifically considering the phase-only synthesis of wide nulls for arrays of arbitrary geometry are addressed in [10], by taking into account the power azimuth spectrum of the interferers. In [11], a beamforming architecture based on neural networks is designed to introduce phase-only adaptive nulling in phased arrays. The synthesis of cylindrical arc antenna arrays controlled by the excitation phases is analyzed in [12], where the null steering capabilities are experimentally proved through a prototype consisting of

microstrip patches. In [13], the authors propose a differential search optimization algorithm for linear arrays, which includes the phase-only control and wide null synthesis options. The same options are enabled in [14], which alternatively applies a metaheuristic backtracking search optimization approach based on an iterative process controlled by a single parameter. The method of alternate projections is adopted in [15], to develop an iterative algorithm for null synthesis problems in arrays of arbitrary geometry, which includes the possibility of imposing the phase-only constraint. In [16], a phase-only beamformer based on the adaptive bat algorithm is designed to impose nulls in a certain number of undesired directions when a uniformly spaced linear array of half-wave dipoles is employed. In [17], a powerful method for the reconfigurability and beam scanning with phase-only control for arbitrary antenna arrays is presented, which also allows to form wide deep nulls. The method is based on a smart application of the alternate projection approach. A versatile solution is presented in [18], where a weighted cost function is used to impose multiple synthesis requirements, including phase-only control and null steering for conformal antenna arrays.

With reference to this overview, some common aspects may be highlighted. First, the phase-only requirement during the null synthesis process implies the maintenance of a unity dynamic range ratio (DRR), representing the ratio between the maximum and the minimum excitation amplitudes of the array. The problem of forming nulls while imposing an upper bound on the DRR has been discussed in detail in [19], where a theorem providing a necessary and sufficient condition for the joint application of null and DRR constraints on any array geometry is mathematically proved. In particular, it has been statistically shown that, if the number of required nulls is small compared to the number of array elements, such condition is typically satisfied also for unity DRR, thus allowing to form exact nulls (and a fortiori deep nulls) by phase-only control, which is here the case of main interest.

Second, many of the proposed solutions are applicable just to specific configurations [6, 8, 9, 11–14, 16], usually having a linear geometry. However, planar structures are expected to be used for 5G MTs and BSs [2], thus an algorithm capable to operate on conformal structures may be highly preferable.

Third, the recent channel measurements carried out at the mmWave frequencies have revealed that the angular dispersion may be often described by a Gaussian distribution [20, 21], thus better identifying the shape of the wide nulls that may be necessary to suppress the undesired sources. Currently, none of the existing methods combines wide nulling and Gaussian shaping together with a low computation time, which represents a not negligible requirement for 5G cellular systems, whose operations will have to necessarily adopt fast beamforming solutions. The specific case of Gaussian null shaping is considered in [10], but passing through a numerical estimation of the position of the nulls, which does not rely on closed-form expressions, and hence may become computationally cumbersome.

For these reasons, in this paper, we present a phase-only control algorithm for arbitrary linear and planar arrays capable of managing Gaussian-distributed nulls. Moving from the

formulation of the problem developed in the next section, the algorithm is mathematically derived in Section 4.

3. Problem Formulation

With reference to a Cartesian system $O(x, y, z)$, consider an antenna array of N elements lying on the yz -plane, where the position of the n -th element is specified by the vector $\bar{\mathbf{d}}_n = x_n \hat{\mathbf{x}} + y_n \hat{\mathbf{y}} + z_n \hat{\mathbf{z}}$ (denoting by $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ the unit vectors of the Cartesian coordinate axes x , y , and z , respectively). In the generic space direction $\hat{\mathbf{r}}$, the radiation pattern of this array is given by

$$P(\mathbf{i}; \hat{\mathbf{r}}) = \sum_{n=1}^N i_n p_n(\hat{\mathbf{r}}) \exp(j\beta \bar{\mathbf{d}}_n \cdot \hat{\mathbf{r}}), \quad (1)$$

where $\mathbf{i} = [i_1, \dots, i_N]^T$ is the column vector of the complex excitations, $p_n(\hat{\mathbf{r}})$ is the n -th array element pattern, $\beta = 2\pi/\lambda$ is the wave number, with λ denoting the wavelength, j is the imaginary unit, and $\hat{\mathbf{r}} = \sin \theta \cos \phi \hat{\mathbf{x}} + \sin \theta \sin \phi \hat{\mathbf{y}} + \cos \theta \hat{\mathbf{z}}$, with $\phi \in [-\pi, \pi]$ and $\theta \in [0, \pi]$ denoting the azimuth and zenith angles, respectively. In the sequel, for more simplicity, we will assume that the direction of observation belongs to the xy -plane, so that (1) can be expressed as a function of the azimuth angle as follows:

$$P(\mathbf{i}; \phi) = \sum_{n=1}^N i_n p_n(\phi) \exp(j\beta y_n \sin \phi). \quad (2)$$

Consider now a reference pattern $P_0 = P(\mathbf{i}_0; \phi)$, with the main beam pointing at a desired direction ϕ_0 , and an interferer whose angle of arrival (AoA) is statistically described by a truncated Gaussian distribution having probability density function (pdf):

$$f_g(\phi) = \frac{Q_L}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\phi - \bar{\phi})^2}{2\sigma^2}\right] \mathbf{1}_{[-\pi, \pi]}(\phi), \quad (3)$$

where Q_L is a constant that imposes the normalization condition $\int_{-\pi}^{\pi} f_g(\phi) d\phi = 1$, σ is the standard deviation, which can be inferred from experimental measurements in the mmWave channel [20, 21], $\bar{\phi}$ is the mean AoA, and $\mathbf{1}_{\mathbf{X}}(\mathbf{x})$ denotes the indicator function (that is, $\mathbf{1}_{\mathbf{X}}(\mathbf{x}) = 1$ if $\mathbf{x} \in \mathbf{X}$ and $\mathbf{1}_{\mathbf{X}}(\mathbf{x}) = 0$ if $\mathbf{x} \notin \mathbf{X}$).

According to this scenario, the objective of this study is to find a pattern $P = P(\mathbf{i}; \phi)$ that solves the following minimization problem:

$$\min_{\mathbf{i}} \|P - P_0\|^2 \quad (4)$$

$$\text{subject to } |i_n| = |i_{0n}|, \quad n = 1, \dots, N \quad (5)$$

$$|P(\mathbf{i}; \phi)| \leq Q f_g(\phi), \quad \phi \in [\phi_l, \phi_u], \quad (6)$$

where $\mathbf{i}_0 = [i_{01}, \dots, i_{0N}]^T$, $\|\cdot\|$ is a suitable norm, Q is a constant taking into account the amplitude of the interferer,

and $[\phi_l, \phi_u]$ identifies the angular region of interest. Note that the constraint in (5) means that only the excitation phases of \mathbf{i}_0 are modified.

4. The Solving Procedure

The adopted synthesis strategy starts by substituting the wide null constraint in (6) with the following M single null constraints:

$$P(\mathbf{i}; \phi_m) = 0, \quad m = 1, \dots, M, \quad (7)$$

that is, imposing that the pattern vanishes at the directions ϕ_1, \dots, ϕ_M . In matrix form, (7) can be expressed as follows:

$$\mathbf{E}\mathbf{i} = \mathbf{0}, \quad (8)$$

where $\mathbf{E} = [E_{mn}]$, with $E_{mn} = p_n(\phi_m) \exp(j\beta y_n \sin \phi_m)$ for $m = 1, \dots, M$ and $n = 1, \dots, N$. As shown in [10], a proper choice of the null positions allows one to satisfy the initial constraint in (6) by reformulating the original problem to an equivalent one solvable in a simple and very fast way. The derived solution is illustrated in the sequel of this section. More precisely, the strategy to suitably select the M null directions is described in the following subsection, while the procedure of phase-only M null synthesis is presented in Section 4.2.

4.1. Null Positioning. The localization of the nulls is performed here by a density tapering technique, which consists in finding the M null directions by imposing an equiareal requirement as follows:

$$\int_{-\pi}^{\phi_1} f_g(\phi) d\phi = \int_{\phi_1}^{\phi_2} f_g(\phi) d\phi = \dots = \int_{\phi_M}^{\pi} f_g(\phi) d\phi. \quad (9)$$

Since $\int_{-\pi}^{\pi} f_g(\phi) d\phi = 1$, each integral in (9) must be equal to $1/(M+1)$. Hence, imposing $\int_{-\pi}^{\phi_m} f_g(\phi) d\phi = m/(M+1)$ for $m = 1, \dots, M$, after some manipulations, one obtains [22]

$$\phi_m = \sqrt{2}\sigma \operatorname{erf}^{-1}\left[\frac{2m}{Q_L(M+1)} + \operatorname{erf}\left(-\frac{\pi + \bar{\phi}}{\sqrt{2}\sigma}\right)\right] + \bar{\phi}, \quad (10)$$

where $\operatorname{erf}(x)$ denotes the error function and $\operatorname{erf}^{-1}(x)$ represents its inverse. In this way, the nulls, whose position is available in analytical form, are more dense near the mean value $\bar{\phi}$, where the pdf is higher, and less dense far from $\bar{\phi}$, where the pdf is lower. Once the null directions are evaluated by (10), the reformulated problem is then solved by the alternating projection approach, as described in detail in the next subsection.

4.2. Phase-Only Null Synthesis by Alternating Projections. In order to model the synthesis problem as an intersection finding problem, first denote \mathcal{P} as the set of all the patterns that can be generated by the considered array. Then, in \mathcal{P} , introduce a set \mathcal{H} , composed by all the patterns that are produced by an excitation vector that satisfies constraint (5), and a set

\mathcal{L} , composed by all the patterns that satisfy constraint (7). It is evident that a radiation pattern belonging to both sets \mathcal{H} and \mathcal{L} (if any), and having the minimum distance from P_0 , is a solution to our (reformulated) problem, since it satisfies (4), (5), and (7). Thus, a point very close to P_0 and belonging to the intersection of the two sets is sought. If $\mathcal{H} \cap \mathcal{L}$ is empty, such a point does not exist, and we search for a point of \mathcal{H} having the minimum distance from \mathcal{L} , so as to obtain deep nulls by phase-only control. To this aim, adopting the alternating projection approach [17, 23], an iterative procedure is performed by starting from the reference pattern $P_0 \in \mathcal{H}$ and then following the scheme:

$$P_n = T_{\mathcal{H}} T_{\mathcal{L}} P_{n-1}, \quad n = 1, 2, 3, \dots, \quad (11)$$

where $P_{\mathcal{H}}$ and $P_{\mathcal{L}}$ denote the projection operators onto the sets \mathcal{H} and \mathcal{L} , respectively. Each pattern of the sequence $\{P_n\}$ belongs to the set \mathcal{H} ; thus, it satisfies the phase-only requirement, and is closer and closer to the set \mathcal{L} , due to a well-known property of the alternating projection method, $\|P_n - T_{\mathcal{L}} P_n\| \geq \|P_{n+1} - T_{\mathcal{L}} P_{n+1}\|$ for each $n \geq 1$. The iterations are stopped at a point $P_n \in \mathcal{H}$ such that

$$\begin{aligned} d_n &< \varepsilon_1 \\ \text{or } \frac{d_{n-1} - d_n}{d_n} &< \varepsilon_2, \end{aligned} \quad (12)$$

where d_n is the distance between the pattern P_n and the set \mathcal{L} , while ε_1 and ε_2 are proper positive thresholds.

It is worth noting that the iteration is stopped at a point of \mathcal{H} ; thus, the constraint in (5) is rigorously satisfied, and hence phase-only control is achieved. In general, $P_n \notin \mathcal{L}$; thus, the null constraints in (7) are not exactly satisfied, but sufficiently deep nulls are achieved after a suitable number of iterations. Hence, the constraints in (7) and, in turn, the original one in (6), are approximated very well. Since the starting point is P_0 , the alternating projection approach provides a point close to P_0 , approximately satisfying (4) and simultaneously satisfying (5) (exactly) and (7) (approximately). This objective is achieved by properly defining the projectors $T_{\mathcal{H}}$ and $T_{\mathcal{L}}$ in (11), whose mathematical derivation is described in detail in the next paragraphs.

4.2.1. The Projector $T_{\mathcal{H}}$. The algorithm for this projector is described here following the procedure in Section V of [19]. Projecting an array pattern $P(\mathbf{i}_x; \phi)$ onto the set \mathcal{H} requires to find an array pattern $P(\mathbf{i}; \phi) \in \mathcal{H}$ minimizing the distance:

$$\delta(\mathbf{i}) = \|P(\mathbf{i}; \phi) - P(\mathbf{i}_x; \phi)\|. \quad (13)$$

The distance in (13) is defined by means of the norm $\|f(\phi)\| = \sqrt{\langle f, f \rangle}$, involving the scalar product $\langle f, g \rangle = \int_{-\pi}^{\pi} f(\phi) g^*(\phi) d\phi$, where the asterisk denotes the complex

conjugate. Substituting (2) into (13), after some manipulations, one obtains

$$\delta^2(\mathbf{i}) = (\mathbf{i} - \mathbf{i}_x)^H \mathbf{G} (\mathbf{i} - \mathbf{i}_x) = \sum_{m=1}^N (i_m^* - i_{xm}^*) \sum_{n=1}^N G_{mn} (i_n - i_{xn}), \quad (14)$$

where the superscript H denotes the transposed conjugate and $\mathbf{G} = [G_{mn}]$ is an $N \times N$ matrix whose generic element is given by

$$G_{mn} = \langle p_n(\phi) \exp(j\beta y_n \sin \phi), p_m(\phi) \exp(j\beta y_m \sin \phi) \rangle, \quad (15)$$

for $m, n = 1, \dots, N$. Since the unknown pattern has to belong to the set \mathcal{H} , it must satisfy constraint (5). Thus, we set

$$i_n = \rho_{0n} \exp(j\alpha_n), \quad (16)$$

where $\rho_{0n} = |i_{0n}|$ for $n = 1, \dots, N$. Substituting (16) into (14) and putting the generic p -th term into evidence, yields

$$\begin{aligned} \delta^2(\mathbf{i}) &= G_{pp} \left| \rho_{0p} \exp(j\alpha_p) - i_{xp} \right|^2 + H_p \left[\rho_{0p} \exp(-j\alpha_p) - i_{xp}^* \right] \\ &\quad + H_p^* \left[\rho_{0p} \exp(j\alpha_p) - i_{xp} \right] + K_p, \end{aligned} \quad (17)$$

where

$$\begin{aligned} H_p &= \sum_{n \neq p} G_{pn} [\rho_{0n} \exp(j\alpha_n) - i_{xn}], \\ K_p &= \sum_{m \neq p} [\rho_{0m} \exp(-j\alpha_m) - i_{xm}^*] \times \sum_{n \neq p} G_{mn} [\rho_{0n} \exp(j\alpha_n) - i_{xn}]. \end{aligned} \quad (18)$$

The unknown phases are found following the single coordinate method (SCM) [7], beginning from the phases of the excitations of the reference pattern. Accordingly, at each step the $N - 1$ variables $\alpha_1, \dots, \alpha_{p-1}, \alpha_{p+1}, \dots, \alpha_N$ are considered as known, and the value of the unknown α_p that minimizes (17) is obtained by imposing the vanishing of the derivative of $\delta^2(\mathbf{i})$ with respect to α_p . This yields

$$\alpha_p = \arg(S_p) + k\pi, \quad k = 0, \pm 1, \pm 2, \dots, \quad (19)$$

where

$$S_p = \rho_{0p} (G_{pp} i_{xp} - H_p). \quad (20)$$

Calculating (19) for $p = 1, \dots, N$ and repeating the iteration for a sufficient number of times yield the unknown phases, which, once substituted in (16), give the required excitation vector. Substituting the latter into (2) gives the radiation pattern $P(\mathbf{i}; \phi)$, projection of $P(\mathbf{i}_x; \phi)$ onto \mathcal{H} .

4.2.2. *The Projector $T_{\mathcal{Z}}$* . The algorithm for this projector has been illustrated in [24] for the case of nulls in the near-field region. It is described here in detail for the case of nulls in the far-field region, which is of interest for the presently addressed scenario. Given any pattern $P(\mathbf{i}_x; \phi)$, the pattern $P(\mathbf{i}; \phi) = T_{\mathcal{Z}}P(\mathbf{i}_x; \phi)$, projection of $P(\mathbf{i}_x; \phi)$ onto \mathcal{Z} , is the pattern that minimizes the squared distance (14) subject to the constraints (7). Here, we note that the matrix \mathbf{G} in (14) is Hermitian. Therefore, it can be written as follows:

$$\mathbf{G} = \mathbf{U}^H \mathbf{\Lambda} \mathbf{U}, \quad (21)$$

where $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal elements are the eigenvalues of \mathbf{G} , while \mathbf{U} is a unitary matrix whose rows are the (orthonormal) eigenvectors corresponding to the eigenvalues. Substituting (21) into (14) and setting $\mathbf{w} = \mathbf{\Lambda}^{1/2} \mathbf{U} \mathbf{i}$ and $\mathbf{w}_x = \mathbf{\Lambda}^{1/2} \mathbf{U} \mathbf{i}_x$ yield

$$\delta^2(\mathbf{i}) = (\mathbf{w} - \mathbf{w}_x)^H (\mathbf{w} - \mathbf{w}_x). \quad (22)$$

On the other hand, being $\mathbf{w} = \mathbf{\Lambda}^{1/2} \mathbf{U} \mathbf{i}$ we can write $\mathbf{i} = \mathbf{K} \mathbf{w}$, where $\mathbf{K} = \mathbf{U}^H \mathbf{\Lambda}^{-1/2}$. Thus, the constraints in (7) and (8) can be written as $\mathbf{S} \mathbf{w} = 0$, where $\mathbf{S} = \mathbf{E} \mathbf{K}$. Therefore, the problem of projecting $P(\mathbf{i}_x; \phi)$ onto \mathcal{Z} reduces to that of determining the column vector \mathbf{w} that minimizes (22) subject to condition $\mathbf{S} \mathbf{w} = 0$. As it is well known, this problem has the solution $\mathbf{w} = (\mathbf{I}_N - \mathbf{S}^\dagger \mathbf{S}) \mathbf{w}_x$, where \mathbf{I}_N is the $N \times N$ identity matrix and \mathbf{S}^\dagger is the pseudoinverse of \mathbf{S} . Being $\mathbf{w}_x = \mathbf{\Lambda}^{1/2} \mathbf{U} \mathbf{i}_x$, it results

$$\mathbf{i} = \mathbf{P}_{\mathcal{Z}} \mathbf{i}_x, \quad (23)$$

where

$$\mathbf{P}_{\mathcal{Z}} = \mathbf{K} (\mathbf{I}_N - \mathbf{S}^\dagger \mathbf{S}) \mathbf{\Lambda}^{1/2} \mathbf{U}. \quad (24)$$

The column vector \mathbf{i} in (23) provides the complex excitations, which, once inserted in (2) give the projected array pattern $P(\mathbf{i}; \phi) = T_{\mathcal{Z}}P(\mathbf{i}_x; \phi)$.

Now that the projectors have been derived, the iterative procedure in (11) with starting point P_0 can be carried out to solve the original phase-only synthesis problem with Gaussian-shaped null in (4)–(6). The performances achievable by this method are investigated in the next section.

5. Numerical Results

In this section, two numerical examples are proposed to assess the validity of the proposed procedure. The two examples are also solved with the first method developed in [7], which can perform phase-only null synthesis for arbitrary arrays and is hence suitable for comparison with that proposed in this paper. Both the synthesis methods are implemented in MATLAB R2015b, and all results are obtained using a laptop equipped with an Intel® Core™ i5-5300U CPU @ 2.30 GHz with 8 GB RAM.

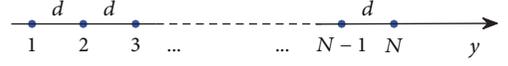


FIGURE 1: First example: geometry of the linear array with $N = 40$ elements and interelement distance $d = \lambda/2$.

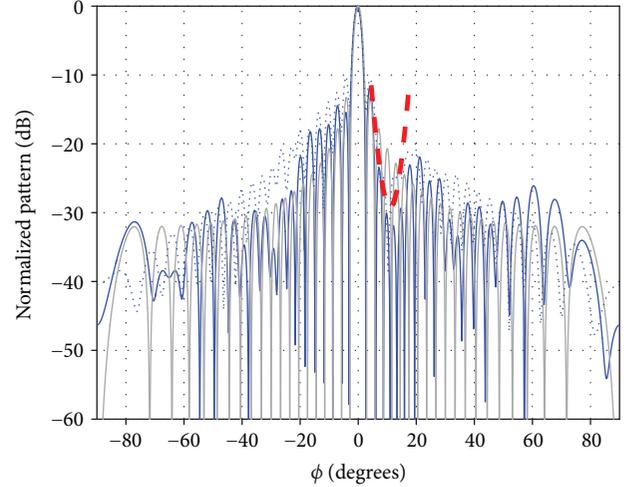


FIGURE 2: First example: linear array in Figure 1. Thin gray line: reference pattern. Thick solid blue line: pattern synthesized using the proposed algorithm. Thick dotted blue line: pattern synthesized using the algorithm proposed in [7]. Dashed red line: local upper bound of constraint (6).

5.1. *First Example: Linear Array of Omnidirectional Elements*. The first numerical example refers to a phase-only synthesis problem for a linear array consisting of $N = 40$ elements equally spaced along the y -axis (Figure 1). The interelement distance is $d = \lambda/2$. The element patterns $p_n(\phi)$ are assumed to be equal and omnidirectional on the xy -plane (this can be the case of an array consisting of half-wavelength dipoles parallel to the z -axis), that is, $p_n(\phi) = 1$ for $n = 1, \dots, N$ and $\phi \in [-\pi, \pi]$. The reference pattern, which is represented by the gray line in Figure 2, is produced by this array with uniform excitations, that is, by setting, in (2), $i_n = 1$ and $y_n = n d$ for $n = 1, \dots, N$. An interferer is assumed to act very close to the pointing direction ($\phi_0 = 0^\circ$) and is characterized by the truncated Gaussian pdf in (3) with $\bar{\phi} = 11^\circ$ and $\sigma = 9^\circ$. The solving procedure described in the previous section is applied adopting $M = 7$ null directions evaluated by (10).

The obtained pattern is represented in Figure 2 by the blue thick line, while the dotted blue line represents the pattern synthesized by the algorithm in [7], and the dashed red line identifies the upper bound corresponding to the pdf of the Gaussian interferer. An enlargement of the synthesized patterns in the null region is reported in Figure 3. This first example shows that, with respect to the starting reference pattern, approximately a 10 dB reduction in the region of interest is achieved by the proposed method. Besides, this reduction is obtained by properly shaping the wide null region according to the desired Gaussian pdf, thus confirming the suitability of the replacement of the original constraint in (6) with the M null one in (7). Of course, this reduction (as any imposed constraint) may cause

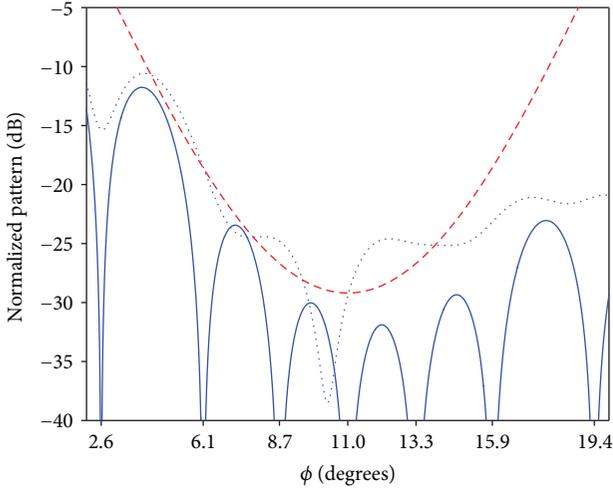


FIGURE 3: Magnification of the pattern synthesized in Figure 2 in the interference region with the evaluated null positions. To be noted that the constraint (6) is satisfied rigorously with the presented procedure and approximately with the method proposed in [7]. Also, it is to be noted that the pattern amplitude on the $M = 7$ null directions evaluated according to (10) is extremely low.

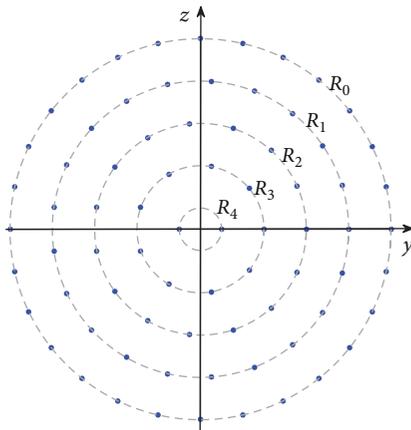


FIGURE 4: Second example: geometry of the circular array with five rings of $N = 75 \cos \phi$ elements, represented by the blue dots on the yz -plane.

modifications in other angular regions of the reference pattern. In particular, in this first example, the most significant changes for the proposed method with respect to the reference pattern are a slight increase of the side lobe between the main beam and the imposed Gaussian null (which is not surprising) and an increase of the side lobes in the intervals $[60^\circ, 80^\circ]$, both balanced by a decrease of the side lobes in the opposite directions. However, we may notice that the pattern modification obtained with the method in [7] is much more relevant on the entire angular domain. It is also worth to observe that the synthesis of the final pattern has required slightly less than half second of computational time, corresponding to just 22 iterations with the proposed algorithm here, and 992 s, corresponding to 1000 iterations, with the method in [7].

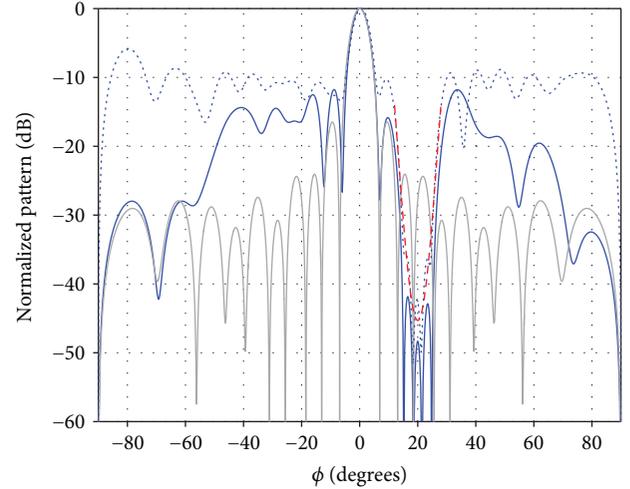


FIGURE 5: Second example: circular array in Figure 4. Thin gray line: reference pattern. Thick solid blue line: pattern synthesized using the proposed algorithm. Thick dotted blue line: pattern synthesized using the algorithm proposed in [7]. Dashed red line: local upper bound of constraint (6).

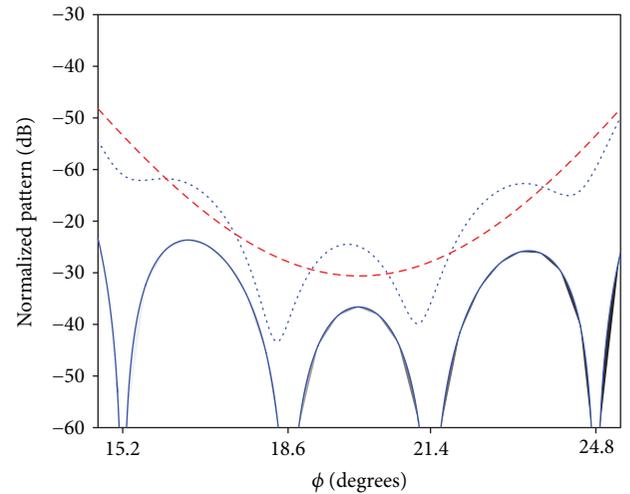


FIGURE 6: Magnification of the synthesized pattern in Figure 5 in the interference region with the evaluated null positions. Also, in this case, the original constraint (6) is satisfied rigorously by the pattern synthesized with the presented method and the pattern amplitude on the $M = 4$ null directions evaluated according to (10) is extremely low.

5.2. Second Example: Planar Circular Array with Multiple Rings and $\cos \Phi$ Element Patterns. The second numerical example deals with the phase-only synthesis of a Gaussian null region using a planar array involving $N_r = 5$ concentric rings lying on the yz -plane. The radius of the outermost ring is $R_0 = 4.5\lambda$. The other radii are $R_i = R_0 - i\lambda$ for $i = 1, 2, 3, 4$. On the i -th ring, N_i elements are equally spaced so as to give an interelement distance not lower than λ . Precisely, the rings consist of $N_0 = 28$, $N_1 = 21$, $N_2 = 15$, $N_3 = 9$, and $N_4 = 2$ elements, resulting in a total number $N = 75$ elements. The

TABLE 1: Comparison of the results.

		Iteration number	CPU time (s)	Constraint (6) satisfied
Example 1	Proposed algorithm	22	0.5	Rigorously
	Algorithm in [7]	1000	992	Approximately
Example 2	Proposed algorithm	75	0.7	Rigorously
	Algorithm in [7]	1000	2236	No

geometry of this planar array is shown in Figure 4. The element patterns are assumed to be equal to $p_n(\phi) = \cos \phi$ for $n = 1, \dots, N$ and $\phi \in [-\pi, \pi]$, which models the case of microstrip patches lying on the yz -plane. The reference pattern, corresponding to equal excitations, is still reported in gray in Figure 5, while the Gaussian interferer is characterized by a pdf with $\bar{\phi} = 20^\circ$ and $\sigma = 7^\circ$. The results are derived by adopting $M = 4$ null directions in (10).

The synthesized pattern is denoted in Figure 5 by the blue thick line, while the dotted blue line represents the pattern synthesized by the algorithm in [7], and the dashed red line represents the truncated Gaussian pdf. The magnification of the null region is shown in Figure 6. This second example confirms the satisfactory results provided by the proposed algorithm, also in comparison with the method in [7]. In particular, in this case, with respect to the starting pattern, a 20 dB reduction is achieved, while properly shaping the region of interest. Moreover, similarly to the previous example, also for this planar geometry, the computational time remained lower than one second (precisely 0.7 s) with the proposed algorithm, whereas 2236 s (1000 iterations) were necessary with the method in [7]. Table 1 finally summarizes the computational times and the satisfied constraints for the two compared methods.

6. Conclusions

An efficient technique has been proposed for the phase-only synthesis of antenna arrays with wide null regions characterized by a Gaussian shape. The synthesis algorithm has been derived by exploiting the alternating projection approach, in which all the operations are carried out in closed form. Thus, despite the iterative nature of the alternating projections, the overall synthesis procedure results extremely fast, so that the proposed algorithm can be suitable for 5G beamforming applications, where fast synthesis algorithms are recommended. Furthermore, the degrees of freedom of the problem seem to be exploited quite satisfactorily. In fact, the phase-only synthesis has been realized by moving from a uniform amplitude distribution of the array excitations, thus modifying just the phases of the element excitations. This results in a very simple feeding network, requiring only phase shifters, which are cheap and fast components. The effectiveness of the developed synthesis strategy has been verified by numerical examples involving linear and planar arrays, which have proved the low computational time required by the developed algorithm and its significant performance in terms of null region shaping and deepening.

Data Availability

All the data necessary to obtain the results presented in Section 5 (Numerical Results) can be found by the reader in the manuscript.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

- [1] J. G. Andrews, S. Buzzi, W. Choi et al., "What will 5G be?," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [2] M. R. Akdeniz, Y. Liu, M. K. Samimi et al., "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1164–1179, 2014.
- [3] J. F. Monserrat, S. Inca, J. Calabuig, and D. Martín-Sacristán, "Map-based channel model for urban macrocell propagation scenarios," *International Journal of Antennas and Propagation*, vol. 2015, Article ID 172501, 5 pages, 2015.
- [4] H. Jung and I.-H. Lee, "Connectivity analysis of millimeter-wave device-to-device networks with blockage," *International Journal of Antennas and Propagation*, vol. 2016, Article ID 7939671, 9 pages, 2016.
- [5] S. Singh, R. Mudumbai, and U. Madhow, "Interference analysis for highly directional 60-GHz mesh networks: the case for rethinking medium access control," *IEEE/ACM Transactions on Networking*, vol. 19, no. 5, pp. 1513–1527, 2011.
- [6] H. Steyskal, R. Shore, and R. Haupt, "Methods for null control and their effects on the radiation pattern," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 404–409, 1986.
- [7] A. D. Khzmalyan and A. S. Kondrat'yev, "Fast iterative methods for phase-only synthesis of antenna array pattern nulls," *Electronics Letters*, vol. 31, no. 8, pp. 601–602, 1995.
- [8] R. L. Haupt, "Phase-only adaptive nulling with a genetic algorithm," *IEEE Transactions on Antennas and Propagation*, vol. 45, no. 6, pp. 1009–1015, 1997.
- [9] M. Mouhamadou, P. Vaudon, and M. Rammal, "Smart antenna array patterns synthesis: null steering and multi-user beamforming by phase control," *Progress In Electromagnetics Research*, vol. 60, pp. 95–106, 2006.
- [10] M. Comisso and R. Vescovo, "Exploitation of spatial channel model for antenna array synthesis," *Electronics Letters*, vol. 42, no. 19, pp. 1079–1080, 2006.
- [11] R. Ghayoula, N. Fadlallah, A. Gharsallah, and M. Rammal, "Phase-only adaptive nulling with neural networks for antenna

- array synthesis,” *IET Microwaves, Antennas & Propagation*, vol. 3, no. 1, pp. 154–163, 2009.
- [12] H. Rammal, C. Olleik, K. Sabbah, M. Rammal, and P. Vaudon, “Synthesis of phased cylindrical arc antenna arrays,” *International Journal of Antennas and Propagation*, vol. 2009, Article ID 691625, 5 pages, 2009.
- [13] K. Guney, A. Durmus, and S. Basbug, “Antenna array synthesis and failure correction using differential search algorithm,” *International Journal of Antennas and Propagation*, vol. 2014, Article ID 276754, 8 pages, 2014.
- [14] K. Guney and A. Durmus, “Pattern nulling of linear antenna arrays using backtracking search optimization algorithm,” *International Journal of Antennas and Propagation*, vol. 2015, Article ID 713080, 10 pages, 2015.
- [15] Y. Han and C. Wan, “Scalable alternating projection and proximal splitting for array pattern synthesis,” *International Journal of Antennas and Propagation*, vol. 2015, Article ID 915293, 13 pages, 2015.
- [16] T. van Luyen and T. Vu Bang Giang, “Interference suppression of ULA antennas by phase-only control using bat algorithm,” *IEEE Antennas and Wireless Propagation Letters*, vol. 16, pp. 3038–3042, 2017.
- [17] R. Vescovo, “Reconfigurability and beam scanning with phase-only control for antenna arrays,” *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 6, pp. 1555–1565, 2008.
- [18] M. Comisso, G. Buttazzoni, and R. Vescovo, “Reconfigurable antenna arrays with multiple requirements: a versatile 3D approach,” *International Journal of Antennas and Propagation*, vol. 2017, Article ID 6752108, 9 pages, 2017.
- [19] R. Vescovo, “Consistency of constraints on nulls and on dynamic range ratio in pattern synthesis for antenna arrays,” *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 10, pp. 2662–2670, 2007.
- [20] M. K. Samimi and T. S. Rappaport, “3-D millimeter-wave statistical channel model for 5G wireless system design,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 7, pp. 2207–2225, 2016.
- [21] C. Ling, X. Yin, R. Müller et al., “Double-directional dual-polarimetric cluster-based characterization of 70–77 GHz indoor channels,” *IEEE Transactions on Antennas and Propagation*, vol. 66, no. 2, pp. 857–870, 2018.
- [22] G. Buttazzoni and R. Vescovo, “Density tapering of linear arrays radiating pencil beams: a new extremely fast Gaussian approach,” *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 7372–7377, 2017.
- [23] G. Buttazzoni and R. Vescovo, “Power synthesis for reconfigurable arrays by phase-only control with simultaneous dynamic range ratio and near-field reduction,” *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 2, pp. 1161–1165, 2012.
- [24] R. Vescovo, “Power pattern synthesis for antenna arrays with null constraints in the near-field region,” *Microwave and Optical Technology Letters*, vol. 44, no. 6, pp. 542–545, 2005.

Research Article

Large-Scale Channel Modeling and Measurements for 10GHz in Indoor Environments

Iury S. Batalha , **Andréia V. R. Lopes**, **Jasmine P. L. Araújo**, **Fabrcio J. B. Barros**,
Bruno L. S. Castro, **Gervásio P. S. Cavalcante**, and **Evaldo G. Pelaes**

LCT, UFPA, Belem 66045-110, Brazil

Correspondence should be addressed to Iury S. Batalha; iurybatalha@gmail.com

Received 31 August 2018; Accepted 11 November 2018; Published 23 January 2019

Guest Editor: Peter Excell

Copyright © 2019 Iury S. Batalha et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the advent of 5G mobile communication and researches into the propagation of large-scale channel modeling for frequencies above 6 GHz, measurement investigation was performed at 10 GHz with horn-type directional antennas in a corridor and a computer room within the Electrical and Computer Engineering Laboratories' first floor, at Federal University of Pará (UFPA), Brazil. This paper presents data obtained through experimental work, channel modeling with co-polarization V-V and H-H and cross-polarization V-H in line-of-sight (LOS) or non-line-of-sight (NLOS) conditions. The large-scale close-in reference is sustained by a comprehensive analysis, considering propagation mechanisms such as reflection and diffraction. Results demonstrate that the established model had inferior standard deviation in relation to measured data, proving itself more significant to propagation in indoor environments.

1. Introduction

The considerable quantity of bandwidth available in frequency bands above 6 GHz is an attractive resource to provide multi-Gigabit per second (Gbps) data rates. Such resource would alleviate the traffic flow of mobile data in lower frequency bands, such as those below 6 GHz [1, 2]. With the advent of the new generation of mobile communication (5G), there is a great deal of research taking place into the development to provide recommendations. One of its branches of study lies within the millimetric waves, resulting in works above 6 GHz [3–6]. Frequencies within 3 to 30 GHz (SHF: Super High Frequency) and 30 to 300 (EHF: Extremely High Frequency) bands present similar propagation characteristics, classifying them as millimetric wave (mmWave) bands [4, 5]. Therefore, it is fundamental to know the mmWave channel's propagation characteristics for the development of fifth generation mobile and wireless communication systems—5G [7].

These channel propagation characteristics can be defined through path loss models, in large scale, that predict the propagation signal's attenuation according to distance. They

are important for developing more efficient communication systems, to optimize the positioning of transmitting antennas and assist on future telecommunication system projects [8].

For indoor environments, there are several structural questions that influence signal behavior, such as construction materials utilized in the building, size of building/rooms/corridors, number of people circulating inside the room, furniture types and placement, and interference with other systems. They confine the waves within the environment, leading to more reflective components and multiple paths for the signal to propagate, as well as crossing of walls (diffraction) and other types of obstacles [9].

A wireless system requires extensive research and comprehension of its propagation channels. Even though a large number of investigations have been undertaken into channels below 6 GHz, there are campaigns aiming at measuring and modeling of 5G channels at 10 to 28 GHz and 30 to 72 GHz bands [5, 10–15], presented by institutes such as New York University (NYU) and the Mobile and Wireless Community Enablers for the Twenty-Twenty Information Society (METIS)—the former created with the sole purpose to standardize 5G systems [16, 17]. The METIS report

identifies probable bands for 5G application; 10 GHz is listed as a promising frequency band.

As an example of a study based in measuring frequencies above 6 GHz, Deng et al. developed models for diffraction and signal strength measurements around sharp-edged objects such as corners, pillars, and irregular surfaces, for the 10, 20, and 26 GHz spectrum. The diffraction measurements were made in closed and exterior environments, using a continuous wave (CW) probe with three pairs of identical horn-type directional antennas at the transmitter and receptor [18].

In 2014, Kim et al. studied large-scale parameters based on various indoor environments at 11 GHz, comprising path-loss polarization characteristics, shadowing, cross-polarization power ratio, delay spread, and coherence bandwidth [19].

Reference [20] presents radio channel characteristics on carrier frequencies 11.2 and 14.6 GHz, with a 1 GHz bandwidth in environments simulated as a corridor, a laboratory, offices, and a conference room inside a building in Beijing, China. There, investigations regarding large- and small-scale fading properties were conducted, based on realistic measurements. Propagation and delay correlations were also discussed.

Ling et al. produced a measurement campaign of channels implementing directional sweeping to capture temporal and spatial propagation characteristics for upper frequencies, from 13 to 17 GHz, in classroom environments [21].

In [13], there are large-scale models and measuring of frequencies 14 and 22 GHz within line-of-sight (LOS) and non-line-of-sight (NLOS) conditions. A dual-slope model is suggested to distinguish the channel—its validity is proven through the close-in free space reference model, and its parameters defined by MMSE (Minimum Mean Square Error). Another approach is a model inspired in waveguide measurements.

Properties of a channel at the 20 GHz band for an office environment are shown in [22]. Measurements were made using a channel probe at the 20 GHz spectrum, with an orthogonal frequency division multiplexing signal and a bandwidth of 50 MHz. Characteristics of omnidirectional antennas were obtained by turning directional antennas through the azimuthal axis and elevated to the receptor's side.

Reference [23] exposes several large-scale measurements utilizing high-gain directional antennas at 28 and 73 GHz. A model that simulates an omnidirectional antenna in which high-gain antennas were rotated a whole 360 degrees in both Tx and Rx is also described in this paper. This took place in offices with line-of-sight (LOS) and non-line-of-sight (NLOS), with copolarization (V-V) and cross-polarization (V-H). For large-scale models, templates were utilized, such as CI, CIX, CIF, CIFX, FI, ABG, and ABGX, in which the path loss exponent (n) is defined through MMSE (Minimum Mean Square Error). The model's random variable has an average value of zero, and standard deviation of measured data is given in decibels. Results show that the models have great accuracy when compared to measured data.

Lei et al. have proposed a measurement system at 28 GHz in 3 different scenarios, including an office, a

corridor, and a generic hall, using 26 dBi directional antennas and a VNA to analyze the channel at a maximum distance of 30 meters. The three properties which were analyzed are path loss, RMS delay spread, and the power angular profiles (PAPs). Results illustrate that indoor environments can enhance the received signal power in the case of LOS. However, in case of NLOS, the penetration loss caused by walls and doors can bring considerable attenuation, which implies that minor cells will have an important role increasing the probability of LOS links for future communication systems. Multipath components (MPCs) can be detected from many directions, albeit needing to use high-directivity antennas [24].

MacCartney et al. compiled a comprehensive study on path loss models of simple or multifrequency mmWave, based on layouts for polarization of separate antennas, and combined in LOS and NLOS environments for three common office layouts. The frequency channels used are 28 GHz and 73 GHz [23].

In [25], Geng et al. made propagation measurements at 60 GHz in several indoor environments, with continuous route (CR) and direction-of-arrival (DOA) measurement campaigns. The propagation mechanisms were studied based on DOA data, indicating that direct waves and the first-order reflected waves on smooth surfaces were sufficient in a LOS environment. Whereas in NLOS cases, diffraction was highly significant, and the propagation loss through walls was very high.

The objective of our study is to present a statistical large-scale analysis of measured data in indoor environments at the 10 GHz spectrum. Signal behavior for LOS and NLOS is specified in two different rooms, with antenna copolarization (V-V and H-H) and cross-polarization (V-H).

The paper is organized as follows. Section 2 presents all measurement equipment utilized during the campaigns, the mounted scenarios, and all measurement procedures; Section 3 elucidates the modeling analysis, the proposed large-scale model for path loss and the obtained results in comparison to other models found in the literature; Section 4 provides discussions and conclusions on the obtained data.

2. Measurement Investigations

This section describes the transmitting and receiving equipment, the scenarios considered for the measurement campaigns, and line-of-sight (LOS)/non-line-of-sight (NLOS) measurements. It also contains information about copolarization (V-V) and (H-H) and cross-polarization for both studied environments.

2.1. Equipment and Setup. Two identical horn-type directional antennas were utilized, fabricated by MCS Industries, with gain of 15 dBi, apertures of 29 degrees on the horizontal plain and 29.3 degrees on the vertical plain, and 1.7 meters above the floor. For the transmitting equipment (Tx), a continuous wave (CW) signal generator (Hewlett Packard® Synthesized Sweeper 83752A) was chosen, sweeping frequencies from 1 to 20 GHz. The Tx was fixed, the environment, remaining immobile, and configured with transmission

powers of 0 dBm for V-V and H-H and 15 dBm for V-H. The selected receptor (Rx), however, was an Anritsu® Signal Analyzer MS2692A.

For the first iteration, Tx and Rx were positioned with a distance of 1 meter from each other. And for the subsequent iterations, the distance between Tx and Rx is increased by another 1 meter. It is important to note that these experiments occurred with no objects or persons obstructing the transmitted beam. Table 1 provides specifications of equipment and parameter configurations for the campaigns.

2.2. Corridor. The corridor, depicted in Figure 1, has dimensions of 15.22 m × 1.37 m. It is built out of brick walls and smooth concrete on the floor, also containing wooden doors, metallic grilles, and containers with power supplies and circuit breakers. The transmitter Tx was fixed at a distance of 1.3 meters from the windowed wall, whereas Rx was progressively moved 1 meter away from Tx for every iteration of the experiment.

For the corridor, approximately 10 power values were measured for each of the 14 measuring points. For each one, 10,001 power data were collected for each meter variation. The procedure was repeated 10 times for each point, resulting in 100,010 power data for each iteration (10 × 10,001). Therefore, 1,400,140 power data were collected at the corridor environment (100,010 × 14 points).

The quantity of measured points by polarization type is distinct for each one. For vertical polarization (V-V) all 14 points were measured. For horizontal polarization (H-H) only 10 measured points were used, and with cross-polarization (V-H) only 8 points were measured. The amount of measured points for H-H and V-V was recorded for each distance until the signal strength reached noise level.

2.3. Computer Room. The computer room's dimensions are 14.97 m × 7.67 m. As configured in the corridor, Tx was placed in a fixed position 1.3 meters away from the wall, also moving Rx away from Tx by 1 meter every iteration. This room is constituted of brick/concrete walls and drywalls, containing glass windows, wood tables, metallic chairs, and evenly disposed desktops, as shown in Figure 2.

The laboratory was chosen because it presents different characteristics comparing with the corridor; the corridor presents a length much larger than the wavelength of the frequencies used, providing greater reflection in this environment; on the other hand the lab presents a smaller length and walls with different compositions; it is formed for masonry walls and wood, for the walls of masonry the reflections are close to that seen in the corridor, and the wooden ones there have less reflection compared to the masonry wall. Therefore, the laboratory is an interesting environment to determine the loss behavior because it is formed by different types of materials and that directly influence the behavior of the signal received in the Rx.

In this environment, a radial component was defined at 3.8 meters away from the lateral walls. In each radial, 10 power values for each of the 12 traced points were measured. Given that 100,010 power data are obtained in every iteration, as previously noted, the total amount of power data

extracted reaches 1,200,120. In contrast to the corridor, all 12 points distributed along the room were measured for all polarization configurations (V-V, H-H, and V-H).

For the measurement campaign utilizing copolarized antennas (V-V and H-H), horn-type directional antennas were also used, with 15 dBi gain on both Tx/Rx and a transmission power of 0 dBm. For cross-polarization, the gain is also 15 dBi on Tx/Rx; however, the transmission power is 15 dBm.

3. Large-Scale Path Loss Models

Large-scale path loss models predict loss according to the distance between the transmitter and the receptor. These models are important for wireless telecommunication projects [8], aiding the analysis of wave propagation through indoor environments, as there are multiple wave deviation events like multipath, reflection, diffraction, penetration, and shadowing—that is, attenuators of received power at Rx [26]. They also provide realistic propagation characteristics for a wireless channel [5, 10–28]. Most models assume loss in a logarithmic scale. Parameters for the models utilized in this study were defined through MMSE, and the close-in free space model (with its extensions) was utilized for approximation.

The close-in (CI) free space path loss model as seen in (1) and given in dB is uniquely parametrized by a sole variable known as the path loss exponent (PLE), or n . This CI model has a physical distance reference represented by d_0 . The loss model parameters are determined by MMSE, minimizing error between measured and simulated data. In this work, the distance reference is set to 1 meter, and the model can be applied to estimate path loss in both cross- and copolarization conditions.

$$PL^{CI}(f, d)[\text{dB}] = \text{FSPL}(f, d_0) + 10n_{V-V} \log_{10} \left(\frac{d}{d_0} \right) + X_{\sigma}^{CI}, \quad (1)$$

$$\text{FSPL}(f, d_0) = 10 \log_{10} \left(\frac{4\pi d_0}{\lambda} \right)^2, \quad (2)$$

where X_{σ}^{CI} is a random variable of the Gaussian type with an average value of zero and standard deviation in dB, n_{V-V} is the path loss exponent for copolarization and FSPL is the initial loss found in the model defined by (2). λ is the wavelength and d_0 is the previously explained reference distance, set to 1 m.

The CI model possesses an extension for the case of cross-polarization (3), adding a factor called cross-polarization discrimination (XPD). This parameter is defined via MMSE [28–30]. The equation then becomes

$$PL^{CIX}(f, d)[\text{dB}] = \text{FSPL}(f, d_0) + 10n_{V-H} \log_{10} \left(\frac{d}{d_0} \right) + X_{\sigma}^{CIX} + \text{XPD}[\text{dB}], \quad (3)$$

TABLE 1: Channel sounding system specifications for the 10 GHz indoor measurement campaign.

Parameters	Configurations	Units
Central frequency	10	GHz
Signal transmission	Continuous wave	—
Antennas	Horn-type, directional	—
Transmitted copolarization power	0	dBm
Transmitted cross-polarization power	15	dBm
Antenna height for Tx/Rx	1.7	m
Antenna gain for Tx/Rx	15	dBi
Azimuthal HPBW for Tx/RX	29	Degrees
HPBW elevation for Tx/Rx	29.3	Degrees

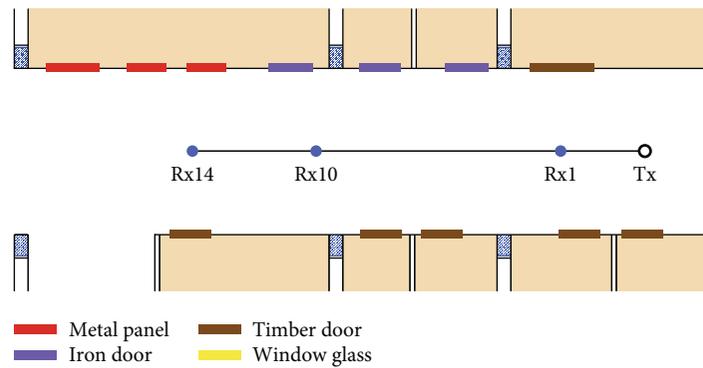


FIGURE 1: Distance for Tx as done.

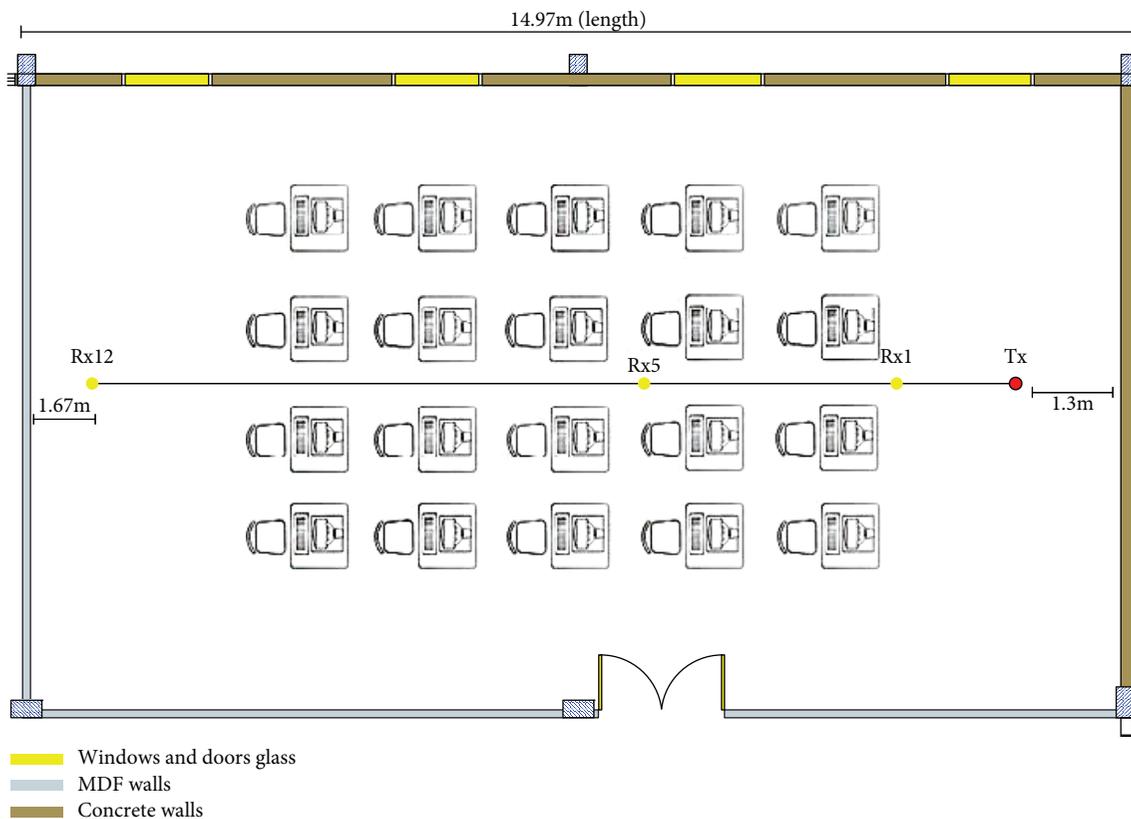


FIGURE 2: The computer room's blueprint.

in which XPD is represented in dB, and X_{σ}^{CIX} is a Gaussian random variable with average value of zero and standard deviation in dB. n used in this calculation has to be defined in the measured data in the cross-polarization configuration. Maccartney et al. [23] utilize $n_{\text{V-V}}$ for the model that considers cross-polarized antennas (CIX), in this paper, $n_{\text{(V-H)}}$ is used to adjust to a slope of the curve.

Another independent extension for the CI model is the case of copolarized H-H configuration, adding a path loss parameter for H-H copolarization (4). This parameter is similar to XPD and also defined through MMSE.

$$\text{PL}^{\text{CIHH}}(f, d)[\text{dB}] = \text{FSPL}(f, d_0) + 10n_{\text{H-H}} \log_{10}\left(\frac{d}{d_0}\right) + X_{\sigma}^{\text{CI}}. \quad (4)$$

In (4), $n_{\text{H-H}}$ represents the copolarized H-H path loss exponent (PLE), and X_{σ}^{CIHH} is a Gaussian random variable with average value of zero and standard deviation in dB.

The last proposed extension illustrated for the CI model is the obstacle path loss exponent (OPLE), in which an attenuation parameter signifying obstruction by same or distinct obstacles found in the propagation's path.

$$\text{PL}^{\text{CI}}(f, d)[\text{dB}] = \text{FSPL}(f, d_0) + 10n_{\text{v-v}} \log_{10}\left(\frac{d}{d_0}\right) + X_{\sigma}^{\text{CI}} + \sum_{i=1} \text{OPLE}_i(np_i), \quad (5)$$

in which OPLE_i is the obstacle path loss exponent, and i corresponds to the quantity of same-type obstacles. This parameter is based on the COST-231 Motley-Keenan model.

4. Modeling and Result Analysis

In this section, there is an analysis of signal attenuation behavior in relation to the measured path loss for both studied scenarios. The complexity of indoor environments generates multiple paths—reflections, diffractions, penetration effects, and shadowing—that have significant impact on received power [31]. Radio signal attenuation in waveguides (or for instance, tunnels and corridors) is generally much lower than in free space and is inversely proportional to frequency due to the wave orientation effect.

About the measured data, PLE, standard deviation, and XPD values were defined through MMSE. Table 2 presents the said values. Table 3 presents values for OPLE for different types of walls.

4.1. Corridor Analysis. Figure 3 presents data modeled through the CI model for V-V and H-H polarizations and CIX for V-H, in which the PLE is the gradient of the straight line that shows how the signal degrades with distance, also directly influencing the standard deviation values of measured data in the corridor. Figure 4 illustrates the shadowing effect defined by the random variable for directional antennas, where it is observed that the model has a good

TABLE 2: PLE values for each polarization in the corridor and in the laboratory.

Scenario	Polarization	PLE	σ (dB)	XPD (dB)
Laboratory	V-V	1.55	6.56	
	H-H	1.64	6.12	22.30
	V-H	0.08	4.23	
Corridor	V-V	1.96	5.6	
	H-H	1.81	5.48	21.30
	V-H	0.65	2.39	

TABLE 3: Loss per wall for 10 GHz indoor measurement campaign.

Frequency	Masonry	Wood	Glass
10 GHz	10 (dB)	2 (dB)	0 (dB)

approximation compared to measured data and checking the simulated signal's variability between Tx and Rx. Figure 5 shows point-to-point standard deviation data, verifying signal variation for each travelled meter in the environment for all 3 polarizations.

Figure 3 presents measured data for the 3 polarization types (V-V, H-H, and V-H), and for the close-in model with extensions for the 10 GHz band, data in blue and red represent the values for copolarized antennas, in which PLE values are very close to 1.96 dB and 1.81 dB for V-V and H-H configurations, respectively (the major difference between them is the initial loss). In Figure 3 the data depicted in green are the cross-polarized measured results, where initial loss is substantially higher and lower PLE is observed compared to copolarization.

The channel's variability is represented by a Gaussian random variable, influenced by the shadowing effect and defined for the close-in free space reference path loss model (1). Figure 4 shows the shadowing effect defined by the random variable for directional copolarized antennas (V-V) in the corridor at 10 GHz. It is possible to denote that the loss model (1) has a very good approximation in relation to measured data, checking the simulated signal's variation between Tx and Rx.

An important approach to measured data is the statistical survey with standard deviation. For this purpose, a point-to-point standard deviation was calculated to verify signal variation for each travelled meter between Tx and Rx for both environments utilizing copolarized and cross-polarized configurations. Figure 5 illustrates point-to-point standard deviation simultaneously for all polarizations.

For the data in red in Figure 5, it is possible to observe a low variability in standard deviation values (on the y -axis) with values between 0 and 0.7 dB. This indicates that there is low variability of received power values for a same point utilizing horn-type V-V copolarized antennas in the corridor. In black, standard deviation values for H-H copolarization are shown, presenting a variability equivalent to V-V configuration and confirming that signal variation is also low in this case. The same approach has been applied for cross-polarization data in the corridor (blue-colored triangles), however yielding higher signal variability ranging from 0 to 3.5 dB.

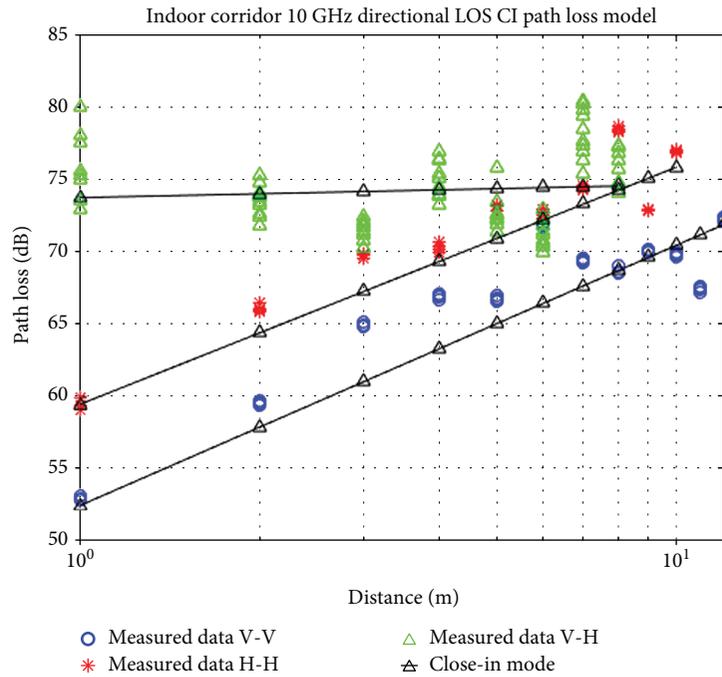


FIGURE 3: 10 GHz directional LOS large-scale CI path loss model with copolarization V-V, H-H, and V-H in the corridor.

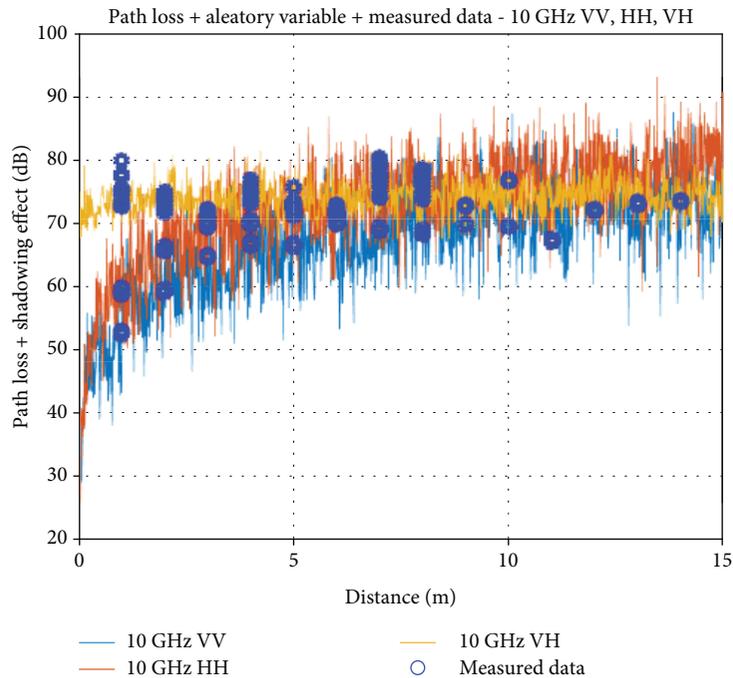


FIGURE 4: Path loss model CI with shadowing effect at measured data for 10 GHz band in corridor scenario.

It is important to acknowledge that the whole behavior's standard deviation presented higher values for copolarization (V-V and H-H) and lower ones for cross-polarization (V-H). This is directly linked to PLE, in which V-H antennas had PLE values below 1, whilst copolarized V-V and H-H presented values above 1. Consequently, point-to-point standard

deviation analysis is important, as it reveals how much the received power varies through measured points.

4.2. Computer Room Analysis. Using the same approach as for the corridor of analysis of the corridor, the computer room measurements have modeled data with CI and CIX

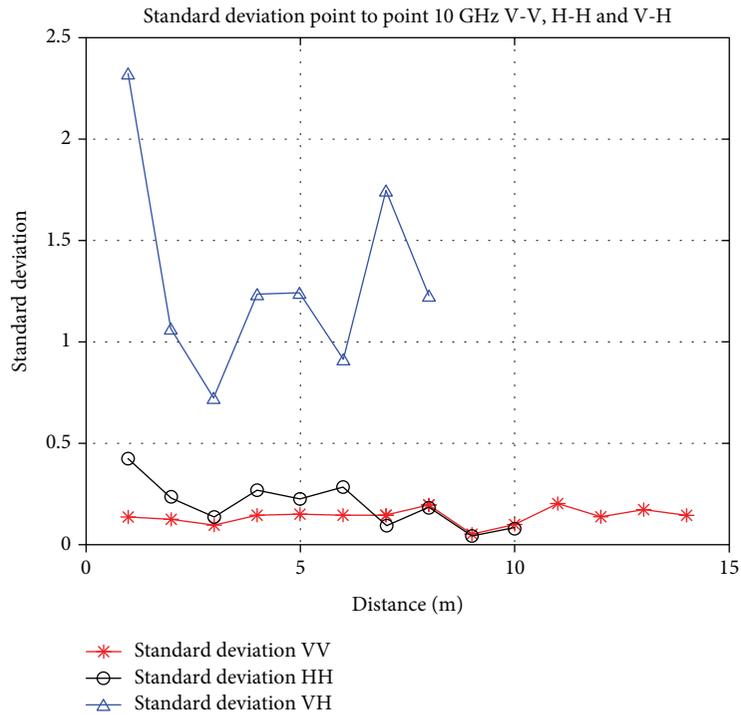


FIGURE 5: Standard deviation point-to-point data measured in the corridor to V-V, H-H, and V-H.

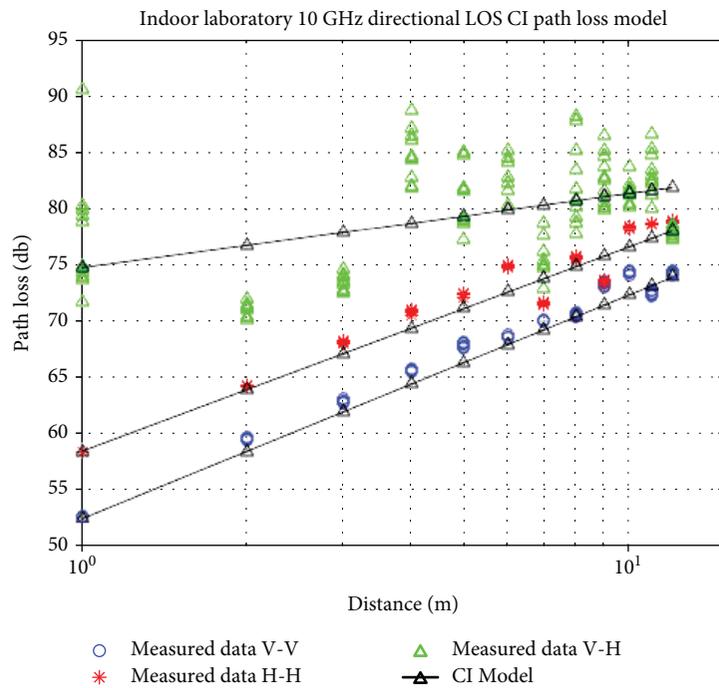


FIGURE 6: 10 GHz directional LOS large-scale CI path loss model with copolarization V-V, H-H, and V-H in the computer room.

models in logarithmic scale with all polarizations. The PLE also represents the slope's inclination and demonstrates how the signal attenuates with distance and directly influences standard deviation values measured in the computer room.

Figure 6 depicts measured data for all three polarizations and the close-in model with its extensions for 10GHz propagation. Data in blue and red present data for copolarized antennas in which PLE values are close to 1.55 dB and 1.64 dB for V-V and H-H, respectively. In

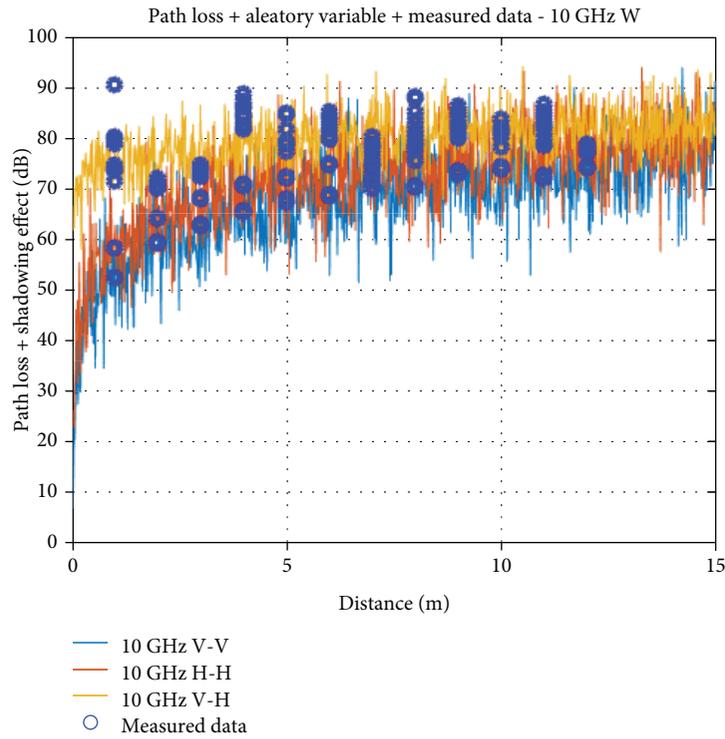


FIGURE 7: Path loss model CI with shadowing effect at measured data for 10 GHz band in laboratory scenario.

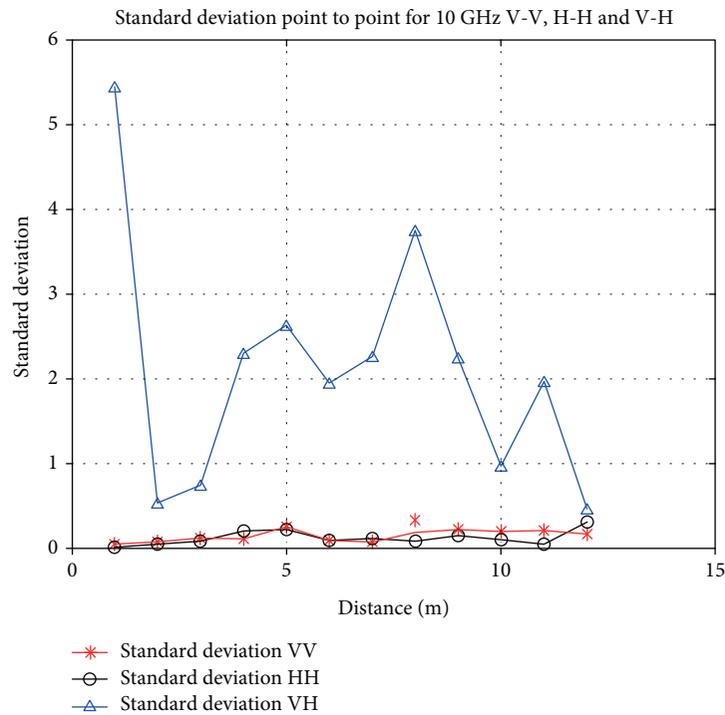


FIGURE 8: Standard deviation for point-to-point data measured in the computer room to polarizations V-V, H-H, and V-H.

Figure 6 again, the difference is mostly noted on initial losses, for the cross-polarized antenna data set, shown in green, PLE values are lower and the initial losses higher in relation to the copolarized counterparts.

The shadowing effect defined by the random variable for directional antennas and signal variability between Tx and Rx is presented in Figure 7; Figure 8 shows point-to-point measured standard deviation data for signal variation

checking in each meter travelled in the environment for all three polarizations, respectively.

The analysis of the data measured by the standard deviation point-to-point is made with the objective of verifying the variation of the signal for each point measured in all the polarizations. Figure 8 depicts standard variation simultaneously for V-V, H-H, and V-H configurations. The red set of data presents the point-to-point approach at 10 GHz with V-V polarization, and the minimum/maximum observed values in the Y axis range from 0.04 to 0.26 dB. This indicates that there is low variability on received power values for the same point whilst using horn-type V-V copolarized antennas in the computer room. For the black points in the graph, it is the standard deviation for H-H configuration, presenting variability equivalent to V-V, elucidating that measured data vary very little on a point-to-point basis with minimum/maximum values of approximately 0.04–1.05 dB. For the cross-polarization data in the computer room (in green), the standard deviation varies from 0.5 and 5.5 dB, which is evidently greater than the values for copolarized antennas.

5. Conclusions

This paper had the objective to demonstrate channel modeling with path loss model close-in and its extensions, to characterize future wireless network channels. The close-in model presents good approximations in LOS and NLOS conditions with directional horn antennas in cross- and copolarization for the corridor and the computer room at 10 GHz, one of METIS-determined frequencies. PLE values in both environments utilizing directional V-V antennas are close to the ones seen in the literature [14, 27]; however, the values for cross-polarized (V-H) configurations are less than the copolarized values. Another important remark is the analysis of the standard deviation, both in average value or point-to-point. For the average standard deviation, the values favor copolarization (V-V and H-H) in comparison to cross-polarization (V-H) due to low PLE value, as shown in Table 2. When the analysis is point-to-point, the copolarizations have less variation in relation to cross-polarization.

On the point-to-point standard deviation for the corridor environment, it was verified that copolarized antennas V-V and H-H provide low variation, between 0.04 and 1.05 dB, as depicted in Figure 5. For cross-polarization there is more variability, in values ranging from 0 to 3.5 dB, also shown in Figure 5. For the measurements made in the computer room, it is valid to pinpoint that the H-H polarization's standard deviation yields more variability than its V-V counterpart. In a point-to-point analysis, the V-V configuration varies between 0.04 and 0.26 dB, and H-H gives a standard deviation ranging from 0.04 to 1.05 dB, as evident in Figure 8.

The walls also had different influences in each obstacle—its attenuation changes according to the used frequency, being directly proportional to the suffered loss. This is an important point, and future access technologies must work around this problem, given that the new generations of

telecommunication systems are bound to evolve to mmWave bands.

Another significant point is the difference in wave propagation within distinct indoor environments. Inside the corridor, wave reflections are heightened due to the proximity of the walls, the material used inside the walls, floor, and ceiling, in addition to grilles and metallic panels. The computer room is a larger environment, with a different composition. It contains drywall and masonry walls as well as chairs, tables, and desktops. These characteristics affect the path loss of electromagnetic waves inside these environments, as this paper displayed its distinctions through the close-in free space loss models and its extensions beyond the standard deviation.

Future works will be carried out in this area, such as large-scale modeling for adjacent frequencies to 10 GHz, such as 8 GHz, 9 GHz, 11 GHz, and 12 GHz, for the same environments cited in this work using different measurement, with the transmitter (Tx) with different heights, with 360° Rx rotation to collect multipath signals using co- and cross-polarization antennas. Moreover, we will carry out small-scale modeling for the 8 to 12 GHz frequency band. Finally, a large-scale and small-scale modeling will be executed for millimeter wave bands in the 28 GHz to 93 GHz band indoors.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

We would like to thank the Electrical Engineering Post-Graduation Program (PPGEE) at Federal University of Pará (UFPA) for giving the opportunity and materials to produce this study, the National Council for Scientific and Technological Development (CNPq) for the financing, and the National Institute of Science and Technology (INCT) for its support on the research.

References

- [1] A. Ghosh, T. A. Thomas, M. C. Cudak et al., "Millimeter-wave enhanced local area systems: a high-data-rate approach for future wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1152–1163, 2014.
- [2] G. R. MacCartney and T. S. Rappaport, "Study on 3GPP rural macrocell path loss models for millimeter wave wireless communications," in *2017 IEEE International Conference on Communications (ICC)*, pp. 1–7, Paris, France, July 2017.
- [3] T. S. Rappaport, S. Sun, R. Mayzus et al., "Millimeter wave mobile communications for 5G cellular: it will work!," *IEEE Access*, vol. 1, pp. 335–349, 2013.

- [4] Z. Pi and F. Khan, "An introduction to millimeter-wave mobile broadband systems," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 101–107, 2011.
- [5] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband millimeter-wave propagation measurements and channel models for future wireless communication system design," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3029–3056, 2015.
- [6] METIS, *Deliverable D1.4, METIS Channel Models*, 2015.
- [7] J. G. Andrews, S. Buzzi, W. Choi et al., "What will 5G be?," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [8] M. F. Iskander and Zhengqing Yun, "Propagation prediction models for wireless communication systems," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 3, pp. 662–673, 2002.
- [9] H. K. Rath, S. Timmadasari, B. Panigrahi, and A. Simha, "Realistic indoor path loss modeling for regular WiFi operations in India," in *Twenty-third National Conference on Communications (NCC)*, pp. 1–6, Chennai, India, 2017.
- [10] J. Huang, C. X. Wang, R. Feng, J. Sun, W. Zhang, and Y. Yang, "Multi-frequency mmWave massive MIMO channel measurements and characterization for 5G wireless communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 7, pp. 1591–1605, 2017.
- [11] K. Saito, J. I. Takada, and M. Kim, "Characteristics evaluation of dense multipath component in 11GHz-band indoor environment," in *2016 10th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–3, Davos, Switzerland, 2016.
- [12] C. Ling, X. Yin, H. Wang, and R. S. Thomä, "Experimental characterization and multipath cluster modeling for 13–17 GHz indoor propagation channels," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 12, pp. 6549–6561, 2017.
- [13] N. O. Oyie and T. J. O. Afullo, "Measurements and analysis of large-scale path loss model at 14 and 22 GHz in indoor corridor," *IEEE Access*, vol. 6, pp. 17205–17214, 2018.
- [14] Q. Liao, Z. Ying, and C. Gustafson, "Simulations and measurements of 15 and 28 GHz indoor channels with different array configurations," in *2017 International Workshop on Antenna Technology: Small Antennas, Innovative Structures, and Applications (iWAT)*, pp. 256–259, Athens, Greece, 2017.
- [15] C. Ling, X. Yin, H. Wang, and X. Zhang, "Comparison of parametric and nonparametric characterization of 15 GHz propagation channels in indoor environments," in *2016 10th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–5, Davos, Switzerland, 2016.
- [16] H. Tullberg, P. Popovski, Z. Li et al., "The METIS 5G system concept: meeting the 5G requirements," *IEEE Communications Magazine*, vol. 54, no. 12, pp. 132–139, 2016.
- [17] H. Droste, G. Zimmermann, M. Stamatelatos et al., "The METIS 5G architecture: a summary of METIS work on 5G architectures," in *2015 IEEE 81st Vehicular Technology Conference (VTC Spring)*, pp. 1–5, Glasgow, Scotland, 2015.
- [18] S. Deng, G. R. MacCartney, and T. S. Rappaport, "Indoor and outdoor 5G diffraction measurements and models at 10, 20, and 26 GHz," in *2016 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–7, Washington, DC, USA, 2016.
- [19] M. Kim, Y. Konishi, Y. Chang, and J. Takada, "Large scale parameters and double-directional characterization of indoor wideband radio multipath channels at 11 GHz," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 1, pp. 430–441, 2014.
- [20] X. Zhou, Z. Zhong, X. Bian et al., "Indoor wideband channel measurements and analysis at 11 and 14 GHz," *IET Microwaves, Antennas & Propagation*, vol. 11, no. 10, pp. 1393–1400, 2017.
- [21] C. Ling, X. Yin, H. Wang, and X. Zhang, "Comparison of characteristics of 13–17 GHz propagation channels in indoor environments with different measurement configurations," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 1–5, Valencia, Spain, 2016.
- [22] N. Tran, T. Imai, and Y. Okumura, "Measurement of indoor channel characteristics at 20 GHz band," in *2015 International Symposium on Antennas and Propagation (ISAP)*, pp. 1–4, Hobart, TAS, Australia, 2015.
- [23] G. R. Maccartney, T. S. Rappaport, S. Sun, and S. Deng, "Indoor office wideband millimeter-wave propagation measurements and channel models at 28 and 73 GHz for ultra-dense 5G wireless networks," *IEEE Access*, vol. 3, pp. 2388–2424, 2015.
- [24] M. Lei, J. Zhang, T. Lei, and D. Du, "28-GHz indoor channel measurements and analysis of propagation characteristics," in *2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*, pp. 208–212, Washington, DC, USA, 2014.
- [25] S. Geng, J. Kivinen, X. Zhao, and P. Vainikainen, "Millimeter-wave propagation channel characterization for short-range wireless communications," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 1, pp. 3–13, 2009.
- [26] A. Hrovat, G. Kandus, and T. Javornik, "Path loss analyses in tunnels and underground corridors," *International Journal of Communication*, vol. 6, no. 3, pp. 136–144, 2012.
- [27] "Mobile and wireless communications enablers for the twenty-twenty information society (METIS): METIS channel models," July 2017, https://www.metis2020.com/wp-content/uploads/METIS_D1.4_v3.pdf.
- [28] T. S. Rappaport and D. A. Hawbaker, "Wide-band microwave propagation parameters using circular and linear polarized antennas for indoor wireless channels," *IEEE Transactions on Communications*, vol. 40, no. 2, pp. 240–245, 1992.
- [29] C. M. P. Ho, T. S. Rappaport, and M. P. Koushik, "Antenna effects on indoor obstructed wireless channels and a deterministic image-based wide-band propagation model for in-building personal communication systems," *International Journal of Wireless Information Networks*, vol. 1, no. 1, pp. 61–76, 1994.
- [30] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice-Hall, Upper Saddle River, NJ, USA, 2nd edition, 2002.
- [31] A. Hrovat, T. Javornik, and K. Guan, "Analysis of radio wave propagation at millimeter-wave band in tunnels for 5G communications," in *2016 22nd International Conference on Applied Electromagnetics and Communications (ICECOM)*, pp. 1–5, Dubrovnik, Croatia, 2016.